Nonsystematic Errors of Monthly Oceanic Rainfall Derived from SSM/I

ALFRED T. C. CHANG AND LONG S. CHIU*

Hydrological Sciences Branch, Laboratory for Hydrospheric Processes, NASA/Goddard Space Flight Center, Greenbelt, Maryland

(Manuscript received 8 April 1998, in final form 4 August 1998)

ABSTRACT

About 10 yr (July 1987–December 1997 with December 1987 missing) of oceanic monthly rainfall based on data taken by the Special Sensor Microwave/Imager (SSM/I) on board the Defense Meteorological Satellite Program satellites have been computed. The technique, based on the work of Wilheit et al., includes improved parameterization of the beam-filling correction, a refined land mask and sea ice filter. Monthly means are calculated for both 5° and 2.5° latitude–longitude boxes.

Monthly means over the latitude band of 50°N–50°S and error statistics are presented. The time-averaged rain rate is 3.09 mm day^{-1} (std dev of 0.15 mm day^{-1}) with an error of 38.0% (std dev of 3.0%) for the 5° monthly means over the 10-yr period. These statistics compare favorably with 3.00 mm day^{-1} (std dev of 0.19 mm day^{-1}) and 46.7% (std dev of 3.4%) computed from the 2.5° monthly means for the period January 1992–December 1994. Examination of the different rain rate categories shows no distinct discontinuity, except for months with a large number of missing SSM/I data.

An independent estimate of the error using observations from two satellites shows an error of 31% (std dev of 2.7%), consistent with the 38% estimated using (A.M. and P.M.) data from one satellite alone. Error estimates (31%) based on the 5° means by averaging four neighboring 2.5° boxes are larger than those (23%) estimated by assuming the means for these neighboring boxes are independent, thus suggesting spatial dependence of the 2.5° means.

Multiple regression analyses show that the error varies inversely as the square root of the number of samples but exhibits a somewhat weaker dependence on the mean rain rate. Regression analyses show a power law dependence of −0.255 to −0.265 on the rain rate for the 5° monthly means using data from a single satellite and a dependence of −0.366 for the 5° monthly means and −0.337 for the 2.5° monthly means based on two satellite measurements. The latter estimate is consistent with that obtained by Bell et al. using a different rainfall retrieval technique.

1. Introduction

In the estimation of space–time rainfall from satellite measurements, one can distinguish three types of errors: systematic, random, and sampling (Wilheit 1988). Errors in the algorithm assumptions (algorithm error) and the sensor calibration are the main contributors to systematic errors. For example, the nonuniform field of view of microwave sensors introduces a beam-filling bias in rain rate retrievals (Chiu et al. 1990). Nonsystematic errors consist of random and sampling errors. Random error refers to error due to noise in the sensor measurements, such as stray thermal emission. Since the systematic and random errors are associated with either the retrieval algorithm or with instrument noise, together, they are also referred to as the retrieval error (Bell et al. 1990). In principle, the systematic error may be reduced by improving on the retrieval algorithm and sensor calibration. The random error may be reduced by improving on the instrument noise, such as reducing the sensor noise equivalent temperature differential or radiometer sensitivity (NEΔT). The sampling error arises because only snapshots of the rain fields are measured, and the total space–time rainfall is estimated by interpolation or extrapolation. The sampling error is determined by the orbit of the satellite, the swath width of the instrument, and the space–time characteristics of the rainfall fields that one intends to measure in the first place.

Algorithm errors are dependent on the algorithm physics and can be large as evidenced in some of the algorithm intercomparison results [e.g., the Third Precipitation Intercomparison Project (PIP-3) of the Wet Net Project]. To the extent that the systematic error can

* Current affiliation: Center for Earth Observing and Space Research, Institute for Computational Sciences and Informatics, George Mason University, Fairfax, Virginia.

Corresponding author address: Dr. Alfred T. Chang, Hydrological Sciences Branch, Laboratory for Hydrospheric Processes, Code 974, NASA/GSFC, Greenbelt, MD 20771.
E-mail: achang@rainfall.gsfc.nasa.gov

be minimized by correcting the bias to ground truth and the random errors are independent from footprint to footprint, the sampling error is by far the most important error source for monthly mean estimates (Bell et al. 1990).

The sampling error of space–time rainfall has been examined by a number of investigators. Laughlin (1981) derived an expression for the sampling error in terms of the sampling frequency and the autocovariance of the rain field. Kedem et al. (1990, 1997) used different sampling strategies and radar rain rates collected during the Global Atmospheric Research Program Atlantic Tropical Experiment and the Tropical Ocean Global Atmosphere Coupled Ocean–Atmosphere Response Experiment to show the sampling error for the Tropical Rainfall Measuring Mission (TRMM) to be about 10%. North and Nakamoto (1989) presented a formalism that relates the sampling error to the rain field space–time spectrum and sampling function. Bell and Kundu (1996) developed an optimal weighting scheme for estimation that relates to the rain structure and sampling volume.

As part of the Global Precipitation Climatology Project (GPCP), the Polar Satellite Precipitation Data Center at the National Aeronautics and Space Administration/Goddard Space Flight Center (NASA/GSFC) has been producing oceanic rainfall from Special Sensor Microwave/Imager (SSM/I) measurements using the technique developed by Wilheit et al. (1991). Huffman et al. (1997) merged this monthly rain rate product together with rain gauge data, other satellite products, and model results to produce a merged global rainfall product for the GPCP.

