# Dependence of Satellite Sampling Error on Monthly Averaged Rain Rates: Comparison of Simple Models and Recent Studies

THOMAS L. BELL AND PRASUN K. KUNDU*

*Laboratory for Atmospheres, NASA Goddard Space Flight Center, Greenbelt, Maryland*

(Manuscript received 19 May 1998, in final form 5 April 1999)

ABSTRACT

Considerable progress has been made in recent years with using satellite data to generate maps of rain rate with grid resolutions of 1°–5° square. In parallel with these efforts, much work has been devoted to the problem of attaching error estimates to these products. There are two main sources of error, the intrinsic errors in the remote sensing measurements themselves (retrieval errors) and the lack of continuity in the coverage by low earth-orbiting satellites (sampling error). Perhaps a dozen or so studies have attempted to estimate the sampling-error component. These studies have been based on rain gauge and radar-derived data, and the estimates vary so much that it is clear that the sampling error cannot be represented satisfactorily by a single value.

These studies are reviewed. Some of the results reported in these studies are based on a method referred to in this paper as ''resampling by shifts.'' The authors find that the method unfortunately tends to produce estimates that are subject to too much uncertainty to be used quantitatively. After setting these results aside, the authors find that the variability in the remaining sampling-error estimates can be explained to a considerable extent using assumptions common to many statistical models of rain. All such models predict that sampling error relative to the average rain rate $R$ is proportional to $R^{-1/2}$. Although the sampling error at any given site seems (to the extent that data have been examined) to change with $R$ in the way predicted by the model, the proportionality constant in this relationship seen in the various studies appears to change from site to site. This constant can be obtained from the satellite estimates themselves if retrieval errors are not correlated over scales of the order of the grid-box size.

## 1. Introduction

Global maps of monthly rainfall are now routinely produced using data from satellites and a variety of techniques for retrieving rainfall estimates from the data. These maps can in principle be further combined with each other and with ground-based measurements to produce still better products (e.g., Huffman et al. 1995; Xie and Arkin 1996).

The maps can be used in a number of ways. They may be compared with the output of climate models to help evaluate the ability of the climate models to produce realistic distributions of precipitation. The algorithms used to convert the satellite data into rainfall estimates can be validated by comparing the maps to surface measurements. The maps can be used to look for signs of climate change and to obtain energy and moisture budgets to help understand climate timescale dynamics.

In all of these cases, quantitative work with the datasets requires that the datasets be accompanied by an error estimate for each gridded value. Valid error estimates almost certainly depend on location, season, type of rain, etc. Some of the reasons for this are investigated here using theoretical models and analysis of rain-rate data derived from ground-based radars and rain gauges. A simple formula is proposed for root-mean-square (rms) error as a function of rainfall amount and satellite sampling characteristics.

This work had its origins in error studies undertaken in preparation for the launch of the Tropical Rainfall Measuring Mission (TRMM) satellite now in orbit. The satellite, described by Simpson et al. (1996), was specifically designed to provide more accurate rain estimates and vertical latent heating profiles than have been possible heretofore. It is the first satellite to carry a meteorological radar. Orbiting the earth at 350 km altitude at a 35° inclination, it offers extra tropical coverage, higher resolution, and changing local observation times to help map out the diurnal cycle of rainfall.

There are a number of sources of error in the monthly averages of satellite-estimated rainfall. Since the sat-

* Additional affiliation: SM&A Corporation (East), Arlington, Virginia.

*Corresponding author address:* Dr. Thomas L. Bell, Code 913, Goddard Space Flight Center, Greenbelt, MD 20771.
E-mail: Thomas.L.Bell.1@gsfc.nasa.gov

ellite is unable to provide continuous coverage of a given area, averages of its observations will suffer from sampling error. Even when the satellite is viewing a given area, rain rates within the area must be inferred with remote sensing methods and are therefore subject to retrieval errors. Various mechanical and electronic problems, geolocation and data transmission problems, computer software problems, etc. can all contribute to the uncertainties. Astin (1997) reviews recent studies of sources of error in global, gridded averages of satellite-derived fields.

Early studies of the potential accuracy of TRMM monthly averages of rainfall made plausible assumptions to simplify the task. One important assumption was that the retrieval errors in rain-rate estimates for each field of view (FOV) of the satellite instrument are independent (uncorrelated from FOV to FOV). A back-of-the-envelope calculation (Wilheit 1988; Bell et al. 1990) suggests that even if rain-rate estimates for each FOV are accompanied by a random factor-of-2 retrieval error, the net error averaged over a month will be small under this assumption when compared with the sampling error. This assumption of uncorrelated retrieval errors requires further examination, however.

Laughlin (1981) was the first to attempt an estimate of the sampling error relevant to TRMM monthly averages. Microwave data from satellites are most easily converted to rain-rate estimates when the data are taken over the open ocean; and it is over the oceans that rain data are particularly scarce. Estimates of TRMM accuracies over the ocean are therefore of special interest. Using radar-derived rain data taken over the open ocean near the intertropical convergence zone during the Global Atmospheric Research Program (GARP) Atlantic Tropical Experiment (GATE), Laughlin (1981) was able to show that monthly averages of satellite-observed rain in the GATE area would have rms sampling errors of the order of 10% of the mean. Subsequent studies (McConnell and North 1987; Shin and North 1988; Bell et al. 1990; North et al. 1993; Bell and Kundu 1996) using the same data but more realistic representations of both rain statistics and satellite sampling arrived at a similar conclusion.

Studies of sampling error using data from other regions followed. Seed and Austin (1990) estimated sampling error with radar-derived rain rates from Patrick Air Force Base in Florida. They were the first to extend sampling studies to datasets different from the GATE data, and the first to raise the possibility that the GATE data may have given lower relative sampling-error estimates than might generally be the case. Soman et al.(1995, 1996) and Li et al. (1996) further enlarged the catalog of sites for which sampling-error estimates have been studied by using data from the vicinity of Darwin, Australia. Steiner (1996) carried out an extensive study of sampling error based on rain-gauge and some radar data from Darwin, as well as rain-gauge data from Melbourne, Florida. Oki and Sumi (1994) describe a sophisticated study using rain data over southern Japan. Weng et al. (1994) also looked at data from Japan. The error estimates obtained in these studies vary substantially and are often much larger than the earlier GATE-based estimates.

Chang et al. (1993) and Weng et al. (1994) examined the errors in averages of rain rates obtained from Special Sensor Microwave/Imager (SSM/I) instruments on the Defense Meteorological Satellite Program satellites. Berg and Avery (1995) develop a careful error budget analysis for SSM/I averages. All of these studies suggest that sampling error, as a fraction of the mean, is larger for regions with less rain.

Such a dependence is intuitively plausible and can be shown to follow quantitatively from simple assumptions about the statistical characteristics of rainfall. One of the purposes of this paper is to explore this relationship to see how well it can be used to describe the variations in sampling error found in studies such as those mentioned above. Ideally, such a relationship could be used to estimate the error in satellite-derived rain maps without requiring ground-based data from each location, which are, practically speaking, impossible to obtain.

In section 2 it will be shown that if rain events occur in well-separated places in a grid box more or less independently, with varying intensities and spatial extents but with the same statistical distribution, then the sampling error in satellite averages can be predicted knowing just the sampling provided by the satellite and the mean rain rate. If $A =$ the grid-box area, $R =$ the mean rain rate, and $S =$ the number of satellite visits during a month, the relative sampling error is predicted to be $\sigma_{\mathrm{samp}}/R = k(RAS)^{-1/2}$, where $k$ is a site-dependent constant but is not expected to depend strongly on location or season. A somewhat similar relationship has been suggested by Huffman (1997). Although quite a few assumptions are made in order to reach this point, these assumptions, though not always articulated, are often found in discussions of error estimates for satellite data. The relationship $\sigma_{\mathrm{samp}}/R = k(RAS)^{-1/2}$ provides a framework for evaluating datasets generated by satellites. At the very least, the error estimates made using these assumptions are almost certainly lower bounds for the true rms error and should serve as a foundation for more elaborate attempts to grapple with the problem.

In section 3 the satellite sampling-error studies mentioned above will be discussed in this framework. Section 4 suggests some reasons for changes in the constant $k$ and discusses methods for estimating it directly from satellite data. Our conclusions are presented in section 5. Appendix A fills in some mathematical details of the simple model omitted from the text. Appendix B describes in detail the problems associated with the method of ''resampling by shifts'' used in some studies to estimate satellite sampling error.

## 2. Simple theory

### a. Definitions

A statistical framework for characterizing the errors in monthly averages of rainfall over 5° × 5° grid boxes on the earth will be developed first. Producing maps of rainfall at this resolution is one of the principal goals of TRMM. The averaging time and area are somewhat arbitrary, but the accuracy of the averages will decrease if the box area or time span for the averages is decreased. The choice made for TRMM averages anticipates that, at this resolution, grid-box averages in the rainiest areas of the Tropics should have accuracies in the neighborhood of 10%, which will make meaningful quantitative comparisons with climate model behavior possible (Simpson et al. 1988).

