Sampling Errors in Rainfall Estimates by Multiple Satellites

GERALD R. NORTH
Climate System Research Program, College of Geosciences, Texas A&M University, College Station, Texas

SAMUEL S. P. SHEN
Mathematics Department, University of Alberta, Edmonton, Canada

ROBERT UPSON
Climate System Research Program, College of Geosciences, Texas A&M University, College Station, Texas

(Manuscript received 6 July 1991, in final form 27 May 1992)

ABSTRACT

This paper examines the sampling characteristics of combining data collected by several low-orbiting satellites attempting to estimate the space–time average of rain rates. The several satellites can have different orbital and swath-width parameters. The satellite overpasses are allowed to make partial coverage snapshots of the grid box with each overpass. Such partial visits are considered in an approximate way, letting each intersection area fraction of the grid box by a particular satellite swath be a random variable with mean and variance parameters computed from exact orbit calculations. The derivation procedure is based upon the spectral minimum mean-square error formalism introduced by North and Nakamoto. By using a simple parametric form for the space–time spectral density, simple formulas are derived for a large number of examples, including the combination of the Tropical Rainfall Measuring Mission with an operational sun-synchronous orbiter. The approximations and results are discussed and directions for future research are summarized.

1. Introduction

Several satellite missions intended to measure rainfall are in the planning and execution phase. For example, the Tropical Rainfall Measuring Mission (TRMM) is an earth probe currently being developed by the National Aeronautics and Space Administration (NASA) and the Japanese space agency, NASDA (cf. Simpson et al. 1988). Also under development is the Earth Observing System (cf., e.g., Baker 1990). A primary goal of these missions is the delivery of a space–time-smoothed time series of rain rates, especially over the tropical oceans. Typically, the duration of such a time series should be years, the temporal smoothing filter should be of the order of 1 month, and spatial smoothing should be over nominal square grid boxes of about 500 km on an edge. Since rain rates exhibit such irregular statistical behavior, it is a formidable problem to understand the logic of the error budget in such measurement configurations. Since virtually all feasible designs can only sample the field in the grid box at intervals of several hours, one must try to understand the magnitudes of the sampling errors and their interaction with the inherent measurement errors characteristic of the individual sensors. It is especially interesting to contemplate the combination of data from several satellites with differing orbital parameters and therefore differing intrinsic error structures.

The problem is sufficiently confusing that simplified treatments of the error problem that lead to "back of the envelope" formulas can be very useful. North and Nakamoto (1989, hereafter referred to as NN) introduced a technique to help sort out the important features that contribute to the total error variance due to a gappy sampling design. The method led to a formula for the sampling-error variance that consisted of an integral over the space–time spectrum of the rain-rate field weighted by a filter that depends only on the sampling design. In particular, they analyzed the cases of a single satellite overpassing the grid box at regular intervals making flush (100% coverage on each crossing) visits and separately a regular array of point raingages located at the surface. Other studies have been performed using simulated flights over a real dataset (McConnell and North 1987; Kedem et al. 1990), computed overflights of simulations of the area-averaged rain-rate field by an autoregressive process (Laughlin 1981; Shin and North 1988), and flights over a stochastic space–time rain field produced by a

Corresponding author address: Gerald R. North, College of Geosciences and Maritime Studies, Texas A&M University, College Station, TX 77843-3150.

model tuned to tropical oceanic rain data (Bell 1987; Bell et al. 1990).

North et al. (1991) have studied the possibility of combining raingages with the satellite overpass data. They found that this particular design combination has some simplifying features that are reminiscent of the estimation problem based upon statistically independent measurements on a population. This finding for obtaining an estimate in the present case where all data are collected from a highly dependent sample base is somewhat surprising at first glance. The orthogonality principle found seems to apply to any gauges (discrete points in space but continuous in time) when combined with snapshot data (continuous in space but gappy in time) such as that taken from satellite overpasses. The rule allows the data to be combined after the fact, weighting them inversely by the error variances that would be obtained as if each subsystem were operating alone.

The present paper augments the list of solved problems in this area by considering the case of multiple satellites and allowing the individual satellites to make realistic partial coverage visits at each overpass. The aim remains to provide simple formulas that can lead the planner to an approximate error analysis for this class of configurations without having to resort to tedious computer intensive case-by-case simulations of random rain-rate fields overflown by concurrent exactly computed orbits. That a brute force approach is inconvenient can be demonstrated by consideration of the number of parameters that need to be varied in a typical study: 1) satellite and instrument parameters—inclination, altitude, instrument scanning swath width for each satellite in the constellation; 2) grid-box parameters—box dimensions and latitude of box center; and 3) rain-rate field parameters used in the simulation algorithm—characteristic length and time scales, both of which may be geographically and seasonally dependent (cf. Bell 1987; North and Nakamoto 1989; Bell et al. 1990).

The plan of the paper is to first review the North and Nakamoto spectral formalism along with the definitions of terms. A simple example of a space–time spectrum is introduced that is reasonably accurate and extremely useful in obtaining approximate analytical results. Before introducing the general case of multiple satellites with partial coverage visits, we present two simple special cases: two identical satellites with flush visits and a single satellite with partial coverage visits. Finally, the general problem is introduced, treating multiple satellites with arbitrary attributes, including the possibility that each makes only a partial coverage visit on each overpass. The partial coverages are still treated in this paper in an approximate way, weighting each visit proportional to the fraction of the grid box's area intersected by the swath on each particular visit. Comparisons of simulation studies have shown that this is a good approximation to the much more laborious task of integrating the rain rate over each intersection by overpass (Shin and North 1988; Bell et al. 1990). As a further simplification to the exact calculation of each visit fraction as used by Shin and North (1988), we take the visit fraction to be a random number drawn from a probability distribution characterized by the satellite sensor attributes (orbit, swath, etc.). In the present paper, we take these random numbers to be independent from one visit to another. In a future paper, we intend to study the parametric form of the distribution of visit fractions as a function of satellite parameters. Simple and useful formulas are presented at each stage, with tedious derivations abbreviated when possible.

An important example frequently alluded to in the paper is the combination of data from the TRMM satellite with that from a typical operational satellite such as DMSP (Defense Meteorological Satellite Program). We will see that it is very beneficial to combine data from the two satellites and that appreciable reduction in error variance occurs if the two datasets are optimally weighted. It does appear that the data can be collected and converted to rain products separately and combined later to form monthly averages.

2. Mean-square error

The aim is to estimate the space–time average of a field such as rain rate $\psi(\mathbf{r}, t)$. The true space–time average is denoted as

$$\Psi = \frac{1}{L^2 T}\int_B d^2\mathbf{r}\,dt\,\psi(\mathbf{r}, t), \qquad (1)$$

where $\mathbf{r}$ is a point in the plane tangent to the surface of the earth, t is the time, $B = D \times [0, T]$ is the space–time box in the above integral, $D = (-L/2, L/2) \times (-L/2, L/2)$ is the averaging grid box, and T is the averaging period. Typically, $L \sim 500$ km and $T \sim 1$ month. A diagram is shown in Fig. 1.

