Verifying Satellite Microwave Rainfall Estimates over the Open Ocean

MARK L. MORRISSEY*

Oklahoma Climate Survey and Department of Geography, University of Oklahoma, Norman, Oklahoma

YIPING WANG

Department of Meteorology, University of Oklahoma, Norman, Oklahoma

(Manuscript received 24 January 1994, in final form 1 July 1994)

ABSTRACT

Satellite rainfall estimates from a microwave emission–based algorithm by Wilheit et al. are verified using the noncontiguous rain gauge method incorporating monthly Pacific atoll rain gauge data. The results are compared with those obtained using an infrared-based satellite algorithm, the GOES precipitation index. Comparisons between satellite estimates with simple spatial averages of point rain gauge data are shown to be ineffective at identifying statistically significant differences between the two algorithms due to substantial amounts of spatial sampling error in the rain gauge spatial averages. By effectively reducing this error, the noncontiguous rain gauge method reveals distinctive differences in the ability of each of the algorithms to accurately estimate monthly rainfall over the open ocean. The results indicate that the microwave algorithm, while slightly biased, is significantly less biased than the infrared, which tends to overestimate high rainfall values and underestimate low rainfall values. However, the random error associated with both algorithms is essentially the same.

1. Introduction

The measurement of precipitation is important to monitor and detect global climate change. The advent of meteorological satellites provided the only means for obtaining estimates of large-scale convective precipitation over the greater portion of the earth and, in particular, over the tropical oceans where other types of instruments are sparsely distributed.

A wide variety of satellite-based algorithms transforming received radiance to rainfall have been developed (Barrett and Martin 1981). Rain-rate estimation techniques using satellite information are based on the empirical or physical relationships between rain rate and cloud variables, such as cloud height, cloud-top temperature, high-cloud amount or scattering, and emissivity properties from raindrops. This information is obtained from visible (VIS), infrared (IR), and microwave (MW) measurements, or a combination of these wavelengths.

An early approach to estimating tropical convective rainfall using satellite imagery from National Oceanic and Atmospheric Administration polar-orbiting satellites was developed by Kilonsky and Ramage (1976). The technique was based upon the relationship between the number of highly reflective clouds within 1.0° × 1.0° latitude–longitude boxes per month and monthly Pacific island rain gauge measurements. The number of highly reflective clouds was determined subjectively from visual imagery. By introducing IR data into their analysis scheme, Garcia (1985) constructed the highly reflective cloud index, which relates the number of occurrences of highly reflective clouds per box to monthly rainfall.

Using satellite IR measurements, Arkin (1979) and Richards and Arkin (1981) found that the fractional amount of high clouds within an area was strongly correlated with the areal rain amount observed during the Global Atmospheric Research Program (GARP) Atlantic Tropical Experiment (GATE). Arkin and Meisner (1987) used these results to develop a simple IR thresholding technique for estimating areal rainfall [i.e., the GOES (Geostationary Operational Environmental Satellite) precipitation index (GPI)].

Wilheit et al. (1991) developed a physical approach to estimating rain rate using MW measurements based upon the earlier work of Wilheit et al. (1977), who established a relationship between columnar rain rate and received MW radiance. By assuming a form for the probability density function (PDF) of rainfall intensity, the algorithm retrieves the monthly rainfall from MW emission of hydrometeors. Herein, monthly average rainfall estimates produced by the PDF method are termed WCC rain rate estimates (Chiu et al. 1993).

---

* Also a visiting scientist in Planetary Geosciences, University of Hawaii, Honolulu, Hawaii.

Corresponding author address: Dr. Mark L. Morrissey, Oklahoma Climatological Survey, University of Oklahoma, Sarkeys Energy Center, Suite 1210, 100 East Boyd, Norman, OK 73019-0628.

To be of practical use, satellite algorithms must be calibrated and their estimates verified using samples of independent data accurately taken over time- and space scales corresponding to those of the satellite estimates. A common validation method is to apply regression analysis to corresponding pairs of surface and satellite volume rainfall estimates. Biases and random errors associated with the satellite algorithm can then be identified. Crucial to the success of the comparison is the validity of the assumption that the surface volume rainfall estimates are accurate. This depends upon rain gauge instrument error and spatial sampling error. The magnitude of the spatial sampling error depends upon the relationship among the rain gauge density, the network geometry, and the spatial covariance structure of rainfall. In addition, temporal sampling error arising from the discrete satellite visiting times within the sampling space can produce significant differences between the surface and satellite estimates.

Efforts to quantify the temporal sampling error associated with different satellite systems have been made (Laughlin 1981; McConnell and North 1987; Shin and North 1988). For the soon to be launched Tropical Rainfall Measuring Mission (Simpson et al. 1988) satellite, the temporal sampling error based upon GATE rainfall statistics is about 8%–12% of estimated monthly mean rain rates over a grid box of 5° × 5° (Shin and North 1988). However, since the sampling configuration of satellites is difficult to alter once they are in orbit and accumulated total rainfall is the quantity to be estimated, the variance due to temporal sampling error remains in the satellite estimates, and, thus, the focus of validation efforts should be placed in removing the spatial sampling error in the areal rain gauge estimates. One such effort by Morrissey (1991) has been to create a hybrid validation method that effectively reduces the spatial sampling error. The method, referred to as the noncontiguous rain gauge method (i.e., NCR method), incorporates the theoretical relationship between the point and areal rainfall distribution to construct statistical models from widely scattered rain gauge measurements. The models provide values that have the statistical distribution of areal rainfall conditional on the satellite value. The values are paired with satellite rainfall estimates using a scattered diagram. Regression analysis can then be applied to assess the bias and random error associated with the satellite algorithm with the scatter due to spatial sampling error minimized. The NCR method is directly applicable to validating satellite estimates over open ocean regions where only widely scattered rain gauge measurements are available. Morrissey and Greene (1993) used the NCR method to verify two simple satellite rainfall algorithms within the different rainfall climate regimes within the tropical Pacific.