The nonsystematic (random plus sampling) error components of the oceanic monthly rainfall have been examined by Chang et al. (1993). The error estimation scheme is based on the existence of pairs of independent estimates of the monthly mean. This paper updates the work of Chang et al. (1993) and presents the total oceanic rainfall (between the 50° latitude bands) and the nonsystematic error for the period July 1987–December 1997 for the 5° boxes and the 2.5° boxes for January 1992–December 1994. The data are described in section 2. The technique for error estimation is described in section 3. Section 4 presents the error estimates for different sampling strategies. In section 5, the error dependence on the mean rain rate and sample size are investigated. Section 6 summarizes and discusses our results.

2. Data

The oceanic monthly average rainfall rates are derived from the data collected by the SSM/I on board the Defense Meteorological Satellite Program (DMSP) satellites using the Wilheit et al. (1991) technique. The production of this dataset is a result of a NASA/Goddard Space Flight Center effort to meet the GPCP requirement for microwave-based oceanic precipitation estimates.

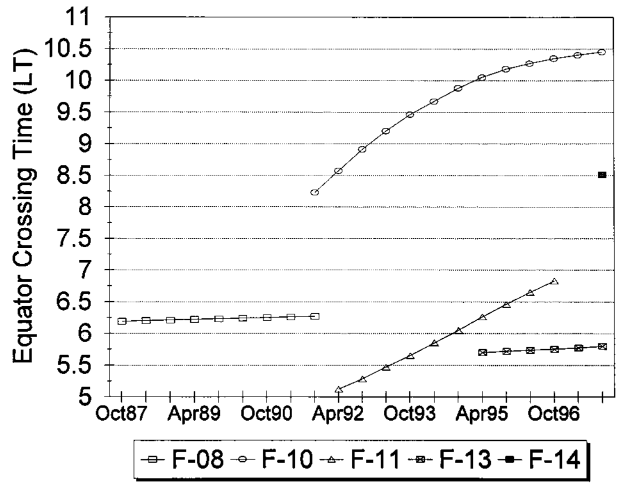

FIG. 1. Equatorial crossing times of the DMSP *F-8, F-10, F-11, F-13,* and *F-14* satellites. The first tick mark on the abscissa shows the value for Oct 1987. The following tick mark show that for Apr 1988, Oct 1988, Apr 1989, . . . , and so on, showing a 6-month interval between tick marks.

Figure 1 shows the equatorial crossing times of the various DMSP satellites launched thus far. From July 1987 to December 1990, the *F-8* satellite was operational. From January 1991 and on, SSM/I data are available from both the *F-10* and *F-11* satellites. The *F-10* was launched into a slightly elliptic orbit. The *F-10* overpass time drifted slowly and stabilized in late 1996. Since May 1995, *F-13* was launched into orbit with an equatorial crossing time that is close to the *F-8* satellite. The latest, *F-14,* was launched in May 1997 with a 9 A.M. (LT) equatorial crossing time.

The SSM/I sensor on board the DMSP *F-8* satellite has an equatorial crossing time of about 6 A.M. and 6 P.M. (LT). The SSM/I microwave brightness temperature (TB) histograms of a combination channel are computed for the A.M. and P.M. data, denoted, respectively, *H*(AM) and *H*(PM). The combination channel, twice the 19-GHz minus the 22-GHz vertical channel of the SSM/I, is used to minimize the effect of water vapor on the rain signal. Over the course of a month, the rain rate distribution is assumed to follow a lognormal distribution. A relation between rainfall and TB, based on the radiative transfer calculations of Wilheit et al. (1977), is used to relate the parameters of the rainfall distribution to the TB histogram. A monthly mean freezing height, which is a parameter in the TB–rain rate relation, is derived from scattergrams of the 19- and 22-GHz SSM/I brightness temperatures. The parameters of the rain rate distribution are solved iteratively by matching the moments of the TB histograms to that computed from the rain rate distributions. The goodness of fit is determined by chi-square statistics. If the chi-square statistics exceed minimum threshold, the monthly rainfall estimates are rejected. Based on A.M. and P.M. TB histograms, monthly estimates for the A.M. and P.M. data are calculated; namely,

$$\mathrm{RR(AM)} = f[H(\mathrm{AM})] \quad \text{and} \quad \mathrm{RR(PM)} = f[H(\mathrm{PM})],$$

where f represents the nonlinear process by which the monthly rainfall based on the A.M. data and RR(AM), the rain rate based on the A.M. data, are calculated from the histogram of H(AM). The derived SSM/I rain rate indices are then multiplied by a correction factor to account for the beam-filling bias (Wilheit et al. 1991; Chiu et al. 1990).

A number of modifications to the original algorithm have been implemented since the initial processing. A refined land–sea mask that includes many small islands and high-resolution coastline has been incorporated. Improvement on the land–sea mask is based on the Defense Mapping Agency Nautical Charts and Publications (1994–96) published by the National Oceanic and Atmospheric Administration (NOAA 1994–96). Monthly means are computed according to the calendar month, instead of the pentad month done previously. (In the initial GPCP data processing, a normal year consisting of 365 days is divided into 73 pentads. A pentad month has six pentads, except for August, which has seven pentads.) A sea ice filter, jointly developed with personnel of the Sea Ice group at GSFC, has also been implemented. The beam-filling correction factor has been modified to vary as a linear function of the freezing height (FL in km), namely, beam-filling correction = (1 + 0.088)FL as suggested by Wang (1995). The addition of a refined land–sea mask greatly improves the data quality of the monthly rain rates, especially near coastal regions and around islands in the west Pacific. The number of grid boxes excluded due to insufficient data or poor chi-square goodness of fit is reduced.

