Physical Validation of TRMM TMI and PR Monthly Rain Products Over Oklahoma

Brad L. Fisher

Summary

The Tropical Rainfall Measuring Mission (TRMM) provides monthly rainfall estimates using data collected by the TRMM satellite. These estimates cover a substantial fraction of the earth's surface. The physical validation of TRMM estimates involves corroborating the accuracy of spaceborne estimates of areal rainfall by inferring errors and biases from ground-based rain estimates. The TRMM error budget consists of two major sources of error: retrieval and sampling. Sampling errors are intrinsic to the process of estimating monthly rainfall and occur because the satellite extrapolates monthly rainfall from a small subset of measurements collected only during satellite overpasses. Retrieval errors, on the other hand, are related to the process of collecting measurements while the satellite is overhead. One of the big challenges confronting the TRMM validation effort is how to best estimate these two main components of the TRMM error budget, which are not easily decoupled.

This four-year study computed bulk sampling and retrieval errors for the TRMM microwave imager (TMI) and the precipitation radar (PR) by applying a technique that sub-samples gauge data at TRMM overpass times. Gridded monthly rain estimates are then computed from the monthly bulk statistics of the collected samples, providing a sensor-dependent gauge rain estimate that is assumed to include a TRMM equivalent sampling error. The sub-sampled gauge rain estimates are then used in conjunction with the monthly satellite and gauge (without sub-sampling) estimates to decouple retrieval and sampling errors. The computed mean sampling errors for the TMI and PR were 5.9% and 7.7%, respectively, in good agreement with theoretical predictions. The PR year-to-year retrieval biases exceeded corresponding TMI biases, but it was found that these differences were partially due to negative TMI biases during cold months and positive TMI biases during warm months.

Physical Validation of TRMM TMI and PR Monthly Rain Products Over Oklahoma

Brad L. Fisher*

NASA/GSFC Laboratory for Atmospheres and Science Systems and Applications, Inc. Greenbelt, Maryland

Submitted To

Journal of Applied Meteorology

*Corresponding Author:
Brad Fisher
NASA/Goddard Space Flight Center/912.1
Greenbelt, MD 20771
Phone: 301-286-8912
Email: fisher@radar.gsfc.nasa.gov

ABSTRACT

This paper reports on the results of a regional validation study of monthly rainfall products generated from sensor measurements aboard the Tropical Rainfall Measuring Mission (TRMM) satellite. The study analyzed four-years of rainfall estimates (1998-2001) produced from data collected by the TRMM microwave imager (TMI) and the precipitation radar (PR), and compared them with corresponding rain estimates computed from 66 rain gauges in the Oklahoma Mesonet. The rain estimates from all sensors were gridded at 1° x 1° and 2° x 5°. The aim of this study was to estimate an annual error budget for the TMI and PR using a methodology designed to decouple retrieval and sampling error.

The methodology of the study sub-sampled the gauge data at satellite overpass times, and estimated monthly rainfall amounts using a simple method of bulk statistical integration. This approach yielded two distinct rain parameters: 1) G_0-integration at all times during observation period (5-minute accumulations) and 2) G_S-bulk statistical integration (sub-sampled during overpasses). These two gauge-inferred rain parameters provided enough information to determine the sampling and retrieval errors for the satellite.

The sampling errors computed for the TMI and PR using this methodology were closely correlated, with overall sampling biases of 5.9% and 7.7%, respectively. The overall retrieval biases were about 10% for the TMI and 29% for the PR. Since retrieval errors were computed in bulk, some of this difference was attributed to an offsetting seasonal bias observed in the TMI estimates, which was negative during the winter months and positive during the summer months.

1. Introduction

Precipitation plays a fundamental role in the global water cycle and in forcing the large-scale dynamics of the general circulation of the atmosphere. Satellite observing platforms seem to offer the best possibility for accurately estimating the mean climatological distribution and variability of global precipitation. In addition to having a larger sampling domain, satellites have a distinct advantage over ground sensors in that they can gather data over oceanic and mountainous regions, where in-situ measurements of rainfall are extremely sparse or altogether non-existent. However, space-borne sensors collect rain information remotely and then estimate rain intensity at the earth's surface indirectly from radiance measurements obtained aloft. As satellite remote-sensing technology continues to improve, a greater emphasis is being placed on the quantitative accuracy of satellite retrievals of rainfall as inferred from regional validation data sets obtained from ground based sensors.

The Tropical Rainfall Measuring Mission (TRMM) satellite was launched in November 1997, as a joint scientific initiative between the National Aeronautics and Space Administration (NASA) and the National Space Development Agency (NASDA) of Japan. TRMM collects information used to study the distribution and transport of latent heat in the tropical and sub-tropical atmosphere. One of the major goals of TRMM is to produce quantitatively accurate mean monthly rain estimates from a space-borne instrument platform with errors not exceeding 10-15% of the true rainfall (Simpson et al. 1988).

The TRMM satellite carries several remote sensors that collect rainfall information between 40° N to 40° S. The aim of this study is to validate rainfall estimates from the 13.8 GHz precipitation radar (PR), and the 9-channel TRMM microwave imager (TMI). The TMI collects passive radiance information at 10.7, 19.4, 21.3, 37.0 and 85.5 GHz (see Kummerow 1998). All TMI frequency channels are horizontally and vertically polarized, except for the 21.3 GHz channel which is only vertically polarized. Each sensor collects instantaneous areal observations during satellite overpasses, with local satellite revisit times

varying between once and twice daily depending on the latitude. Rainfall estimates for each sensor are subsequently determined from the bulk statistics associated with the collection of a month of discrete instantaneous observations.

An important part of the TRMM ground validation strategy uses independently derived estimates from ground-based sensors to estimate the uncertainties and biases in the satellite estimates of surface rainfall. Wilheit (1988) describes the TRMM error budget as consisting of three types of errors: retrieval, sampling and random errors. Random errors are associated with random instrument noise, such as statistical effects of thermal emission. Wilheit points out that these errors do not generally average to zero, but that any residual random error will be considerably smaller than the other two error terms and so can be effectively absorbed into the retrieval component of the error budget (see Wilheit 1988 for more details). Physical validation of TRMM rain estimates, therefore, reduces to an analysis of the retrieval and sampling errors.

Retrieval errors are caused by calibration and algorithm errors, causing the rain estimate to be systematically biased. Calibration errors are linked with real time measurements by the sensor hardware, whereas algorithm errors occur in the post real-time processing of sensor measurements and depend on the physics of the algorithmic model. Retrieval errors can be corrected by determining the magnitude and sign of the bias. This information can also lead to improvements in the algorithm physics (Chang and Liu 1999).

Sampling errors are the other major contributor to the TRMM error budget. TRMM estimates of the mean monthly rainfall for any given area in the sampling domain are determined based on the discrete rain statistics collected during a month of overpasses. Sampling errors are due to non-continuous sampling and occur because the discrete set of satellite observations for a given region are not statistically representative of the long periods between subsequent satellite overpasses (lasting10-12 hours). Temporal sampling errors are intrinsic to the measurement process, since they depend on the fixed orbital characteristics of the satellite. Several researchers prior to the launch of TRMM rigorously studied the

effects of discrete sampling. Most of these studies considered simulated TRMM orbits and statistically modeled the rain fields to be consistent with the Global Atmospheric Research Project (GARP) Atlantic Tropical Experiment (GATE). Studies by Laughlin (1981), Bell (1987), McConnell and North (1987), Shin and North (1988), North and Nakamoto (1989), and Bell et al. (1990) all concluded that sampling errors should average between 8-12% per month.

This paper presents a comprehensive four-year validation study (1998-2001) of TRMM monthly rain estimates produced from the TMI and PR data. The general aim of this study was to estimate the annual error budget for TMI and PR rain estimates over a 2° x 5° region in Oklahoma, using a network of 66 rain gauges from the Oklahoma mesonet. An important part of this study involved the development of a validation strategy designed to estimate the individual contributions of retrieval and sampling errors to the TRMM error budget. The methodology of this study sub-sampled the gauge data at satellite overpass times and produced a secondary sensor-dependent gauge estimate of rainfall based on the month-to-month bulk rain rate statistics. This secondary gauge estimate, which was assumed to contain a TRMM equivalent sampling error, provided an empirical parameter that effectively decoupled the total error into sampling and retrieval components.

Section 2 of this paper provides descriptions of the data sets used in the analysis. Section 3 develops the formalism used for the error model. Section 4 describes the procedure for sub-sampling and further uses regressions to illustrate the effects of sub-sampling. Results and analysis are presented in sections 5 and 6, and in section 7 the systematic effects of TRMM sampling at high latitudes near the satellite turning point are discussed.

2. Data products

a. Satellite Products

TRMM satellite rainfall products are processed by the TRMM Science and Data Information System (TSDIS) and released to the general public through the Goddard Earth

Sciences (GES) Distributed Active Archive Center (DAAC). Table 1 summarizes the TRMM rain products used in this study, which were processed using TRMM version 5 rain algorithms. The reference numbers assigned to these different product levels correspond to the different stages of the processing. Detailed descriptions of TRMM products and the algorithms used to process the data can be found at the TRMM Project web site: http://trmm.gsfc.nasa.gov/data_dir/ProductStatus.html

The strategy used to grid the data is illustrated in Figure 1. Using Level 3 products, monthly rainfall estimates were first obtained for the ten 1° x 1° boxes shown in the figure. The 2° x 5° region shown in Fig. 1 covers an area from 34° to 36° N and 100° to 95° W. The longitude boundaries shown in the figure are defined by the east west boundaries of Oklahoma, not including the Panhandle, whereas the latitude boundaries span the lower two thirds of the state. The upper third of the state had to be excluded from the study because the PR swath did not collect data above 36°.

Special Geo-Regional Gridded Level 2 products (RG2A12 and RG2B31) provided gridded instantaneous orbital track data at a resolution of 0.5° for the TMI and 0.1° for the PR. The orbital track information was used to match the gauge time series with TRMM overpasses. It should be noted that the RG2B31 instantaneous rain rates are produced with the combined 2B31 algorithm. These estimates corresponded to the orbital track of the PR. Standard Level 2 products (2A12 and 2A25) were also used to compile rain rate statistics for the TMI and PR at the characteristic resolution of each sensor.

b. Rain gauge data products

The Oklahoma Mesonet is operated and maintained by the Oklahoma Climate Survey (OCS). The Mesonet consists of 118 automated weather stations that measure 10 meteorological variables, including precipitation. Each station reports by radio to a base receiving station every 15 minutes. The tipping bucket gauges in the network accumulate rain in increments of 0.01" and produce a rain record every 5 minutes, even when no rain is measured (Brock, et al. 1995). The mesonet gauges are carefully calibrated in the lab

according to different rain rates, so that data collected in the field can be calibration adjusted. Brock et al. used fairly rigorous criteria in the selection of sites in order to minimize systematic errors caused by wind, obstructions and topography. As noted earlier, the partitioning of gauges into grid boxes was done in accordance with Fig. 1 and Table 2.

One-minute gauge rain rates were computed with a cubic spline algorithm. The interpolation was done at this temporal scale so that the gauge time stamp could be matched with the TRMM overpases. Interpolated rain rates were automatically bias-adjusted for each observed rain event. An independent event was defined as a contiguous time series of tips, where consecutive tips in the time series were separated by less than thirty minutes. Bias adjustments were generally between one and five percent, which ensured that over and under-estimates of rain were not introduced through the interpolating algorithm. The interpolation of rain rates from a discrete time series of tips represents one potential source of error, but this error was significantly reduced by averaging rain rates within a time window. The gridding procedure will be discussed in more detail in the section 4.

3. Error Model

This section provides a description of the error model used in this study. The error corresponding to a given monthly areal rain estimate R_0 is defined as the magnitude of its deviation from the true mean areal rainfall R_T:

$$\varepsilon = R_0 - R_T. \tag{1}$$

In estimating the total error associated with a large number of such estimates, it is more practical to define the error in terms of a statistical variance. In the error model described by Wilheit (1988), the total error is related to two primary sources of variance, retrieval ($\sigma^2_{err,R}$) and sampling ($\sigma^2_{err,S}$), as:

$$\left\langle \left(R_0 - R_T \right)^2 \right\rangle = \sigma^2_{err,R} + \sigma^2_{err,S}. \tag{2}$$

In (2) it is assumed that $\sigma^2_{err,R}$ and $\sigma^2_{err,S}$ are uncorrelated and that R_T is obtained from a perfect rain sensor that incurs no error itself. The first assumption is justified because

retrieval errors are incurred in conjunction with satellite overpasses, whereas sampling errors are incurred when the satellite is not overhead.