The purpose of this study is to apply the NCR method over the tropical Pacific using monthly rain gauge data obtained from the Comprehensive Pacific Rainfall Data Base (CPRDB) (Morrissey et al. 1993b) to verify the WCC satellite algorithm, which uses MW data from the Special Sensor Microwave/Imager (SSM/I). The SSM/I is onboard the polar-orbiting Defense Meteorological Satellite Program (DMSP) satellite. Recently, Chang et al. (1993) compared the performance of the WCC algorithm over the tropical Pacific region using data from rain gauges sited on Pacific atolls (Morrissey and Greene 1991; Morrissey et al. 1993a). Monthly rain gauge measurements within the 5° × 5° latitude–longitude areas represented by the satellite estimates were spatially averaged to produce volume (i.e., volume = area × time) estimates. Based upon a comparison of the rain gauge and satellite estimates, the algorithm produced estimates with random errors of 50%–60% and a low bias of approximately 8%. The correlation coefficient between the rain gauge and the algorithm estimates was 0.70. However, it must be assumed that a significant portion of the random error is due to spatial sampling error in rain gauge–estimated volume rainfall since the rain gauge density in the Pacific is quite low, and the spatial scale of the tropical rainfall process is small (Morrissey and Greene 1993). Thus, the correlation coefficient found by Chang et al. (1993) must be considered as a minimum value with its actual value unknown. It will be demonstrated that, by minimizing the spatial sampling error, the NCR method provides accurate estimates of both the bias and the random error in the satellite algorithm. Hence, the correlation coefficient is accurately determinable.

For comparison purposes, the GPI will also be verified over the same time and space scales. The GPI has often been used as a standard by which other satellite algorithms have been compared (Chiu et al. 1993; Berg 1993; Morrissey and Greene 1993). The NCR method will be described briefly in section 2. The two satellite algorithms are discussed in sections 3 and 4, and the results of the verification are shown and discussed in section 5. Finally, the conclusions are given in section 6.

2. The NCR method

Satellite rainfall estimates are verified by statistically comparing them with surface estimated rainfall over the same time- and space scales. Surface volume rainfall is usually estimated using spatial averages of accumulated rainfall totals obtained from closely spaced rain gauges in order to minimize spatial sampling error. Assuming homogeneous and stationary statistics, rain gauge–estimated volume rainfall has a statistical distribution corresponding to every value of satellite-estimated rainfall. These distributions are thus "conditional" on the satellite estimates. Assuming a rain gauge network density of infinity, the first moments of these distributions characterize the biases inherent in the satellite algorithm, and the second moments characterize its precision. For rain gauge densities less than

infinity, these distributions become sample distributions whose second-order moments are larger than those of the population due to sampling error. This can significantly affect assessments of the bias and random error associated with the satellite algorithm by reducing the statistical significance of the fitted regression parameters. Although the sampling error can be reduced by increasing the rain gauge density, this is a difficult and expensive task over open ocean regions.

Morrissey (1991) developed, and Morrissey and Greene (1993) later applied, an alternative method of verifying satellite rainfall estimates using a statistical modeling approach. The theoretical basis of the NCR method, which lies in the transformation of the point rainfall[1] distribution into a distribution of volume rainfall estimates over the desired scales, is described in detail by Morrissey (1991) and Morrissey and Greene (1993). The NCR method incorporates the functional relationship between the variance of estimated volume rainfall and rain gauge density to construct statistical models whose output have the distribution of volume rainfall estimated from networks of a specified gauge density. Estimates of the model parameters are determined using data obtained from sparsely distributed rain gauges. Since the rain gauge density is a parameter in these models, its value may be selected so as to minimize the sampling error. A different model is constructed for each value of a quantized satellite rainfall estimate. A scatter diagram can then be produced relating model output to the corresponding satellite values. Linear or nonlinear regression analyses may be then applied to the pairs of model and satellite-estimated values. By forcing the number of model values to be a function of the number of rain gauge data available, the number of degrees of freedom is preserved allowing tests of significance of the resulting regression relationship.

To apply this method, it must be assumed that rainfall is a random process and that its statistics are homogeneous conditional on the satellite rainfall value. This is a more realistic assumption than one requiring the rainfall statistics to be homogeneous independent of satellite values. For example, it is known that the mean rainfall rate varies from one climate regime to another. However, the mean rainfall rate associated with a given satellite value may be the same in both climate regimes. This, of course, depends upon the relationship between satellite-estimated and actual rainfall. Since conditional homogeneity is assumed, the model parameters may be determined using rain gauges distributed over a considerably larger region than the spatial dimensions of the satellite rainfall estimate.

The initial step in the NCR method requires that all rain gauge values in a selected region be sorted into several datasets. Each rainfall dataset contains rain gauge measurements taken simultaneously in time and space with a quantized satellite rainfall estimate. The quantization interval of the satellite values depends upon the number of associated rain gauge values available per dataset. A good rule of thumb is that the satellite values should be quantized so that an adequate estimate of the point rainfall distribution can be made for each dataset. The flowchart shown in Fig. 1 illustrates the sequence of steps required to construct a statistical rainfall model for a given quantized satellite value. This procedure should be repeated for each quantized satellite value.

The point rain gauge values from a given dataset are used to estimate the first- and second-order moments of the point rainfall distribution. The conditional spatial correlation function is then estimated by fitting (using a least-squares method) an appropriate function to interstation correlation coefficients plotted as a function of rain gauge separation distance (to simplify the description an isotropic spatial correlation structure will be assumed). To increase the number of rain gauge pairs so that an adequate estimate of the spatial correlation function can be made, an average rainfall correlation structure representing a range of quantized satellite rainfall values can be used (Morrissey and Greene 1993).

The next step involves the determination of the conditional volume rainfall variance using the variance reduction relationship

$$\mathrm{var}(\hat{P}_s) = F_s(N)\,\mathrm{var}(\hat{p}_s), \qquad (1)$$

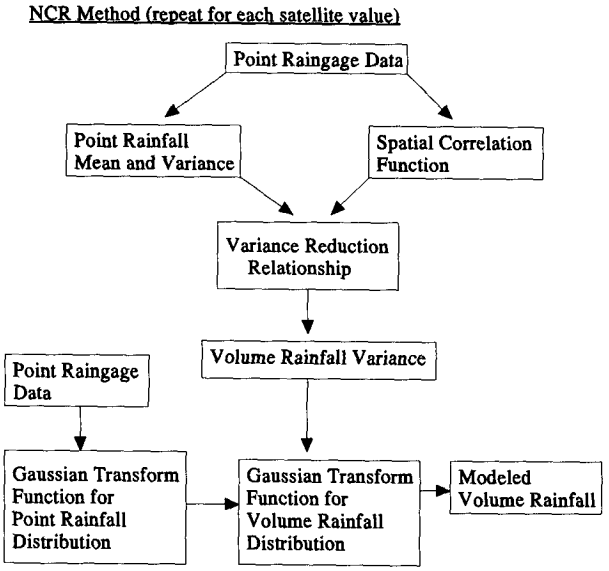


FIG. 1. The flowchart illustrating the sequence of steps required in the NCR method to acquire model rainfall for a given quantized satellite value.