Two products have been produced: monthly oceanic rainfall at 5° × 5° and 2.5° × 2.5° latitude–longitude resolutions. The former covers the region between 50°S and 50°N whereas the latter is extended to the 65° latitudes. For comparison purposes, the spatial coverage is restricted to the 50° latitude bands. We shall refer to the former and the latter as the 5° and 2.5° products, respectively. The 5° products consist of monthly estimates based on A.M. and P.M. SSM/I data. Approximately 10 yr (July 1987–December 1997, with December 1987 missing) of monthly mean rain rate have been processed. In order that the intervals of the histograms are well represented, both the A.M. and P.M. data from a single satellite have been used in computing the histogram for the 2.5° product. Hence for the 2.5° product, we can write

$$\mathrm{RR}(F\text{-}10, 2.5) = f[H(F\text{-}10)] \quad \text{and}$$
$$\mathrm{RR}(F\text{-}11, 2.5) = f[H(F\text{-}11)],$$

where RR(*F-10,* 2.5) and RR(*F-11,* 2.5) are the monthly rain rate for the 2.5° product derived from *F-10* and *F-11* SSM/I data, respectively.

3. Error estimation technique

The technique for estimating the error, described in Chang et al. (1993), requires the existence of a pair of independent estimates. Let x_1 and x_2 represent the pair of independent estimates of the monthly mean; namely,

$$x_1 = \langle x_1 \rangle + e_1 \tag{1}$$

$$x_2 = \langle x_2 \rangle + e_2, \tag{2}$$

where the ensemble averaging, $\langle\ \rangle$, is taken over the appropriate rain rate categories, and e_1 and e_2 are the error associated with the independent estimates. Assuming that the estimates are unbiased with uncorrelated errors,

$$\langle e_i \rangle = 0, \quad \text{for } i = 1, 2; \tag{3}$$

$$\langle e_1 e_2 \rangle = 0. \tag{4}$$

Assuming that the root-mean-square (rms) error is the same for both estimates, or

$$\langle e_1^2 \rangle = \langle e_2^2 \rangle = \langle e^2 \rangle,$$

we get

$$\langle e^2 \rangle = \tfrac{1}{2}[\langle (x_1 - x_2)^2 \rangle - (\langle x_1 \rangle - \langle x_2 \rangle)^2]. \tag{5}$$

The assumption that the rms errors for the two estimates are equal can be relaxed. If we assumed that the rms errors are related as

$$\langle e_1^2 \rangle = r\langle e_2^2 \rangle,$$

where r is the ratio between the rms errors, an optimal estimate can be obtained by a weighted mean of the two estimates, that is,

$$R = (\alpha_1 x_1 + \alpha_2 x_2), \tag{6}$$

where α_1 and α_2 are weighting factors, normalized to 1, that is, $\alpha_1 + \alpha_2 = 1$. The expected mean and rms error of R are, respectively,

$$E(R) = \langle R \rangle = \langle (\alpha_1 \langle x_1 \rangle + \alpha_2 \langle x_2 \rangle) \rangle \quad \text{and} \tag{7}$$

$$s = \langle (\alpha_1 e_1 + \alpha_2 e_2)^2 \rangle^{1/2}. \tag{8}$$

The optimal weights are

$$\alpha_1 = r/(1 + r) \quad \text{and} \quad \alpha_2 = 1/(1 + r).$$

In section 5, we show that the rms error is inversely proportional to the number of samples; that is, r is the ratio of the number of samples for the two estimates. When r is substituted into the above relation, we see that the optimal weights are proportional to the number of samples (see the appendix). We examined the average number of samples for the A.M. and P.M. estimates and for the *F-10* and *F-11* estimates. The difference in sample size between the A.M. and P.M. estimates for a single satellite, and between *F-10* and *F- 11,* for January 1992–December 1994, is generally less than 10%, except for July 1994 (section 4c). In the following analysis, we shall assume that $\alpha_1 = \alpha_2 = 0.5$.

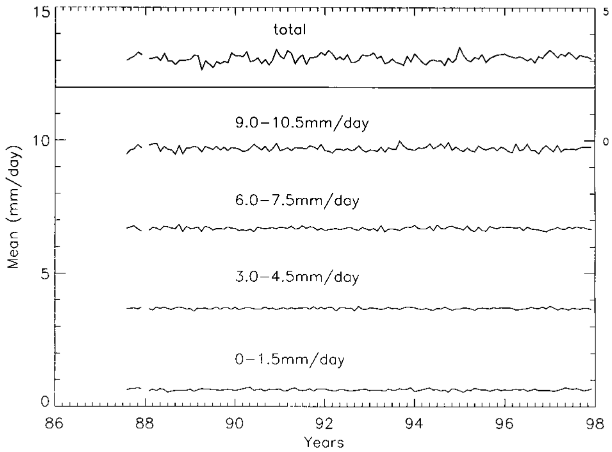

FIG. 2. Time series of the total (50°S–50°N) monthly 5° × 5° mean rain rates (top curve) and mean rain rates for rain rate categories of 0–1.5, 3–4.5, 6–7.5, and 9–10.5 mm day^{-1}, respectively. Monthly means are based on the average of the A.M. and P.M. estimates. Data from *F-8* are used for Jul 1987–Dec 1991, *F-11* for Jan 1992–Apr 1995, and *F-13* for May 1995–Sep 1997. The large tick marks on the abscissa indicate January of the year.

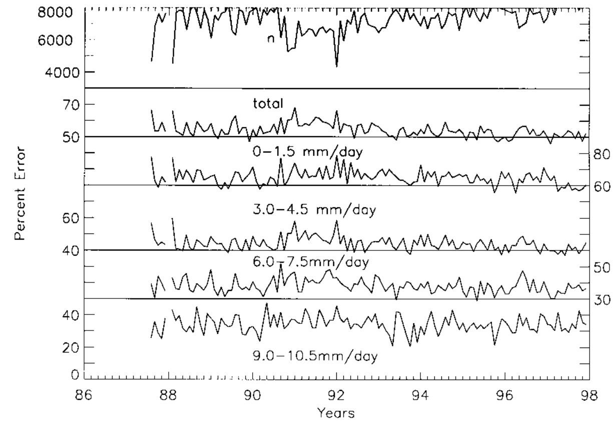

FIG. 3. Time series of the average number of pixels in each 5° × 5° grid between 50°S and 50°N (top curve) and percent error of the total rain rate and for the rain rate categories of 0–1.5, 3–4.5, 6–7.5, and 9–10.5 mm day^{-1}, respectively. Error estimates are based on the average of the A.M. and P.M. estimates.