The monthly averaged rain rate for a grid box with area $A$ is defined as

$$R = \frac{1}{T}\int_0^T dt\,\frac{1}{A}\int_A d^2\mathbf{x}\; R(\mathbf{x}\ t), \tag{2.1}$$

where $R(\mathbf{x}, t)$ is the instantaneous rain rate at location $\mathbf{x}$ and time $t$, $T = 1$ month, and $\int_A$ denotes integration over the grid-box area. The beginning of the month is arbitrarily set at $t = 0$. Although the basic quantity discussed here is always the average rain rate $R$, readers who prefer to work with rain accumulation $TR$ can convert all the results given here to those units by multiplying by the appropriate factor of $T$.

Let the visit times of the satellite during the month be denoted by $\{t_i, i = 1, \ldots, n\}$, and the portion of $A$ seen by the satellite instrument during visit $i$ be $A_i$, $A_i \le A$. It will be convenient to use the notation

$$R_B(t) \equiv \frac{1}{B}\int_B d^2\mathbf{x}\; R(\mathbf{x}, t) \tag{2.2a}$$

for the instantaneous rain rate averaged over an area $B$, in terms of which the true average rain rate seen by the satellite at overpass $i$ is defined as

$$R_i \equiv R_{A_i}(t_i). \tag{2.2b}$$

Satellite estimates of $R$ are typically made by collecting all the instrument footprints or FOVs that fall within the area $A$ during the month, converting the observations to rain-rate estimates, and averaging them (by summing them and dividing by the total number of observations). The average can be adjusted to account for nonuniform spacing of the footprints, if any. Such an estimate is equivalent to the linear average

$$\hat{R} = \frac{1}{n}\sum_{i=1}^{n} w_i \hat{R}_i, \tag{2.3}$$

where $\hat{R}_i$ is defined as the best estimate of the true rain rate $R_i$ occurring in the area $A_i$ that can be made using the satellite data, and where the weights $w_i \propto A_i/A$ are normalized to $n^{-1}\sum_i w_i = 1$. Still better estimates of $R$ can be made with different choices for $w_i$ that take into account the space–time correlations of rain, but these will not be considered here (see Bell and Kundu 1996).

The mean squared error of the satellite estimate (2.3) is

$$\sigma_E^2 = \langle(\hat{R} - R)^2\rangle, \tag{2.4}$$

where the angular brackets denote an average over an ensemble of rain scenarios with a probability distribution representing the local monthly rain climatology in the area $A$. Stated thus, the ensemble average is not an especially well-defined concept, nor is it directly computable from a limited dataset without further simplifying assumptions. The error $\sigma_E$ is abstract, in exactly the same sense that the concept of a rainfall climatology in $A$ is an abstraction. The rain climatology for $A$ is a statistical characterization of what *might* happen in $A$ during a month when all the environmental factors that affect rain statistics in $A$ are specified (e.g., season, sea surface temperature, large-scale wind patterns, etc.). The rain climatology does not tell us what actually *did* happen during that particular month. When $\sigma_E$ is obtained for every grid box, the result is a global ''error climatology'' for the satellite rain estimates. Note that $\sigma_E$ calculated this way is an average over an ensemble in which both $R$ and, more importantly, $\hat{R}$ vary. This error estimate differs from the error estimate $\sigma_E(\hat{R})$ that would be obtained from an ensemble in which $R$ varies but $\hat{R}$ is fixed at its measured value. Such an error estimate would certainly be interesting but is much more difficult to obtain.

Practically speaking, only a limited stretch of data will be available on which to base an estimate of (2.4). One must generally assume that the rain events present in the dataset are typical of what might occur in $A$. Such is the case in all of the studies mentioned in the introduction, where an ensemble of months is manufactured from the available data either by shifting the dataset in time (e.g., the resampling-by-shifts method) or by finding a statistical model to represent the data.

The total error $\sigma_E$ has several components. If the satellite could measure the rain in the area $A_i$ perfectly, so that $\hat{R}_i$ could be replaced by $R_i$ in (2.3), then the ''perfect instrument'' estimate of $R$ would be

$$\hat{R}_s = \frac{1}{n}\sum_{i=1}^{n} w_i R_i, \tag{2.5}$$

where the subscript $s$ indicates that this is the satellite-sampled estimate of the true monthly average $R$. The only error in $\hat{R}_s$ is due to imperfect coverage by the satellite. The mean squared error (2.4) can then be written

$$\begin{aligned}\sigma_E^2 &= \langle[(\hat{R} - \hat{R}_s) + (\hat{R}_s - R)]^2\rangle \\ &= \langle\varepsilon_{\mathrm{retr}}^2\rangle + \langle\varepsilon_{\mathrm{samp}}^2\rangle + 2\langle\varepsilon_{\mathrm{retr}}\varepsilon_{\mathrm{samp}}\rangle,\end{aligned} \tag{2.6}$$

with

$$\varepsilon_{\mathrm{retr}} = \hat{R} - \hat{R}_s \tag{2.7}$$

$$\varepsilon_{\mathrm{samp}} = \hat{R}_s - R, \tag{2.8}$$

so that $\varepsilon_{\mathrm{retr}}$ is the error due to remote sensing, referred to here as the ''retrieval error;'' and $\varepsilon_{\mathrm{samp}}$ is the error due to noncontinuous sampling by the satellite.

In part because it makes further progress so much easier, it is customary and almost irresistible to assume that sampling and retrieval errors are uncorrelated, so that

$$\langle \varepsilon_{\mathrm{retr}} \varepsilon_{\mathrm{samp}} \rangle \approx 0, \tag{2.9a}$$

$$\sigma_E^2 \approx \sigma_{\mathrm{retr}}^2 + \sigma_{\mathrm{samp}}^2, \tag{2.9b}$$

with

$$\sigma_{\mathrm{retr}}^2 \equiv \langle \varepsilon_{\mathrm{retr}}^2 \rangle, \tag{2.10}$$

$$\sigma_{\mathrm{samp}}^2 \equiv \langle \varepsilon_{\mathrm{samp}}^2 \rangle. \tag{2.11}$$

Assumption (2.9a) is at first sight plausible, because it seems to affirm the fairly commonplace assertion that the error made in a measurement should not depend on when the observation is made, something usually enforced in laboratory settings.

It is not immediately obvious why there should be a correlation in the retrieval error of a satellite estimate and whether or not it happens to over- or undersample the rain. There is, however, a particular kind of error that can affect averages of retrieved rain rate when the retrieval has a bias dependent on the type of rain or some other rain characteristic that changes slowly in space or time (the relative amounts of area covered by convective and stratiform rain, for instance, or a slowly changing characteristic of a strong diurnal cycle). The retrieval error will then contain significant biases that vary from month to month or grid box to grid box, even though the net bias over long enough times or large enough areas averages to zero. Such a ''varying bias'' will correlate with whether or not the rain is missed. There are surely other ways that such a correlation can enter. It should be noted, however, that such a correlation could, at worst, double the estimate given by (2.9b).

Equation (2.9b) assumes that neither the sampling nor the retrieval has a long-term bias (i.e., $\langle \varepsilon_{\mathrm{retr}} \rangle = \langle \varepsilon_{\mathrm{samp}} \rangle = 0$). If a nonzero bias were present, its square would have to be added to (2.9b). If such a bias were known, however, it would normally have been added as a correction to the estimates.

As mentioned above, it has been argued (Wilheit 1988; Bell et al. 1990) that if retrieval errors are uncorrelated with each other from one instrument FOV to another, then their contribution to the total error should be relatively small even if individual errors are quite large, because a large number of observations are averaged together to calculate $\hat{R}$; that is, $\sigma_{\mathrm{retr}}^2 \ll \sigma_{\mathrm{samp}}^2$. If this is correct, only the sampling-error component $\sigma_{\mathrm{samp}}$ is needed in order to get a good estimate for $\sigma_{\mathrm{E}}$. Just as before, however, this assumption too might be invalidated by the type of ''variable bias'' error mentioned above. This subject will be revisited at the end of the paper.

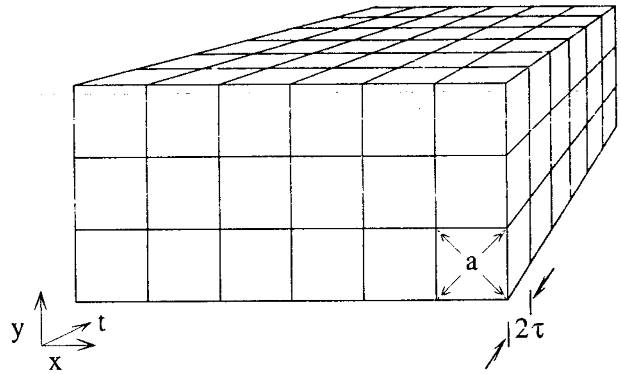


FIG. 1. Sketch of rain model where rain at one location and time is uncorrelated with rain at locations separated by distances of order $a^{1/2}$ and by times of order $\tau$. The space–time volume $AT$ is divided into cells of volume $a(2\tau)$.