[1] Since rain gauges estimate accumulated rainfall, point estimates in this paper refer to individual rain gauge accumulations over one month.

where $\text{var}(p_s)$ is the point rainfall variance conditional of satellite value s. The resultant conditional volume rainfall variance $\text{var}(P_s)$ is equivalent to the variance of volume rainfall estimates obtained from a single rain gauge network (if one were available), consisting of a number of randomly distributed rain gauges of density N (i.e., number of rain gauges per area). The conditional variance reduction factor $F_s(N)$ (Rodriguez-Iturbe and Mejia 1974; Morrissey 1991) is equal to

$$F_s(N) = \frac{1}{AN}\{1 + (AN - 1)E[\bar{\rho}_s|A]\}, \quad (2)$$

where A is the area corresponding to the spatial scale of the satellite rainfall estimates. The expected value of the conditional mean interstation correlation within areas of size A is given by

$$E[\bar{\rho}_s|A] = \int_0^{\text{DMAX}} \rho_s(\nu)f(\nu)d\nu, \quad (3)$$

where the conditional spatial correlation $\rho_s(\nu)$ is written as a function of rain gauge separation distance ν, $f(\nu)$ is the frequency function of the distance between two randomly distributed points in area A (appendix A), and DMAX is the diagonal distance across a square area of size A.

The variance reduction relationship accounts for the reduction in the variance of volume rainfall estimated from rain gauge networks having a density of N in the presence of a spatially correlated field. A value for N is selected analytically by first deciding on an acceptable amount of sampling error (e.g., 5%) (refer to Morrissey and Greene 1993).

The advantage of the NCR method is that values from widely scattered rain gauges can be used to determine values for the parameters in Eq. (1), so long as the spatial correlation function is adequately estimated. An unbiased estimate of the first-order moment of the volume rainfall distribution is the sample mean point rainfall value. Once the first- and second-order moments associated with each of the conditional rainfall distributions are estimated, the construction of the statistical models can begin. This procedure, which is fully described by Morrissey (1991) and Morrissey and Greene (1993), will be described briefly below.

The modeling process begins by defining a random variable p_s representing the conditional point rainfall histogram. The Gaussian transform function $\varphi_s(U)$ is then defined, which expresses the random variable p_s in terms of a standard normal random variable U [i.e., $p_s = \varphi_s(U)$]. Using an expansion of Hermite polynomials, the transform function is fitted to the conditional point rainfall histogram (Journel and Huijbregts 1978; Morrissey 1991). The point rainfall variance appears as one of the parameters defining the transform function. It is assumed that the transform function of the random variable representing volume rainfall is the same as that representing point rainfall with the point variance replaced by the volume variance. Standard normal random numbers are used as input to the model. This results in values having the same mean and variance and similar higher-order moments as would volume rainfall estimated from a network of randomly distributed rain gauges having a density N.

Once the models are constructed, their output is compared with the appropriate quantized satellite value using a scatter diagram. The number of values produced from each model is dependent upon the number of rain gauge values available for a given dataset (i.e., M_s), and is given by M_s/N_s (refer to Morrissey and Greene 1993). Linear or nonlinear regression analysis is then applied to the values on the scatter diagram and significance tests conducted.

The most important factor concerning the application of the NCR method is the estimation of the spatial correlation function. Frequently, there are an insufficient number of rain gauge pairs to estimate a separate spatial correlation function for each quantized satellite estimate. In this case, a spatial correlation function representing a range of satellite estimates may be computed (refer to Morrissey and Greene 1993). This maximizes the number of pairs available without significantly affecting the results of the comparison. For verifying the satellite algorithms over the tropical Pacific, a special weighting scheme was devised (described in appendix B) and applied to the WCC and GPI algorithms.

3. Satellite algorithms

a. The WCC algorithm

Considerable work has been placed in developing rainfall retrieval methods using SSM/I data (e.g., Spencer et al. 1989; Olson 1989; Grody 1991; Liu and Curry 1992). Some algorithms have been specifically designed to estimate large-scale, long-term rainfall estimates suitable for climate research (Wilheit et al. 1991; Berg 1993).

Wilheit et al. (1991) developed an approach to estimating rainfall integrated over the time- and space scales of climatological interest using operationally available SSM/I MW data. The SSM/I is a four-frequency (19.35, 22.235, 37, and 85.5 GHz) microwave radiometer mounted on the near polar-orbiting, sun-synchronous DMSP satellite (WCRP 1986). The spatial resolution of the imagery varies from 69 km × 43 km at 19.35 GHz to 15 km × 13 km at 85.5 GHz. The first two channels are attenuation-based measurements representing the oxygen band (i.e., 50–60 GHz). The last channel is a scattering-based measurement within the water vapor line (i.e., 22.235 GHz). The 37-GHz channel is the intermediate frequency channel in which both measurement approaches must be considered. The emission-based algorithm developed by Wilheit et al. (1991) uses a linear combination of the 19.35- and

22.235-GHz channel observations to minimize the impact of water vapor.

The technique of Wilheit et al. (1991) is based upon the Wilheit et al. (1977) radiative transfer model, which establishes an analytical relationship between brightness temperature and rain rate. Two basic assumptions in the algorithm are that 1) the probability distribution function of rainfall intensity has a lognormal form, and 2) the nonraining portion of the brightness temperature histogram has a normal distribution.