Now consider an estimator of Ψ given by discrete

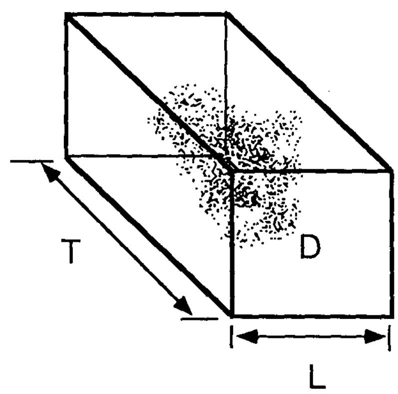

FIG. 1. Schematic illustration of the space–time volume used in averaging the rain rate.

visits consisting of area averaging at times t_n, $n = 1, \cdots, N$ through the period T:

$$\Psi_S = \frac{1}{L^2 N} \sum_{n=1}^{N} \int_D d^2\mathbf{r} \frac{\chi(\mathbf{r}, t_n)}{\mu} \psi(\mathbf{r}, t_n), \quad (2)$$

where $\chi(\mathbf{r}, t_n)$ is a function that is unity in the intersection of the satellite swath and the grid box and zero outside the intersection (see Fig. 2) and $\mu = \sum g_n/N$ is introduced to make the estimate unbiased. The *visit fraction* g_n is given by

$$g_n = \frac{1}{L^2} \int_D d^2\mathbf{r}\, \chi(\mathbf{r}, t_n). \quad (3)$$

A measure of the performance of a particular design is the mean-square error for a particular volume $D \times [0, T]$. Consider then

$$\epsilon_S^2 \equiv \langle (\Psi - \Psi_S)^2 \rangle, \quad (4)$$

where $\langle \ \rangle$ denotes ensemble average. The mean-square error is the average error squared after performing the experiment many times with the same space–time process.

By inserting the Fourier transform of $\psi(\mathbf{r}, t)$ into the error formula,

$$\epsilon_S^2 = \int df \int d^2\nu |H(\nu, f)|^2 S(\nu, f) \quad (5)$$

is obtained (see NN for more details), where f is frequency (cycles per second, c s^{-1}), $\nu = (\nu_1, \nu_2)$ is wavenumber (cycles per meter, c m^{-1}), $H(\nu, f)$ is a complex valued function dependent only on the *design*, and $S(\nu, f)$ is the space–time spectral density of the rain-rate field given by the triple Fourier transform of the covariance lagged in space and time. Here

$$S(\nu, f) = \int d\tau \int d^2\mathbf{r}' \sigma^2 \rho(\mathbf{r}', \tau) \exp[-2\pi i(\nu \cdot \mathbf{r}' + f\tau)] \quad (6)$$

and

$$\sigma^2 \rho(\mathbf{r}', \tau) = \langle \psi(\mathbf{r}, t) \psi(\mathbf{r} - \mathbf{r}', t - \tau) \rangle, \quad (7)$$

such that $\rho(0, 0) = 1$. It is here that the assumptions of stationarity of the time series and homogeneity of the spatial process have been invoked. In principle, these assumptions can be relaxed by use of empirical orthogonal functions or, equivalently, Karhunin–Loéve functions (e.g., North et al. 1982), but the heterogeneous cases are not considered here.

The formula for ϵ_S^2 is noteworthy for two reasons: 1) it separates the factors concerning the rain field (spectral density) from those of the design (the filter function H) and 2) it tells us that the mean-square error depends only on second-moment properties of the rain field. The spectral density may be very difficult to estimate for such a strangely behaved field as rain rate, but nevertheless, the mean-square error as a performance indicator does not depend on any higher moments. This means that issues of non-Gaussian behavior of the rain-rate field do not enter (providing, of course, that all relevant moments exist). Of course, the mean-square error may not be the appropriate indicator of error depending upon how non-Gaussian the errors are distributed and the application.

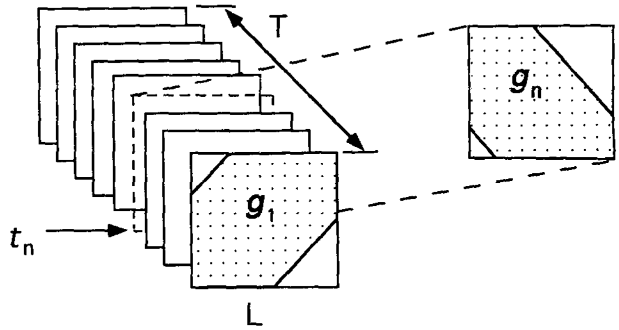

FIG. 2. Schematic illustrating the passages of a satellite swath over slices of the space–time volume at equal intervals. The g_n is the fraction of the grid-box area intersected by the swath at the nth overpass.

A spectral form for rain rates. In NN, North and Nakamoto assumed that the stochastic rainfall process $\psi(\mathbf{r}, t)$ is governed by the following differential equation:

$$\tau_0 \frac{\partial \psi(\mathbf{r}, t)}{\partial t} - \lambda_0^2 \nabla^2 \psi(\mathbf{r}, t) + \psi(\mathbf{r}, t) = F(\mathbf{r}, t), \quad (8)$$

where τ_0 is an inherent time scale of the rainfall process and λ_0 is an inherent length scale. These two scales are sufficient to describe the rain field in this model. The forcing function $F(\mathbf{r}, t)$ is a zero-mean noise process in space and time but is cut off at some high wavenumber ν_c (the spectral density of F vanishes beyond ν_c) to prevent divergences. The last considerations ensure that $\langle \psi(\mathbf{r}, t) \rangle \equiv 0$, which means that the long-term mean has been removed. Our model rain-rate field has the unphysical property that even for large mean values it can become negative. This is not considered to be a significant drawback in the present series of applications. The rain-rate field is then essentially a first-order continuous autoregressive process in time and a special isotropic form of a second-order autoregressive process in space. One interpretation is that rain rates are sporadically produced and destroyed by the source term; they are damped away by the linear term and diffused spatially by the ∇^2 term.

The main reason for using this rain-field model is that it is so easy to analyze spectrally. It is a pleasant surprise that it is reasonably accurate in describing GATE [GARP (Global Atmospheric Research Program) Atlantic Tropical Experiment] data, which were taken over the tropical Atlantic in the summer of 1974

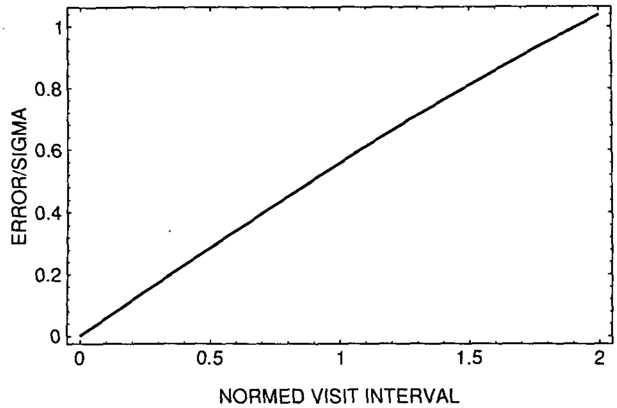

FIG. 3. Graph of the normalized error (ϵ) divided by the standard deviation for grid-box-month averages (σ_{AT}) versus the gap between visits Δt, normalized by twice the intrinsic time scale of the rain field, τ_0. For a typical low-orbiting satellite, such as TRMM or DMSP, flying over tropical oceanic rain, the abscissa is 0.5 ($\Delta t \approx 12$ h, $\tau_0 = 12$ h).