*b. Simple model*

By making some simple assumptions about the statistical behavior of rain, an equation for the dependence of $\sigma_{\mathrm{samp}}$ on the rain statistics in the area can be derived, which helps make some sense of the error estimates obtained for various places and seasons. Let us suppose that over a sufficiently large area $A$ and a long enough period $T$ rain occurs as separate, uncorrelated events. These events have certain mean properties such as area covered, intensity, and duration. If that is the case, the amount of rain that falls in the area $A$ during the time $T$ is largely determined by the number of rain events that occur, if they are sufficiently numerous. Although each event may produce rain amounts that differ from the mean, these differences will tend to average out. A picture such as this underlies the area–time-integral (ATI) methods of estimating rainfall from the fraction of the area covered by rain (Chiu 1988; Kedem et al. 1990; Kedem and Pavlopoulos 1991; Short et al. 1993).

In this picture, a rain event will, on average, occupy an area $a$ and last for a time $2\tau$, where $\tau$ is the correlation time of the rain and the factor of 2 appears following Leith's (1973) result that the mean of a correlated series of length $T$ behaves as if it were the sum of $T/(2\tau)$ independent samples, at least for the purpose of placing confidence limits on the mean. Rain rate during an event will be, on average, $r$. Divide the space–time volume $AT$ into cells each of size $a(2\tau)$, the ''size'' of one rain event, as sketched in Fig. 1. There is a total of

$$N = AT/(2a\tau) \tag{2.12}$$

cells in the space–time volume. Suppose $N'$ of these cells have rain in them. Then if the mean rain rate during an event is $r$ [i.e., the mean rain rate conditional on

$R(x, t) > 0$], then the monthly averaged rain rate over the area $A$ will be

$$R = (N'/N)r. \tag{2.13}$$

The satellite views the area at certain instants $\{t_i, i = 1, \ldots, n\}$ and sees portions $A_i$ of the area $A$ at time $t_i$. Because of the space–time correlation of the rain, if the satellite swath intersects any part of a cell in $AT$ it ''knows'' whether an event has occurred in that cell. The number of cells the satellite knows about from its overflights at times $\{t_i\}$ is

$$N_{\text{obs}} = \sum_{i=1}^{n} A_i/a. \tag{2.14}$$

Among the observed cells, $N'_{\text{obs}}$ have rain in them, and so the satellite sample estimate of the mean rain rate is

$$\hat{R} = (N'_{\text{obs}}/N_{\text{obs}})r. \tag{2.15}$$

The sampling error (2.11), then, is

$$\sigma^2_{\text{samp}} = \left\langle \left( \frac{N'_{\text{obs}}}{N_{\text{obs}}} - \frac{N'}{N} \right)^2 \right\rangle r^2. \tag{2.16}$$

Given the assumptions made here, appendix A shows that (2.16) implies

$$\sigma^2_{\text{samp}} \approx \left( \frac{1}{N_{\text{obs}}} - \frac{1}{N} \right) Rr, \tag{2.17}$$

where $R/r$ has been substituted for the probability $p = N'/N$ that one of the observed cells has rain in it, using (2.13). Upon substituting (2.12) and (2.14) into (2.17), one obtains

$$\frac{\sigma^2_{\text{samp}}}{R^2} \approx \frac{r}{R} \left( \frac{a}{\sum_i A_i} - \frac{2a\tau}{AT} \right) = \frac{ar}{AR} \left( \frac{1}{S} - \frac{2\tau}{T} \right)$$

or

$$\frac{\sigma_{\text{samp}}}{R} \approx c(\tau) \left( \frac{r}{R} \frac{a}{A} \frac{1}{S} \right)^{1/2}, \tag{2.18}$$

where $S$ is the ''effective number of visits'' with full coverage of $A$ by the satellite,

$$S \equiv \sum_i A_i/A, \tag{2.19}$$

and

$$c(\tau) = \left( 1 - \frac{2\tau}{T/S} \right)^{1/2}. \tag{2.20}$$

The ratio $T/S$ in (2.20) may be thought of as the effective time interval between full visits by the satellite. The factor $c(\tau)$ depends on grid-box latitude, since $S$ does. When $T/S$ is large compared with the independent sample time $2\tau$ (i.e., poor sampling), the coefficient $c(\tau) \approx 1$ and does not depend much on the correlation time. For effective sampling intervals comparable to $\tau$, however, $c(\tau)$ differs significantly from 1, decreasing to 0 for $N_{\text{obs}} = N$, as it should, since all cells are seen by the satellite in this limit. The cell model is no longer a good approximation in this regime, however, and one must return to the basic definition (2.11) using a more accurate representation of the space–time correlation and sampling pattern.

### c. Laughlin model

Laughlin (1981) takes a step in this direction. He assumes that the variance

$$\sigma^2_A = \text{var}[R_A(t)] \tag{2.21}$$

of area-averaged rain rate $R_A(t)$ [defined in (2.2a)] is known and that its lagged autocorrelation in time decreases exponentially as $\exp(-|t - t'|/\tau_A)$, where $t - t'$ is the time lag and $\tau_A$ is the correlation time. The correlation time $\tau_A$ corresponds to the correlation time $\tau$ introduced in the simple model discussed above. Assuming also that the area $A$ is seen at equally spaced intervals $\Delta t$, he found

$$\sigma^2_{\text{samp}} = \sigma^2_A \frac{\Delta t}{T} c^2_L(\Delta t/\tau_A) + \mathcal{O}(1/T^2), \tag{2.22}$$

where

$$c^2_L(x) = \coth(x/2) - 2/x \tag{2.23}$$

(subscript $L$ for Laughlin). Similar derivations are given in Shin and North (1988) and Bell et al. (1990).

Equation (2.22) has the advantage that it yields a precise definition of $a$ in (2.18) in terms of the spatial correlation of the rain rate, but at the price of more assumptions. This can be seen by defining the spatial correlation of point rain rate

$$\rho(|\mathbf{x} - \mathbf{y}|) = \sigma_0^{-2} \langle R'(\mathbf{x}, t) R'(\mathbf{y}, t) \rangle, \tag{2.24}$$

where the primes indicate deviations from $\langle R \rangle$ and $\sigma^2_0 = \langle R'(x, t)^2 \rangle$ is the variance of rain rate at a point. We can then write $\sigma^2_A$ (with $A = L^2$) as

$$\sigma^2_A = \frac{1}{A} \int_0^L dx_1 \int_0^L dx_2 \, \frac{1}{A} \int_0^L dy_1 \int_0^L dy_2 \, \langle R'(x,t) R'(y,t) \rangle, \tag{2.25}$$

which, after some algebra, can be written

$$\sigma^2_A = \sigma^2_0 \, a/A, \tag{2.26}$$

where

$$a = 4 \int_0^L dx_1 \int_0^L dx_2 \left( 1 - \frac{x_1}{L} \right) \left( 1 - \frac{x_2}{L} \right) \rho(|\mathbf{x}|). \tag{2.27}$$

[Bell and Kundu (1996) give a more detailed discussion of this.] The area $a$ is the size of a ''statistically independent'' rain event. If the area $A = L^2$ is large compared with distances over which spatial correlations are significant, $a$ depends very little on $A$. It is a measure of the area over which rain rates are significantly correlated. (Note that if $\sigma_0^2$ is not well defined, as occurs in some idealized models, the above discussion must be replaced by one for a slightly smoothed or gridded rain field. The value of $a$ may then depend on the smoothing.)

Combining Eqs. (2.22) and (2.26), one obtains

$$\sigma_{\mathrm{samp}}^2 \approx \sigma_0^2 \frac{a}{A}\frac{\Delta t}{T} c_L^2(\Delta t/\tau_A). \tag{2.28}$$

If $p$ is the average fraction of time it rains over a point, then

$$R = pr_+, \tag{2.29}$$

$$\sigma_0^2 \approx p\sigma_+^2, \tag{2.30}$$

where $r_+$ is the (conditional) average rain rate $\langle R(\mathbf{x}, t) | R(\mathbf{x}, t) > 0\rangle$, and $\sigma_+^2$ is the variance of point rain rate conditional on $R(\mathbf{x}, t) > 0$. Equation (2.30) is approximate because terms of order $p^2$ are neglected, which is reasonable since $p$ is typically no larger than 0.1. Combining (2.28)–(2.30) one can reexpress (2.22) as

$$\frac{\sigma_{\mathrm{samp}}}{R} \approx c_L(\Delta t/\tau_A)\frac{\sigma_+}{r_+}\left(\frac{r_+}{R}\frac{a}{A}\frac{1}{S}\right)^{1/2}. \tag{2.31}$$

Laughlin's (1981) result (2.22) is thus equivalent to (2.18) to the first order in $1/T$, using $\sigma_A^2 \approx (N'/N)r^2$ [the analogue to (2.26) for $a \ll A$] and using $S = T/\Delta t$ and $\tau_A = \tau$, which is consistent with the simple model assumptions in the previous section. Beyond order $1/T$ the expressions differ, and the difference depends on the details of the time correlation.