Retrieving monthly rainfall totals from the SSM/I data begins with the collection and assimilation of SSM/I data into frequency histograms of brightness temperatures (BT). The BT histogram allows the initial estimation of the rain-rate distribution moments assuming a lognormal distribution. The mean height of the rain clouds is also estimated from the BT histogram. These rain rates are then translated into BT using the radiative transfer function. The nonraining part of the BT histogram is treated as a normal distribution. By comparing the computed BT histogram with the observed BT histogram adjustments of the lognormal rain rate, parameters can be made and further iterations performed until the first three moments of the two BT histograms match. Finally, the resulting values of the lognormal parameters determine the monthly rain totals that are retrieved for each 5° × 5° box over oceanic areas.

b. The GPI algorithm

The GPI rainfall estimates are derived from geostationary IR imagery using a simple thresholding technique that was presented by Arkin and Meisner (1987). The technique is based upon the relationship between the fractional coverage of cold cloud within 2.5° × 2.5° boxes and area-averaged rainfall. Richards and Arkin (1981) compared radar rain rates taken during GATE with satellite IR imagery and found that the linear relationship between areally averaged rainfall and fractional coverage by cold clouds is relatively insensitive to the threshold chosen. Furthermore, they indicate that the optimum threshold for estimating convective rainfall in 2.5° × 2.5° boxes is 235 K. For this threshold, the slope of the regression relation between fractional cloudiness and rainfall rate is relatively stable with respect to the temporal averaging scale. For the GPI, this coefficient was determined to be 3 mm h^{-1}. The GPI relation is expressed as

$$\text{rainfall (mm)} = 3.0t\ \text{Fc},$$

where Fc is the fractional cloudiness (a dimensionless number between 0 and 1), and t is the length of the period (h) for which Fc was the mean fractional cloudiness. The GPI was found to work best in tropical oceanic regions where convectively produced rainfall predominates.

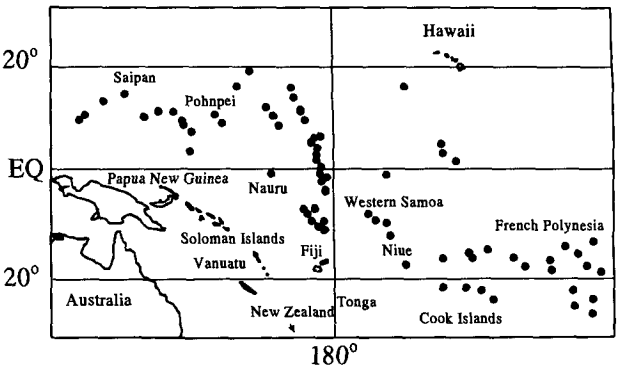


FIG. 2. Locations of the Pacific atoll rain gauge sites used for this study.

4. Data

This study uses rain gauge data extracted from the CPRDB, which consists of daily rain gauge data from over 250 Pacific island and atoll stations (Morrissey et al. 1993b). Due to the questionable representation of large island rainfall measurements, only rainfall from gauges sited on atolls were in this study. This amounted to approximately 99 sites across the Pacific (Fig. 2). On the average, about 70% of the stations reported on any given month.

The WCC rainfall estimates were obtained from the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center.[2] Because the DMSP satellite is in a sun-synchronous orbit, the observations are restricted to twice a day over a given location. Thus, two datasets are obtained separately and are referred to as AM and PM, respectively. The monthly rain rate estimates are the average of the AM and PM values and the results are corrected for beam-filling error. The equation used is

$$\text{rainfall (mm month}^{-1}) = \frac{(\text{AM} + \text{PM})}{2.0}\,1.5.$$

The dataset covers the global belt from 50°N to 50°S from July 1987 to December 1990 with some records during 1987 missing.

The GPI data were obtained from the Global Precipitation Climatology Project Geostationary Satellite Precipitation Center.[3] The GPI rainfall estimates for the Pacific region were derived using IR data from both the Japanese Geostationary Meteorological Satellite (GMS) and the GOES satellite. The IR data constituting the GPI have a temporal resolution of three hours and a spatial resolution of approximately 11 km (at nadir). The spatial resolution of the GPI data is

[2] Courtesy of Alfred Chang, NASA Goddard Space Flight Center, Greenbelt, Maryland.

[3] Courtesy of John Janowiak, Climate Analysis Center.

2.5° × 2.5° latitude–longitude, and the time resolution is five days. The period of record is five years (1986–90). The GPI rainfall estimates represent the region between 40.0°N and 40.0°S. To match the scales of the WCC estimates, the original pentad GPI estimates were averaged into monthly values and spatially into 5° × 5° resolution.

5. Results and discussion

a. Comparisons of satellite-estimated areal rainfall with point surface measurements

To provide a baseline comparison, point rainfall measured concurrently in time and space with the satellite estimates were compared. It will be demonstrated that, although some conclusions may be drawn from such a comparison, concrete statements concerning the statistical validity of the results cannot be made due to the influence of spatial sampling error and that the NCR method must be applied under these circumstances. Figures 3 and 4 show the resulting comparison between spatially averaged point rain gauge data and the satellite-estimated areal rainfall from the GPI and the WCC data, respectively. There are 2496 pairs of such comparisons for the GPI and 2567 pairs for the WCC estimates. The thick solid lines, which are the regression lines computed using the least-squares method, provide an estimate of the functional relationship between the surface point data and the areal satellite estimates. The thin solid lines are the 45° lines drawn to illustrate the bias in the relationship with respect to rainfall amount.

Overall, the satellite values increase with increasing point rainfall values. However, the error variance of point rainfall, indicated by the scatter about the regression lines, is substantial for both algorithms. The slopes of the regression lines are 0.81 for the WCC values and 0.73 for the GPI data. Although both algorithms produce biased values (i.e., the slopes are different than 1.0), it appears the WCC estimates are somewhat less biased than the GPI values. The correlation between the WCC estimates and rain gauge data is 0.66, which is quite close to that found by Chang et al. (1993). The correlation between the GPI estimates and rain gauge data, while higher at 0.69, is not statistically significantly higher (at the 95% level). A positive intercept is indicated by both algorithms. The intercept is 48.3 mm month^{-1} for the WCC estimates and 57.4 mm month^{-1} for the GPI.

Residual analysis showed that the data pairs associated with both algorithms were nonnormally distributed and the error variance was not constant. A Box–Cox transformation (Draper and Smith 1981) was applied to the data pairs that satisfactorily stabilized the variance and produced a quasi-normal distribution. Statistical significance tests (not shown) were then applied to the regression parameters associated with the transformed data. The results of the tests indicated that