(Hudlow and Patterson 1977; for the comparison see Nakamoto et al. 1990). It was shown by NN that

$$S(\nu, f) = \frac{\alpha}{4\pi^2\tau_0^2 f^2 + (1 + 4\pi^2\lambda_0^2\nu^2)^2}, \qquad (9)$$

where α takes the value such that the integral of $S(\nu, f)$ over the cylinder $\{\nu = |\nu| < \nu_c\} \times [-\infty, \infty]$ is equal to σ^2, the point variance.

North and Nakamoto (1989) studied the case $D = [0, L] \times [0, L]$ for a single satellite returning at regular intervals Δt and with $g_n \equiv 1$ (flush visits). Let σ_A^2 be the variance of $\int_D d^2\mathbf{r}\psi(\mathbf{r}, t)$. The quantity σ_A^2 is a convenient parameter to use instead of σ^2, the point variance, since the latter is often very large and hard to measure. When L is sufficiently large (much larger than λ_0), the relation between the two quantities in the above rain-field model is

$$\frac{\sigma^2\alpha}{L^2} = 2\tau_0\sigma_A^2. \qquad (10)$$

Another convenient variance to refer to is the variance of space–time volume averages, σ_{AT}^2. This is the variance of grid-box-month averages of the exact rain rate. It is also easily expressed in terms of the space–time point variance, σ^2, and the variance of instantaneous box averages, σ_A^2:

$$\sigma_{AT}^2 = \sigma_A^2\left(\frac{2\tau_0}{T}\right). \qquad (11)$$

Then North and Nakamoto's method leads to

$$\epsilon^2 = \frac{\sigma_A^2 2\tau_0}{N\Delta t}\left[\left(\frac{\Delta t}{2\tau_0}\right)\coth\left(\frac{\Delta t}{2\tau_0}\right) - 1\right] \qquad (12)$$

$$= \sigma_{AT}^2\left[\left(\frac{\Delta t}{2\tau_0}\right)\coth\left(\frac{\Delta t}{2\tau_0}\right) - 1\right]. \qquad (13)$$

Here $\Delta t = T/N$ is the time interval between two successive flights over the area D, and N is the total number of flights of the satellite over D in $[0, T]$. Time interval Δt is called the *satellite revisit interval.* Hence, when $\Delta t \rightarrow 0$, then $\epsilon^2 \rightarrow 0$ as expected. The above formula is an approximation that is more accurate when both T and N are larger, say $T > 15 \times 24$ h and $N > 30$.

The ratio of ϵ to σ_{AT} is a measure of the sampling error to the standard deviation of grid-box-month averages as taken in the simple rain-field model. Figure 3 shows the dependence of this index on the visit interval in units of $2\tau_0$. For a low-orbiting satellite, the normalized visit interval is about 0.5; hence, the normalized error is about 0.25. One can also ask about the portion of the measured variance that is accounted for by sampling error. As is shown in Fig. 4, the portion accounted for by the sampling error is only about 7%. If the abscissa is reduced to 0.25 (roughly equivalent to a pair of satellites with optimal phasing, that is, with visits every 6 h), the portion of unexplained variance can be reduced to only about 3%. In both cases it must be borne in mind that the computation of σ_{AT}^2 presented here does not include such low-frequency phenomena as El Niño, the inclusion of which could easily enlarge our value of the natural variability and make our satellite estimates even better when referred to this index.

There are further useful approximations of formula (13) for cases of rare and frequent satellite visits. If the visits are very rare (say $\Delta t/2\tau_0 > 3.5$), notice that $\coth(3) = 1.0049$ and $\coth(x) \rightarrow 1$ as $x \rightarrow \infty$. Therefore, we have

$$\epsilon^2 = \frac{\sigma_A^2}{N}\left(1 - \frac{2\tau_0}{\Delta t}\right). \qquad (14)$$

In this case, we have the estimate of a mean by a fixed number of independent drawings N. The samples drawn are not independent and hence the effective number of independent drawings is reduced by the factor in brackets. Otherwise the formula looks like the

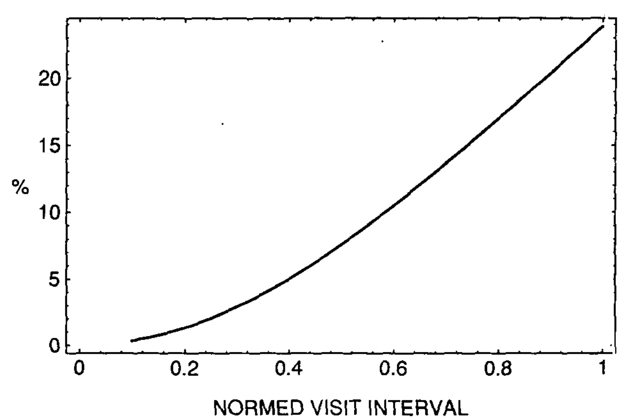

FIG. 4. Percentage of variance of a time series of month averages due to sampling error for a single satellite making flush visits. The calculation is based upon results of Fig. 3.

usual standard-error formula familiar from elementary statistics.

If the visits are very frequent (say $\Delta t/2\tau_0 < 2$), using the Taylor expansion (not the Laurent expansion usually listed in mathematical handbooks) of $\coth(\Delta t/2\tau_0)$, we have

$$\epsilon^2 = \frac{1}{6}\frac{\sigma_A^2}{N}\frac{\Delta t}{\tau_0} = \frac{1}{12}\left(\frac{\Delta t}{\tau_0}\right)^2 \sigma_{AT}^2. \quad (15)$$

The conditions of the last case often hold in practical cases of low-altitude earth-orbiting satellites flying over tropical oceanic rain. For derivation details of the above formulas (9)–(12), the reader is referred to NN. The exact summation form of (12) (which replaces π^2 by 6) was not given in NN, in which only the first term of the expression was retained.

These formulas can also be used for two identical satellites whose visits are spaced at equal intervals (phase difference is one-half period) by considering N as the total number of visits by both satellites and Δt as the interval between successive visits. The formula for I satellites returning at equal intervals for a given T can also be used. Consider, for example, I identical sun-synchronous satellites, equally phased, and for a grid box at the equator with tropical oceanic rain characteristics [see NN: $\tau_0 \approx 12$ h and $\Delta t = (12\text{ h})/I$, $T = 1$ month]; finally, the variance of area-average rain rate σ_A is about equal to the mean rain rate μ_A (Shin and North 1988). The sampling error as percentage of the mean is $5.3\%/\sqrt{I}$. It must be kept in mind that the assumption of flush visits has been used and relaxation of this assumption will enlarge the error considerably as will be shown in subsequent sections.