In most practical situations, R_T is replaced by an independent estimate of surface rainfall, derived from measurements obtained from surface instruments. In this study, a gridded network of rain gauges were used to physically validate satellite rain estimates. Treating R_T as an observed quantity rather than an ideal quantity contributes an additional error variance in (2) related to the validation sensor (Bell and Kundu, in press),

$$\left\langle \left(R_0 - G_0 \right)^2 \right\rangle = \sigma^2_{err,R} + \sigma^2_{err,S} + \sigma^2_{err,G}, \tag{3}$$

where G_0 corresponds to gridded gauge estimates of mean rainfall, representing the same areal region and temporal period as R_0. G_0 is assumed to incurr an error designated as $\sigma^2_{err,G}$, which is uncorrelated with the other two error terms in (3). The motivation for the above equation is developed in a recent paper by Bell and Kundu (in press).

Part of the validity in using G_0 as an estimate of R_T is based on the sometimes implicit assumption that gauge errors associated with estimating the mean areal rainfall are at least one of order of magnitude smaller than the satellite errors shown in (3). Three different sources of error contribute to the variance $\sigma^2_{err,G}$, which can be represented by

$$\sigma^2_{err,G} = \sigma^2_{rand} + \sigma^2_{SS} + \sigma^2_{syst}, \tag{4}$$

where σ^2_{rand}, σ^2_{SS} and σ_{syst} correspond to random instrument error, spatial sampling error and systematic gauge error, respectively. Random instrument errors associated with tipping bucket measurements vary with rain rate, but are typically only about ±0.5% per 0.01". Spatial sampling errors are incurred when point gauge estimates are used to estimate areal rainfall. Gauge sampling errors will be qualitatively addressed later in this section. Systematic gauge errors due to wind, topography, obstructions such as trees and buildings, and mechanical and maintenance-related issues can be significant in some cases and so generally cannot be entirely ignored (Habib et al. 2001). Although it is hard to quantify

systematic errors, greater qualitative precision can be gained by understanding whether rain estimates are positively or negatively affected by their influence.

The aim of this study is to determine $\sigma^2_{err,R}$ and $\sigma^2_{err,S}$ for the TMI and PR rain estimates. The Mesonet gauges used in this study are sampling continuously and record rainfall at a high temporal resolution. Hence, the gauges are assumed to contain no temporal sampling errors due to discrete sampling. The rainfall estimate of an individual gauge g_i can then be expressed in terms of the integral

$$g_i = \frac{1}{T}\int_0^T r_G(t)dt, \tag{5}$$

where $r_G(t)$ represents the complete recorded time series of rainfall for the gauge and T represents the duration of the validation period. The mean areal rainfall defined as G_0 in (3) is obtained by first summing over the index i for all gauges located within a gridded area and then dividing by the number of gauges in the grid

$$G_0 = \frac{1}{N_g}\sum_{i=1}^{N_g} g_i . \tag{6}$$

Even though the individual gauges g_i are assumed to contain negligible temporal sampling error, G_0 contains spatial sampling errors associated with using a gridded ensemble of point rain estimates to represent the mean rainfall over a large area.

The rain estimate G_0 does not provide enough information to decouple the sampling and retrieval errors associated with R_0. The decoupling of these two error terms is accomplished by first sub-sampling the gauge data during overpass times and then using a simple method of bulk statistical integration to obtain a sensor dependent gauge-inferred estimate of rainfall, G_S. This methodology introduces a TRMM equivalent sampling error into the gauge-inferred estimate of rainfall. A satellite sampling error variance can now be defined by considering the relative statistical differences between the monthly estimates of G_S and G_0 defined as

$$\sigma^2_{err,S} = \left\langle \left(G_s - G_0 \right)^2 \right\rangle. \quad (7)$$

The error variance in (7) is used to approximate the TRMM sampling error, since the main differences between G_S and G_0 are due to the discrete sampling of G_0 during satellite overpasses.

Assuming that G_S and R_0 contain equivalent temporal sampling errors, any remaining variance is attributed to the retrieval error of the satellite sensor. A satellite retrieval error variance can be then obtained in a similar manner using G_S and R_0,:

$$\sigma^2_{err,R} = \left\langle \left(R_0 - G_S \right)^2 \right\rangle. \quad (8)$$

In (7) and (8), G_S provides a rain parameter that effectively decouples the retrieval and sampling error components of the TRMM error budget (Bell et al. 2000).

The error model described in this section assumes that the gauge and satellite spatio-temporal references are statistically equivalent. However, since individual rain gauges only represent point measurements, a spatial sampling error is incurred when a point estimate is taken to represent the mean of a larger area (Rodriquez-Iturbe and Mejia 1974b; Barnston 1991; Hulme and New 1997). Spatial sampling errors with respect to the gauge estimates are analogous to the temporal sampling errors with respect to the satellite estimates. In the case of TRMM, temporal sampling errors occur because no rain information exists in the long time interval between contiguous observations. Similarly in the case of gauges, spatial sampling errors occur because no rain information exists in the large amount of space between the gauges used to compute G_0.

Prior studies have shown that spatial sampling errors can be significantly reduced by carefully planning the relative distribution and density of gauges in the network. By optimizing these two network parameters, the error associated with estimating the mean rainfall over a test area can be effectively minimized (Rodriquez-Iturbe and Mejia 1974a; North and Nakamoto 1989; Morrissey et al. 1995). There are other studies that make specific recommendations about the number of gauges needed to adequately sample a given

area. Xie and Arkin (1995) recommend using a minimum of at least 5 gauges for sampling a 2.5° x 2.5° area; Krajewski et al. (2000), however, suggests an even more stringent requirement of about 25 gauges for the same area.

The Oklahoma Mesonet gauge network configuration approximates a quasi-homogeneous distribution and so it satisfies the first minimum error condition. Table 2 shows that there were 66 gauges distributed across a 2° x 5° area. This distribution easily satisfied the Krajewski et al. criteria. Therefore, in this study, spatial sampling errors associated with G_0 and G_S are assumed to be minimized and at least one order of magnitude less than the TRMM sampling error. It should be pointed out that collecting more satellite observations, would also in principle reduce temporal sampling error, but since the orbital parameters of TRMM are fixed, the temporal sampling error cannot be further minimized through the simple collection of more data.

4. Gauge data: methodology and analysis

a. Formal procedure for estimating G_S for the TMI and PR.

This validation strategy takes advantage of the high temporal sampling resolution of the gauges by sub-sampling the gauge time series during coincident satellite overpasses using orbital track information obtained from the gridded Level 2 products described in Table 2 (RG2A12 & RG2B31). The TRMM satellite collects about 80 observations per month for each sensor over the study area in Oklahoma. Each observation, however, only results in partial coverage of the region shown in Fig. 1, thus in an average month, the full region is sampled about 30 times. In this study, coverage was determined at the 1° x 1° scale. If any part of the grid box was sampled by the satellite, then all the gauges in the box were included in the estimation of the mean rainfall for that box.

A time-averaged rain rate is computed for each gauge by centering a constant time window Δt_W on a given overpass time t_k

$$\hat{R}_{ijkm}^{(G)}(t_k) = \frac{1}{\Delta t_W} \sum_{t=t_k-\Delta t_w/2}^{t_k+\Delta t_w/2} R_{ij}^{(G)}(t), \tag{9}$$

where subscripts i, j, k, and m correspond to the i^{th} gauge, located in the j^{th} cell, at the k^{th} overpass and in the m^{th} month, respectively. Using (9), an estimate of the mean monthly rainfall for the i^{th} gauge is then computed as

$$\overline{R}_{ijm}^{(G)} = \frac{1}{N_0} \sum_{k=1}^{N_0} \hat{R}_{ijkm}^{(G)}(t_k), \tag{10}$$

where N_0 is the number of overpasses in the m^{th} month, for a gauge located in the jth cell cell. Note that in general, N_0 is not a fixed quantity and can vary for different months and different cells. The mean gauge monthly rainfall for the j^{th} cell and the m^{th} month can then be expressed as

$$\hat{G}_{jm} = \frac{1}{N_g} \sum_{i=1}^{N_g} \overline{R}_{ijm}^{(G)}, \tag{11}$$

where N_g is number of gauges in the grid cell, j.

The satellite-coincident sub-sampled mean annual rainfall for the 2° x 5° cell is obtained from (11) by summing over all ten 1° x 1° cells for the whole year as

$$G_S = \sum_{j=1}^{10} \sum_{m=1}^{12} \hat{G}_{jm}. \tag{12}$$

Differences in the swath geometry of the TMI and PR require that the sub-sampling procedure described above be performed independently for each sensor, which makes G_S a sensor dependent rain parameter (e.g., G_s_TMI and G_s_PR). In general, G_s_TMI and G_s_PR are not equal, but should be highly correlated.

b. Sensitivity of G_S to changes in Δt_w and empirical formulas for estimating biases.

Time averaging the gauge data about the overpass time takes advantage of the coupling between the time and space autocorrelation (Laughlin 1981; Bell 1990) by extending the effective representativeness of each point estimate to a larger areal domain. Zawadzki (1975) used this technique to match the gauge time domain with the radar areal domain, and Doneaud et al. (1984) used the technique in their formal development of the area-time

integral approach. The optimal time averaging scale defined earlier as Δt_W, depends on sensor resolution, orbital characteristics and storm properties. Although the sensor resolution and orbital characteristics are fixed, storm parameters such as rain type, cloud coverage and environmental wind speed can change from one event to the next, and are affected by climatological factors such as time of day, season, and synoptic and planetary scale climate anomalies such as El Niño-La Niña, droughts and monsoons .

In this study, G_S was treated as a statistical variable that depended on the averaging time interval Δt_W computed at 2-minute intervals over a dynamic range between 3 and 61 minutes. Figures 2a and 2b show the results for both G_S_TMI and G_S_PR and illustrate the sensitivity of G_S to changes in Δt_W. From the figures it can be seen that G_S values for both sensors were fairly insensitive to the window size for the range of intervals selected. The difference between the minimum and maximum values of G_S in Fig. 2 was less then 4% for both sensors during 1998, 1999 and 2001, but was about 8.5% for the TMI and about 6% for the PR in year 2000.

In all subsequent estimates of the sampling and retrieval errors, G_S was represented by the mean of the data displayed in Fig. 2. Table 3 provides estimates of $\overline{G_s}$ and $\overline{\Delta t_W}$ (i.e., $G_s(\overline{\Delta t_W}) = \overline{G_s}$) for each year of the study. This method of determining G_S allows the effective time window to adjust in response to climate variability from year to year.

A sampling bias can now be computed as the relative difference between G_0 and $\overline{G}_s$ as

$$b_s = \frac{\overline{G}_s - G_0}{G_0}. \qquad (13)$$

The sign of b_S is especially important because it indicates whether the rainfall is being statistically over-sampled or under-sampled. For example, a negative sampling bias indicates that the satellite is undersampling the actual rainfall; a positive total bias in this case will subsequently lead to a retrieval bias that exceeds the total relative bias.

A retrieval bias can be similarly computed as the difference between R_0 and $\overline{G}_s$ as,

$$b_R = \frac{R_0 - \overline{G}_S}{G_0}. \qquad (14)$$

Whereas (13) only involves two rain parameters, the calculation of the retrieval bias in (14) involves G_0, G_S and R_0, with the bias in both cases expressed as a fraction of G_0. The results of these equations will be presented in section 6.

c. G_0 vs. G_S

The variance due to the discrete temporal sampling of the gauge data was apparent in the regressions of G_0 and G_S.The six panels in Figure 3 show a regression analysis of G_0 versus G_S_TMI and G_S_PR, using all four years of gauge data (1998-2001) at resolutions of 1° x 1° and 2° x 5°. At the 2° x 5° spatial resolution, scattergrams are shown at two different temporal resolutions: 1) monthly and 2) 4-year monthly averaged. The corresponding correlation coefficients for each case are shown in Table 4. Note that Table 4 also includes regressions comparing corresponding satellite and gauge data, which regressions will be discussed in Section 5a.