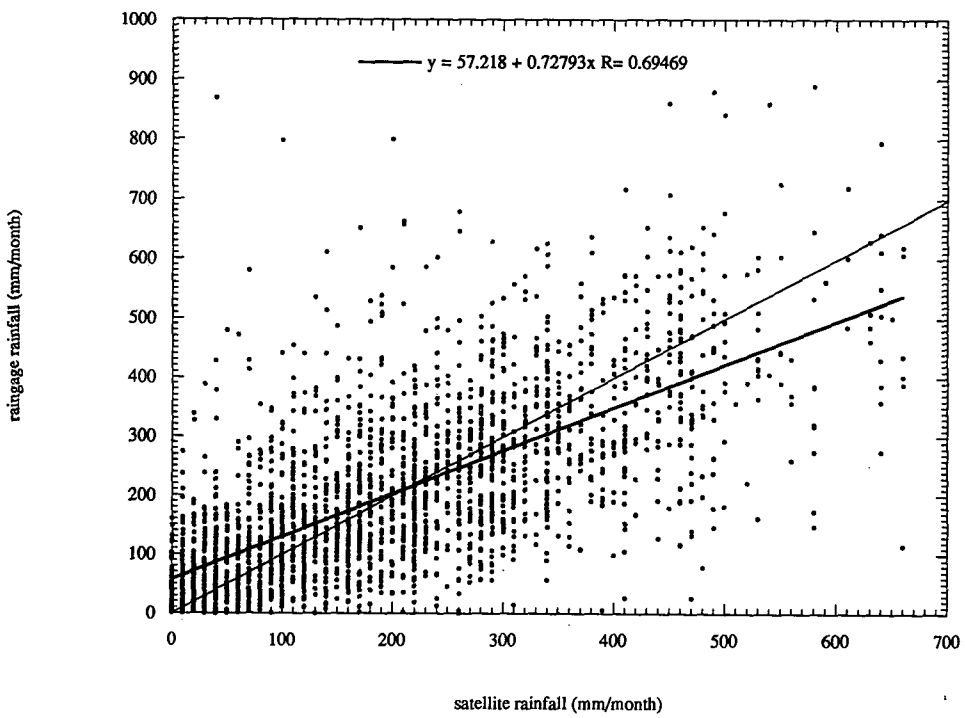


FIG. 3. GPI-estimated rainfall versus rain gauge–measured point rainfall. The thick solid line is the least-squares linear regression analysis line, and the thin solid line is the 45° line.

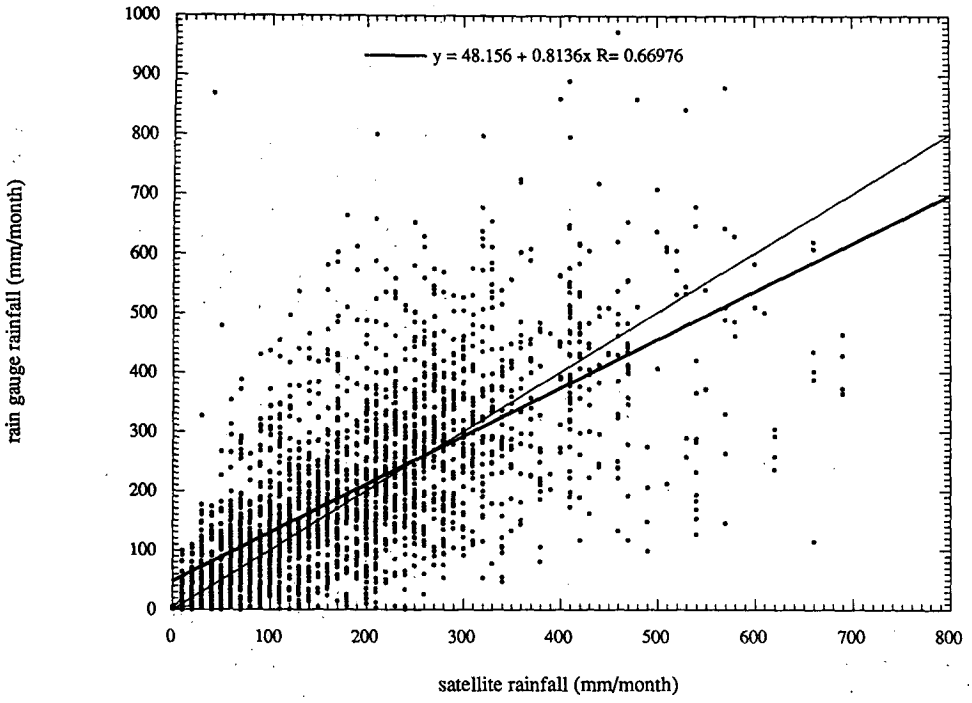


FIG. 4. As in Fig. 3 but for the WCC estimates.

the slopes, intercepts, and the correlation coefficients associated with both algorithms were statistically different from zero at the 95% confidence level. Unfortunately, the very large confidence intervals about the regression parameters (not shown) prevented the identification of a statistical distinction between the corresponding regression parameters associated with the two algorithms (i.e., the null hypothesis of similar regression parameters could not be rejected at the 95% confidence level). Also, a bias could not be firmly established in either algorithm (i.e., the null hypothesis that the slopes were equal to 1.0 could not be rejected). Thus, the large amount of scatter about both regression lines prevented statistically sound and conclusive statements being made about the differences between the algorithms and their respective amounts of bias and random error.

A large percentage of the scatter may be attributed to spatial sampling error in the rain gauge estimates, instrument error, and temporal sampling error resulting from the difference in the temporal sampling characteristics between the two satellites. A certain percentage of this scatter cannot be reduced since the temporal sampling characteristics of the satellites are fixed and the amount of rain gauge instrument error is unknown. However, it is likely that if the contribution to the scatter from spatial sampling error can be significantly reduced using the NCR method, concrete statements could be made concerning the accuracy of the algorithms and intercomparisons of the regression parameters.

b. Application of the NCR method

The satellite rainfall estimates were first quantized into 80 intervals of 10 mm each (i.e., 10, 20, . . . , 800 mm) for both WCC and GPI data. All point rain gauge measurements taken simultaneously in time and space with the quantized satellite values were sorted into two groups of 80 datasets (i.e., one group for each algorithm). The point rainfall variance for each dataset was then calculated individually. Due to an insufficient number of pairs of observations existing for each dataset, one spatial correlation function representing the range of satellite values was computed for each algorithm (appendix B).

