Random Errors of Oceanic Monthly Rainfall Derived from SSM/I Using Probability Distribution Functions

ALFRED T. C. CHANG
Hydrological Sciences, NASA/Goddard Space Flight Center, Greenbelt, Maryland

LONG S. CHIU
General Sciences Corporation, Laurel, Maryland

THOMAS T. WILHEIT
Meteorology Department, Texas A&M University, College Station, Texas

(Manuscript received 29 September 1992, in final form 19 February 1993)

ABSTRACT

Global averages and random errors associated with the monthly oceanic rain rates derived from the Special Sensor Microwave/Imager (SSM/I) data using the technique developed by Wilheit et al. are computed. Accounting for the beam-filling bias, a global annual average rain rate of 1.26 m is computed. The error estimation scheme is based on the existence of independent (morning and afternoon) estimates of the monthly mean. Calculations show overall random errors of about 50%–60% for each 5° × 5° box. The results are insensitive to different sampling strategy (odd and even days of the month). Comparison of the SSM/I estimates with raingage data collected at the Pacific atoll stations showed a low bias of about 8%, a correlation of 0.7, and an rms difference of 55%.

1. Introduction

The advent of satellite observations allows the estimation of many geophysical parameters from space observations. The diversity of algorithms to retrieve geophysical parameters results in varied estimates for the same parameter, despite the fact that identical observations are used. Since these parameters will be used in other scientific investigations, uncertainties in these estimates must be known and understood. The error can be estimated by comparing the estimates with in situ measurements or with other estimates with a known or higher accuracy.

The spatial and temporal variability of rainfall renders the estimation of space–time-averaged rainfall difficult. In addition to the error associated with retrieval, there are sampling errors associated with space and time sampling. The ultimate hurdle of validation lies in the absence of "ground truth" for rainfall. Gauge measurements are hampered by effects such as wind and interference from nearby structures (Sevruk 1985). They also suffer from the sparsity of gauge stations, especially over the open ocean regions.

Ground-based radar probably provides the most comprehensive spatial coverage of rainfall estimates at radar sites. However, the radar reflectivity must be converted to rain rates via a reflectivity–rain-rate relation, which depends, among other variables, on the drop size distribution. Surface radars, calibrated with disdrometers and networks of raingages, offer the best strategy for obtaining rain-rate "ground truth" (Simpson 1988). However, over the oceans the coverage by surface radar is limited.

We have developed a technique for estimating oceanic monthly average rainfall rates over 5° longitude × 5° latitude boxes using microwave observations taken by the Special Sensor Microwave/Imager (SSM/I) on board the Defense Meteorological Satellite Program (DMSP) satellite (Wilheit et al. 1991). Approximately three years (July 1987–June 1990, with December 1987 missing) of monthly mean rain rate have been compared to Jaeger's (1983) climatology and Arkin and Meisner's (1987) GOES precipitation index (GPI) (Chiu et al. 1993). The Wilheit et al. (1991) rain rates compare well with the GPI in regions where the GPI is expected to work well, namely, in the tropics and warm extratropics. Regions of disagreement are associated with known problems of the GPI, such as the presence of nonraining high cirrus over the Indian Ocean region, discontinuities of coverage in the mid-Pacific when GOES-West failed, and in the extratropics where nonconvective rainfall dominates. However, this is only a qualitative comparison between these two rainfall estimates.

Corresponding author address: Dr. Alfred T. C. Chang, NASA/Goddard Space Flight Center, Code 974, Greenbelt, MD 20771.

Our rain estimation technique uses a combination of SSM/I channels, namely, twice the 19-GHz minus the 22-GHz vertical channel to minimize the effect of water vapor, assumes a lognormal rain-rate distribution, and estimates the freezing height from scattergrams of the SSM/I brightness temperature. The derived SSM/I rain-rate indexes are then multiplied by a correction factor to account for the beam-filling error (Wilheit et al. 1991; Chiu et al. 1993). The SSM/I makes two measurements per day, respectively, at about 0600 and 1800 LST. Estimates based on the morning and afternoon passes have been computed separately. Monthly means are computed as the average of the morning and afternoon estimates. Since these estimates are taken 12 h apart, they represent a good estimate of the monthly mean if the semidiurnal cycle is negligible compared to the diurnal cycle. Except for the difference due to the presence of the diurnal cycle, the morning and afternoon estimates represent independent estimates of the monthly means. The random error can, therefore, be estimated. It must be emphasized that this technique estimates only the random errors; estimates of any systematic biases remain a difficult and elusive goal.