As observed in Fig. 3, sub-sampling the gauge data at overpass times produced a sharp decorrelation at the 1° x 1° scale. Since no other data manipulations were performed, the large amount of variance suggested by the scatter at the 1° x 1° scale could only be attributed to the discrete sub-sampling of the gauge data. When the precipitation was averaged over the whole 2° x 5° grid space, the correlation coefficients significantly improved. It is interesting to note that the additional temporal averaging did not lead to any large improvement in the correlation coefficients computed at 2° x 5°. In both cases at the 2° x 5° scale, the G_0 and G_S data points showed high correlations. Based on this result, it is not expected that extending the study area to a more standard size of 5° x 5° will produce much additional improvement in these correlations.

Figure 4 shows the distributions of gauge-inferred sampling errors for the TMI and PR, computed relative to the monthly differences between G_S and G_0 over four study years at a

resolution of 1° x 1°. If the process of sampling were purely random, the distribution of errors should be quasi-normal about zero. The two distributions shown in Fig. 4 are approximately normal. The median, mean and standard deviation values were –0.29, 0.13 and 2.3 for the TMI, and –0.29, 0.27 and 2.7 for the PR, respectively. The negative median was due to a higher frequency of negative $G_S - G_0$ differences, whereas the mean was positively skewed due to the presence of some large positive $G_S - G_0$ differences. G_S estimates in some cases far exceeded the natural upper bound of G_0, which was about 8 mm/day. High G_S estimates were presumably due to statistical oversampling of rainy periods.

5. Validation of level 3 satellite products with G_0 and G_S.

a. Monthly regression analysis: TMI and PR vs. G_0 and, G_S

Figures 5a and 5b show the four-year time series of mean monthly rain estimates at 2° x 5° resolution for: 1) G_0, G_S_TMI and TMI, and 2) G_0, G_S_PR and PR, respectively. Figures 6 and 7 show scattergrams of the TMI and PR satellite rain estimates matched against corresponding G_0 and G_S gauge rain estimates at the following spatio-temporal resolutions: 1) 1° x 1° using monthly estimates (i.e., 480 points), 2) 2° x 5° using monthly estimates (i.e., 48 points), and 3) climatological monthly estimates (i.e., 4-year monthly averages or 12 points). Table 4 presents a summary of the computed correlation coefficients resulting from this analysis.

The correlation between gauge and satellite rain estimates improved markedly as the spatial and temporal averaging scale was increased. As shown in Figures 6a-b and 7a-b, low correlation coefficients were observed at the 1° x 1° scale between the satellite estimates and both gauge estimates G_0 and G_S. Random errors were an important reason for the low correlations at the higher spatial resolution. At 1° x 1° scale, satellite random errors were distributed between adjacent grid boxes, whereas gauge random errors were related to using a smaller number of gauges to estimate the mean rainfall for each grid box. The G_S

correlations with respect to the TMI and PR improved significantly as the averaging area was expanded to the 2° x 5° scale that included all 66 gauges. Although Table 4 results suggest that G_0 was more highly correlated with the satellite estimates, the observed improvements in going from the 1° x 1° to the 2° x 5° resolution seem to indicate that G_S and G_0 correlations were converging, and additional convergence might have been realized if the averaging area could have been expanded to the more standard size of 5° x 5°. Laughlin (1981) and Bell (1987) showed that the auto-correlation time at the 1° x 1° scale is only about 3.5 hours, significantly less than the satellite revisit time, which is about 10.5 hours at the satellite's turning point. At the 2° x 5° scale, the auto-correlation time is about 8 hours, which is a time more consistent with the mean revisit time of the satellite, resulting in the higher gauge-satellite correlations observed in Fig. 6c-f and 7c-f.

b. Probability Density Functions (PDF) for TMI, PR, 3B42, G_S and G_0.

The effects of discrete sampling were further investigated by examining the probability density functions (PDF) compiled from four years (1998-2001) of monthly rain events at the 1° x 1° scale. Figure 8a displays PDFs for the TMI, G_S and G_0, and similarly, Fig. 8b displays the PDFs for the PR with G_S and G_0. At this spatial scale, each PDF consisted of 480 points (10 cells x 12 months x 4 years), which were then binned at 0.5 mm intervals, expressed in units of mm/day.

Note that each of the three panels shown in Fig. 8 consists of two plots, a main plot and an inner plot. The inner plot is associated with high monthly accumulations of low probability. It extends the dynamic range of the distribution shown in the main plot and adjusts the vertical axis so as to amplify the fine structure in the tail. The main outer plot covers a range from 0 and 10 mm/day, which spans the full dynamic range of G_0. The inner plot extends this range from 10 and 25 mm/day. Although the points in the outer tail of the distribution only represent a small percentage of the cumulative probability, they play an important role in determining the mean of each distribution. The high frequency of low rain

rates (bins near zero) is balanced against high rain rates that have a very low frequency of occurrence.

In Fig. 8a and 8b, the PDFs constructed from the TMI, PR and G_S reveal a similar structure, showing a maximum probability in the zero rain bin (i.e., 0 to 0.5 mm/day), followed by an exponential decay. The effects of discrete temporal sampling can be clearly observed by comparing TMI, PR and G_S with G_0 near extremes of each spectrum. G_0 showed a relatively low probability of observing trace monthly rainfall (~8%), with zero probability of observing gridded monthly rain totals that exceeded 8 mm/day. G_S, on the other hand, closely matched the PR spectrum in the first bin (~26%) and had an exponential tail that extended well beyond the observed threshold for G_0 (as shown in the inner plot) The PDF for the TMI revealed a similar exponential distribution, but displayed a higher frequency of monthly estimates in the first bin (~31%), suggesting that the TMI had a greater likelihood of observing trace amounts of monthly rainfall than either the PR or G_S. This feature was presumably due to the TMI not observing weaker rain systems with ice particle sizes and concentrations below detectable levels (i.e., small ice-scattering cross section).

Since the data used to construct the PDF for G_S was computed from the data extracted directly from G_0, it was assumed that the statistical differences in the PDF spectrums of G_0 and G_S were due to the effects of discrete temporal sampling. Since the autocorrelation time at 1° x 1° scale was about 3.5 hours, the structural characteristics of the PDFs for G_S, TMI and PR could also be related to statistical effects of undersampling at this spatial scale. For example, Fig. 8c compares the PDF of G_0 with the 3B42 rain product. The 3B42 is a special level 3 TRMM product that provides daily rainfall estimates by adjusting the Global Precipitation Index (GPI) to the TRMM combined rain product (2B31) during coincident TRMM overpasses (Adler at al. 1994). The 3B42 uses eight GOES IR observations per day separated by 3-hour intervals to compute daily rainfall amounts, which results in a time separation between subsequent observations that is more compatible with the autocorrelation

time at the 1° x 1° scale. As a result, the PDF of 3B42 shown in Fig. 8c appears structurally similar to G_0 and so does not show pronounced effects of discrete sampling as do the PDFs of TMI, PR and G_S.

6. Results

a. TMI and PR error statistics: annual means, standard errors and biases

Table 5 presents a summary of the four-year mean annual rain rates (μ) together with the sampling and retrieval standard errors ($\sigma_{err,S}$, $\sigma_{err,R}$) using the same spatio-temporal averaging scheme described at the beginning of section 5a: 1) monthly 1° x 1° resolution, 2) monthly 2° x 5° resolution, and 3) climatological monthly 2 x 5 resolution. The standard errors shown in Table 5 were computed for the TMI and PR using (7) and (8). The higher standard errors observed at 1° x 1° scale were consistent with the low correlations observed at this scale in section 5a, relating to a fundamental mismatch between the size of the grid box, the autocorrelation time and the satellite revisit time. Note that in going from 1° x 1° to 2° x 5° scale, the means remained constant while the errors were significantly reduced, due to the reduction of random errors between 1° x 1° grid boxes.

Satellite sampling and retrieval biases for the TMI and PR were computed using (13) and (14). Figures 9a and 9b show plots of the sampling and retrieval biases, respectively, for each year of the study. During 1998 and 1999, the sampling biases were negative for both sensors, but switched to positive values during 2000 and 2001. The highest sampling biases observed for the TMI were ~20% in 2000 and 2001, whereas for the PR a sampling bias of ~30% was observed in 2000. The sampling biases for the other years of the study fell within the nominal range of 8-12% predicted by prior theoretical simulations (Laughlin 1981, McConnell and North; 1987, Shin and North 1988; North and Nakamoto 1989; and Bell et al. 1990). As expected, overall, the PR sampling biases were larger than the TMI sampling biases because the PR's swath covered only one-third the area of the TMI swath. Interestingly, in1998 and 2001, the PR actually revealed slightly smaller sampling biases

than the TMI, however, these differences were extremely marginal and non-significant. The sense of the bias (i.e., positive or negative) was the same for both sensors.

The retrieval biases shown in Fig. 9b illustrated a different trend from the sampling biases shown in Fig. 9a. Whereas the sampling errors for the TMI and PR appeared to be closely correlated, the retrieval errors revealed more disparity between the two sensors. Retrieval biases for the TMI were slightly negative in 2000 and 2001, with a maximum bias of about +40% observed in 1999. The PR retrieval biases, on the other hand, remained positive and were greater than 10% for all four years. The PR retrieval bias also reached a maximum of about 40% in 1999.

b. Instantaneous rain rate profiles: TMI, PR and G_0.

Monthly rainfall accumulations are statistically coupled to rain retrievals obtained during instantaneous satellite overpasses. Figure 10 shows lognormal rain rate distributions for the TMI, PR and G_0, with rain rates presented in units of dBR. The gauge rain rates were compiled from all available data between 1998 and 2001 to approximate a climatological distribution. The anomalous spikes observed in the central part of the gauge distribution shown in Fig. 10 were introduced by the cubic spline transformation used to generate a continuum of rain rates from a discrete time series of tips. The satellite profiles were produced from 56 rain cases over Oklahoma, selected to be nominally representative of the overall rain rate climatology for the region. Standard Level 2 products were used (2A12 and 2A25) to pick a diverse grouping of rain cases from all four seasons that covered all four years of the study.

As shown in Fig. 10, the TMI rain rate profile was structurally discrete and did not exceed a threshold of 50 mm/hr (~17 dBR). This result can be explained by examing the way TMI estimates surface rainfall over land. The TMI algorithm estimates surface rainfall by matching brightness temperature depressions from all available TMI channels to a database of brightness temperature-rain rate profiles generated by the Goddard Cumulous Ensemble Model (Tao and Simpson, 1993; Kummerow et al. 2001). Data collected over

land is processed using only the brightness temperature information in the 85 GHz channel. Information retrieved by the other TMI channels are not used because of difficulties in resolving the rain signal from the highly variable emissions coming from the land surface beneath the cloud column (Spencer et al.1989, Ferraro and Marks 1995, Conner and Petty 1998). In the current version 5 TMI "land" algorithm, 85 GHz brightness temperatures are matched against a database of only 28 modeled cases. As a result, the TMI land rain rate profile in Fig. 10 appears discrete. These quantized effects are not observed in the TMI oceanic rain rate profiles. It should be noted that the number of land cases in the database will be expanded in the version 6 algorithm. This algorithm improvement will distribute the rain profile across a greater number of bins. With only the 85 GHz channel, the TMI land algorithm infers rain rate-brightness temperature relations based on the assumed correlation between processes of ice scattering and surface precipitation. The physical connection between these two cloud processes is fairly well established in the literature (Spencer et al.1989, Ferraro and Marks 1995, Conner and Petty 1998), but these same studies also suggest that the connection is physically indirect and highly non-linear.