For the two sets of spatial correlation coefficients (one set for each algorithm), an exponential function of the following form provided an adequate fit using a least-squares method

$$\rho(v) = a + be^{(-v/c)},$$

where a, b, and c are nonlinear regression coefficients determined by a least-squares method, and v is the separation distance. Figures 5 and 6 show the correlation coefficients and the resulting function of each algorithm. Statistical models, constructed using the previously described procedure, were then used to produce two scatter diagrams, one for each satellite algorithm. The selected rain gauge densities (i.e., N) obtained using Eq. (18) in Morrissey (1991) were 14 and 15 per 308 025 km^2 (i.e., 5° × 5° latitude–longitude areas) for the GPI and the WCC algorithms, respectively (refer

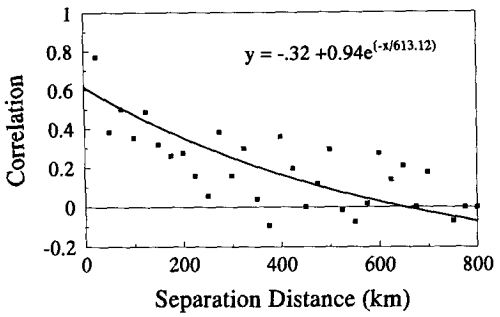
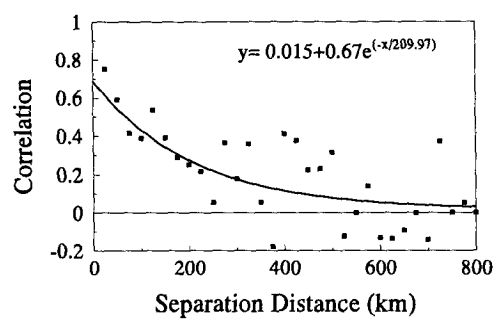


FIG. 5. The spatial correlation among rain gauge stations for GPI data. The least-squares fit exponential function is also given.

to Table 1). It was assumed that 5% sampling error remained in the areal averages [i.e., CV in Eq. (18) in Morrissey 1991]. The statistics required to derive the volume rainfall variances are given in Table 1. It should be noted that since a single correlation function was used to represent the range of quantized algorithm values, a single value of $E[\rho|A]$, N, and, consequently, $F(N)$ was computed, one for each algorithm (Morrissey and Greene 1993).

The results of the NCR method applied to the GPI is shown in Fig. 7. Overall, the variance about the regression line has been reduced by approximately 67% from that using the point data. While the parameters of the regression equation have not significantly changed, the correlation has increased to 0.88. It should be assumed that this is a minimum value since rain gauge instrumentation errors and satellite temporal sampling error undoubtably account for some additional variance about the regression line. The standard error about the regression line indicates that approximately 68% of the values fall within ±47 mm.

For the WCC estimates, the regression parameters have changed somewhat (Fig. 8) from these found using point data. The slope has increased to 0.91 and the intercept has decreased to 33.9 mm month^{-1}. This probably results from an insufficient number of model values greater than 480 mm month^{-1}, which reduced the range of WCC values tested. The amount of rain



'FIG. 6. As in Fig. 5 but for the WCC estimates.

TABLE 1. Values for the statistical parameters used to determine the variance of volume rainfall using Eq. (1).

Algorithm	$E[\rho\|A]$	N	$F(N)$
GPI	0.27	14	0.33
WCC	0.22	15	0.27

gauge data associated with the larger WCC values was insufficient to produce model values greater than 480 mm month^{-1}. Although inconclusive, the change in the slope resulting from the decrease in the range of WCC estimates suggests that a slight nonlinear relationship may exist between the WCC estimates and the surface-measured volume rainfall. The standard error about the regression line suggests that roughly 68% of the values fall within ±53 mm of the value predicted by the regression line. The correlation coefficient has increased to 0.87.

Analysis conducted on the residuals (not shown) indicated that the error variance was relatively stable given changes in abscissa value. The NCR method produced model values that were quasi-normally distributed in accordance with the central limit theorem that states that a distribution converges to normality with averaging. Also, the NCR method inherently produces independent model values, so that multicolinearity in the residuals was not a factor. Thus, standard statistical significance tests could be appropriately applied to the regression parameters and the correlation coefficients associated with both algorithms.

Significance tests applied at the 95% confidence level indicate that the regression parameters and the correlation coefficients associated with both algorithms were significantly different from zero. Also, the slope values for both algorithms were significantly different from one. This indicates that both algorithms are biased to some degree. Table 2 shows that the GPI is significantly biased with the population value for the slope between 0.80 and 0.67, while the same value for the WCC al-

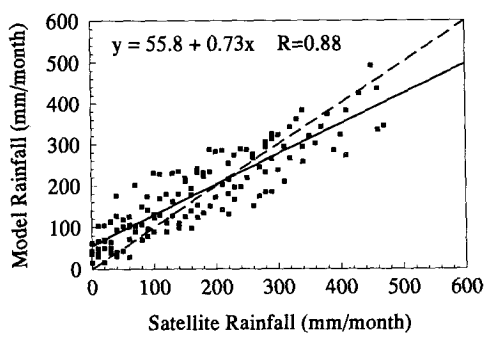


FIG. 7. The modeled rainfall using the NCR method versus the GPI-estimated rainfall. The solid line represents the best-fit linear function getting by least-squares nonlinear regression analysis. The dashed line is the 45° line.

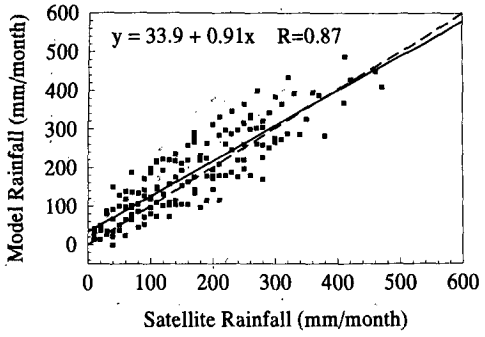


FIG. 8. As in Fig. 7 but for the WCC estimates.

gorithm lies between 0.99 and 0.83. The slope and intercept values associated with the GPI algorithm were also significantly different from the corresponding parameters associated with the WCC algorithm.

While the correlation coefficients for the two algorithms were significantly different from zero, they were not significantly different from each other, suggesting that the level of random error in the algorithms estimates is essentially the same. The similarity in the amounts of random error is, at first, surprising since the more direct rainfall measurements come from the MW technique. However, the WCC algorithm suffers to some extent from the relatively low temporal sampling rate associated with the DMSP satellite.

The overall bias in the GPI, computed using one minus the ratio of the mean rainfall estimated by the GPI algorithm to that from the model estimates, is approximately 5.4%. The value of this ratio for the WCC algorithm is approximately 8.5%, which is essentially the same as that found by Chang et al. (1993). Although the overall bias in the GPI is less than that in the WCC algorithm, comparisons of the GPI regression line with the 45° line in Fig. 7 illustrate that the GPI tends to substantially overestimate high rainfall values and underestimate low rainfall values. This is in agreement with what was found by Morrissey and Greene (1993). This "range-dependent bias" is much less for the WCC algorithm (Fig. 8).