2. Random error analysis

Let the morning and afternoon estimates be denoted a and p, respectively. Assuming that there are errors associated with these estimates, we can write

$$a = \langle a \rangle + e_a$$
$$p = \langle p \rangle + e_p, \qquad (1)$$

where $\langle\ \rangle$ represent ensemble averaging taken over different rain-rate categories within a month, as has been done in Wilheit (1988). Since we are concerned with morning and afternoon pairs at each box, temporal and geographical variations—as well as sampling and retrieval biases—are included in the $\langle a \rangle$ and $\langle p \rangle$ terms. By partitioning the average morning and afternoon estimates into rain-rate categories, the error within each category can be estimated for each month. Equation (1) is valid if

$$\langle e_a \rangle = \langle e_p \rangle = 0 \qquad (2)$$

holds. The error term contains errors due to sampling, retrieval, and possibly residue of the diurnal component from the ensemble averaging. We can write the mean square difference of morning and afternoon estimates as

$$\langle (a-p)^2 \rangle = (\langle a \rangle - \langle p \rangle)^2 + \langle (e_a - e_p)^2 \rangle. \qquad (3)$$

Equation (3) states that the mean square difference between the morning and afternoon rain estimate is the variance due to the random errors and the diurnal difference. Assuming that the errors are homogeneous and uncorrelated, we can write

$$\langle e_a^2 \rangle = \langle e_p^2 \rangle = \langle e^2 \rangle \qquad (4)$$

and

$$\langle (e_a - e_p)^2 \rangle = 2\langle e^2 \rangle.$$

After rearranging, the error term can be written as

$$2\langle e^2 \rangle = \langle (a-p)^2 \rangle - (\langle a \rangle - \langle p \rangle)^2. \qquad (5)$$

This estimate is probably an upper bound for the error term since it may contain residues of the diurnal cycle.

Table 1 shows the number of observations N, the mean rain rate over the rain-rate category [mean $= (\langle a \rangle + \langle p \rangle)/2$], the difference between the morning and afternoon estimates (difference $= \langle a \rangle - \langle p \rangle$), root-mean-square of the difference [rmsd $= \langle (a-p)^2 \rangle^{1/2}$], and percent error ($\langle e^2 \rangle^{1/2}$/mean) for the different rain-rate categories for May 1991. Rain rates for 5° × 5° cells are classified according to the mean of the morning and afternoon values. These estimates have not been corrected for the beam-filling error discussed earlier. Both morning and afternoon values must exist for the data point to be included. The percent error decrease from 69% for the 0–50 mm month^{-1} category to 26% at the 300–350 mm month^{-1} category. The magnitude of the random errors increases for the high rain-rate categories, probably due to fewer samples at these rain-rate categories. Overall, the percent error is 59%. The mean difference is positive for all rain-rate categories, suggesting higher morning rain rates than the afternoon rain rates over the global ocean. This is consistent with the findings of Sharma et al. (1991) based on 10 months of data.

The morning and afternoon estimates represent independent estimates of the monthly mean; they are estimated from the brightness temperature T_B histograms of satellite measurements grouped according to the time of the passage of the satellite. Independent estimates of the monthly mean can be obtained by grouping the data by odd and even days of the month. Table 2 shows the random error computed from the odd–even day estimates. Since there is no a priori reason to expect variability on a two-day scale, their dif-

TABLE 1. Error statistics for different rain-rate categories for May 1991 estimated from morning and afternoon histograms.