In contrast with the TMI, the PR distribution showed structural continuity and revealed a larger dynamic range of observations, extending beyond 100 mm/hr (20 dBR). The distribution of PR rain intensities was found to be structurally more similar to the climatological gauge distribution. Although the PR distribution of instantaneous rain rates appeared to be more physically realistic, the TMI interestingly showed lower retrieval biases when the instantaneous rain rates were used to estimate rainfall over longer integration periods (e.g., one month). As shown in section 6a, bulk PR retrieval biases were consistently positive, suggesting that retrieval errors were being incurred at the instantaneous scale.

PR rain estimates are affected by three potential sources of retrieval error (Iguchi 2000): 1) attenuation correction 2) non-homogenous beam-filling of the PR footprint, (i.e., presence of reflectivity gradients) and 3) stratiform-convective classification. The PR

transmits a beam at 13.8 GHz, which is strongly attenuated by precipitation. The return power is first subjected to an attenuation correction that results in the calculation of an effective reflectivity factor. In heavy rain cases, this correction can change the estimated rain rate by a factor of 10 or more (Iguchi, et al. 2000, Meneghini et al., 2000). After correcting for attenuation, further corrections account for non-uniform beam filling associated with reflectivity gradients, for as Iguchi points out, the attenuation correction is made under the assumption that the rain is homogenously distributed. Each rainy pixel must then be matched to an assumed drop-size distribution using a dual Z-R stratiform-convective classification scheme.

Based on the analysis of PR instantaneous observations it was suspected that part of the PR (positive) retrieval bias could be due to excessive rainfall in the extreme tail of the PR distribution. Though difficult to observe in Fig. 10, it was found that the PR estimated more rain in the far ranges of the distribution. Further examination of convective cells in the 56 selected rain cases revealed the existence of correlated clusters of rain rates exceeding 200 mm/hr. These cases were typical of the rain regimes observed during the late spring and summer months. These particular cases not only required a large attenuation correction, but these cases also displayed strong reflectivity gradients. The Z-R relation itself has no fundamental limiting constraints and will generate arbitrarily high rain rates for a corresponding high Z value. Beam smearing may also be important in the spreading of higher reflectivity values into adjacent bins (Iguchi et al. 2000). High rain rate cases corresponding to individual overpasses have an especially large statistical impact on bulk monthly rain estimates, since the determination of a mean unconditional monthly rain rate depends on a relatively small number of observations.

c. Seasonal and diurnal rainfall climatology

Seasonal and diurnal variations over land at higher latitudes can affect the error statistics. In particular, it is important to understand how such variations affect retrievals and whether the TRMM rain algorithms are biased by cyclical climatological variability. This

part of the study investigated the secondary effects of Oklahoma's seasonal and diurnal cycles on the TRMM sampling and retrievals. Figure 11a displays a plot showing a normalized monthly climatology for Oklahoma constructed from the four years of gauge data used in the study. Figures 11b and 11c show four-year mean monthly sampling , retrieval and total biases as a fraction of G_0 computed at 2° x 5° for the TMI and PR using (13) and (14). The total bias is simply the sum of those two equations.

As shown in Fig. 11b and 11c, the monthly sampling and retrieval biases were about the same order of magnitude, with the retrieval biases showing an overall larger range of variability. Monthly sampling biases were both positive and negative, but revealed a positive tendency in the case of both sensors. Large positive retrieval biases of over one hundred percent were observed for both the TMI and the PR during the summer months of July and August. Fig. 11a indicates that July and August are relatively dry compared with other months. Large statistical deviations were therefore further amplified because these deviations are considered relative to G_0. It is also suspected that these retrieval biases are associated with the way the retrieval algorithms handle observations from deep convective systems, commonly observed in the Midwestern Plains during the summer. The PR algorithm must apply a complicated two-way mixed phase attenuation correction, while the TMI, algorithm must infer rain rates based on the scattering signature from the 85 GHz channel.

The TMI displayed a systematic reversal in the sign of its retrieval bias during the winter and summer months. The retrieval bias was negative during the winter months and positve during the summer months, suggesting a fundamental connection between the algorithm physics and the seasonal changes in the distribution of ice above the freezing level (e.g., size, shape and number density). The negative winter retrieval bias of the TMI may partly explain the difference in the overall annual retrieval bias observed between the two sensors. In the case of the TMI, the negative seasonal winter bias partly canceled the positive summer bias leading to a lower annual retrieval bias. The PR, on the other hand, shows a generally

positive monthly retrieval bias for all observed months. The large negative June PR retrieval bias upon further investigation was related to a large negative anomaly in 2001 (-87%).

Figure 12 shows the normalized diurnal profiles for the gauges and the TMI, plotting the hourly rain accumulations relative to the total diurnal rainfall. Hourly rainfall totals for the TMI were computed from the gridded Level 2 products. As seen in Fig. 12, the TMI shows a warm bias in the afternoon (12 noon to 6 pm CST), relative to the gauges. The timing of this feature suggests a direct correlation with the development of afternoon convection, associated with the production of large ice particles. Mohr et al. (1996) classified mesoscale convective systems (MCS) over land and ocean using the TMI 85 GHz channel, and showed the existence of strong correlations between the TMI hourly rain accumulations and the late afternoon development of MCS over land as indicated by the 85 GHz scattering signature. This result suggests that the TMI algorithm is overestimating the late afternoon component of the diurnal rain budget, resulting in an observed bias relative to the gauges diurnal profile.

7. Discussion of systematic effects of discrete sampling at high latitudes on TRMM rain estimates.

Prior investigations of the TRMM sampling error were mostly focused on the effects of sampling over the equatorial oceans, where the actual satellite revisit times are nearly equally spaced in time from one orbit to the next (Laughlin 1981; McConnel and North 1987; Shin and North 1988). Nonetheless, Bell et al. (1990) showed that TRMM sampling statistics varied substantially between the equator and the satellite's turning point at ~35° N. Bell et al. used simulations of the TRMM orbit to compare two sampling errors using two test 500 x 500 km^2 grid boxes at latitudes: 1) 5° and 2) at 25° N. In their study, it was demonstrated that satellite overpasses close to the equator were nearly equally spaced in time, but at high latitudes, the time interval between overpasses systematically fluctuated between very short and very long periods. Although there were a greater number of observations at higher

latitudes, these observations were not all statistically independent when considering the revisit times in correspondence with the autocorrelation time.

The TRMM satellite orbits precesses through the diurnal cycle with a period of about 46 days at an inclination angle of 35°. Oklahoma is situated at the satellite turning point where the satellite switches between ascending and descending orbits. It is uncertain how much the sampling error at high latitudes is affected by systematic factors. An examination of the orbital characteristics for Oklahoma revealed that the time separation between contiguous overpasses subdivided into two repetitive modes: 1) Δt_1=21.6 hours and 2) Δt_2=1.6 hours. When averaged together over the 4-year study period, these two modes produce a mean revisit time of 10.5 hours. These results are consistent with Bell (1990). Although the mean revisit time determines the number of overpasses, the variation between actual revisit times suggests that this variability will systematically increase the sampling error variance.

The ratio of the revisit time to the autocorrelation time provides an approximate criterion for evaluating statistical independence (i.e., $\Delta t/\tau_0 \geq 1$). For the 2° x 5° area shown in Fig. 1, the autocorrelation time of about 8 hours. Comparing this autocorrelation time with the satellite revisit times over Oklahoma suggests that TRMM sensors are routinely either under-sampling or over-sampling the region. For example, for Δt_1=21.6, there is little question whether contiguous observations are independent; however, there is nearly a factor of 2 difference between Δt_1 and $\overline{\Delta t}$. Laughlin (1981) clearly showed that the sampling error rapidly increased when the sampling interval exceeded 12 hours. In going from 12 to 24 hours, Laughlin showed that sampling errors increased by about a factor of 2. Salby and Callaghan (1997) in another study, probed the effects of satellite under-sampling of the diurnal cycle and found that even in the case of precessing satellites like TRMM, an empirical determination of the mean diurnal cycle was still biased because the observations collected were too far about in time. The opposite situation occurs in the case of the second mode with Δt_1=1.6. In this case, the revisit time is less than the autocorrelation time by a

factor of 6.5. In this scenario, contiguous observations are no longer independent in any statistical sense, indicating that the region is being over-sampled. Over-sampling, just like under-sampling, can skew the rainfall statistics, either positively or negatively. More rigorous study will be needed to determine the net effect of sampling near the satellite point where contiguous observations are routinely being either under-sampled or over-sampled. In better understanding the sampling error component of the error budget at high latitudes, it is also important to probe the statistical properties of rainfall in association with the structure and life cycles of the rain fields over land, for the orbital characteristics of TRMM are closely coupled with the natural variability of rain.

8. Conclusions

This paper tested a validation strategy designed to estimate sampling and retrieval errors associated with TRMM rain estimates. Four years (1998-2001) of gauge and satellite monthly rain estimates were analyzed in this study. An error model was proposed that estimated bulk sampling and retrieval errors by sub-sampling the validation data set only at TRMM overpass times. The method was then applied to gauge estimates that provided good quasi-homogeneous coverage for a 2° x 5° gridded region of Oklahoma near the turning point of the TRMM satellite.

The results of this study were reasonably consistent with prior theoretical simulations of the TRMM sampling error. Averaged over all four years, the PR had a sampling bias of about 7.7% versus a sampling bias of about 5.9% for the TMI. These percentages differed by a factor of 1.3, as expected based on the swath differences between the two sensors. The overall TMI retrieval bias of approximately 10% was shown to be considerably lower than the 29% overall retrieval bias for the PR. Meaningful insights were gained into the nature of these differences through further data analysis of retrievals using TRMM Level 2 and Level 3 rain products.

An examination of instantaneous rain rate profiles compiled for the TMI and the PR revealed a better statistical match between the PR and the gauges, both in the shape and

dynamic range of the distributions. The TMI distribution, on the other hand, was structurally discrete due to inherent physical limitations in the TMI algorithm, particularly with regards to the database matching over land. However, PR retrieval errors computed over longer integration times generally exceeded those calculated for the TMI. This analysis of instantaneous rain rates suggested that the PR maybe overestimating rain rates for the cores of convective systems over land. The TMI retrievals, on the other hand, appeared to be affected by the seasonal cycle. The mean monthly retrieval biases for the TMI changed sign during winter and summer, with negative winter biases partially canceling positive summer biases when the integration was performed over an annual time scale. In the case of the PR, the sign of the retrieval bias remained consistently positive, which contributed to the PR retrieval biases (without accounting for sampling biases) exceeding the TMI retrieval biases in all four years of the study. Further examination of the diurnal effects on the TMI as compared against a gauge-inferred climatology suggested the presence of a strong warm bias in the late afternoon..

Based on the results of this study, to apply the presented methodology to other regions, the following conditions and constraints are recommended:

- Gauge data should be obtained from a quasi-homogenous network of rain gauges covering a large area of at least 2° x 5° in size (ideally 5° x 5° in size);
- Gauge density should be sufficiently high to minimize the gauge spatial sampling error variance, (Krajewski et al. criterion);
- Gauges should sample rainfall at a high temporal resolution so that time stamps can be optimally matched with instantaneous TRMM overpasses;
- Satellite error statistics should be evaluated over a temporal period T that is much greater than the TRMM granule (e.g., in this study, T>1-month).

The first two conditions attempt to minimize the gauge spatial sampling error. The third condition applies a logistical requirement that ensures that gauge-inferred rain rates can be matched in time with satellite overpasses. Many gauges do not meet this requirement

because they record rain amounts over larger time intervals (i.e., hourly, daily, etc.). The last condition is a statistical requirement that requires the sample sizes for both satellite and gauge rain estimates be large enough to compute representative bulk statistics.

Satellite and ground based estimates of rainfall are both important in improving the accuracy of global precipitation estimates. Ground estimates are not only valuable for validating satellite estimates, but also can be used to minimize errors in global estimates through modern techniques of data assimilation. The Global Precipitation Mission (GPM), a follow-up mission to TRMM, is already in the planning stages. With more attention focused on global rain estimation and latent heating budgets, it is important that ground and spaceborne measurements be collected in a manner that best achieves this important scientific goal. Additional large-scale gauge telemetry networks like the Oklahoma Mesonet are needed to provide more accurate gauge estimates of rainfall that can be used to minimize errors in spaceborne estimates of rainfall. This scientific objective requires good coordination and careful planning among scientists, government agencies such as NASA and NOAA, and other countries.