4. Analysis

a. The 5° A.M.–P.M. pair for a single satellite

Let x_1 and x_2 described in section 3 be the A.M. and P.M. estimates for the 5° product, respectively; that is, x_1 = RR(AM) and x_2 = RR(PM). We partitioned the monthly rain rates into 1.5 mm day^{-1} rain rate categories and computed the rms error using (5) for these rain rate categories and for all the rain rate categories (hereafter referred to as the total rain rate). Figure 2 shows the time series of the total (50°N–50°S) mean (not weighted by latitude) and the average for the different rain rate categories computed from (7). The mean (July 1987–December 1997) is 3.09 mm day^{-1} with a standard deviation (std dev) of 0.15 mm day^{-1}. Figure 3 shows the average number of pixels within each 5° box $\langle n \rangle$, percent error for the total, and the percent error for different rain rate categories. The *F-8* SSM/I data are used for the period July 1987–December 1991, *F-11* for the period January 1992–April 1995, and *F-13* for the period May 1995–December 1997. The ensemble average of n, $\langle n \rangle$, and the total error are taken over all rainfall categories for each month; that is, $\langle\ \rangle$ is taken over all grid boxes for the total. The total percent error, defined as $\langle s \rangle / \langle R \rangle$, is relatively constant over the entire period. The time-averaged (over all 125 months) nonsystematic percent error for the A.M. or P.M. estimate is 53.8% with a standard deviation (std dev) of 4.3%. Assuming independence of the A.M. and P.M. estimates, the combined estimate using (7) and (8) would reduce the error to about 53.8%/$\sqrt{2}$, or about 38%. Since the A.M. and P.M. estimates are approximately 12 h apart, the combined (A.M. + P.M.) estimate eliminates the bias due to the first harmonic of the diurnal cycle.

b. The 5° F-10–F-11 *pair*

Taking advantage of the presence of two satellites (*F-10* and *F-11*), another independent estimate of the error can be provided. In this case, the estimates are the monthly means from *F-10* and *F-11*, respectively. Monthly estimate for a single satellite is simply the arithmetic average of the A.M. and P.M. estimates; that is,

$$x_1 = \mathrm{RR}(F\text{-}10) = \tfrac{1}{2}[\mathrm{RR}(F\text{-}10\ \mathrm{AM}) + \mathrm{RR}(F\text{-}10\ \mathrm{PM})],$$

and

$$x_2 = \mathrm{RR}(F\text{-}11) = \tfrac{1}{2}[\mathrm{RR}(F\text{-}11\ \mathrm{AM}) + \mathrm{RR}(F\text{-}11\ \mathrm{PM})].$$

We chose the period January 1992–December 1994 when data from both *F-10* and *F-11* are available.

We examined the bias between the *F-10* and *F-11* estimates using a paired t test proposed by Chang et al. (1995). The paired t statistic (t) is defined as

$$t = (Y/m^{1/2})\mathrm{SD}, \tag{9}$$

where Y and SD are the expected mean and standard deviation of the difference of the two variables (i.e., monthly mean from the *F-10* and *F-11* datasets, respectively) and m is the number of data pairs (i.e., $m = 36$ for the number of months between 1992 and 1994). For $m > 30$, t follows approximately a normal distribution. The paired t statistics have been computed for all 5° grid boxes from these 3 yr of data. Figure 4 shows the distribution of the paired t statistics. Inspection of the distribution of t shows no systematic pattern, and the number of grid points with $|t| > 1.96$ is about 5%. The difference between the *F-11* and *F-10* estimate is 0.024 mm day^{-1}.

We estimated the monthly error associated with the *F-10* and *F-11* measurements using (5). The average monthly mean (*F-10* and *F-11*) is 3.10 mm day^{-1}, with a std dev of 0.18 mm day^{-1}. The average percent error

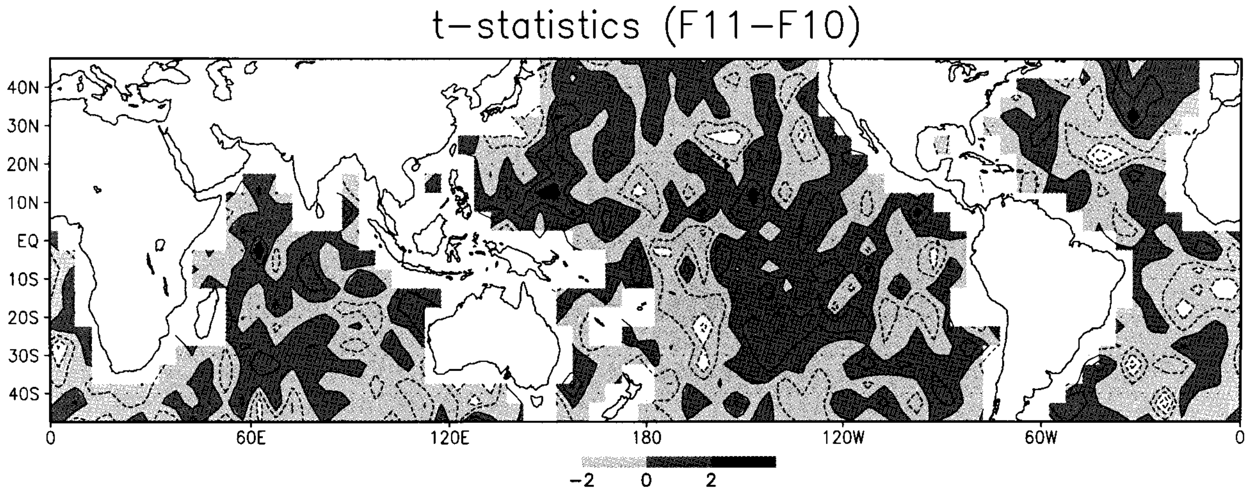

FIG. 4. Distribution of paired *t* statistics between *F-10* and *F-11* estimates. The *t* statistics are computed using monthly data for the period Jan 1992–Dec 1994.