Category	N	Mean	Difference	rmsd	Error (%)
0–50	473	21.5	6.6	21.0	65.6
51–100	276	72.6	19.5	53.9	48.9
101–150	113	119.6	7.9	72.1	42.4
151–200	54	171.9	52.7	103.1	36.5
201–250	32	223.7	75.5	130.8	33.8
251–300	14	271.0	141.3	218.2	43.4
301–350	10	313.4	18.3	118.8	26.5
351–400	0	0.00	0.0	0.0	0
401–450	1	435.6	12.1	12.1	0
Total	964	69.3	17.3	60.3	58.9

Mean = ½($\langle a \rangle + \langle p \rangle$); difference = $\langle a \rangle - \langle p \rangle$; rmsd = $\langle (a - p)^2 \rangle^{1/2}$; and error (%) = $\langle e^2 \rangle^{1/2}$/mean. All units are in mm month^{-1} except the number of observations N, which is nondimensional, and percent error, which is in percent. These rain rates have not been corrected for the beam-filling bias.

ference is expected to be smaller. This can be seen by comparing column 4 (difference) in Tables 1 and 2. The error structures between these tables (error %) are quite similar. The overall error is 56%, compared to 59% for the morning–afternoon estimates.

To examine the stability and trends in this dataset, the procedure is applied to all monthly morning–afternoon estimates for the period July 1987–December 1991. Figure 1 shows the time series of the overall mean and mean monthly rain rate for the 0–50, 101–150, 201–250, and 301–350 mm month^{-1} categories. The global average of the morning and afternoon estimates is about 70 mm month^{-1}. If we apply a beam-filling correction factor of 1.5, as discussed in the next section, an annual global oceanic average of 1260 mm is obtained.

Figure 2 shows the time series of the overall percent error and the percent error for the corresponding rain-rate categories. The overall error and error for different rain-rate categories are relatively constant for the period of our study.

3. Comparison with Pacific atoll gauge data

The SSM/I estimates can be validated against raingage data. Morrissey (1991) compiled a set of raingage data collected at the Pacific atoll stations. The dataset consists of monthly mean gauge rain rates at 88 surface stations for the period up to December 1990. Initial quality control has been performed by Morrissey (1991). The gauge stations are first grouped according to their location into 5° latitude × 5° longitude boxes. The number of station within each 5° × 5° box ranges from a maximum of 11 to as few as 1. If there is more than one station within a box, the gauge rain rates are simply averaged arithmetically. Coincident and collocated SSM/I (morning and afternoon average) estimates are selected for the validation. Altogether there are 1056 monthly estimates in forty-two 5° × 5° boxes.

TABLE 2. Error statistics for different rain-rate categories for May 1991, estimated from odd–even-day histograms.

Category	N	Mean	Difference	rmsd	Error (%)
0–50	480	21.9	−0.6	21.2	68.5
51–100	264	73.3	1.6	47.7	46.0
101–150	122	121.0	−11.8	68.4	39.1
151–200	44	173.0	−6.7	107.9	44.0
201–250	38	221.7	−32.9	132.7	41.0
251–300	15	272.4	−1.3	115.2	29.9
301–350	9	318.1	−43.0	145.2	30.8
351–400	2	375.4	−171.7	172.9	3.9
401–450	0	0	0	0	0
Total	973	70.1	−3.8	55.6	55.9

Mean = $\frac{1}{2}(\langle \text{odd} \rangle + \langle \text{even} \rangle)$; difference = $\langle \text{odd} \rangle - \langle \text{even} \rangle$; rmsd = $\langle (\text{odd} - \text{even})^2 \rangle^{1/2}$; and error (%) = $\langle e^2 \rangle^{1/2}$/mean. All units are in mm month^{-1} except the number of observations N, which is nondimensional, and percent error, which is in percent. These rain rates have not been corrected for the beam-filling bias.