3. Two simple special cases

This section is primarily a preparation for the next section. Here two simple special cases are considered before the complicated general situation described in section 4 is reached.

a. Two identical satellites with flush visits

Consider the measurement of the same rain field in the space–time box $D = [-L/2, L/2] \times [-L/2, L/2] \times [-T/2, T/2]$ in section 2. Instead of using one satellite, two satellites that have the same orbital attributes are used. Both of them make flush visits to the square $[-L/2, L/2] \times [-L/2, L/2]$. Hence, from the sampling viewpoint, the two satellites are identical and fly on the same orbits. The distinction between the satellites visits to a given grid box is the phase separation between their periodic overpasses. Because of nonzero τ_0, the information in the rain field at a time is correlated to that at another close time. Visits that are bunched too close together will tend to be redundant, while too sparse visits will miss important variations. On this basis we conclude that the least sampling error is realized when the phase lag between the satellites is adjusted for equally separated visits. This section will show how important this effect is.

Next the error estimation formula is derived for two identical satellites with an arbitrary phase separation. Let Ψ_1 and Ψ_2 be the unbiased estimators of the space–time volume average based upon the discrete samplings of the first satellite and the second satellite, respectively. Then

$$\Psi_i = \frac{1}{L^2 T}\int_D d^2\mathbf{r}\int_T dt K_i(t)\psi(\mathbf{r}, t), \quad (16)$$

where

$$K_i(t) = \Delta t \sum_{n=0}^{N-1} \delta\left[-\frac{T}{2} + t - (n + \theta_i)\Delta t\right] \quad (17)$$

and $i = 1, 2$. Here $\theta_i \Delta t$, $i = 1, 2$ are the phase lags. Let $\theta_1 = 0$, $\theta_2 = \theta$ and $0 < \theta < 1$. The unbiased linear estimator from the combined data of the two satellites is

$$\Psi_s = \alpha_1\Psi_1 + \alpha_2\Psi_2, \quad 0 \leqslant \alpha_1, \quad \alpha_2 \leqslant 1,$$
$$\alpha_1 + \alpha_2 = 1. \quad (18)$$

The mean-square error is then

$$\begin{aligned}\epsilon^2 &= \langle(\Psi - \Psi_s)^2\rangle \\ &= \langle[(\alpha_1 + \alpha_2)\Psi - \alpha_1\Psi_1 - \alpha_2\Psi_2]^2\rangle \\ &= \langle[\alpha_1(\Psi - \Psi_1) + \alpha_2(\Psi - \Psi_2)]^2\rangle \\ &= \alpha_1^2\epsilon_1^2 + \alpha_2^2\epsilon_2^2 + 2\alpha_1\alpha_2\epsilon_{12}. \quad (19)\end{aligned}$$

In the above, ϵ_i^2, $i = 1, 2$ are the mean-square errors of a single-satellite sampling as studied in section 2. Term ϵ_{12} is the interference term that depends on the phase lag. One can derive [assuming (15)] that

$$\begin{aligned}\epsilon_{12} &= \langle(\Psi - \Psi_1)(\Psi - \Psi_2)\rangle \\ &= \sigma_{AT}^2 \sum_{n\neq 0} \frac{\cos(2n\pi\theta)}{1 + 4\pi^2\tau_0^2(n/\Delta t)^2}. \quad (20)\end{aligned}$$

Of course, it is easy to show that for identical satellites, $\alpha_{1,2} = 0.5$ independent of θ.

Figure 5 shows a graph of the error (normalized by the standard deviation of grid-box-month averages, σ_{AT}) as a function of the phase separation θ for typical values of the parameters for tropical rain ($\Delta t = 12$ h, $\tau_0 = 12$ h, $T = 30$ days, $L \gg \lambda_0 \approx 40$ km). As expected, the error is least when the satellites have equal phasing ($\theta = 0.5$).

A satellite such as TRMM only makes equal interval visits to grid boxes along the equator. At other latitudes, the visits are such that ascending overpasses are at periods of about 24 h and descending overpasses are at intervals of about 24 h, but the two have a phase relationship that ranges from equal interval at the equator to redundancy at the turning point (35°). Actually,

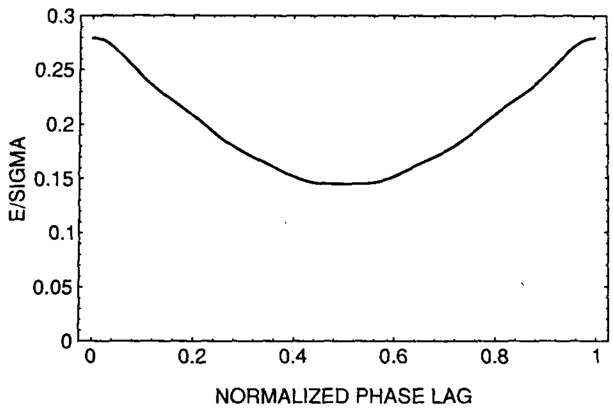

FIG. 5. Normalized error (ϵ/σ_{AT}) versus normalized phase lag for two identical satellites making flush visits. The least error occurs when the satellites are equally spaced in time ($\theta = 0.5$). The largest error occurs when the phase is zero or unity when the two satellites are redundant.

TRMM may make several crossings of a grid box near its turning point within a few hours of each other. Our approximations no doubt break down near these peculiar zones. The previous formula can be used, pretending that TRMM is two satellites each with a period of 24 h. Figure 5 shows the error (again normalized by the volume standard deviation σ_{AT}) that might be expected for tropical oceanic rain for two satellites with periods of 24 h making flush visits but with a phase difference θ; $0 \leq \theta \leq 1$. For TRMM, the equator corresponds to $\theta = 0.5$ and the turning point corresponds to $\theta = 0, 1$. Of course, these calculations have been made for flush visits. The strong latitude dependence of the fractional coverage will tend to reduce the latitude dependence shown in Fig. 5 (Shin and North 1989; Bell et al. 1990).

b. Single satellite with partial-coverage visits

Shin and North (1988) considered the grid-box-averaged rain-rate field as a univariate time series in their calculations of sampling error. Partial coverage visits to the grid box were taken into account by weighting a visit by the fraction g_n of the box intersected by the swath on the nth visit (Fig. 2). In their study, they calculated the sequence g_n ($n = 0, 1, \cdots, N-1$) from exact orbit calculations. Unfortunately, the computation of exact visit fractions is tedious and must be repeated for each set of orbit and grid-box parameters. Here we continue to make the approximation of weighting the visit by g_n but instead of computing the sequence explicitly, the g_n is modeled as random variables with first and second moments computed from the exact orbital calculations. This procedure simplifies the formulation and brings out the important dependences. Unfortunately, an additional approximation is necessarily introduced that the visit fractions are statistically independent from one visit to another. A later study will address the issue of just how good an approximation the weighting by fraction and their corresponding statistical independence really is in these computations. In its defense we point to the favorable comparison of the Shin and North calculations with those of Bell et al. (1990), which were performed without use of the visit fraction approximation.