Acknowledgments. I would like to thank George Huffman for reviewing the manuscript. I am also indebted to him for his scientific expertise in the area satellite rain estimation over the last couple of years.

REFERENCES

Adler, R. F., G. J. Huffman, P. R. Keehn, 1994: Global tropical estimates from microwave-adjusted geosynchronous IR data. *Rem. Sens. Rev.*, **11**, 125-152.

Bell, T. 1987: A space-time stochastic model of rainfall for satellite romote-sensing studies. *J. of Geophys. Res.*, **92**, 9631-9643.

Bell, T., L., A. Abdullah, R. L. Martin and G. North, 1990: Sampling errors for satellite-derived tropical rainfall: Monte Carlo study using a space-time stochastic model. *J. Geophys. Res.*, **95**, 2195-2205.

Bell, T., L, and P. K., Kundu, 1996: A study of the sampling error in satellite rainfall estimates using optimal averaging of data and a stochastic model. *J. Climate*, **9**, 1251-1256.

Bell, T., L, P. K. Kundu, and C. D. Kummerow, 2001: Sampling errors of SSM/I and TRMM rainfall averages: Comparison with error estimates from surface data and a simple model. *J. Appl. Meteor.*, **40**, 938-954.

Bell, T., L, P. K. Kundu, 2003: Comparing satellite rainfall estimates with rain-gauge data: Optimal Strategies Suggested by a Spectral Model. *J. Geophys. Res.*, accepted for publication.

Brock, F. V. K. C. Crawford, R. L. Elliot, G. W. Cuperus, S. J. Stadler, H. J. Johnson, M. D. Eilts, 1995: The Oklahoma Mesonet: A technical overview. *J. Atmos. Oceanic Technol.*, **12**, 5-18.

Chang, A. T. C., and L. Chiu, 1999 : Nonsystematic errors of monthly oceanic rainfall derived from SSM/I. *Mon. Wea. Rev.*, **127**, 1630-1638.

Conner M.D., and G. W. Petty, 1998: Validation and intercomparison of SSM/I rain-rate retrieval methods over the continental United States. *J. Appl. Meteor.*, **37**, 679-700.

Doneaud, A, S. Ionescu-Niscov, D. L. Priegnitz, P. L. Smith, 1984: The area-time integral as an indicator of convective rain volumes. *J. App. Meteor.*, **23**, 555-561.

Ferraro and Marks, 1995: The development of SSM/I retrieval algorithms using ground-based radar measurments. *J. Atmos. Oceanic Technol.*, **12**, 755—770.

Habib, E., W. F. Krajewski, A. Kruger, 2001: Sampling Errors of Tipping-Bucket Rain Gauge Measurements. *J. of Hydr. Eng.*, **6**,159-166

Huff, F. A., 1970: Sampling errors in measurement of mean precipitation. *J. App. Meteor.*, **9**, 35-45.

Huffman, G. J.,1997: Estimates of root-mean-square random error for finite samples of estimated precipitation, *J. Appl. Meteor.*, **36**, 1191-1201.

Iguchi, T., T. Kozu, R. Meneghini, J. Awaka, K. Okamoto, 2000:Rain-profiling algorithm for the TRMM precipitation radar. *J. Appl. Meteor.*, **39**, 2038-2052.

Krajewski, W. F., G. J. Ciach, J, R. McCollum, C. Bacotiu, 2000: initial validation of the Global Precipitation Climatology Project monthly rainfall over the United States. *J. of Appl. Meteor.* **39**, 1071–1087.

Kummerow, C. W. Barnes, T. Kozu, J. Shiue, J. Simpson, 1998: The tropical rainfall measuring mission (TRMM) sensor package. *J. Atmos. and Oceanic Technol.*, **15**, 809-817.

Kummerow, C. W., et al., 2000: The status of the tropical rainfall measuring mission (TRMM) after two years in orbit. *J. Appl. Meteor.*, **39**, 1965-1982.

Kummerow, C. W., et al. 2001: The Evolution of the Goddard Profiling Algorithm (GPROF) for Rainfall Estimation from Passive Microwave Sensors.*J. Appl. Meteor.*, **40**, 1801–1820.

Laughlin, C. R., 1981: On the effect of temporal sampling on the observation of mean rainfall. Precipitation Measurements from Space, *Workshop Report.* D. Atlas and O. Thiele, Eds., NASA Publication, D59-D66.

McConnell, A., and G. North, 1987: Sampling errors in satellite estimates of tropical rain. *J. Geophys. Res.*, **92**, 9567-9570.

Meneghini, R., T. Iguchi, T. Kozu, L. Liao, K. Okamoto, J. A. Jones, J., Kwiatkowski, 2000: Use of the surface reference technique for path attenuation estimates from the TRMM precipitation radar. *J. Appl. Meteor.*, **39**, 2053-2070.

Mohr, K. I., and E. Zipser, 1996: Mesoscale convective systems defined by their 85 GHz ice scattering signature: size and intensity comparison over tropical oceans and continents. *Mon. Wea. Rev.*, **124**, 2417-2437.

Morrissey, M. L., 1991: Using sparse raingages to test satellite-based rainfall algorithms. *J. Geophys. Res.*, **31**, 18561-18571.

Morrissey, M. L., and J. S. Greene, 1993: Comparison of Two Satellite-based Rainfall Algorithms Using Pacific Atoll Raingage Data. *J. Appl. Meteor.*, **32**, 211-425.

Morrissey, M. L., J. A. Maliekal, J S. Greene, and J. Wang, 1995: The uncertainty of simple spatial averages using rain gauge networks. *Water Resour. Res.*, **31**, 2011-2017.

Morrissey, M.L. and J.S. Greene, 1998. Using the Oklahoma Mesonet to develop and test a sampling error statistic for meteorological time series. *J. Geophys. Res.*, **103**, 8979-8984.

North, G., and S. Nakamoto, 1989: Formalism for comparing rain estimation designs. *J. Atmos. and Oceanic Technol.*, **6**, 985-992.

North, G., 1992: Characteristics of tropical precipitation important for its estimation by satellites. *The Global Role of Tropical Rainfall*, J. S. Theon and N. Fugono, Eds., A. Deepak Publishing,169-181.

Rodriguez-iturbe, I., and J. M. Mejia, 1974a: The design of rainfall networks in time and space, *Water Resour. Res.*, **10**, 713 –728.

Rodriguez-Iturbe, I., J. M. Mejia, 1974b: On the transformation of point rainfall to areal rainfall. *Water Resour. Res.*, **10**, 729-735.

Salby, M. L., P. Callaghan, 1997:Sampling error in climate properties derived from satellite measurements: consequences of undersampled diurnal variability. *J. Climate*, **10**, 18-36.

Shin, K., and G. R. North, 1988: Sampling error study for rainfall estimate by satellite using a stochastic model. *J. Appl. Meteor.*, **27**,1218-1231.

Shin, K., G. R. North, Y. Ahn, P. A. Arkin, 1990: Time scales and variability of area-averaged tropical oceanic rainfall. *J. Appl. Meteor.*, **118**, 1507-1516.

Shin, D., L. S. Chiu, M. Kafatos, 2001: Comparison of the monthly precipitation derived from the TRMM satellite. *J. Geophy. Res. Let.*, **28**, 795-798.

Simpson, J., R. F. Adler, G. R. North, 1988: A proposed tropical rainfall measuring mission (TRMM) satellite. *Bull. Amer. Meteor. Soc*, **69**, 278-294.

Spencer, R., W. H., Goodman, R. E. Hood, 1989: Precipitation retrieval over land and ocean with the SSM/I: identification and characteristics of the scattering signal. *J. Appl. Meteor.*, **6**, 254-273.

Tao, W.-K., and J. Simpson, 1993: Goddard Cumulus Ensemble Model. Part I: Model description. *Terr. Atmos. Oceanic Sci.,* **4**, 35—72.

Wilheit, T. T., 1988: Error analysis for the Tropical Rainfall Measuring

Mission (TRMM). *Tropical Rainfall Measurements*, J. S. Theon

and N. Fugono, Eds., A. Deepak Publishing, 377—385.

Wilson, J. W, and E. A. Brandes, 1979: Radar measurement of rainfall summary. *Bull. Amer. Meteor. Soc*, **9**, 1048-1058.

Xie, P., and P. A. Arkin, 1995: An intercomparison of gauge observations and satellite estimates of monthly precipitation. *J. Appl. Meteor.*, **34,**1143-1160.

Zawadzki, I. I., 1975: On radar-raingage comparison, *J. Appl. Meteor.*, **14,** 1430-1436.

Figure Captions

FIG. 1. A schematic of Oklahoma showing the geographical partitioning of the 2° x 5° grid space into 1° x 1° cells. Cell numbers correspond to numbering in Table 1.

FIG. 2. Distribution of the sub-sampled rain parameter G_S for a) TMI and b) PR as a function of the time window, Δt_w.

FIG. 3. Regressions of G_0 versus G_S monthly rain estimates for a) TMI and b) PR at 1° x 1° resolution using monthly estimates (480 points); c) TMI and d) PR a 2° x 5° resolution using monthly estimates (48 points); and e) TMI and f) PR at 2° x 5° resolution using climatological 4-year monthly estimates (12 points).

FIG. 4. Distribution of sampling errors for the TMI (solid) and PR (dashed) for all four study years (1998-2001) computed at a resolution of 1° x 1° (480 data points).

FIG. 5. Four-year time series (1998-2001) of monthly rainfall estimates for a) G_0, G_S_TMI and TMI and b) G_0, G_S_ PR and PR.

FIG. 6. Regressions of monthly rain estimates for TMI vs. G_0 and G_S at different spatiotemporal resolutions: a) G_0 and b) G_S_TMI at 1° x 1° resolution using monthly estimates (480 points); and c) G_0 and d) G_S_TMI at 2° x 5° resolution using monthly estimates (48 points); and e) G_0, and f) G_S_TMI at 2° x 5° resolution 4-year monthly means (12 points).

FIG. 7. Regressions of monthly rain estimates for PR vs. G_0 and G_S at different spatiotemporal resolutions: a) G_0 and b) G_S_PR at 1° x 1° resolution using monthly

estimates (480 points); and c) G_0 and d) G_S_PR at 2° x 5° resolution using monthly estimates (48 points); and e) G_0, and f) G_S_PR at 2° x 5° resolution 4-year monthly means (12 points).

FIG. 8. Monthly probability density functions computed at 1° x 1° resolution for a) TMI, G_S and G_0; b) PR, G_S and G_0 and c) 3B42 (AGPI) and G_0. The dynamic range on the outer plot extends to 10 mm/day. The inner plot adjusts the scaling on the ordinate axis and extends the range from 10 to 25 mm/day. The inner plot displays the fine structure in the tail of each distribution.

FIG. 9. TMI and PR a) sampling biases and b) retrieval biases plotted for each year of the study.

FIG 10. Rain rate distributions for the TMI, PR and Oklahoma Mesonet gauges. TMI and PR profiles were constructed from 56 rain cases. Gauge profile was compiled from all data available over four study years (1998-2001).

FIG. 11 a) Bar plot of normalized monthly accumulations for the Oklahoma Mesonet rain gauges averaged over all four study years, with each month representing the fractional contribution to the total annual rainfall; 4-year monthly sampling, retrieval and total biases for b) TMI and c) PR computed at 2° x 5°.

FIG. 12. Normalized rainfall accumulations for the TMI and Oklahoma Mesonet gauges. Each hour represents the fractional contribution to the total diurnal rainfall.

TABLE 1. TRMM satellite products

TRMM Product	Sensor	Data Resolution	Data Description*
2A12	TMI	~6.9 km (85 GHZ)	Instantaneous surface rain rates
2A25	PR	~4 km	Instantaneous surface rain rates
RG2A12	TMI	0.5°	Instantaneous orbital track info, gridded instantaneous rain rates
RG2B31	Combined	0.1°	Instantaneous orbital track info
3A25	PR	0.5°	Gridded monthly rain information
3B31	TMI	1°	Gridded monthly rain information
3B42	Geo-IR	1°	Gridded monthly rain information

* only describes file information used in this study.