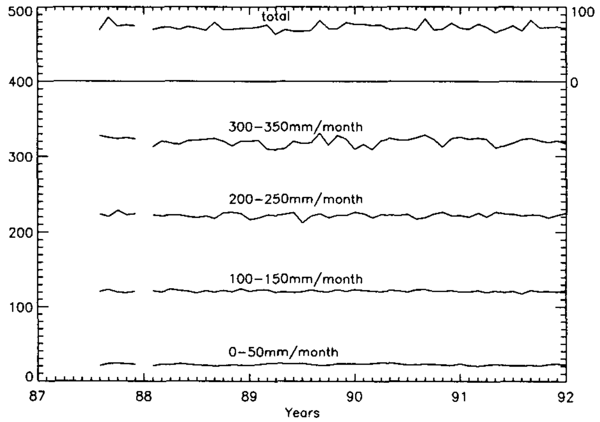

FIG. 1. Time series of the overall mean and mean monthly rain rate for the 0–50, 101–150, 201–250, and 301–350 mm month^{-1} categories.

Studies have shown that microwave estimates tend to underestimate the rain rate due to the beam-filling problem (Chiu et al. 1990; Short and North 1990). The beam-filling error arises because of two factors: inhomogeneity within the field of view of the microwave sensor and nonlinearity in the brightness temperature–rain rate (T_B–R) relations. Chiu et al. (1990) showed that the beam-filling bias can, to first order, be expressed as a product of a function of the T_B–R relation and the rain-rate variance. Based on the rain-rate statistics from GATE and T_B–R relation for the 19 GHz on board *Nimbus-5*, the percent bias, expressed as the fraction of the difference between the true and estimated rain rate of the true rain rate, is about 45%. Our method uses a combination of the 19- and 22-GHz channel, and the percent bias is reduced to about 35%. A similar bias is obtained by examining the rain-field statistics compiled by the Japan Meteorological Agency in support of the Global Precipitation Clima-

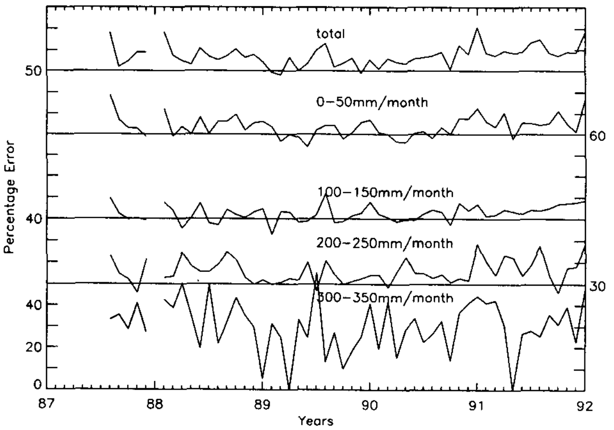

FIG. 2. Time series of the percent error for the 0–50, 101–150, 201–250, and 301–350 mm month^{-1} rain-rate categories and the overall error for all rain-rate categories. The 0–50, 101–150, and 201–250 mm month^{-1} categories and the overall error have been shifted by 60%, 40%, 20%, and 100% in the display for clarity. The abscissa represents December 1986.

tology Project—Algorithm Intercomparison Program (GPCP-AIP/1) using gauge-corrected radar rain-rate data for convective rain (Chiu and Chang, unpublished manuscript). The SSM/I rain rates are multiplied by a factor of 1.5 [or $(1 - 0.35)^{-1}$] to account for the beam-filling bias.

Figure 3 shows the scattergram of monthly rain rate derived from SSM/I and from station gauge data. The SSM/I and gauge mean rain rate for all 1056 months are 185 and 199 mm month^{-1}, respectively, or a low bias of 8% of the gauge mean. Linear regression analysis shows a correlation of 0.70, significant at the 95% level of confidence, and an rms difference of 109.7 mm month^{-1}, or 55% of the gauge mean. If we substitute these values into (5), an error of about 40% is derived.

Linear regression analyses have been performed separately for each season. The results show slight seasonality in the biases, correlation coefficients, and rms differences. Biases of 4.4%, 7.8%, 7.8%, and 6.8% and rms differences of 63%, 53%, 56%, and 52% of the respective seasonal means are computed for the December–February, March–May, June–August, and September–November seasons. The corresponding correlation coefficients are 0.65, 0.71, 0.75, and 0.70, and sample sizes are 234, 245, 285, and 313, respectively.