### d. *Summary of model predictions*

Equation (2.18) has interesting implications for satellite sampling error. To the extent that rain has the statistical characteristics assumed in the model, and to the extent that sampling is, in a sense, not very good, the sampling error depends very little on the time correlation of rain. To the extent that some geographical locations have more rain in them than others chiefly because rain events occur more frequently in those areas rather than because the intensity or coverage of the individual events is greater, average relative error is proportional to $R^{-1/2}$. This increase with decreasing average rain rate is of course plausible, since a satellite is more likely to miss the rare rain events in a dryer area and so make larger relative errors in estimating the total amount of rain when compared with its performance over a grid box with a lot of rain. It would explain, at least qualitatively, the error dependence on rain rate noted by Chang et al. (1993) and Weng et al. (1994).

Equation (2.18) predicts that sampling error decreases with grid-box size $A$ and satellite visits $S$ as $(AS)^{-1/2}$. Ultimately, both this dependence and the dependence on $R$ are manifestations of the central limit theorem, which, in its simplest form, states that the average of $N$ identically distributed random variables with finite mean and variance differs from the true mean with rms difference proportional to $N^{-1/2}$. Any statistical model of rainfall postulating that rainfall during a month is the sum of the rainfall from a collection of independent rain events, each event having similar statistical properties, will give similar ''square root'' predictions.

If areas with more rain tend also to have larger spatial extent $[a = a(R)]$ and/or more intense rain $[r_+ = r_+(R)]$, the $R^{-1/2}$ dependence on $R$ predicted by (2.18) may be altered. As mentioned before, however, the success of ATI methods of estimating area-averaged rain rate suggests that such deviations might be small.

To what extent, then, can the relationship (2.18) explain more quantitatively the variability in sampling error estimates mentioned in the introduction? Let us turn now to those studies.

## 3. Sampling error estimates from data

Error studies have been carried out with data from different locations and time periods, for different satellite sampling patterns, and for different grid-box sizes. This section will attempt to present the sampling-error estimates from these studies in terms of what they would be for a common satellite sampling pattern and grid-box size. Once reduced to the same basis, their agreement with the prediction (2.18) can be examined. Results from some of these studies have unfortunately had to be discarded.

### a. *Methods used*

Three approaches have been used to estimate satellite sampling errors.

a) Direct method. The ensemble average in (2.11) would suggest that the mean-squared difference between the sample average (2.5) and the true average (2.1) be calculated using a sequence of ''statistically similar'' months from a dataset. Since the area $A$ is large, radar data with its wide coverage would be the first choice for such a study, although a dense rain gauge array might be used as well. For a given satellite sampling pattern, one month of data provides one value of the error (2.8). If enough months of data are available, $\langle \varepsilon_{\mathrm{samp}}^2 \rangle$ can be estimated. A straightforward application of this approach requires large amounts of data.

b) Resampling by shifts. An alternative method followed in some studies uses only a few months of data. A difference (2.8) is calculated using a sampling time sequence $\{t_i, i = 1, \ldots, n\}$ typical of the satellite visit times and falling within the period for which there are data. Then another difference (2.8) is obtained with a sequence of visit times $\{t_i + \delta, i = 1, \ldots, n\}$ shifted by $\delta$ from the previous sequence, then another with the sequence $\{t_i + 2\delta\}$, and so on, stopping when a shifted sequence requires data outside the period covered by the dataset. The value of $\delta$ is usually chosen to be the

TABLE 1. Satellite sampling-error studies. Geographical areas from which data were analyzed in the studies are indicated by crosses. The method(s) of estimating sampling error are indicated in the final column: method a, average over large dataset; method b, resampling by shifts; method c, model-based.

| | Geographical location | | | | | |
|---|---|---|---|---|---|---|
| Study | GATE | Darwin, Australia | Florida | Southern Japan | TOGA COARE | Method |
| Oki and Sumi (1994) | | | | X | | a |
| Steiner (1996) | | X | X | | | a, b* |
| McConnell and North (1987) | X | | | | | b |
| Seed and Austin (1990) | | | X | | | b |
| Soman et al. (1995) | | X | | | | b |
| Cosgrove and Garstang (1995) | | | X | | | b |
| Li et al. (1996) | | X | | | | b |
| Laughlin (1981) | X | | | | | c |
| North et al. (1993) | X | | | | | c |
| Bell and Kundu (1996) | X | | | | | c |
| Soman et al. (1996) | | X | | | | c** |
| Bell and Kundu (this paper) | | | | | X | c |

* Statistics from rain gauge dataset for 1 yr, corrected for unknown effects due to poor spatial sampling by rain gauges using results from a 2-month radar dataset.

** Confidence intervals for estimates based on the spectral method were difficult to determine. In addition, a diurnal cycle contributed significantly to the estimates and so made comparison with the results of the other studies difficult.

smallest time interval the dataset allows. The method will be referred to as resampling by shifts.

A large number of differences (2.8) can be manufactured this way. The mean-squared error $\langle \varepsilon^2_{\rm samp} \rangle$ is then estimated from the average of the squares of these differences. These differences can be highly correlated, though, when $\delta$ is small (i.e., $\delta \ll \tau$), and the procedure may generate few effectively independent differences. The sampling error in the estimate of $\sigma_{\rm samp}$ may itself therefore be quite large. This problem is examined in more detail in appendix B. It is shown there that if only a single month or less of data are used, as has frequently been the case, resampling by shifts could easily give answers off by a factor of 2 or more from what they should be, simply due to the time correlations and the small size of the dataset.

This, at least, may be an explanation for why the results of studies that employed resampling by shifts with only a few months of data seemed to exhibit no coherent pattern when we plotted them together with the results of other studies in the manner described in the next section. Because our analysis indicates that random error could overwhelm the results obtained using resampling by shifts, we have chosen to exclude these results from the analysis that follows. (This is by no means meant as a criticism of the studies as a whole, since they have all contributed much valuable insight to this difficult area of research.)

It should be noted that, despite the problems inherent in resampling by shifts, studies using only a few months of data are possible. The solution to this dilemma when only a small amount of data are available would be to chop the data into small sections, each of which is sufficiently long to preserve a good bit of the rain's time correlation (these sections would be about 16 h long in GATE), and then use randomized resampling techniques (Zwiers 1990; Wilks 1997; see also Solow 1985) to create an ensemble of simulated months from which to estimate $\sigma_{\rm samp}$.

c) Model estimates. Still another method for estimating $\sigma_{\rm samp}$ is to devise a statistical model of rain whose parameters are obtained from rain data, as Laughlin (1981) did with the GATE data. Satellite sampling error can then be calculated assuming that rain conforms to the model. The accuracy of the calculation depends on the model. The model must capture certain aspects of the true rain statistics well, and its parameters must be accurately estimated from the dataset.

### *b. Sampling-error studies*

Table 1 lists many of the sampling studies mentioned in the introduction. The method(s) used in each are indicated in the last column.

The study by Oki and Sumi (1994) over southern, coastal Japan used approximately 4 yr of rain gauge–adjusted radar data. Sampling-error estimates were made for five different $5° \times 5°$ grid boxes at approximate latitude 33.5° and for four different TRMM-like observation sequences. Confidence intervals for the average sampling error they report for each month of the year were estimated by us, based on the number of independent cases for each month (estimated to be 48–64) and using chi-squared statistics.

The study by Steiner (1996) used a long dataset from a rain gauge network near Darwin as well as a shorter radar dataset collected over the same region and data from Melbourne, Florida. These results will be discussed more later.

Results from the studies using resampling by shifts with datasets of order 1 month long (labeled ''b'' in Table 1) had to be discarded because of the large error bars for their results, as explained in appendix B. The results for Soman et al. (1996) for Darwin were not used because it was unclear how to evaluate the accuracy of the spectral method used, and because the results were reported for sampling at exact 24-h intervals, so that the diurnal cycle played a large role. TRMM sampling tends to yield statistics averaged over different hours of the day.

Model-based sampling-error estimates are labeled ‘‘c’’ in Table 1. Many of the model-based estimates used data from GATE. These rain data were taken at a site about 1000 km west of the African coast in the Atlantic Ocean (8°30′N, 23°30′W) during the summer of 1974. Gridded maps over a 400-km-diameter circular area derived from the data from two 18-day periods, phase I (28 June–16 July 1974) and phase II (28 July–15 August 1974), supplied the required statistics.