The linear extimator of the space–time box average of the rain rate is given by (2). Since the evaluation of (2) is extremely complicated as we go through a sequence of visits, we adopt here the simple rule that it may be approximated by

$$\Psi_S \approx \frac{1}{L^2 N} \sum_{n=1}^{N} \frac{g_n}{\mu} \int_D \psi(\mathbf{r}, t_n) d^2\mathbf{r}, \qquad (21)$$

where g_n, the visit fraction, is defined by (3) and μ is its average value. This is of course the approximation used by Shin and North (1989), but here we go a step further and take the g_n to be a sequence of independent identically distributed random variables whose moments depend only on orbit–swath–footprint geometry. This will be very convenient since in the final formulas the only properties of the g_n will be their mean and variance, which depend on the averaging box, the satellite orbit, and the swath width. Several rather drastic assumptions have been made here and we need to be careful about their implications. First, the individual intersections are smaller than the entire box, and area rain-rate averages for such intersections will have larger variances than for the entire box. Second, the intersections will have shorter autocorrelation times than for the entire box. Each of these will lead to estimates of the sampling error that are smaller than the actual sampling error. While an exact analysis is complicated

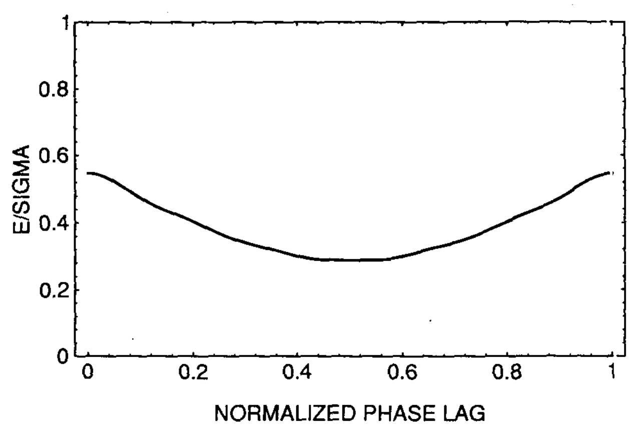

FIG. 6. Same as Fig. 4 except that the revisit period is twice as large. This graph applies when one considers a single satellite but the ascending and descending nodes (each with revisit period 24 h) separately. These crossings are equally spaced when the grid box is located at the equator but shift their phase toward zero as the turning point of the orbit (35°N for TRMM) is approached. The graph is for flush visits. Inclusion of fractional visit effects will tend to compensate for the latitude dependence.

and breaks the spirit of our present endeavor, it is important to carry out and will be done at a later time. It is comforting to note that the calculations of Shin and North (1989) agree well with the simulations of Bell et al. (1990), which computed these forms more exactly:

$$\Psi_S = \frac{1}{L^2T}\int_D d^2\mathbf{r}\int_T dtK(t; g_0, \cdots, g_{N-1})\psi(\mathbf{r}, t). \tag{22}$$

The function $K(t)$ is defined by

$$K(t; g_0, \cdots, g_{N-1}) = \Delta t\sum_{n=0}^{N-1}\frac{g_n}{\mu}\delta\left(-\frac{T}{2}+t-n\Delta t\right). \tag{23}$$

In what follows, g_n is taken to be a random variable. Let μ_g denote its expectation value, which is assumed to be independent of n.

Following the steps in NN, one can derive

$$\epsilon^2 = \langle(\Psi - \Psi_s)^2\rangle = \int d^2\nu\int dfS(\nu, f)G^2(\nu_1L)G^2(\nu_2L)G^2(fT) \times B(f; g_0, \cdots, g_{N-1}), \tag{24}$$

where

$$B(f; g_0, \cdots, g_{N-1}) = 1 - \frac{2\Delta t}{TG(fT)}\sum_{n=0}^{N-1}\frac{g_n}{\mu}\cos\{\pi f[(2n-1)\Delta t - T]\} + \left(\frac{\Delta t}{T}\right)^2\frac{1}{G^2(fT)}\sum_{m,n=0}^{N-1}\frac{g_mg_n}{\mu^2}\exp[2\pi if(m-n)\Delta t]. \tag{25}$$

Here the function $G(x)$ is defined by

$$G(x) = \frac{\sin(\pi x)}{\pi x}. \tag{26}$$

We assume $g_0, g_1, \cdots, g_{N-1}$ to be N independent and identically distributed random variables. Let $P_g(g_n)$ be the PDF (probability distribution function) of g_n, $n = 0, 1, \cdots, N-1$. Then the multivariate PDF for $\{g_0, \cdots, g_{N-1}\}$ is

$$\mathcal{P}(\{g_0, \cdots, g_{N-1}\}) = \prod_{n=0}^{N-1}P_g(g_n). \tag{27}$$

Let

$$\mu_g = \int dg_ng_nP_g(g_n) \quad \text{(first moment)}, \tag{28}$$

$$\gamma_g^2 = \int dg_ng_n^2P_g(g_n) \quad \text{(second moment)}, \tag{29}$$

$$\sigma_g^2 = \gamma_g^2 - \mu_g^2 \quad \text{(variance of } g_n\text{)}. \tag{30}$$

Following NN and carrying out some tedious algebraic manipulations, one can obtain

$$E^2 = \prod_{n=0}^{N-1}\int dg_n\epsilon^2(g_0, \cdots, g_{N-1})\mathcal{P}(g_0, \cdots, g_{N-1}) = \frac{\sigma_A^2}{N}\left(-\frac{2\tau_0}{\Delta t}+\frac{\gamma^2}{\mu^2}+\frac{2}{e^{\Delta t/\tau_0}-1}\right) - \frac{2\sigma_A^2}{N^2}\left[1+\frac{e^{(\Delta t/\tau_0)(1-N)}-1}{1-e^{\Delta t/\tau_0}}\right]. \tag{31}$$

After neglecting the $O(1/N^2)$ term, the above formula can be rearranged as

$$E^2 = \sigma_{AT}^2\left[\left(\frac{\Delta t}{2\tau_0}\right)\coth\left(\frac{\Delta t}{2\tau_0}\right) - 1 + \frac{\sigma_g^2}{\mu_g^2}\frac{\Delta t}{2\tau_0}\right]. \tag{32}$$

This formula clearly indicates that the sampling-error variance due to the fractional visit sampling may be decomposed into two parts. One part is purely due to the time gap between flush visits. The other part is an additional sampling-error variance due to the fact that the visits cover variable fractions of the grid box.

We note that a deficiency of our approximate formula is apparent when σ_g is small, namely, the portion of the error variance due to fractional visits approaches 0 even though μ_g is less than 1. This means that visiting exactly one-half the box on each visit is as good as visiting the whole box—clearly a shortcoming of our simple weighting approximation. This error comes about because we used the large-grid-box approximation earlier and even applied it to the fractional visit. Hence, implicit in the formulation is the assumption that the fractional visit areas are much larger than λ_0^2, where λ_0 is the inherent length scale in the rain-rate field.