TABLE 2. Number of gauges per 1° x 1° cell

Cell No.	No. Gauges
1	5
2	6
3	8
4	9
5	7
6	5
7	6
8	7
9	5
10	8

TABLE 3. $\langle \Delta t_w \rangle$ and corresponding estimate of for TMI and PR

	TMI		PR	
Year	$\langle G_S \rangle$	$\langle \Delta t_w \rangle$	$\langle G_S \rangle$	$\langle \Delta t_w \rangle$
1998	261.0	29	268.4	29
1999	298.9	29	297.6	37
2000	373.4	31	401.6	31
2001	330.9	43	313.6	43

TABLE 4. Correlation coefficients for TMI, PR and G_S at grid resolutions of 1° x 1° and 2° x 5°.

	1° x 1° (all months)		2° x 5° (all months)		2° x 5° (4-yr monthly avg.)	
Sensor	G_0	G_S	G_0	G_S	G_0	G_S
TMI	0.44	0.26	0.55	0.53	0.73	0.71
PR	0.46	0.25	0.65	0.54	0.85	0.76
G_S_TMI	0.65	1.0	0.83	1.0	0.84	1.0
G_S_PR	0.63	1.0	0.84	1.0	0.90	1.0

Table 5. 4-year annual means and standard errors (sampling and retrieval) for the TMI and PR. Mean annual rainfall also shown for G_0.

	1° x 1° (all months)			2° x 5° (all months)			2° x 5° (4-yr monthly avg.)		
Sensor	μ (mm/day)	$\sigma_{err,S}$ (mm/day)	$\sigma_{err,R}$ (mm/day)	μ (mm/day)	$\sigma_{err,S}$ (mm/day)	$\sigma_{err,R}$ (mm/day)	μ (mm/day)	$\sigma_{err,S}$ (mm/day)	$\sigma_{err,R}$ (mm/day)
TMI	2.9	2.31	3.93	2.9	1.31	2.37	2.5	0.59	1.05
PR	3.4	2.71	4.62	3.4	1.42	2.42	2.8	0.90	0.99
G_0	2.5			2.3			1.9		

Figures

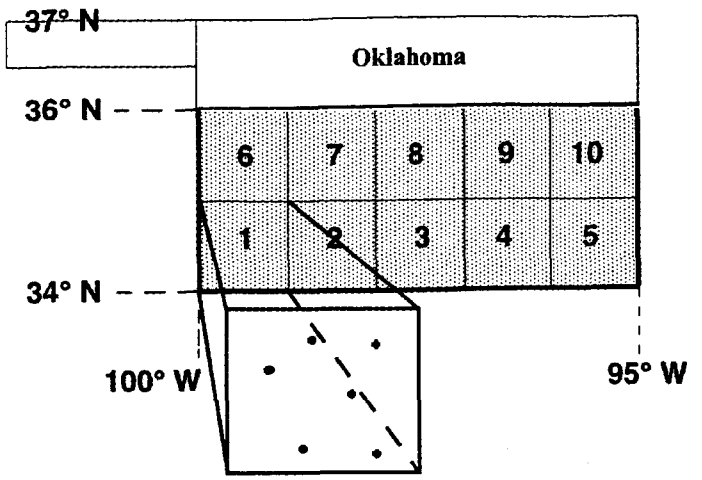

FIG. 1. A schematic of Oklahoma showing the geographical partitioning of the 2° x 5° grid space into 1° x 1° cells. Cell numbers correspond to numbering in Table 1.

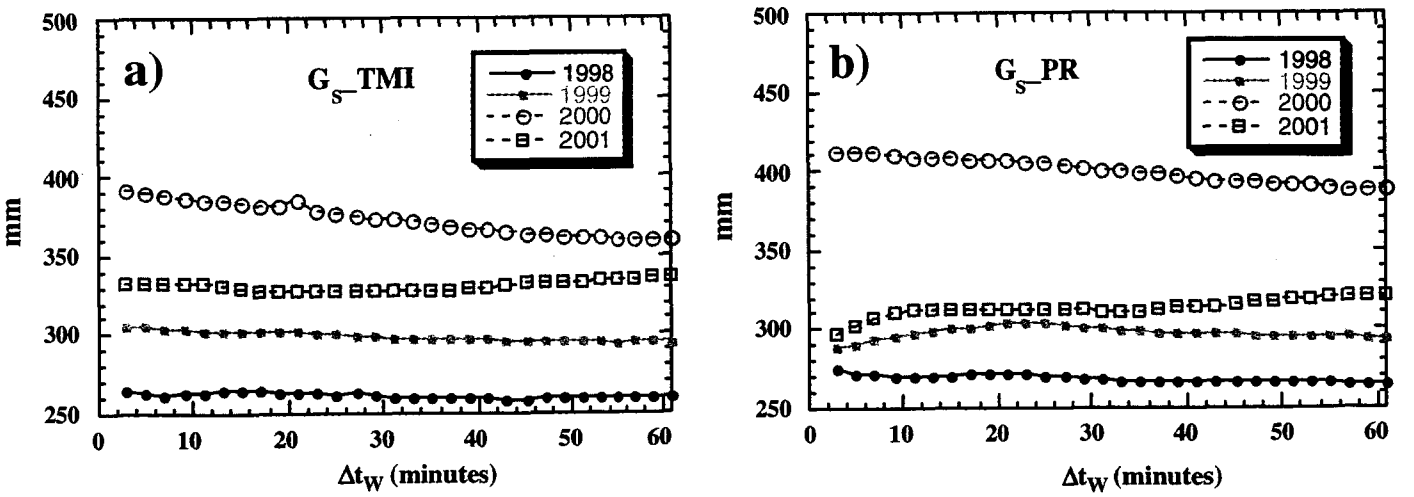

FIG. 2. Distribution of the sub-sampled rain parameter G_S for a) TMI and b) PR as a function of the time window, Δt_W.

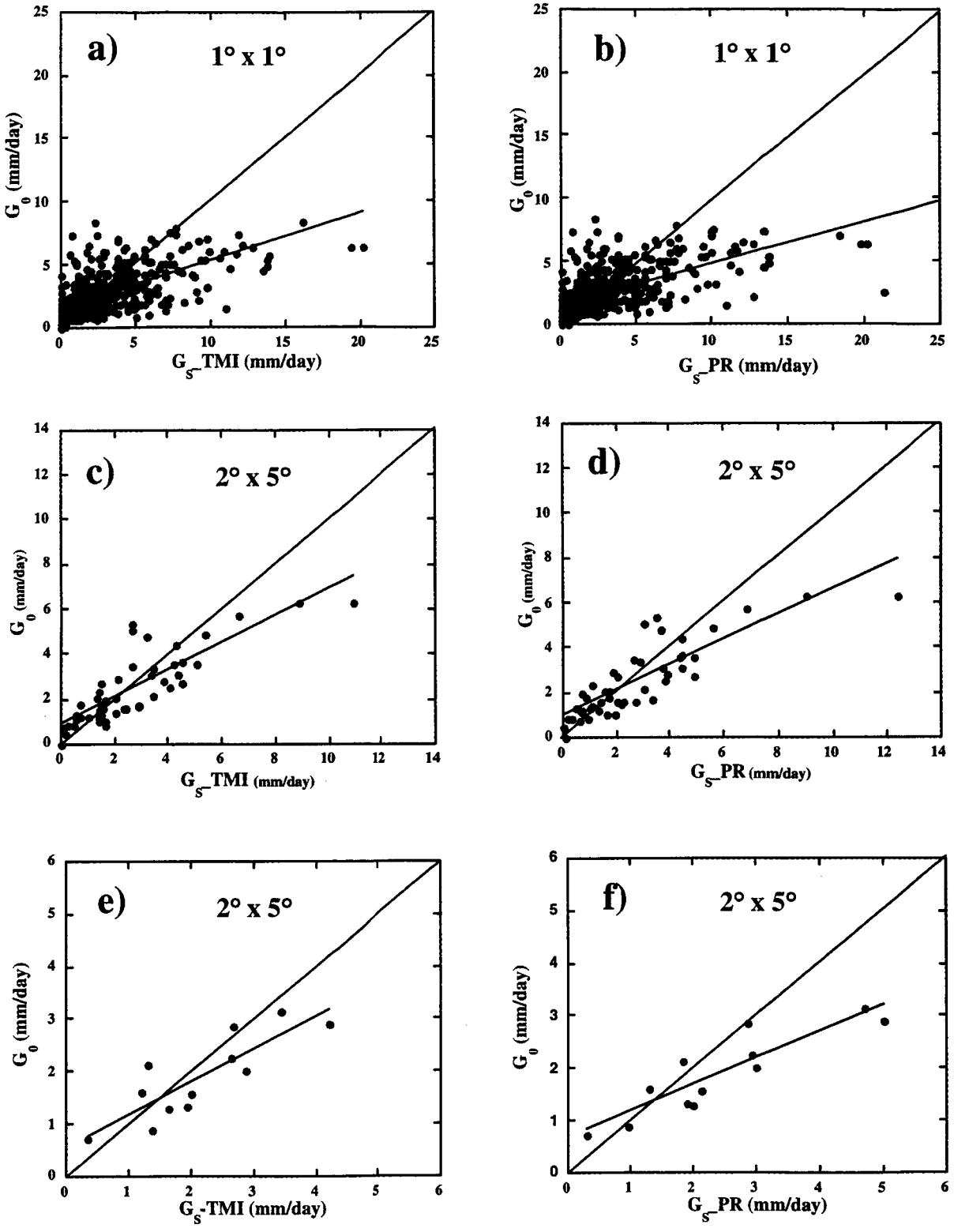

FIG. 3. Regressions of G_0 versus G_S monthly rain estimates for a) TMI and b) PR at 1° x 1° resolution using monthly estimates (480 points); c) TMI and d) PR a 2° x 5° resolution using monthly estimates (48 points); and e) TMI and f) PR at 2° x 5° resolution using climatological 4-year monthly estimates (12 points)..

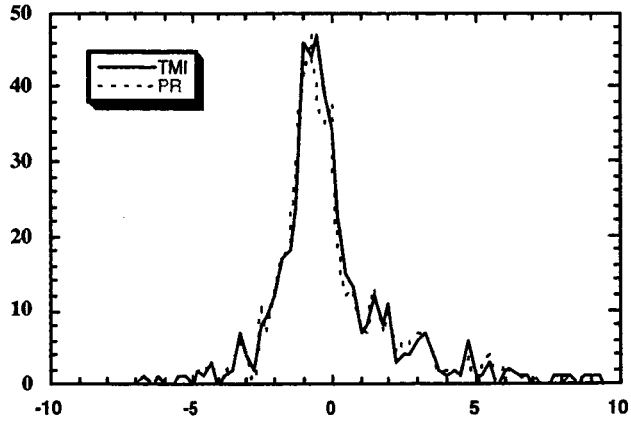

FIG. 4. Distribution of sampling errors for the TMI (solid) and PR (dashed) for all four study years (1998-2001) computed at a resolution of 1° x 1° (480 data points).