A model estimate by us is also listed in Table 1 that is based on data taken in the Tropical Ocean Global Atmosphere (TOGA) Coupled Ocean–Atmosphere Response Experiment (COARE). The statistics needed for the model estimates were obtained from the radar-derived rain-rate maps produced as part of the experiment. The preparation of the dataset is described by Short et al. (1997). The rain data were collected during three cruises of two ships in the western Pacific (near 2°S, 156°E) during the 4-month period November 1992–February 1993. Since the correlation time of the rain rate averaged over a 288-km-diameter circular area was generally 5 h or longer, comparable to the 6–8-h correlation time seen in the GATE data, the sampling-error calculations previously done using GATE statistics were simply scaled by us to give the corresponding TOGA COARE estimates.

For the model-based estimates using GATE data and TOGA COARE data, error bars were obtained by repeatedly resampling the datasets using randomly selected segments of the data 16 h in length reassembled to equal the total length of the dataset, a procedure similar to one suggested by Wilks (1997). (The segments were given lengths about twice the 8-h ‘‘correlation time’’ of the area-averaged rain rate.)

### *c. Translation of study results to same grid-box area and satellite sampling pattern*

After eliminating the studies in Table 1 whose quantitative results are problematical, sampling-error estimates over three, or perhaps four geographical areas remain. Since most of these studies use data covering areas 2.5° × 2.5° or smaller and focus on TRMM sampling error, the results of these studies have been extrapolated to what they would be for a 2.5° × 2.5° box located on the equator observed approximately once per day ($\Delta t \approx 1$ day). Such a satellite visits the box approximately 30 times per month, so that $S = 30$ in (2.18). (The SSM/I satellites, incidentally, have similar coverage, but the times of day of the observations fall in two clusters 12 h apart.)

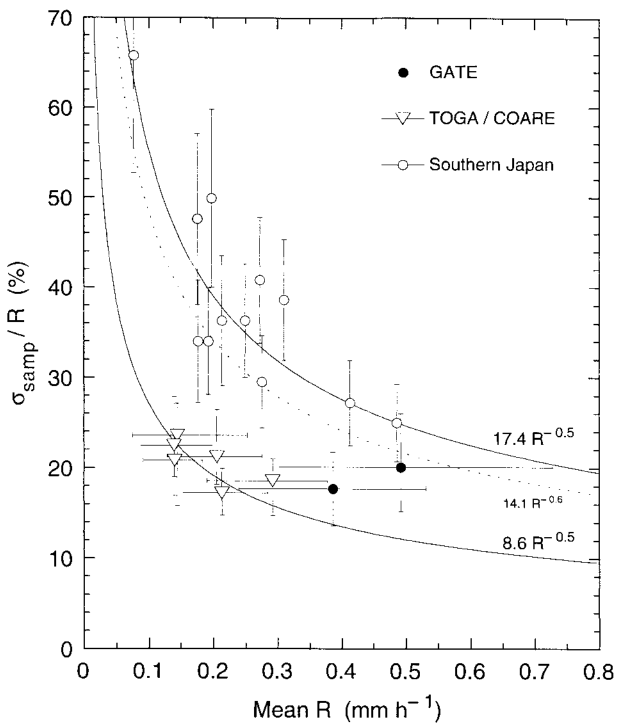


FIG. 2. Plot of relative sampling error $\sigma_{samp}/R$ found in various studies whose quantitative accuracy could be estimated (three locations altogether). The sampling errors are all for monthly averaged rain rate over a 2.5° × 2.5° grid box at the equator seen by the TRMM microwave instrument (equivalent to about 30 visits per month). Similar values would be predicted for SSM/I-based estimates (for a single satellite), aside from diurnal biases introduced by SSM/I’s sampling at the same times of day. Separate fits to $R^{-1/2}$ are shown for the southern Japan and TOGA COARE estimates. Error bars are 95% confidence intervals. The dashed line shows a curve fitted by Steiner (1996) to sampling errors he obtained using data from Darwin, Australia, and Melbourne, Florida.

In most cases, the extrapolation required is small. The extrapolation of the Oki and Sumi (1994) results relied most on the dependence on $A$ and $S$ predicted by (2.18) because of the size and latitude of the area studied by them. Oki and Sumi (1994) presented some results that seem to confirm the $A^{-1/2}$ dependence. The sampling in time of the TRMM satellite at 33.5°N that they used, however, is quite different from once-per-day sampling. Results obtained by Bell and Kundu (1996) for scaling of sampling error with $S$ suggest that the extrapolation using (2.18) should be accurate enough, but it would be preferable to sample the southern Japan dataset with visit patterns more similar to what the other studies used.

The extrapolated error estimates from all of the studies are plotted together in Fig. 2 with 95% confidence intervals (error bars) shown, except for Steiner’s (1996), which will be discussed next. The curve fitted to the TOGA COARE error estimates is added in order to make visual comparison of the various results easier.

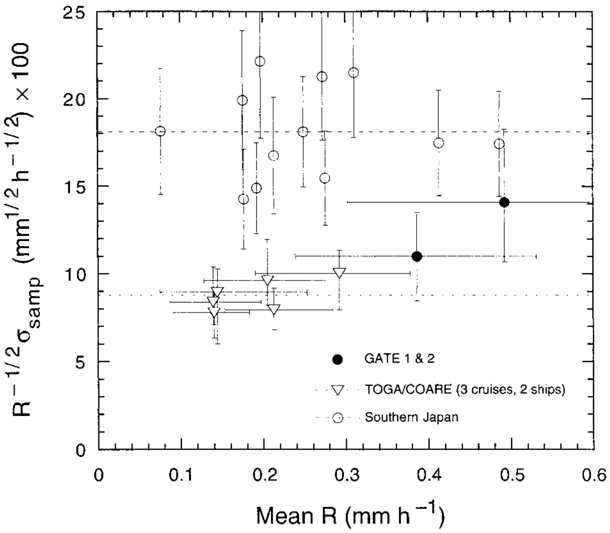


FIG. 3. Results in Fig. 2 multiplied by $R^{1/2}$ to extract coefficient $k$ in (3.1). The value of $k$ is predicted by Eq. (4.2).

The range of values of the TOGA COARE results by themselves is too small to provide a significant test of $R^{-1/2}$ behavior. While it might be argued that the error estimates of each study fall on curves

$$\frac{\sigma_{\mathrm{samp}}}{R} = kR^{-1/2}, \tag{3.1}$$

with the value of $k$ depending on the site, it is clear that the points from GATE and TOGA COARE do not fall on the same curve as the ones from southern Japan and seem unlikely to fall on a single curve themselves. This can be seen more clearly in the plot of $k = R^{1/2}(\sigma_{\mathrm{samp}}/R)$ in Fig. 3 (Steiner's fit not shown).

Steiner's (1996) sampling-error estimates are represented in Fig. 2 by the dashed curve. From his study of a large rain gauge dataset over Darwin, Australia, and Melbourne, Florida, he found a rain-rate dependence of sampling error which, if extrapolated to the conditions assumed in Fig. 2, would give a curve

$$\sigma_{\mathrm{samp}}/R = 0.141 \times R^{-0.6}, \tag{3.2}$$

if $R$ is measured in mm h$^{-1}$. This falls quite close to the results for southern Japan. Because of the spatial sampling characteristics of rain gauges and time averaging of the gauge data, however, it was difficult for Steiner (1996) to establish the effective area $A$ observed by the rain gauges. He found an overall correction factor for his rain gauge–derived relationship based on a resampling-by-shifts study using two months of radar data. Because of the issues discussed in appendix B, this correction is subject to some uncertainty. The agreement with the southern Japan values is nevertheless intriguing. The data from Japan, Florida, and Australia are all from coastal regions, whereas the data from GATE and TOGA COARE were taken over open ocean.

## 4. Discussion

If rain events, when they occur, had similar distributions everywhere with average strength $r_+$ and areal coverage $a$ [see Eq. (2.18) or (2.31)], sampling-error estimates obtained at one site could be extrapolated using (2.18) to other sites on the globe. The results of the sampling-error studies as depicted in Fig. 2, though far from covering a satisfactory range of environments, suggest that Eq. (2.18) might be valid at a single site but that the factor multiplying $(RAS)^{-1/2}$ can differ from site to site.

Before abandoning a single relationship (2.18) with a unique coefficient covering all situations, the possibility that the variations from site to site seen in Fig. 2 might be due to differences in measuring systems needs to be considered. Radar data are usually converted to rain rates using a power-law relation between radar reflectivity $Z$ and rain rate. The coefficient of the $Z$–$R$ relation is not always well known, but the ratio $\sigma_{\mathrm{samp}}/R$ is in any case independent of the choice of this coefficient. The value of $R$ for the points plotted in Fig. 2 (the abscissa) might be shifted up or down by changes in this coefficient or by calibration changes, but the factor-of-2 shift required to bring the curves more nearly into alignment would be uncomfortably large. Other uncertainties in the $Z$–$R$ relation could explain some of the variability but are unlikely to explain all of it.