Now we return to the case of I identical sun-synchronous satellites as at the end of section 2, only this time the fractional-visit term is allowed to enter. Using the same data as in that case along with $\sigma_g^2 \approx 0.50\,\mu_g^2$, we find

$$\text{percentage error} \approx 100\left[\frac{(0.053)^2 + 0.0042}{I}\right]^{1/2} \tag{33}$$

$$= \frac{10.5\%}{\sqrt{I}}. \tag{34}$$

Two interesting numerical examples follow.

1) The TRMM satellite with altitude 350 km, inclination 35°, nominal swath width 600 km: for grid boxes along the equator of dimension 500 km, $\mu_g = 0.439$ and $\sigma_g^2 = 0.112$ are found from orbit calculations. We are led to $\epsilon_{\mathrm{TRMM}}/\sigma_A \sim 0.112$, which corresponds to an error of about 11.2%, since $\sigma_A \approx \mu_A$ for a 500-km box over the tropical oceans (Shin and North 1988).

2) The Special Sensor Microwave/Image SSM/I on board the DMSP (sun-synchronous) satellite with altitude 800 km, inclination 98.6°, swath width 1400 km: for the same grid boxes as above orbit computations show, $\mu_g = 0.671$ and $\sigma_g^2 = 0.132$. In this case we find $\epsilon_{\mathrm{DMSP}}/\sigma_A \sim 0.0876$, which corresponds to an error of about 8.76%. The higher altitude with its wider swath width for the sun-synchronous orbiter leads to a somewhat smaller sampling error than for TRMM.

Later the question of how the data can be combined with optimal weighting is considered.

4. Combining data from I satellites

In this section, we approach the complicated problem of I satellites each with its own orbit characteristics. Since the computations are complicated, they are presented in the Appendix. Here we present the notation and define the problem. The treatment will make use of the approximation scheme introduced earlier for fractional visits.

a. Notations and the general formulation

The area we are interested in is $D = [0, L] \times [0, L]$. The number of satellites flying over D to sample the rain field $\psi(\mathbf{r}, t)$, $\mathbf{r} \in D$, and $t \in [0, T]$ is I. Each time a satellite visits D, it takes an area average over the intersection of the swath width and the grid box D. The percentage of the coverage of the area D at the nth visit of the satellite (i) is denoted by $g_n^{(i)}$. The following notations are adopted:

T:	total sampling time;
$D = [0, L] \times [0, L]$:	sampling area;
N_i:	number of visits of satellite (i) over D during $[0, T]$,
$\Delta t^{(i)} = T/N_i$:	sampling interval (also called the revisit period) of satellite (i);
θ_i:	the phase lag of satellite (i) measured as a fraction of $\Delta t^{(i)}$;
$t_n^{(i)} = (n - 1/2 - \theta_i)\Delta t^{(i)}$, $-1/2 < \theta_i < 1/2$:	sampling time for satellite (i) at its nth visit;
$\Psi_S^{(i)}$:	sample estimate of the volume average for satellite (i) in $[0, T]$;
Ψ_S:	combined sampling average of all the I satellites in $[0, T]$;
$g_n^{(i)}$:	random percentage of the area D for satellite (i) to cover at its nth visit $[0 < g_n^{(i)} < 1]$; and
Ψ:	space–time average of the random field $\psi(\mathbf{r}, t)$ in the box $D \times [0, T]$.

In the above list, $n = 1, 2, 3, \cdots, N_i$; $i = 1, 2, 3, \cdots, I$. In our computations a choice was made about how to compute N_i for satellite i. One possibility is to take an overpass to occur at approximately every 100 min, in which case N_i is very large and most of the g_n are zero. Another choice is to make T/N_i close to 12 h, and when there are multiple visits in one 12-h period, the individual g_n are added and they are lumped into a single visit. This latter leads to a sequence of g_n that is mostly nonzero and therefore easier to deal with. Hence, the latter method was chosen in our numerical presentations.

As shown in the Appendix, the result for the mean-square error (taken over the ensemble of realizations of the field and over the ensemble of fractional visits) for the I nonidentical satellite combination is given by the sum

$$E^2 = \sum_{i=1}^{I} \alpha_i^2 E_i^2 + \sum_{i \neq j} 2\alpha_i \alpha_j E_{ij}, \tag{35}$$

where E_i^2 is the mean-square error for satellite i as though it were acting alone, $\alpha \equiv (\alpha_1, \cdots, \alpha_I)$ is a weight vector such that $\sum_1^I \alpha_i = 1$, and E_{ij} are cross or interference terms. The cross terms are a measure of the statistical interdependence of the different satellite measurement subsystems. As formulated, the unit sum of the α_i keep the estimate from being biased; however, consistent with this constraint, a choice for them can be made so as to minimize the mean-square error, as will be discussed in the next section. The general formulas for the mean-square error for I nonidentical satellites are presented in the Appendix under the assumptions mentioned above.

b. Optimal weights

The problem of optimal weighting can be solved by minimizing the last formula for $D^2(\alpha)$ with respect to each α_i subject to the constraint, $\sum \alpha_i = 1$. First, consider the problem for two satellites and suppose the values of E_1^2, E_2^2, and E_{12} have been computed. The total error variance is given by

$$E^2(\alpha_1, \alpha_2) = \alpha_1^2 E_1^2 + 2\alpha_1\alpha_2 E_{12} + \alpha_2^2 E_2^2 \tag{36}$$

subject to

$$\alpha_1 + \alpha_2 = 1, \quad \alpha_{1,2} > 0. \tag{37}$$

The last equation may be solved for α_2 and substituted in the equation just previous to find a function only of α_1, which may be minimized by setting the derivative of $E^2(\alpha_1)$ to zero and solving the result for the α_1, which achieves the minimum in E^2:

$$\alpha_1^{(\mathrm{best})} = \frac{E_2^2 - E_{12}}{E_1^2 - 2E_{12} + E_2^2}, \tag{38}$$

which leads to the other coefficient

$$\alpha_2^{(\text{best})} = \frac{E_1^2 - E_{12}}{E_1^2 - 2E_{12} + E_2^2}. \qquad (39)$$

Note that when $E_{12} = 0$, we get the familiar result that independent measurements lead to weights that are inversely proportional to the error variances $\alpha_i \propto 1/E_i^2$ that would have been obtained had the measurements been done singly.

Now turn to the general case of I satellites with individual error variances E_i^2 and cross (interference) terms E_{ij} as indicated in (35). Treat the constraints

$$\sum_{i=1}^{I} \alpha_i = 1 \qquad (40)$$

by the method of Lagrange multipliers (Arfken 1985). That is, minimize

$$\sum \alpha_i E_i^2 + 2 \sum_{i \neq j} \alpha_i \alpha_j - \lambda\left(\sum \alpha_i - 1\right) \qquad (41)$$

with respect to the α_i and the Lagrange multiplier λ. The first set of derivatives set to zero leads to

$$\sum_j M_{ij}\alpha_j = \lambda u_i, \qquad (42)$$

where $u_i = 1$, $i = 1, \cdots, I$. The above set of I equations is to be solved for α with the parameter λ chosen so that the constraint is ultimately satisfied. This is easily and uniquely accomplished by finding the inverse of the symmetric matrix M_{ij}.