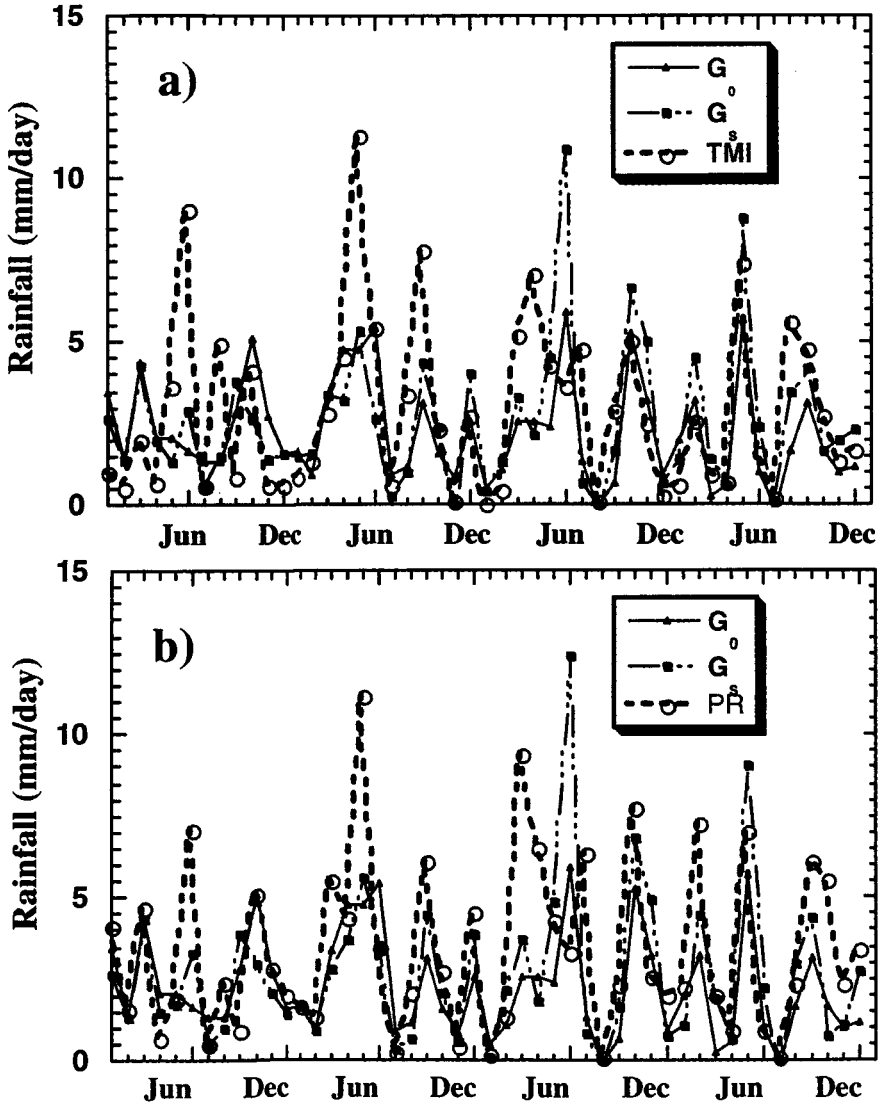

FIG. 5. Four-year time series (1998-2001) of monthly rainfall for a) G_0, G_S_TMI and TMI and b) G_0, G_S_PR and PR.

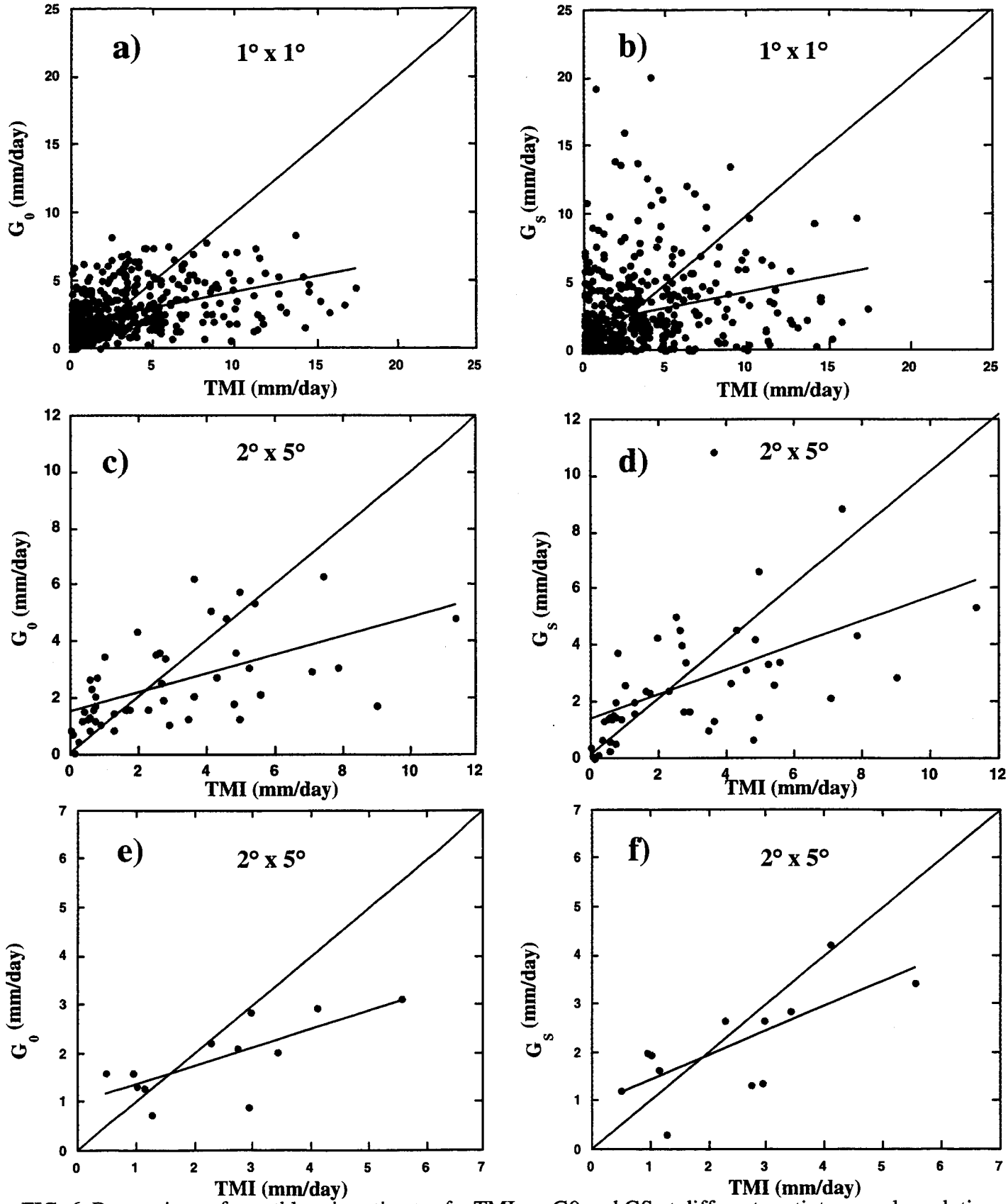

FIG. 6. Regressions of monthly rain estimates for TMI vs. G0 and GS at different spatiotemporal resolutions: a G0 and b) GS_TMI at 1° x 1° resolution using monthly estimates (480 points); and c) G0 and d) GS_TMI at 2 x 5° resolution using monthly estimates (48 points); and e) G0, and f) GS_TMI at 2° x 5° resolution 4-year monthly means (12 points).

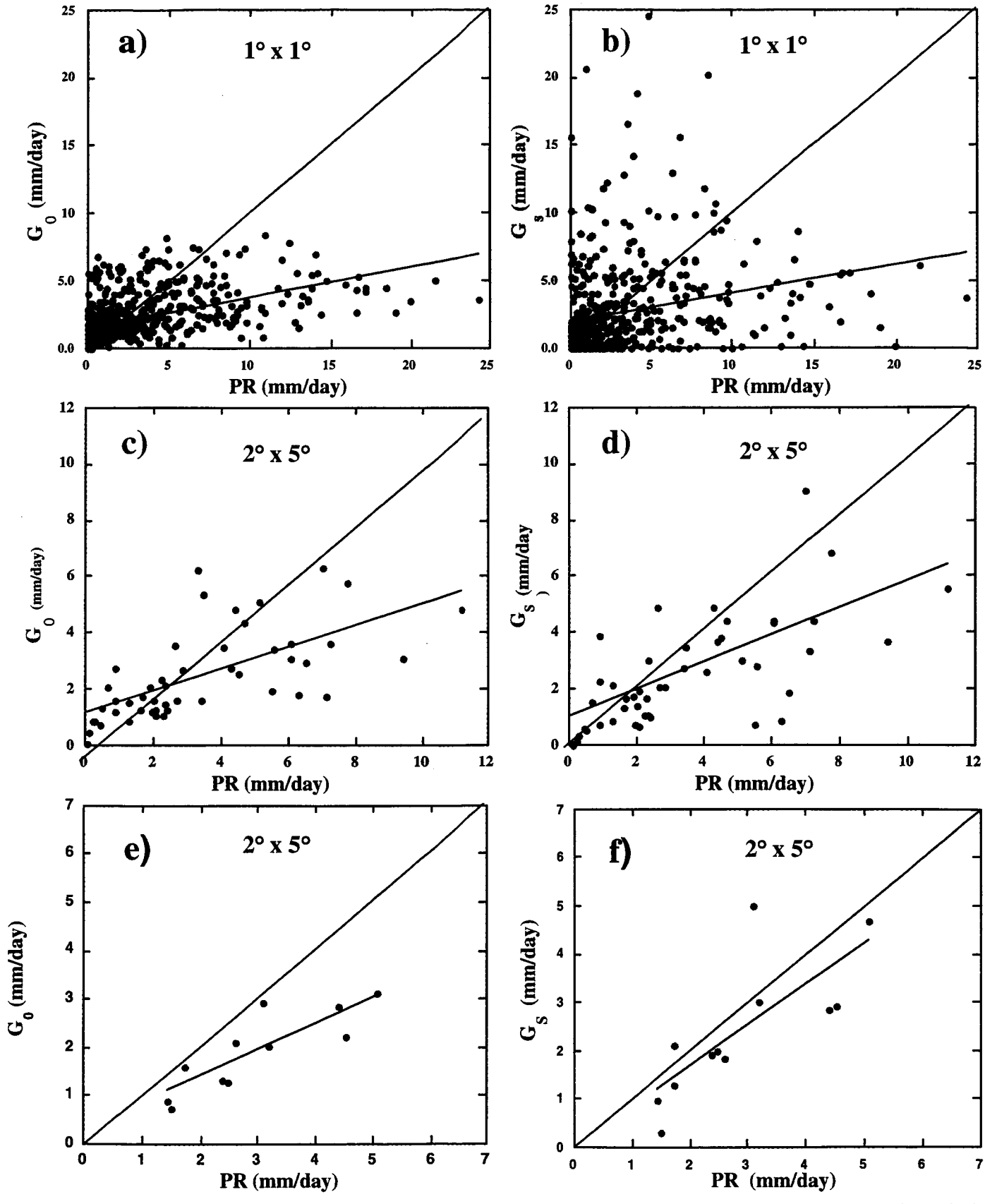

FIG. 7. Regressions of monthly rain estimates for PR vs. G0 and GS at different spatiotemporal resolutions: a) G0 and b) GS_PR at 1° x 1° resolution using monthly estimates (480 points); and c) G0 and d) GS_PR a 2° x 5° resolution using monthly estimates (48 points); and e) G0, and f) GS_PR at 2° x 5° resolution 4-year monthly means (12 points).