Relations (2.22) and (3.1) can be used to rewrite $k$, defined as

$$k = R^{-1/2}\sigma_{\mathrm{samp}}, \tag{4.1}$$

as

$$k = [S^{-1/2}c_L(\Delta t/\tau_A)](R^{-1/2}\sigma_A). \tag{4.2}$$

The quantity in square brackets depends entirely on the sampling pattern of the satellite and the time correlation of the area-averaged rain rate. Figure 3 is a plot of 100 $\times$ $k$ as defined in (4.1), adjusted so that $S = 30$ for all of the points. From (4.2), variations in $k$ seen there are due either to changes in $R^{-1/2}\sigma_A$ or to the time correlations of $R_A(t)$ as reflected in $\tau_A$.

Correlation times in GATE for areas $A = 2.5° \times 2.5°$ were found to be of the order of 7–8 h (Laughlin 1981; Bell et al. 1990). Correlation times (as measured by the lag at which correlation falls to $1/e$) were found in TOGA COARE to range from 4 to 7 h during the course of the experiment. Oki and Sumi (1996) reported correlation times for $5° \times 5°$ grid boxes over southern Japan ranging from 8 h in winter to 14 h in the summer (when the highest rain rates occur). They argued that these longer correlation times for the $5° \times 5°$ box averages are probably consistent with the shorter times found for the smaller $2.5° \times 2.5°$ boxes. For the satellite sampling assumed in Figs. 2 and 3 (with $\Delta t \approx 1$ day or $S = 30$),

$$c_L(\Delta t/4\text{ h}) = 0.82, \qquad c_L(\Delta t/8\text{ h}) = 0.66,$$

and so variations in correlation time seen in the data could result in changes in the sampling-dependent constant $c_L(\Delta t/\tau_A)$ in (4.2) by about 25%. The remainder of the changes in $k$ seen in Fig. 3 would have to be explained by changes in $R^{-1/2}\sigma_A$ from site to site.

For southern Japan, a value of $k \approx 0.18$ seems to fit the data. If we accept Steiner's (1996) fit for the Darwin area as quantitatively accurate, his fit (3.2) would predict $k = 0.14R^{-0.1}$ or $k \approx 0.16$ over the range of rain rates studied by him. In the case of rain with the statistics of TOGA COARE and GATE, $k \approx 0.10$ seems to describe most of the data, except for the highest rain rates, where it seems to be larger.

## 5. Summary

When faced with determining sampling error for any given grid box, then, the sampling error for monthly averaged satellite estimates can be obtained from (3.1). Based on the somewhat limited evidence so far, it appears that the factor $k$ in (3.1) can be treated as approximately constant in a given area, and estimates of $k$ are available in a number of instances. If we fix the time correlation–dependent factor $c_L$ to be about 0.75 $\pm$ 0.10, our present knowledge might be summarized as

$$\frac{\sigma_{\text{samp}}}{R} \approx kR^{-1/2}, \qquad k \approx 0.75S^{-1/2}(R^{-1/2}\sigma_A). \tag{5.1}$$

For $S = 30$ (one visit per day by the satellite),

$$k \approx 0.18, \quad \text{[southern Japan, Darwin (?)]} \tag{5.2a}$$

$$k \approx 0.10, \quad \text{[GATE, TOGA COARE].} \tag{5.2b}$$

These conclusions are, of course, based on taking radar and rain gauge data at face value, as ‘‘truth.’’

Most of the variability of $k$ is likely to be due to changes in $R^{-1/2}\sigma_A$ rather than changes in the correlation time of the area-averaged rain. The value of $R^{-1/2}\sigma_A$ is governed by $r_+$ and $\sigma_+^2$, the conditional mean and variance of point rain rate, and by $a$, the typical areal extent of a rain event in that locale. Short et al. (1993) have noted that the ratio $\sigma_+/r_+$ seems to be relatively constant over a range of averaging areas, types of data (rain gauge or radar), and climates. As a result, it may be the areal coverage of a typical rain event in an area that has the greatest influence on $k$.

It is important to note that if retrieval error does not contribute significantly to the total error $\sigma_E$, the quantity $R^{-1/2}\sigma_A$ can be estimated directly from the data, since $\sigma_A^2$ is just the variance of grid-box averaged, instantaneous rain rate. Statistically stable estimates of $R$ and $\sigma_A$ using long enough time averages and/or spatial averages over data from each location are all that are required. This plus the number of samples per month $S$ is all that is needed to obtain $k$ [see Eq. (5.1)].

It is usually argued that the contribution of retrieval errors to averages of rain estimates over large areas will be small because the number of satellite-instrument footprints is large and estimates for each footprint have independent retrieval errors. As mentioned in section 2, however, there is an insidious kind of error that can creep into averages of retrieved rain rate when the retrieval error depends on some rain characteristic that changes slowly in space or time. An example might be the relative amounts of area covered by convective and stratiform rain, for instance. The average retrieval error for each month may not be small in such a situation, even though the retrieval error averaged over long enough times or large enough areas is zero. Such a situation could invalidate the argument for $\sigma_{\text{retr}}^2 \ll \sigma_{\text{samp}}^2$ as well as cause sampling errors and retrieval errors to be correlated. Estimates of $R^{-1/2}\sigma_A$ from satellite data would then also include such bias effects.

Since ground-based radar, rain gauge measurements, and remote sensing by satellite differ in nature so profoundly, it would not be surprising to find such slowly changing biases in all of these methods. Discovering these effects and trying to estimate them and eventually correct for them will be an important challenge in the coming years.

*Acknowledgments.* We wish to thank Dr. C. Kummerow for very stimulating discussions. This research was supported by the Office of Earth Science of the National Aeronautics and Space Administration as part of the Tropical Rainfall Measuring Mission.

## APPENDIX A

### Sampling Error of Simple Model

Some details concerning the evaluation of the average on the right-hand side of Eq. (2.16),

$$\left\langle \left( \frac{N'_{\text{obs}}}{N_{\text{obs}}} - \frac{N'}{N} \right)^2 \right\rangle, \tag{A1}$$

are discussed here. Assign a binary random variable $y_j$ ($y_j = 0$ or 1) to each of the $N$ cells in the volume $AT$ sketched in Fig. 1, with $j = 1, \ldots, N$. The occurrence of a rain event in cell $j$ is indicated by $y_j = 1$. By definition,

$$N' = \sum_{j=1}^{N} y_j. \tag{A2}$$

The rain events occur randomly in the volume, and so the value of each $y_j$ is unpredictable. The satellite views $N_{\text{obs}}$ cells. Since the index $j$ can be assigned arbitrarily to a cell, arrange the labeling so that the observed cells are labeled by $j = 1, \ldots, N_{\text{obs}}$ and the unobserved cells are labeled by $j = N_{\text{obs}} + 1, \ldots, N$. With this choice,

$$N'_{\text{obs}} = \sum_{j=1}^{N_{\text{obs}}} y_j. \tag{A3}$$

The average (A1) may be defined in two ways, with slightly different results: (a) The number of rainy cells

$N'$ can be fixed, so that (A2) acts as a constraint on the random variables $y_j$; or (b) the *average* number of rainy cells, $\langle N' \rangle$, can be fixed instead, so that the probability of any cell having rain in it is some chosen fraction of $N$. If the first choice is made, then (A2) requires that

$$\langle y_j \rangle = N'/N \equiv p, \tag{A4}$$

and since $y_j^2 = y_j$ (because $y_j = 0, 1$),

$$\langle y_j^2 \rangle = \langle y_j \rangle = p. \tag{A5}$$

Since the location of the rain events is unpredictable in this model, $\langle y_j y_k \rangle = \langle y_j \rangle \langle y_k \rangle = p^2$ if $j \neq k$, so that

$$\langle y_j y_k \rangle = p^2 + \delta_{jk}(p - p^2), \tag{A6}$$

where $\delta_{jk}$ is the Kronecker delta. Using Eqs. (A2)–(A6) in (A1), it is easy to show that

$$\left\langle \left( \frac{N'_{\text{obs}}}{N_{\text{obs}}} - \frac{N'}{N} \right)^2 \right\rangle = p(1 - p)\left( \frac{1}{N_{\text{obs}}} - \frac{1}{N} \right). \tag{A7}$$

If the second choice for the ensemble average in (A1) had been made instead, the factor $(1 - p)$ would not be present in (A7). Since rainy events typically occur less than 10% of the time, $p$ is small, and the two ensemble averages have nearly the same outcome. This is interesting, since it suggests that at least one of the ambiguities in defining the ensemble average in (2.4), whether to fix the total rainfall or only the ensemble average rainfall, may not make much quantitative difference.

Equation (2.17) follows immediately from (A4), (A7), and (2.13), neglecting the factor $(1 - p)$.

## APPENDIX B

## Rms Error in Sampling Error Estimated with Resampling by Shifts

So, naturalists observe, a flea
Hath smaller fleas that on him prey;
And these have smaller fleas to bite ’em,
And so proceed ad infinitum.

Jonathan Swift (1667–1745)

Numerous estimates of the size of TRMM’s sampling error have appeared in the literature, but some have had to be discarded here because the method and dataset size used in obtaining them may have given inaccurate results. This appendix discusses these issues in more detail.