There are other potentially useful optimal weighting schemes, which will be explored in a later paper. For example, one might consider a weight $\alpha_{i,n}$ for the ith satellite on its nth pass.

5. Weights for the TRMM–DMSP combination

As a very practical example we now consider the case of data from the TRMM satellite combined with that of a sun-synchronous satellite such as the microwave radiometer on the DMSP satellite. In a previous section, TRMM and DMSP were shown to separately have sampling errors of 11.2% and 8.76%, respectively. A major difference between the satellite orbits is their period of revisit to a box. For DMSP, this is 12.0 h since it is sun-synchronous. For TRMM at the equator, this period is about 11.75 h, which allows cycling through the local clock in 24 days. Hence, in the several-satellite formalism given above, $(\Delta t)_{\text{DMSP}} = 12.00$ and $(\Delta t)_{\text{TRMM}} = 11.75$, which leads to $T = 720$ h, $N_D = 60$, $N_T = 61.28$, and $r = 0.0213$. Figure 7 shows the (normalized) interference term E_{12}/σ_{AT}^2 as a function of the TRMM period $(\Delta t)_{\text{TRMM}}$. It is found that in the range of interest, $|E_{12}/\sigma_{AT}^2| < 0.01$, whereas $E_{\text{TRMM}}^2/\sigma_{AT}^2 \approx 0.372$ and $NE_{\text{DMSP}}^2/\sigma_{AT}^2 \approx 0.228$. Figure 8 shows the combined error as a function of α_{TRMM}. The graph clearly shows the importance of including the DMSP data (error reduced by about one-third). On the other hand, the error is insensitive to the exact value of α_{TRMM} chosen. Figure 8 actually shows two curves that bracket the effect of the interference term. The upper curve is for the normalized interference term equal to 0.02 and the lower curve is for the value −0.02, both of which are larger in magnitude than the size of the interference term as shown in Fig. 7. Again, the effect on our choice of optimal α_{TRMM} is slight. In fact, the error is hardly affected by the interference term. This near-statistical orthogonality (null interference) is reminiscent of the result found when data from point gauges are combined with satellite overpass data (North et al. 1991).

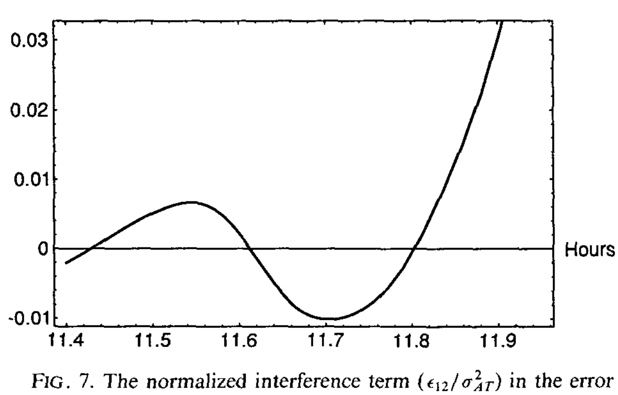

FIG. 7. The normalized interference term $(\epsilon_{12}/\sigma_{AT}^2)$ in the error variance for the combined TRMM and DMSP datasets. The nominal TRMM period is 11.50 h. Note that the magnitude of this term is below 0.02 in the range of interest.

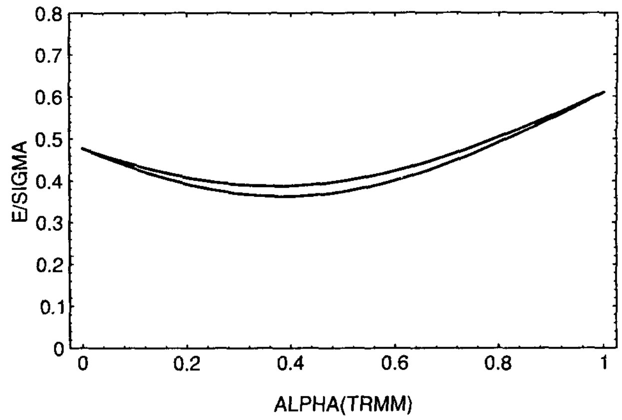

FIG. 8. Normalized error $(\epsilon^2/\sigma_{AT}^2)$ versus weight given to the TRMM data in combination with the DMSP data (whose weight is $1 - \alpha_{\text{TRMM}}$). Shown are graphs for the interference term above and below expected limits (cf. Fig. 7). It is clear that the choice of optimum α_{TRMM} is hardly affected by the value of ϵ_{12}. It is also clear that adding the DMSP data is beneficial, but the exact choice of weighting (±0.2) is not critical to the reduction of error.

6. Concluding remarks

In this paper, we have presented a framework for the approximate assessment of the random sampling errors incurred in estimating space–time-averaged rain rates based upon aggregation of data from several low earth-orbiting satellites. We have crudely included the effects of partial coverage visits in a way that leads to relatively simple formulas that can be used in preliminary decision making. In our approximate formulas the partial visit terms tend to roughly double the error in cases of interest. As a pedagogical introduction, we illustrated our procedure for a system of two satellites making flush visits and for a single satellite making fractional visits to a grid box. Several useful I-satellite results can be found when the symmetry of the orbits is high.

For a system of I satellites with different individual orbital attributes, we find that the total error variance can be written as a sum of the individual error variances plus interference terms contributed from all combinations of pairs of subsystems. After presenting the general formulas, we presented explicit numerical results for the TRMM satellite combined with a typical sun-synchronous partner. The results of these preliminary calculations show that it is definitely worthwhile to include data from the second orbiter since it roughly halves the sampling error. It is also worthwhile to optimally weight the two datasets approximately inversely to their individual error variances. The error of the combined dataset is insensitive to including the small interference term between these two systems.

We note that this is hardly "the end of the story" in combining satellite data for an optimally estimated time series of space–time-smoothed rain rates. There are a number of biases connected with individual sensors that have yet to be reckoned with. These are questions of the diurnal cycle and how it can be extracted from the two-satellite-combined data. Complicating this last are the so-called 40–50 waves prevalent in tropical wave data. These will alias the diurnal cycle and vice versa. In future papers, we intend to answer some of these questions with improved ideas and more satisfactory models.

Acknowledgments. We are grateful for continued support from the NASA Earth Observation Program (Grant NAG5-869). We also thank T. L. Bell for many helpful suggestions.