is 31.05%, with a std dev of 2.72%. This is smaller than the 38% derived in the previous section. We can combined the *F-10* and *F-11* estimates by applying (7) and (8) and obtain an error of $31.05\%/\sqrt{2}$, or about 22%.

c. The 2.5° F-10–F-11 *pair*

We examined the errors associated with the 2.5° products. In this case, the pair of estimates are

$$x_1 = \mathrm{RR}(F\text{-}10,\ 2.5) \quad \text{and} \quad x_2 = \mathrm{RR}(F\text{-}11,\ 2.5).$$

The same time period and coverage (50° latitude band) as described in section 4b are used. Figure 5 shows the time series of the total and the mean rain rate for the rain rate categories. Figure 6 shows the average number of pixels per 2.5° grid box, the total percent error, and the percent error for the different rain rate categories. The long-term monthly mean rain rate is 3.00 mm day^{-1}, with a std dev of 0.19 mm day^{-1}, and an error of 46.68%, with a std dev of 3.37%.

During this period, the mean rain rate within each rain rate categories is fairly constant. However, the total (top curve) shows a dip (2.15 mm day^{-1}) for July 1994. The time-average (1992–94) contributions to the total rainfall by the different rain rate categories are calculated to be 8.2%, 18.1%, 18.6%, 14.9%, 11.2%, 8.8%, 6.8%, 4.9%, and 8.3%, respectively, for the 0–1.5, 1.5–3, 3–4.5, 4.5–6, 6–7.5, 7.5–9, 9–10.5, 10.5–12, and >12 mm day^{-1} rain rate categories. However, for July 1994, the contributions for these categories are 13%, 25%, 22%, 17%, 11%, 6%, 3%, 2%, and, 1%, respectively. We examined the difference in sample size between *F-10* and *F-11* during this period. The difference is in general less than 10%, except for July 1994 (41%) and April 1993 (17%). For July 1994, 18% of the data were missing for *F-11* (depicted in Fig. 6, top panel) and 45% for *F-10* (not shown). The case was worse for

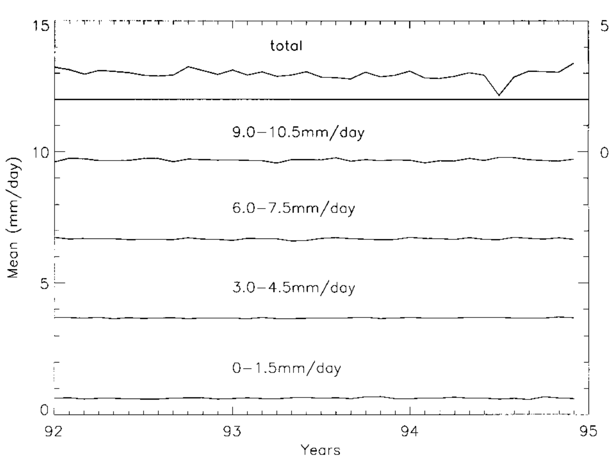

FIG. 5. Same as Fig. 2 except for 2.5° × 2.5° grids between 50°S and 50°N from Jan 1992 to Dec 1994.

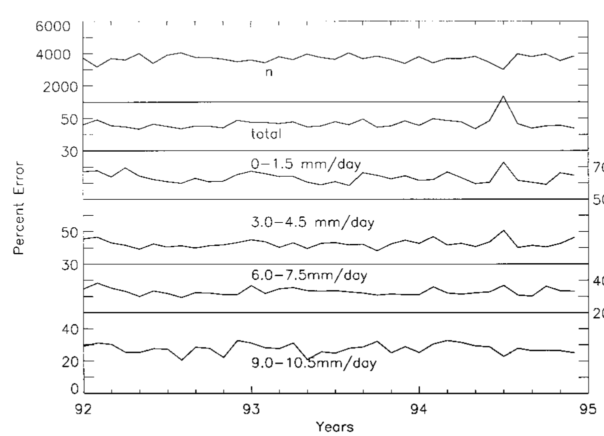

FIG. 6. Same as Fig. 3 except for 2.5° × 2.5° grids between 50°S and 50°N from Jan 1992 to December 1994.

F-10 when no data were available from 17 to 27 July continuously.