The datasets used in these studies are usually of the order of a month long and consist of a series of radar- or rain gauge–derived rain field maps. Spacing in time is once per hour or perhaps more frequent than that. The method involves comparing the ‘‘true’’ average rain rate, taken to be the average over all the rain data during the month, with the average of a sample of the dataset taken with a sampling pattern typical of the satellite.

The satellite sampling pattern is usually idealized,

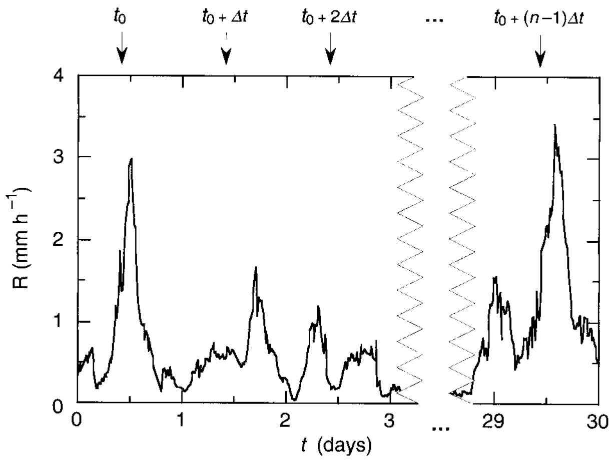


FIG. B1. Illustration of the method of resampling by shifts. The area-averaged rain-rate data are shown as the continuous curve plotted vs time. The data-span $T$ is typically about one month. The satellite observes the rain at intervals $\Delta t$, making $n = T/\Delta t$ observations during the month. The satellite estimate is obtained from the average of the satellite observations. ‘‘New’’ satellite estimates are generated by shifting all the satellite observation times by $\delta$, then $2\delta$, and so on.

with observations assumed to occur at equally spaced intervals $\Delta t$ and covering the entire grid-box area; that is, if the first satellite observation occurs at $t = t_0$, successive observations occur at $t_k = t_0 + k\Delta t$, $k = 1, \ldots, n - 1$, with $n\Delta t \sim 1$ month. Figure B1 illustrates such a pattern. The interval $\Delta t$ is typically of order 1 day for the TRMM studies.

The difference between the true monthly average $\overline{R}(t_0)$ and the sample average

$$\hat{R}(t_0) = \frac{1}{n} \sum_{k=0}^{n-1} R(t_0 + k\Delta t) \tag{B1}$$

is the sampling error

$$\varepsilon(t_0) = \hat{R}(t_0) - \overline{R}(t_0). \tag{B2}$$

For studies using only a single month of data, the true average $\overline{R}(t_0)$ is usually taken to be the average over all of the data and so does not, in fact, depend on $t_0$. To generate an ensemble of sampling errors to approximate the bracket average in (2.4), new error values are generated by shifting the satellite observation times. The first new set is obtained by shifting the original set by $\delta$ to $t_k^{(1)} = t_k^{(0)} + \delta$, with $t_k^{(0)} = t_0 + k\Delta t$. The next set of observation times is $t_k^{(2)} = t_k^{(0)} + 2\delta$, and so on, until an observation time is required that lies beyond the limits of the dataset. This method resamples the dataset, attempting to simulate the results that would be obtained from a straightforward study using a bigger dataset. With one month of data, the number of values of $\varepsilon_{\text{samp}}$ generated by this approach is $\Delta t/\delta$.

Papers that make use of the method are identified in Table 1 by the label ‘‘b.’’ Details of the implementation of the method vary among the papers. In one of the papers, for example, the dataset used is longer than one month, and in that study the period over which the true monthly average is taken can shift along with the sample starting point $t_0$. None of these variations affect the conclusions reached here in a substantial way, however.

The estimate of $\sigma_{\rm samp}$ obtained using the method described above is subject to variability about its true value because of the limited size of the dataset used. We will attempt to estimate this variability using Laughlin's (1981) model of the rain statistics. The time series $R(t)$ of the area-averaged rain rate provided by the dataset is assumed to be available for a period of length $T \sim$ 1 month. The autocorrelation of $R(t)$ at times separated by $\tau$ is assumed to be $\exp(-|\tau|/\tau_A)$. To simplify the algebra in the calculations below, time will be measured in units of $\tau_A$, and rain-rate variance in units of $\sigma_A^2$.

In the papers mentioned above, the value of $\delta$ is usually set to the smallest time interval between successive values of $R(t)$ available in the dataset. In the calculation here, however, we let $\delta \to 0$ so that $t_0^{(i)}$ can be treated as a continuous variable, which we shall denote by $t_0$, $0 \le t_0 < \Delta t$. This change from discrete to continuous variables makes very little difference because $\delta \ll \tau_A$, and the problem raised here is, if anything, underestimated due to this change. For a given starting point $t_0$, the true average is defined to be

$$\overline{R}(t_0) = \frac{1}{T}\int_0^T dtR(t). \tag{B3}$$

It does not depend on $t_0$ because the dataset is assumed to be one month long. If the dataset spans a period longer than $T$, an alternative definition would be

$$\overline{R}(t_0) = \frac{1}{T}\int_{t_0}^{t_0+T} dtR(t). \tag{B4}$$

This alternative would have the advantage that the position of the samples in the sample average relative to the period averaged over would always be the same, no matter what $t_0$ is. It changes the calculation slightly but does not change the conclusions significantly, unless the dataset is many multiples of $T$ in length. This is not the case for any of the studies discussed here.

We assume unbiased measurements and use the result (2.22) by Laughlin (1981), so that

$$\langle \varepsilon(t_0)\rangle = 0 \tag{B5}$$

$$\sigma_{\rm samp}^2 = \langle \varepsilon^2(t_0)\rangle = \frac{\Delta t}{T}\coth\left(\frac{\Delta t}{2T}\right) - \frac{2}{T} \tag{B6}$$

($\sigma_A^2 = \tau_A = 1$ by assumption). Laughlin's formula is used only to order $1/T$ and does not depend on $t_0$.

The resampling technique shifts $t_0$ progressively from 0 to $\Delta t$, getting a set of sampling errors $\varepsilon(t_0)$, $0 \le t_0 < \Delta t$, and then estimates $\hat{\sigma}_{\rm samp}^2$ by averaging over the set:

$$\hat{\sigma}_{\rm samp}^2 = \frac{1}{\Delta t}\int_0^{\Delta t} dt_0 \varepsilon^2(t_0). \tag{B7}$$

We want to know the variance of $\hat{\sigma}_{\rm samp}^2$ about its expected value (B6). [Note that $\langle\hat{\sigma}_{\rm samp}^2\rangle$ as defined in (B7) will differ from Laughlin's result for terms of the order $1/T^2$ and higher.] The variance of $\hat{\sigma}_{\rm samp}^2$ about $\sigma_{\rm samp}^2$ due to having only a finite dataset is

$$\begin{aligned}\sigma_L^4 &\equiv \mathrm{var}(\hat{\sigma}_{\rm samp}^2) = \langle(\hat{\sigma}_{\rm samp}^2)^2\rangle - \langle\hat{\sigma}_{\rm samp}^2\rangle^2 \\ &= \frac{1}{(\Delta t)^2}\int_0^{\Delta t} dt_0 \int_0^{\Delta t} dt_0' [\langle \varepsilon^2(t_0)\varepsilon^2(t_0')\rangle - \langle\varepsilon^2(t_0)\rangle\langle\varepsilon^2(t_0')\rangle].\end{aligned} \tag{B8}$$

Equation (B8) after substituting (B2) contains terms of the form

$$\langle R(t_1)R(t_2)R(t_3)R(t_4)\rangle.$$

If the statistics of $R(t)$ are approximated by those of normally distributed variables, expectations of quartic terms can be written in terms of products of expectations of quadratic terms (see Anderson 1958, p. 39, for instance). The same is therefore true of $\langle\varepsilon^2(t_0)\varepsilon^2(t_0')\rangle$. If we do that, and use the double-integral identity

$$\int_0^T dt_1 \int_0^T dt_2 f(t_1 - t_2) = 2\int_{-T}^T du(T - |u|)f(u), \tag{B9}$$

we obtain

$$\sigma_L^4 = \frac{4}{\Delta t}\int_0^{\Delta t} dt_0\left(1 - \frac{t_0}{\Delta t}\right)\langle\varepsilon(t_0)\varepsilon(0)\rangle^2. \tag{B10}$$

This assumes that $\langle\varepsilon(t_0)\ \varepsilon(t_0')\rangle$ is a function only of $|t_0 - t_0'|$, which is true for the terms of the order $1/T$ but not for the higher-order ones. [It is exact if definition (B4) is used instead of (B3).]