6. Conclusions

Using the standard verification method of comparing rainfall estimates from two satellite algorithms with simple spatial averages of monthly rainfall data from Pacific atoll rain gauge sites proved useful in the sense that a statistical relationship could be established between the algorithm estimates and the surface observations (i.e., the slopes, intercepts, and correlation coefficients were statistically different from one). However, the substantial amount of scatter about the regression lines produced by spatial sampling error prevented conclusive statements being made about the amount of bias and random error in the algorithms. The NCR method, by reducing this scatter, allowed statistically valid conclusions concerning differences between the two algorithms and the accuracy of their estimation capabilities. It was found that the correlation between monthly surface-measured rainfall and the WCC-estimated monthly rainfall (0.87) was essentially the same as that found for the GPI (i.e., 0.88). Thus, both algorithms can account for approximately 77% of monthly rainfall variance over the Pacific. However, the significant differences were found between the regression parameters associated with the algorithms. While the WCC algorithm contained an overall bias slightly higher than the GPI (8.5% compared to 5.4%), its slope value was much closer to one (0.91 as compared to 0.73 for the GPI). The intercept values for both algorithms were significantly different from zero, indicating that positive monthly rainfall values occurred when the algorithms indicated that no rainfall fell during the month. Thus, using the resulting regression relationships, correction factors based upon the computed slope values may enable the algorithms to produce relatively unbiased estimates of monthly rainfall for the tropical Pacific region.

It should be mentioned that since the algorithms were tested using data from the entire tropical Pacific region, regional biases could not be identified and may exist. Insufficient data prevented the use of the NCR method to assess these biases. As the WCC record lengthens, it will soon be possible to test for regional biases in this algorithm.

The regression statistics associated with the two algorithms suggest that the GPI suffers from regional and perhaps seasonal biases in its calibration coefficient of 3 mm h^{-1}. This probably explains some of the scatter observed in Fig. 7. Morrissey and Greene's (1993) work suggests that the overestimation by the GPI of high monthly rainfall stems from a fundamental difference in the relationship between the amount of cold clouds per area and volume rainfall between the Pacific and the Atlantic (where the GPI was calibrated). Since convective systems in the Atlantic intertropical convergence zone (ITCZ) tend to be more isolated than those in the South Pacific convergence zone (SPCZ), it was hypothesized that in the Pacific a greater fraction of nonraining cirrus is advected into a given area from systems located outside this area. This "cirrus contam-

TABLE 2. Confidence limits associated with the regression parameters and the correlation coefficients. The 95% significance level was assumed.

	Slope (mm $month^{-1}$)		Intercept (mm $month^{-1}$)		Correlation	
Algorithm	Lower	Upper	Lower	Upper	Lower	Upper
GPI	0.67	0.80	41.8	69.9	0.85	0.92
WCC	0.83	0.99	17.8	50.1	0.85	0.91

ination" (Morrissey 1986) increases the value of the GPI but does not contribute to the monthly rainfall value. Since the SPCZ is a major rain producer in the Pacific, the overall effect on the GPI is a lowering of the 3 mm h^{-1} calibration coefficient for the tropical Pacific region. This effect is directly related to the size of the averaging area and the climatological mean spacing between convective systems. The microwave algorithms do not suffer from this source of bias since cirrus is relatively transparent to microwave radiation emitted or scattered by raindrops. However, the low temporal sampling rate of polar-orbiting satellites does contribute to error in the regression parameters and is probably responsible for the small amount of bias observed for the WCC algorithm.

Since the GPI is based upon a threshold of 235 K, it cannot detect rain falling from relatively warm clouds. Thus, the intercept value of 55.8 mm $month^{-1}$ for the GPI suggests that a significant amount of warm rain occurs in the tropical Pacific.

Acknowledgments. The authors would like to note the support of the National Oceanic and Atmospheric Administration through Grant NA90 RAH0078. A special thanks to Dr. Alfred Chang from the Goddard Space Flight Center for supplying us with the satellite rainfall data from the microwave-based algorithm, and Mr. John Janowiak for supplying us with the GPI values used in this study.

APPENDIX A

The Frequency Function

The frequency function $f(\nu)$ of the distance ν between two randomly selected points within a square of area A is given by

$$f(\nu) = \frac{1}{A^{0.5}} g\left(\frac{\nu}{A^{0.5}}\right),$$

where

$$g(a) = 2ag_1(a) + 4ag_2(a)$$

$$g_1(a) = \begin{cases} \pi + a^2 - 4a, & 0 < a < \sqrt{2} \\ 0, & \text{otherwise} \end{cases}$$

$$g_2(a) = \begin{cases} 2(a^2 - 1)^{1/2} - 2\cos^{-1}\left(\frac{1}{a}\right) - (a-1)^2, & 0 < a < \sqrt{2} \\ 0, \quad \text{otherwise.} \end{cases}$$

APPENDIX B

Estimation of the Spatial Correlation Function

Equation (1) in the text requires that the expected value of the conditional correlation $E[\rho_s|A]$ between two rain gauges randomly located within a square surface A ($5° \times 5°$ area) be computed. A different value for this quantity is required for each quantized satellite rainfall estimate. For open ocean regions, there are generally insufficient observations to adequately estimate this quantity for each satellite value. In this case, $E[\rho_s|A]$ may be estimated using a rainfall correlation function representing several different satellite values. This replacement will not affect the computed slope of the relationship nor the correlation coefficient computed between the model and satellite rainfall values (Morrissey and Greene 1993).