5. Error structure: Dependence on mean rain rate and sample size

Bell et al. (1990) derived a formula for sampling error for regularly spaced uncorrlelated observations; namely,

$$s/\langle R\rangle = p^{-1/2}[s_a(R > 0)/R_a(R > 0)](L^2_{\mathrm{corr}}/A)^{1/2}(DT/T)^{1/2},$$

where p is the rain probability; $R_a(R > 0)$ and $s_a(R > 0)$ are, respectively, the mean and rms of the conditional rain rate in the rain area a; L_{corr} is a typical spatial dimension of rain events; A is the area of the grid box; and T/DT is the number of samples taken over the sampling period. Based on the works of Short et al. (1993), which indicated that the ratio $s_a(R > 0)/R_a(R > 0)$ is fairly constant, and that p is correlated to the areal average rainfall (Chiu 1988; Kedem et al. 1990), Bell and Kundu (1996) suggest that the rms error varies approximately as the inverse square root of the mean rain rate and sample volume. When the relation is put into our notation, we get

$$s/\langle R\rangle \sim (\langle R\rangle\langle n\rangle)^{-1/2}. \qquad (10)$$

We checked the above relation against estimates we derived. An assessment of this relation can be made using the data depicted in Figs. 2 and 3 for the 5° product and Figs. 5 and 6 for the 2.5° product. The average number of pixels that are included in the TB histograms $\langle n\rangle$ are available for the total mean (top curves, Figs. 3 and 6). Inspection of Fig. 3 shows three months with percent error exceeding three standard deviations from the mean: July 1987, January 1988, and December 1991. Correspondingly, the average number of pixels per grid box, $\langle n\rangle$, is much smaller than the average for these months. Examination of the original data showed more than six consecutive days of data were missing during these three months.

We fitted the rms error, the mean, and the number of samples for the total rain to the form

$$s/\langle R\rangle = C\langle R\rangle^a\langle n\rangle^b. \qquad (11)$$

Taking the logarithm of both sides, we obtain

$$\log(s/\langle R\rangle) = a\log\langle R\rangle + b\log\langle n\rangle + c. \qquad (12)$$

Multiple regressions were performed on the logarithm of the variables. Four cases were considered: (a) A.M. or P.M. estimates for the 5° product from a single satellite for the period July 1987–December 1997 (section 4a), (b) *F-10* or *F-11* estimates for 5° products for January 1992–December 1994 (section 4b), (c) *F-10* or *F-11* estimate for 2.5° products for the same period as b (section 4b), and (d) same as a except for the period January 1992–December 1994.

Table 1 summarizes the results of our multiple regression analyses. All cases show dependence of about −0.5 on the sample size $\langle n\rangle$. The best fit case is c, which

TABLE 1. Summary of multiple regression analyses of mean rain rate and sampling size on error for four different sampling strategies. Multiple regression is of the form $\langle s\rangle/\langle R\rangle = C\langle R\rangle^a\langle n\rangle^b$, where $\langle R\rangle$ is in mm day^{-1}. The standard deviations for the mean appear in parentheses underneath the mean in columns 4–11. Errors estimated from strategy a are based on difference between A.M. and P.M. 5° estimates for a single satellite for the period Jul 1987–Dec 1997; strategy b is based on difference between 5° estimates for *F-10* and *F-11* for Jan 1992–Dec 1994, strategy c is based on 2.5° estimates for *F-10* and *F-11* for the same period as b, and strategy d is similar to a except for the time period, which is the same as b.

Strategy	Grid size	Period (no. of months)	Mean rain rate (mm day^{-1})	Error (%)	Avg no. of pixels	No. pixels per 1° × 1°	Multiple corr. (standard error)	C	A	B
a: One satellite	5° × 5°	July 1987–Dec 1997	3.09	53.80	7238	290	0.82	7956	−0.242	−0.531
A.M. vs P.M.		(125)	(0.15)	(4.26)	(780)		(0.04)		(0.082)	(0.034)
b: *F-10* vs *F-11*	5° × 5°	Jan 1992–Dec 1994	3.10	31.05	14 422	576	0.72	161 468	−0.897	−0.523
		(36)	(0.18)	(2.72)	(906)		(0.06)		(0.156)	(0.175)
c: *F-10* vs *F-11*	2.5° × 2.5°	Jan 1992–Dec 1994	3.00	46.68	3682	589	0.93	6350	−0.540	−0.527
		(36)	(0.19)	(3.37)	(231)		(0.02)		(0.067)	(0.072)
d: *F-11*	5° × 5°	Jan 1992–Dec 1994	3.07	53.81	7225	289	0.65	6519	−0.031	−0.536
A.M. vs P.M.		(36)	(0.16)	(2.90)	(474)		(0.04)		(0.153)	(0.116)

TABLE 2. Time-averaged mean rain rate, percent error, and number of grid boxes for the nine rain rate categories for four different sampling strategies described in Table 1. Units of ⟨R⟩ is in mm day^{-1}.

Rain rate category	0–1.5 mm day^{-1}	1.5–3.0 mm day^{-1}	3.0–4.5 mm day^{-1}	4.5–6.0 mm day^{-1}	6.–7.5 mm day^{-1}	7.5–9.0 mm day^{-1}	9.–10.5 mm day^{-1}	10.5–12.0 mm day^{-1}	>12 mm day^{-1}
a: 5° one-satellite A.M. vs P.M. best power law fit: $\langle s\rangle/\langle R\rangle = 0.598\ \langle R\rangle^{-0.255}$									
Mean	0.63	2.21	3.69	5.18	6.70	8.20	9.70	11.19	14.63
% error	64.87	48.95	44.24	41.17	37.97	35.08	34.25	30.30	29.11
No. of grids	333.08	224.12	138.54	78.71	48.11	30.69	19.90	12.45	18.22
b: 5° *F-10* vs *F-11* best power law fit: $\langle s\rangle/\langle R\rangle = 0.420\ \langle R\rangle^{-0.366}$									
Mean	0.65	2.21	3.67	5.18	6.07	8.20	9.69	11.19	14.30
% error	44.61	32.28	28.47	24.11	22.38	20.56	19.39	15.28	13.90
No. of grids	294.86	222.44	138.17	76.67	43.92	27.75	18.17	11.03	15.36
c: 2.5° *F-10* vs *F-11* best power law fit: $\langle s\rangle/\langle R\rangle = 0.602\ \langle R\rangle^{-0.337}$									
Mean	0.63	2.20	3.68	5.19	6.69	8.19	9.69	11.20	14.15
% error	63.45	49.00	42.59	36.81	33.03	29.29	27.61	25.53	22.23
No. of grids	1405.64	881.75	542.19	308.67	182.22	115.64	76.42	47.42	63.58
d: 5° *F-11* A.M. vs P.M. best power law fit: $\langle s\rangle/\langle R\rangle = 0.604\ \langle R\rangle^{-0.265}$									
Mean	0.62	2.21	3.68	5.19	6.70	8.20	9.70	11.19	14.57
% error	66.02	49.13	44.10	40.79	37.90	33.94	33.68	30.95	28.23
No. of grids	335.31	225.33	134.92	79.17	45.22	30.86	19.44	12.17	18.44