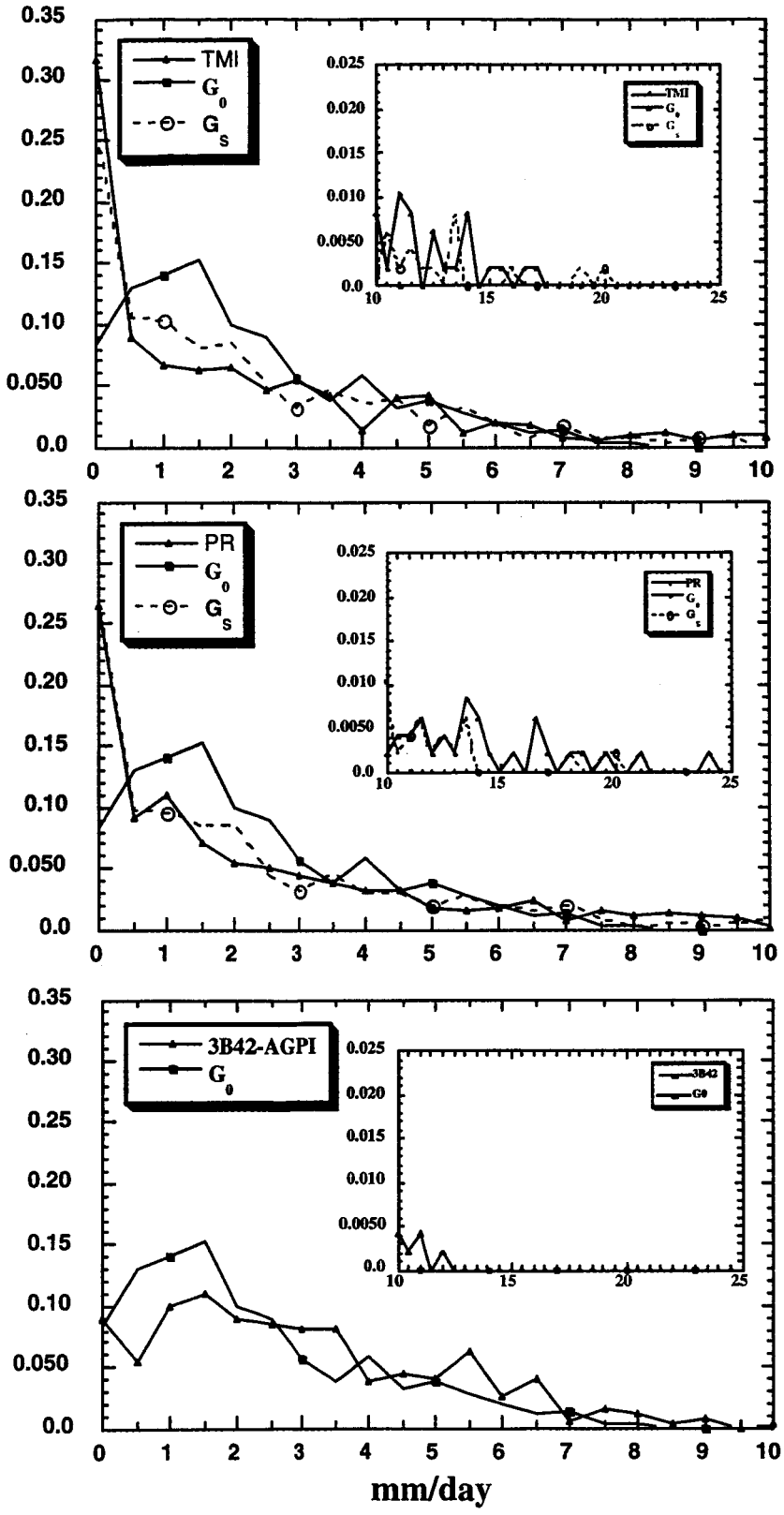

FIG. 8. Monthly probability density functions computed at 1° x 1° resolution for a) TMI, G_S and G_0; b) PR, G_S and G_0 and c) 3B42 (AGPI) and G_0. The dynamic range on the outer plot extends to 10 mm/day. The inner plot adjusts the scaling on the ordinate axis and extends the range from 10 to 25 mm/day. The inner plot displays the fine structure in the tail of each distribution.

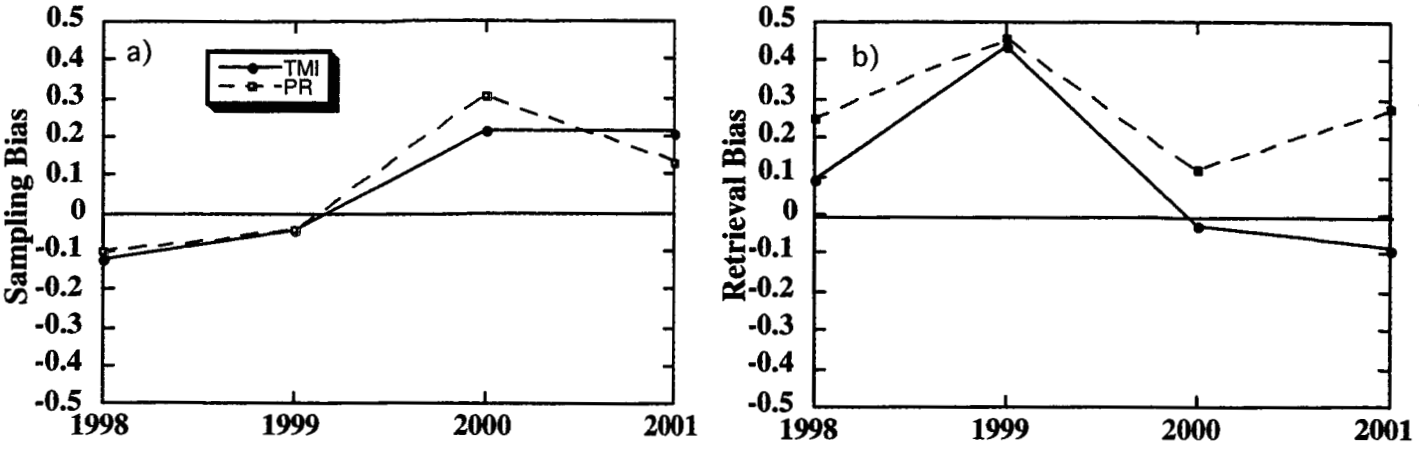

FIG. 9. TMI and PR a) sampling biases and b) retrieval biases plotted for each year of the study.

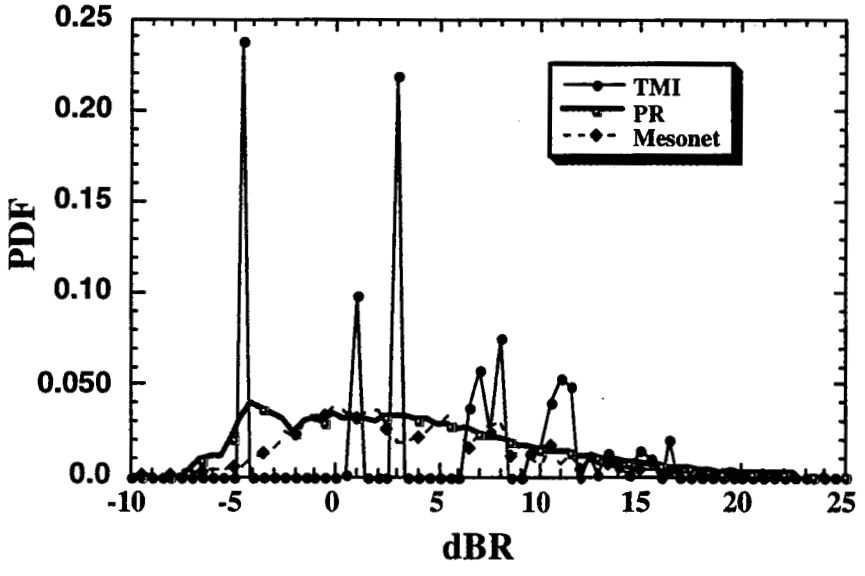

FIG 10. Rain rate distributions for the TMI, PR and Oklahoma Mesonet gauges. TMI and PR profiles were constructed from 56 rain cases. Gauge profile was compiled from all data available over four study years (1998-2001).

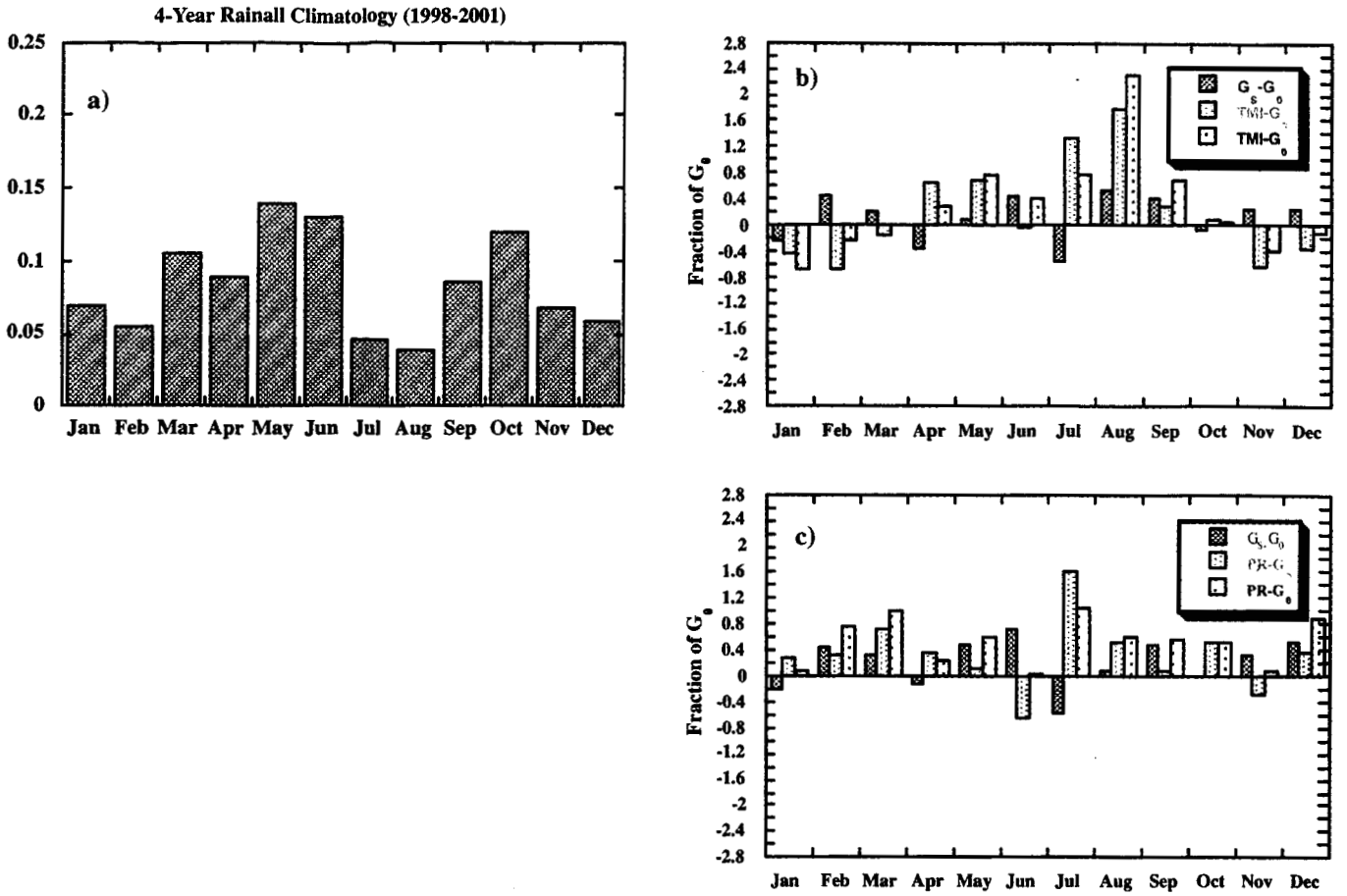

FIG. 11 a) Bar plot of normalized monthly accumulations for the Oklahoma Mesonet rain gauges averaged over all four study years, with each month representing the fractional contribution to the total annual rainfall; 4-year monthly sampling, retrieval and total biases for b) TMI and c) PR computed at 2° x 5°.

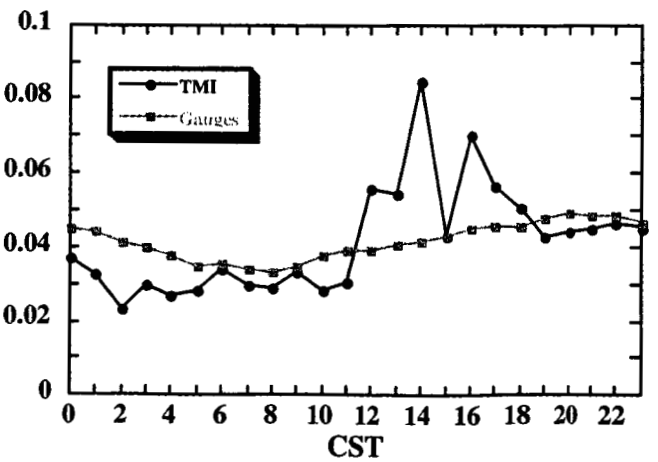

FIG. 12. Normalized rainfall accumulations for the TMI and Oklahoma Mesonet gauges. Each hour represents the fractional contribution to the total diurnal rainfall.