4. Summary

Based on four and half years of monthly rain rates derived from SSM/I data using the technique developed by Wilheit et al. (1991), global oceanic average rain rate of about 1260 mm yr^{-1} and random errors of about 50%–60% are estimated. The global average can be compared to a value of about 1000 mm yr^{-1} previously estimated (Jaeger 1983). The magnitude of the errors must be appreciated in terms of the large uncertainties in the global rain-rate climatology (Simpson et al. 1988). An objective of the Tropical Rainfall Measuring Mission (TRMM), initiated jointly by the United States and Japan, is the production of at least three years of monthly mean rainfall amount over 500-km × 500-km boxes. Error estimates for TRMM show sampling errors of about 10% (Shin and North 1988; Bell et al. 1990) and total errors of about 20% (Wilheit 1988). Since the sampling strategy and retrieval methods used in TRMM will improve upon existing techniques, our set of monthly mean oceanic rain rates may be considered a prelaunch rainfall climatology against which the TRMM estimates can be compared.

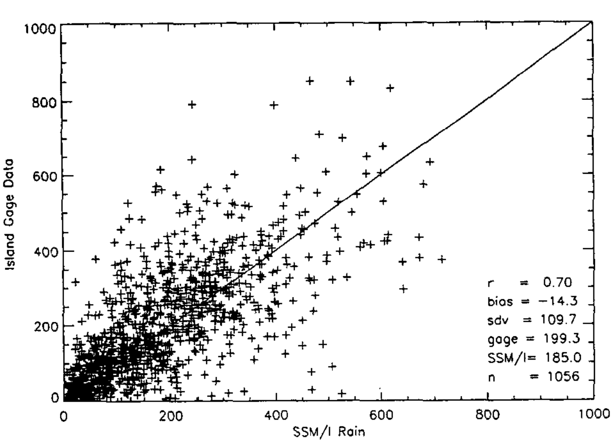

FIG. 3. Scatterplot of SSM/I rain rates versus gauge rain rates over 5° × 5° areas for all available months. Units are in millimeters per month.

Acknowledgments. We thank M. Morrissey of Oklahoma Climate Survey for providing the Pacific atoll raingage dataset, Drs. T. Bell and D. Martin and two anonymous reviewers for comments that helped improve the manuscript, and H. Powell and V. Kuan of General Sciences Corporation for programming and graphic support. This research is partially supported by the National Aeronautic and Space Administration, Climate and Hydrologic System Research and Analysis Program, and the National Oceanic and Atmospheric Administration, Office of Global Programs.

REFERENCES

Arkin, P., and B. 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.

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

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

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

Jaeger, L., 1983: Monthly and areal patterns of mean global precipitation. *Variation in the Global Water Budget,* A. Street Perrott, M. Beran, and R. Ratcliffe, Eds., D. Reidel Publishing, 129–140.

Morrissey, M., 1991: Using sparse rain gages to test satellite-based rainfall algorithms. *J. Geophys. Res.*, **96,** D10, 18 561–18 571.

Sevruk, B., 1985: Correction of monthly precipitation for wetting losses. WMO Instruments and Observing Methods Report No. 22, 7–12.

Sharma, A., A. Chang, and T. Wilheit, 1991: Estimation of the diurnal cycle of oceanic precipitation from SSM/I data. *Mon. Wea. Rev.*, **119,** 2168–2175.

Shin, K.-S., and G. North, 1988: Sampling error study for rainfall estimates by satellite using a stochastic model. *J. Appl. Meteor.*, **28,** 1218–1231.

Short, D. A., and G. R. North, 1990: The beamfilling error in ESMR-5 observations of GATE rainfall. *J. Geophys. Res.*, **95,** 2187–2194.

Simpson, J., 1988: TRMM: A satellite mission to measure tropical rainfall. Report of the science steering group. NASA/Goddard Space Flight Center, Greenbelt, MD, 99 pp.

Wilheit, T., 1988: *Error analysis for the Tropical Rainfall Measuring Mission (TRMM) in Tropical Rainfall Measurements.* A. Deepak Publishing, 377–385.

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