shows dependence of about −0.5 for both the mean rain rate and number of samples. This is consistent with (10). The low multiple correlation and high normal standard errors for the other cases (a, b, and d) suggest poor fits to (11). It can be argued that the dynamic ranges in the total rain rate are not large enough to provide a good estimate of the exponent for $\langle R\rangle$.

Assuming independence of the sampling volume and mean rain rate, we can examine the dependence on the latter using data for the different rain rate categories. Time averaging of the rain rate categories provided percent error and mean rain rate pairs for our analysis. Table 2 shows the time-averaged monthly mean rain rate, [$\langle R\rangle$], percent error [⟨%error⟩], and the number of grid boxes within each rain rate category for four different estimation schemes. The dependence is graphically depicted in Fig. 7 for these cases. The percent error is fitted to a power law of the form $\langle s\rangle/\langle R\rangle = C\langle R\rangle^{a}$. The results of the best power law fit are included in Table 2. The regression is performed on the logarithm of the percent error and mean rain rate. The correlation coefficients for the logarithm of the these variables for the sampling schemes are 0.98, 0.98, 0.99, and 0.97, respectively. There is little difference between strategy a and d, which show roughly a −0.25 dependence on the mean rain rate. The percent error associated with the 5° and 2.5° products using both *F-10* and *F-11* data (strategies b and c) shows a slightly stronger dependence (−0.337) but is still somewhat weaker than the −0.5 dependence suggested by Bell and Kundu (1996).

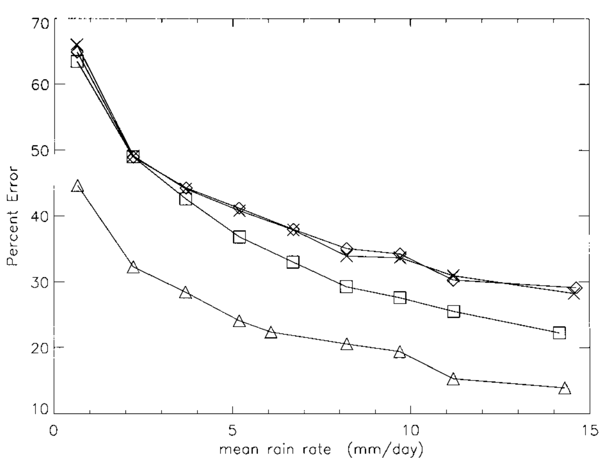

FIG. 7. Percent error as a function of the mean rain rate for (a) 5° × 5° grid rainfall based on A.M. or P.M. SSM/I data for a single satellite for Jul 1987–Dec 1997 (cross), (b) 5° × 5° grid based on *F-10* and *F-11* estimates for Jan 1992–Dec 1994 (triangle), (c) 2.5° × 2.5° grid based on *F-10* and *F-11* estimates for the same period as (b) (square), and (d) same as (a) except for the same period as (b) (diamond).

6. Summary and discussion

About 10 yr of oceanic monthly rain rate data derived from SSM/I have been produced and the error associated with these rain rates estimated based on the Chang et al. (1993) technique. The long-term (July 1987–December 1997) average total rain rate is 3.09 mm day^{-1}, with a std dev of 0.15 mm day^{-1} and an error of 53.8% with a std dev of 4.26% for monthly estimates using only the A.M. or P.M. data at 5° latitude–longitude grid boxes. If we assume that there is no difference in the sample size, simple arithmetic averaging of the A.M. and P.M. monthly pairs reduces the error to about 38%.

We showed that monthly estimates based on *F-11* data are slightly larger than those derived from the *F-10* measurements. This difference may be due to the noctunal to early morning maximum of oceanic diurnal rainfall. The equatorial crossing time of the *F-11* satellite is between 0500 and 0700 UTC, whereas that for *F-10* is between 0800 and 1030 UTC. The data obtained from the TRMM satellite may provide more information for assessing this difference. The error for a single satellite observation derived from the *F-10* and *F-11* pairs is 31%, which is consistent with that (38%) derived from the A.M. and P.M. pair of a single satellite.

When the error estimated using the A.M. and P.M. data of a single satellite were fitted to a power law, a constant of 0.60 (60% error at 1 mm day^{-1}) and a dependence of -0.265 on the rain rate were obtained. A similar analysis for the error estimates based on the *F-10* and *F-11* monthly pairs showed a constant of 0.42 and a dependence of -0.366. Since the sample size of the latter doubles that for the A.M. or P.M. alone, the estimate error is expected to be $1/\sqrt{2}$ smaller than that of either A.M. or P.M., consistent with the ratio of $0.60/0.42 = \sqrt{2}$.