The averaged expected value of correlation over all satellite values $E[\rho|A]$ is computed as follows. Let $r_s(x_i, t)$ equal the rainfall measurement located at x_i, representing time t and associated with satellite value s. Let $\text{cov}_s(x_i, x_i + \nu)$ equal the covariance between two stations with separation distance ν conditional on satellite value s. Under the homogeneous assumption, we can define that

$$E[r_s(x_i, t)] = E[r_s(x_i + \nu, t)]$$

(i.e., the mean rainfall rate associated with a given satellite rainfall value is same as all rain gauge stations). Thus,

$$\text{cov}[r_s(x_i, t), r_s(x_i + \nu, t)] = \frac{1}{T\text{NP}(\nu)}\left[\sum_{i=1}^{T}\sum_{i=1}^{\text{NP}(\nu)} r_s(x_i, t) r_s(x_i + \nu, t)\right] - \bar{r}_s(\nu)^2,$$

where $\text{NP}(\nu)$ equals the number of pairs of values separated by ν. Note that $\text{NP}(\nu)$ may be a function of t if some rainfall values are missing, and $\bar{r}_s(\nu)$ is the mean rainfall rate from pairs separated by ν. The conditional spatial correlation is

$$\rho_s[r_s(x_i, t), r_s(x_i + \nu, t)] = \frac{\text{cov}_s[r_s(x_i, t), r_s(x_i + \nu, t)]}{\sigma_s^2(\nu)},$$

where $\sigma_s^2(\nu)$ is the conditional variance associated rainfall measured at gauges separated by ν. Now an expression exists for the correlation at each separation distance and each quantized satellite value. An average over all s values is taken and a weight applied based upon the number of rainfall pairs at each separation distance s;

$$\bar{\rho_s}[r_s(x_i, t), r_s(x_i + \nu, t)] = \sum_{s=1}^{S} w_s(\nu)\rho_s[r_s(x_i, t), r_s(x_i + \nu, t)],$$

where S is the number of satellite categories, and $w_s(\nu)$ is weight function defined by

$$w_s(\nu) = \frac{H_s(\nu)}{H(\nu)},$$

where $H_s(\nu)$ is total number of rainfall pairs associated with satellite value s and separated by distance ν, and $H(\nu)$ is total number of rainfall pairs separated by distance ν.

REFERENCES

Arkin, P. A., 1979: The relationship between the fractional coverage of high cloud and rainfall accumulations during GATE over the B-scale array. *Mon. Wea. Rev.*, **107,** 1382–1387.

——, and B. N. Meisner, 1987: The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982–1984. *Mon. Wea. Rev.*, **115,** 51–74.

Barrett, E. C., and D. W. Martin, 1981: *The Use of Satellite Data in Rainfall Monitoring.* Academic Press, 340 pp.

Berg, W. K., 1993: Estimation and analysis of climate-scale rainfall over the tropical Pacific. Ph.D. dissertation, Colorado Center for Astrodynamics Research, University of Colorado, Boulder, 202 pp.

Chang, A. T. C., L. S. Chiu, and T. T. Wilheit, 1993: Random error of oceanic monthly rainfall derived from SSM/I using probability distribution function. *Mon. Wea. Rev.*, **121,** 2351–2354.

Chiu, L. S., A. T. C. Chang, and J. Janowiak, 1993: Comparison of monthly rain rates derived from GPI and SSM/I using probability distribution functions. *J. Appl. Meteor.*, **32,** 323–334.

Draper, N. R., and H. Smith, 1981: *Applied Regression Analysis.* John Wiley and Sons, 709 pp.

Garcia, O., 1985: Atlas of highly reflective clouds for the global tropics, 1971–1983. U.S. Dept. of Commerce Publications, Environmental Research Labs, Boulder, CO, 365 pp.

Grody, N. C., 1991: Classification of snow cover and precipitation using the Special Sensor Microwave Imager (SSM/I). *J. Geophys. Res.*, **96,** 7423–7435.

Journel, A. G., and C. H. J. Huijbregts, 1978: *Mining Geostatistics.* Academic, 600 pp.

Kilonsky, B. J., and C. S. Ramage, 1976: A technique for estimating tropical open-ocean rainfall from satellite observation. *J. Appl. Meteor.*, **15,** 972–975.

Laughlin, C. R., 1981: On the effect of temporal sampling on the observation of mean rainfall. Precipitation Measurements from Space, Workshop Report, NASA Publications, Goddard Space Flight Center, Greenbelt, MD, D59–D66.

Liu, G., and J. A. Curry, 1992: Retrieval of precipitation from satellite microwave measurement using both emission and scattering. *J. Geophys. Res.*, **97,** 9959–9974.

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

Morrissey, M. L., 1986: A statistical analysis of the relationships among rainfall, outgoing longwave radiation and the moisture budget during January–March 1979. *Mon. Wea. Rev.*, **114,** 931–942.

——, 1991: Using sparse raingages to test satellite-based rainfall algorithms. *J. Geophys. Res.*, **96,** 18 561–18 571.

——, and J. S. Greene, 1991: The Pacific atoll raingage data set. JIMAR Tech. Rep., No. 648, University of Hawaii, Honolulu, HI, 30 pp.

——, and ——, 1993: Comparison of two satellite-based rainfall algorithms using Pacific atoll raingage data. *J. Appl. Meteor.*, **32,** 411–425.

——, ——, and J. A. Maliekal, 1993a: Pacific atoll raingauge data set. *TOGA-Notes,* **10,** 1–4.

——, M. A. Shafer, S. E. Postawko, and B. Gibson, 1993b: Comprehensive Pacific Rainfall Data Base. Oklahoma Climatological Survey Tech. Rep. No. 93-003, University of Oklahoma, Norman, OK, 19 pp.

Olson, W. S., 1989: Physical retrieval of rainfall rates over the ocean by multispectral microwave radiometer: Application to tropical cyclones. *J. Geophys. Res.*, **94,** 2267–2279.

Richards, F., and P. A. Arkin, 1981: On the relationship between satellite-observed cloud cover and precipitation. *Mon. Wea. Rev.*, **109,** 1081–1093.

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

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

Simpson, J., R. F. Adler, and G. R. North, 1988: A proposed Tropical Rainfall Measuring Mission (TRMM) satellite. *Bull. Amer. Meteor. Soc.*, **69,** 278–295.

Spencer, R. W., M. H. Goodman, and R. E. Hood, 1989: Precipitation retrieval over land and ocean with the SSM/I: Identification and characteristics of the scattering signal. *J. Atmos. Oceanic Technol.*, **6,** 254–273.

WCRP, 1986: Report of the first session of the International Working Group on Data Measurement for the Global Precipitation Climatology Project. World Meteorological Organization, WCP-132 WMO/TD No. 171, 60 pp.

Wilheit, T. T., A. T. C. Chang, M. S. V. Rao, E. B. Rodgers, and J. S. Theon, 1977: A satellite technique for quantitatively mapping rainfall rates over the oceans. *J. Appl. Meteor.*, **16,** 551–560.

——, ——, and L. Chiu, 1991: Retrieval of monthly rainfall indices from microwave radiometric measurements using probability distribution function. *J. Atmos. Oceanic Technol.*, **8,** 118–136.