To calculate $\langle\varepsilon(t_0)\varepsilon(0)\rangle$, expand it using definition (B2):

$$\begin{aligned}\langle\varepsilon(t_0)\varepsilon(0)\rangle &= \langle R(t_0)R(0)\rangle + \langle\hat{R}(t_0)\hat{R}(0)\rangle \\ &\quad - \langle R(t_0)\hat{R}(0)\rangle - \langle\hat{R}(t_0)R(0)\rangle.\end{aligned} \tag{B11}$$

Each of these terms must be separately evaluated. We will assume that $T$ is sufficiently large that we are only interested in the lowest-order term in $1/T$. Mathematica Version 2.2 (Wolfram Research, Inc. 1988) was very helpful here.

The first term in (B11) does not actually depend on $t_0$ because of the way $R(t_0)$ is defined in (B3). This term was calculated by Leith (1973). To order $1/T$ he obtained

$$\langle R(t_0)R(0)\rangle = 2/T. \tag{B12}$$

The cross terms in (B11) are also easy to obtain:

$$\langle R(t_0)\hat{R}(0)\rangle = \langle\hat{R}(t_0)R(0)\rangle = 2/T. \tag{B13}$$

To order $1/T$ they are the same and do not depend on $t_0$.

Finally, the term $\langle\hat{R}(t_0)\hat{R}(0)\rangle$ can be computed:

$$\langle\hat{R}(t_0)\hat{R}(0)\rangle = \frac{e^{-t_0} + e^{t_0-\Delta t}}{n(1 - e^{-\Delta t})}. \tag{B14}$$

The fact that $n = T/\Delta t$ has been used.

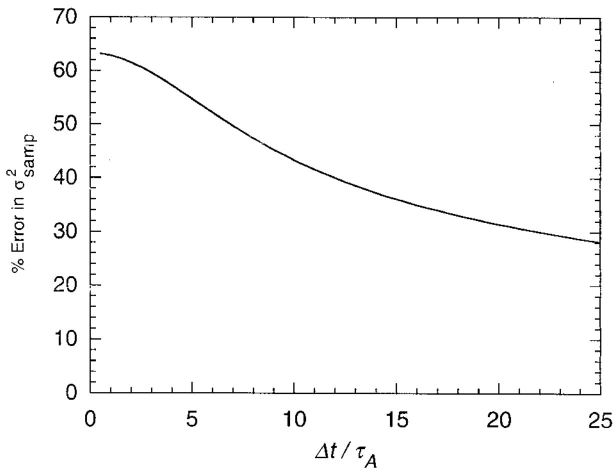


FIG. B2. Plot of percent error in $\hat{\sigma}^2_{\text{samp}}$, given by Eq. (B18), when $\sigma^2_{\text{samp}}$ is estimated using the resampling-by-shifts method from $\Delta t/\delta$ samples.

Substituting these results into (B11), we obtain, to order $1/T$,

$$\langle \varepsilon(t_0)\varepsilon(0)\rangle = \frac{e^{-t_0} + e^{t_0-\Delta t}}{n(1 - e^{-\Delta t})} - \frac{2}{T}. \tag{B15}$$

Note that for $t_0 = 0$ this agrees with Laughlin's result, as it should. Performing the integral in (B10) we obtain

$$\sigma_L^4 = \frac{2}{T^2}\frac{[2ux^2 + (1 + u)x - 4]}{(1 - u)^2}, \tag{B16}$$

with the definitions

$$x \equiv \Delta t/(1 - u), \qquad u \equiv e^{-\Delta t}. \tag{B17}$$

Since (B16) is the variance about Laughlin's sampling error $\sigma^2_{\text{samp}}$ due to the finite size of the dataset used to evaluate it, the ratio of its square root to $\sigma^2_{\text{samp}}$ is the fractional error of the estimate of $\sigma^2_{\text{samp}}$ obtained by the resampling method:

$$\frac{\sigma_L^2}{\sigma^2_{\text{samp}}} = \frac{2^{1/2}[2ux^2 + (1 + u)x - 4]^{1/2}}{x(1 + u) - 2}. \tag{B18}$$

It is plotted in Fig. B2. For the resampling method to estimate $\sigma^2_{\text{samp}}$ with a factor-of-2 accuracy, the fractional error (B18) should be less than about ⅓. This requires that $\Delta t \geq 18\tau_A$. Since the resampling studies discussed in these papers generally use $T \sim 1$ month and $\tau_A \sim 8$ h, a factor-of-2 accuracy for $\hat{\sigma}^2_{\text{samp}}$ could be achieved only for $\Delta t \gtrsim$ 6 days. It is therefore clear that estimates of sampling error for $\Delta t \sim 1$ day cannot be trusted quantitatively.

## REFERENCES

Anderson, T. W., 1958: *An Introduction to Multivariate Statistical Analysis.* Wiley, 374 pp.

Astin, I., 1997: A survey of studies into errors in large scale space-time averages of rainfall, cloud cover, sea surface processes and the earth's radiation budget as derived from low earth orbit satellite instruments because of their incomplete temporal and spatial coverage. *Surv. Geophys.,* **18,** 385–403.

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. R. North, 1990: Sampling errors for satellite-derived tropical rainfall: Monte Carlo study using a space–time stochastic model. *J. Geophys. Res.,* **95** (D3), 2195–2205.

Berg, W., and S. K. Avery, 1995: Evaluation of monthly rainfall estimates derived from the special sensor microwave/imager (SSM/I) over the tropical Pacific. *J. Geophys. Res.,* **100** (D), 1295–1316.

Chang, A. T. C., L. S. 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.

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

Cosgrove, C. M., and M. Garstang, 1995: Simulation of rain events from rain gauge measurements. *Int. J. Climate,* **15,** 1021–1029.

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

——, R. F. Adler, P. Arkin, B. Rudolf, U. Schneider, and P. R. Keehn, 1995: Global precipitation estimates based on a technique for combining satellite-based estimates, rain gauge analysis, and NWP model precipitation information. *J. Climate,* **8,** 1284–1295.

Kedem, B., and H. Pavlopoulos, 1991: On the threshold method for rainfall estimation: Choosing the optimal threshold level. *J. Amer. Stat. Assoc.,* **86,** 626–633.

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

Laughlin, C. R., 1981: On the effect of temporal sampling on the observation of mean rainfall. *Precipitation Measurements from Space, Workshop Report,* D. Atlas and O. W. Thiele, Eds., NASA Publication, D-59–D-66.

Leith, C. E., 1973: The standard error of time-average estimates of climatic means. *J. Appl. Meteor.,* **12,** 1066–1069.

Li, Q., R. L. Bras, and D. Veneziano, 1996: Analysis of Darwin rainfall data: Implications on sampling strategy. *J. Appl. Meteor.,* **35,** 372–385.

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

North, G. R., S. S. P. Shen, and R. Upson, 1993: Sampling errors in rainfall estimates by multiple satellites. *J. Appl. Meteor.,* **32,** 399–410.

Oki, R., and A. Sumi, 1994: Sampling simulation of TRMM rainfall estimation using radar–AMeDAS composites. *J. Appl. Meteor.,* **33,** 1597–1608.

Seed, A., and G. L. Austin, 1990: Variability of summer Florida rainfall and its significance for the estimation of rainfall by gages, radar, and satellite. *J. Geophys. Res.,* **95** (D), 2207–2215.

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.

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

——, P. A. Kucera, B. S. Ferrier, J. C. Gerlach, S. A. Rutledge, and O. W. Thiele, 1997: Shipboard radar rainfall patterns within the TOGA COARE IFA. *Bull. Amer. Meteor. Soc.,* **78,** 2817–2836.

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.

——, C. Kummerow, W.-K. Tao, and R. F. Adler, 1996: On the Tropical Rainfall Measuring Mission (TRMM). *Meteor. Atmos. Phys.,* **60,** 19–36.

Solow, A. R., 1985: Bootstrapping correlated data. *Math. Geol.,* **17,** 769–775.

Soman, V. V., J. B. Valdés, and G. R. North, 1995: Satellite sampling and the diurnal cycle statistics of Darwin rainfall data. *J. Appl. Meteor.,* **34,** 2481–2490.

——, ——, and ——, 1996: Estimation of sampling errors and scale parameters using two- and three-dimensional rainfall data analyses. *J. Geophys. Res.,* **101** (D), 26 453–26 460.

Steiner, M., 1996: Uncertainty of estimates of monthly areal rainfall for temporally sparse remote observations. *Water Resour. Res.,* **32,** 373–388.

Weng, F., R. R. Ferraro, and N. C. Grody, 1994: Global precipitation estimations using *Defense Meteorological Satellite Program F10* and *F11* special sensor microwave imager data. *J. Geophys. Res.,* **99** (D), 14 493–14 502.

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, 377–385.

Wilks, D. S., 1997: Resampling hypothesis tests for autocorrelated fields. *J. Climate,* **10,** 65–82.

Wolfram Research, Inc., 1988: Mathematica. Version 2.2. Wolfram Research, Inc.

Xie, P., and P. A. Arkin, 1996: Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *J. Climate,* **9,** 840–858.

Zwiers, F. W., 1990: The effect of serial correlation on statistical inferences made with resampling procedures. *J. Climate,* **3,** 1452–1461.