The 1992–94 total mean rain rate for the 2.5° product was computed to be 3.00 mm day^{-1} (std dev 0.19 mm day^{-1}), which shows no significant difference from the 5° product of 3.10 mm day (std dev 0.18 mm day^{-1}) for the same period. A similar error analysis performed on the 2.5° products for the *F-10* and *F-11* pairs showed mean error of 46.68% (std dev 3.37%). This can be compared to the error of the 5° product. If we average four neighboring 2.5° grid boxes to form a new 5° grid box, the error associated with such a 5° grid monthly rainfall is expected to be $46.68\%/\sqrt{4}$, or $\sim$23%, if all four grid boxes are assumed independent. We computed such a new 5° mean product and computed the error assicated with it using (5). The error (average of 31%) is larger than the 23% expected from independent 2.5° means but consistent with that based on a single satellite (case b). This result suggests that neighboring 2.5° monthly means are spatially dependent.

Multiple regression analyses show agreement for four different sampling strategies on the power law dependence of -0.5 on sample size. While the power law dependence of both the sample size and mean rain rate (-0.5) are consistent with that proposed by Bell and Kundu for the 2.5° product, the dependence on the mean rain rate is much weaker than predicted for the 5° products. The inverse dependence on the square root of the mean rain rate is based on the observation that the ratio of the standard deviation to the mean for the rainy conditions is approximately constant over a range of area, rain, and climate type (Short et al. 1993). These data are available for the SSM/I monthly estimates and we plan to further examine this ratio for space and time scales typical of the SSM/I monthly sampling.

The objective of GPCP is to produce a long-term (15 yr) global rainfall dataset. The quality of the long-term dataset is critical to the utilization of the monthly rainfall products. For example, the merged GPCP global datasets relies on error estimates of its component rain fields, of which the oceanic SSM/I rainfall is one. Our analysis also showed that over the 10-yr period the data quality is fairly consistent.

Sampling studies based on GATE shipborne radar for an area of 280 km × 280 km lead to errors of about 8% for 30-day averages. When this estimate is extrapolated to 500 km × 500 km grid boxes, sampling errors of 7%–12% are estimated. Our study shows estimated error in the range of 15% for the high rain rate categories ($>$9.5 mm day^{-1}) to about 45% for the low rain rate categories (0–1.5 mm day^{-1}) based on measurements taken by one SSM/I sensor. Monthly means based on two SSM/I satellite estimates (e.g., *F-11* and *F-13*) reduce the sampling error to $<$10% for the high rain rates and $<$20% overall. Additional microwave sensor measurements, such as the SSM/I launched on *F-14* and the TRMM Microwave Imager launched by TRMM, can be used to further validate, and can be combined to reduce, the sampling error on oceanic monthly mean rainfall derived from passive microwave measurements.

Acknowledgments. We thanked J. Meng and H. Powell of SAIC/General Sciences Corporation for programming and graphics support. This research is supported by the National Aeronautic and Space Administration, Office of Earth Sciences.

APPENDIX

Optimal Weights for Combining Monthly Means

From the results of the regression analyses, the errors are inversely proportional to the number of samples; that is,

$$\langle e_1^2 \rangle \sim 1/N_1 \quad \text{and} \quad \langle e_2^2 \rangle \sim 1/N_2.$$

Let r be the ratio of the rms errors; that is, $r = \langle e_2^2 \rangle / \langle e_1^2 \rangle = N_1/N_2$. Equation (8) can now be written as

$$s^2 = \langle (\alpha_1 e_1 + \alpha_2 e_2)^2 \rangle = (\alpha_1^2 + (1 - \alpha_1)^2 r) \langle (e_1)^2 \rangle$$

after substituting in the normalization condition, $\alpha_1 + \alpha_2 = 1$. We optimize s with respect to α_1 and find that s is a minimum when

$$\alpha_1 = r/(1 + r) = N_1/(N_1 + N_2)$$

and

$$\alpha_2 = 1/(1 + r) = N_2/(N_1 + N_2).$$

Hence the optimal weights for combining monthly estimates are proportional to the number of samples.

REFERENCES

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,** 1251–1268.

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

Chiu, L. S., 1988: Estimating areal rainfall from rain area. *Tropical Rainfall Measurements,* J. Theon and N. Fugono, Eds., A. Deepak Publishing, 361–367.

——, D. Short, A. McConnell, and G. North, 1990: Rain estimation from satellites: Effect of finite field of view. *J. Geophys. Res.,* **95,** 2177–2185.

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

——, ——, and G. Yang, 1995: Diurnal cycle of oceanic precipitation from SSM/I data. *Mon. Wea. Rev.,* **123,** 3371–3380.

Huffman, G. J., and Coauthors, 1997: The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *Bull. Amer. Meteor. Soc.,* **78,** 5–20.

Kedem, B., L. S. Chiu, and G. N. North, 1990: Estimation of mean rain rate: Application to satellite observations. *J. Geophys. Res.,* **95,** 1965–1972.

——, R. Pfeiffer, and D. A. Short, 1997: Variability of space–time mean rain rate. *J. Appl. Meteor.,* **36,** 443–451.

Laughlin, C. R., 1981: On the effect of temporal sampling on the observation of mean rainfall. *Precipitation Measurements from Space,* D. Atlas and O. Thiele, Eds., NASA. [Available from NASA/Goddard Space Flight Center, Greenbelt, MD 20771.]

National Oceanic and Atmospheric Administration, 1994–1996: Defense Mapping Agency nautical charts and publications. NOAA. [Available from NOAA Distribution Branch, N/CG33, National Ocean Service, Riverdale, MD 20737-1199.]

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

Short, D. A., D. B. Wolff, D. Rosenfeld, and D. Atlas, 1993: A study of the threshold method utilizing raingage data. *J. Appl. Meteor.,* **32,** 1379–1387.

Wang, S. A., 1995: Modeling the beamfilling correction for microwave retrieval of oceanic rainfall. Ph.D. dissertation, Texas A&M University, College Station, TX, 99 pp. [Available from Dept. of Meteorology, Texas A&M University, College Station, TX 77801.]

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.

——, 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 ocean. *J. Appl. Meteor.,* **16,** 551–560.

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