APPENDIX

The Case of *I* Satellites

This appendix is the continuation of the discussion in section 4a, where notation was established. Let $P_i[\{g_n^{(i)}\}_{n=1}]$ be the pdf (probability distribution function) of the ith random sequence $\{g_n^{(i)}\}_{n=1}^{N_i}$. The first and the second moments of the sequence are defined as

$$\mu_n^{(i)} = \int_0^1 dg_n^{(i)} P_i(\quad) g_n^{(i)} \tag{A1}$$

and

$$(\gamma_n^{(i)})^2 = \int_0^1 dg_n^{(i)} \{P_i(\quad)[g_n^{(i)}]\}^2 \tag{A2}$$

for $n = 1, 2, \cdots, N_i$ and $i = 1, 2, \cdots, I$. The sampling average of the ith satellite is

$$\Psi_S^{(i)} = \frac{1}{TL^2} \int_{D\times[0,T]} d^2\mathbf{r}\, dt \psi(x, t) K_i(t), \tag{A3}$$

where

$$K_i(t) = \frac{T}{N_i} \sum_{n=1}^{N_i} \frac{g_n^{(i)}}{\mu^{(i)}} \delta(t - t_n^{(i)}). \tag{A4}$$

Here $\delta(\quad)$ is the Dirac delta function. The combined sampling average of all the I satellites is

$$\Psi_S = \sum_{i=1}^{I} \alpha_i \Psi_S^{(i)}, \quad 0 \leqslant \alpha_i \leqslant 1 \quad \text{and} \quad \sum_{i=1}^{I} \alpha_i = 1. \tag{A5}$$

The true average of the rain field is

$$\Psi = \frac{1}{TL^2} \int_{D\times[0,T]} d^2\mathbf{r}\, dt \psi(\mathbf{r}, t). \tag{A6}$$

We wish to find the ensemble average of $(\Psi - \Psi_S)^2$, which will serve as a measure of the quality of the particular design under our consideration:

$$\epsilon^2 = \langle(\Psi - \Psi_S)^2\rangle. \tag{A7}$$

Then

$$\epsilon^2 = \frac{1}{L^4T^2} \iint d^2\mathbf{r}\, dt \langle \psi(\mathbf{r}, t)\psi(\mathbf{r}', t')\rangle \times [1 - K(t)][1 - K(t')], \tag{A8}$$

where

$$K(t) = \sum_{i=1}^{I} \alpha_i K_i(t). \tag{A9}$$

Note that $\{\rho(\mathbf{r}, t), S(\nu, f)\}$ are a Fourier transformation pair. Hence,

$$\rho(\mathbf{r}, t) = \iint d^2\nu df S(\nu, f) \exp[-2\pi i(\nu \cdot \mathbf{r} + ft)]. \tag{A10}$$

With the above information, the following can be derived:

$$\epsilon^2 = \sum_{i=1}^{I} \alpha_i^2 \epsilon_i^2 + \sum_{i\neq j} 2\alpha_i\alpha_j\epsilon_{ij}. \tag{A11}$$

Here

$$\epsilon_i^2 = \langle[\Psi - \Psi_S^{(i)}]^2\rangle, \quad i = 1, 2, \cdots, I, \tag{A12}$$

and

$$\epsilon_{ij} = \langle [\Psi - \Psi_S^{(i)}][\Psi - \Psi_S^{(j)}] \rangle, \quad i \neq j, \quad i, j = 1, 2, \cdots, I. \quad \text{(A13)}$$

Statistically we are interested in the expectation value of ϵ^2, which by definition is

$$E^2 = \int\int \cdots \int dg_1^{(1)} dg_2^{(1)} \cdots dg_n^{(1)} \cdots dg_1^{(I)} dg_2^{(I)} \cdots dg_n^{(I)} \prod_{i=1}^{I} P_i \epsilon^2. \quad \text{(A14)}$$

This is rewritten into

$$E^2 = \sum_{i=1}^{I} \alpha_i^2 E_i^2 + \sum_{i \neq j} 2\alpha_i \alpha_j E_{ij}, \quad \text{(A15)}$$

where

$$E_i^2 = \int \cdots \int dg_1^{(i)} dg_2^{(i)} \cdots dg_{N_i}^{(i)} P_i \epsilon_i^2 \quad \text{(A16)}$$

and

$$E_{ij} = \int \cdots \int dg_1^{(i)} dg_2^{(i)} \cdots dg_{N_i}^{(i)} dg_1^{(j)} dg_2^{(j)} \cdots dg_{N_j}^{(j)} P_i P_j \epsilon_{ij}. \quad \text{(A17)}$$

a. Evaluation of E_i^2

Following the steps in section 2 for single satellite, one can derive the following:

$$E_i^2 = \sigma^2 \int\int d^2\nu df S(\nu, f) G^2(\nu_1 L) G^2(\nu_2 L) G^2(fT) \times \int dg_1^{(i)} dg_2^{(i)} \cdots dg_{N_i}^{(i)} B_i[\{g_n^{(i)}\}_{n=1}^{N_i}; f] \times P_i[\{g_n^{(i)}\}_{n=1}^{N_i}]. \quad \text{(A18)}$$

Here $\nu = (\nu_1, \nu_2)$, and

$$B_i[\{g_n^{(i)}\}_{n=1}^{N_i}; f] = 1 - \frac{2\Delta t^{(i)}}{TG(fT)} \times \sum_{n=1}^{N_i} \frac{g_n^{(i)}}{\mu^i} \cos\{\pi f[(2n-1)\Delta t^{(i)} - T]\} + \left[\frac{\Delta t^{(i)}}{T}\right]^2 \frac{1}{G^2(fT)} \times \sum_{m,n=1}^{N_i} \frac{g_m^{(i)} g_n^{(i)}}{\mu^{(i)}\mu^{(i)}} \exp[i2\pi f(m-n)\Delta t^{(i)}]. \quad \text{(A19)}$$

The function $G(x)$ is defined in section (b), and $\mu^{(i)}$ is the average fraction of the satellite i. And in this section, it is also assumed that

$$\theta_i = 0, \quad i = 1, 2, \cdots, I. \quad \text{(A20)}$$

Noticing that

$$G(Lx) \rightarrow \frac{1}{L}\delta(x), \quad \text{as} \quad L \rightarrow \infty, \quad \text{(A21)}$$

we can derive that

$$E_i^2 = \frac{\sigma_A^2}{N_i} \left\{ \frac{2\tau_0}{\Delta t^{(i)}} \left[\frac{\Delta t^{(i)}}{2\tau_0} \coth\left(\frac{\Delta t^{(i)}}{2\tau_0}\right) - 1 \right] + \left(\frac{\sigma^{(i)}}{\mu^{(i)}}\right)^2 \right\} \quad \text{(A22)}$$

after dropping a term of $O(1/N_i^2)$. The quantity $[\sigma^{(i)}]^2$ is the variance of the fractions over all the visits and is equal to $[\gamma^{(i)}]^2 - [\mu^{(i)}]^2$. Note that in the limit of very wide swath widths, the fractions will always be flush (100%) and $\mu^{(i)} \rightarrow 1$ and $\sigma^{(i)} \rightarrow 0$, which leads us to the formula already derived:

$$E_i^2(\text{flush}) = \sigma_A^2 \frac{2\tau_0}{T} \left\{ \frac{\Delta t^{(i)}}{2\tau_0} \coth\left[\frac{\Delta t^{(i)}}{2\tau_0}\right] - 1 \right\}. \quad \text{(A23)}$$

This agrees with the formula (32).

b. Evaluation of E_{ij}

Next consider the cross term E_{ij} when $i \neq j$. By (A17) and (A13)–(A15), one can derive the following:

$$E_{ij} = \sigma^2 \int d^2\nu df S(\nu, f) G^2(\nu_1 L) G^2(\nu_2 L) G^2(fT) \int dg_1^{(i)} dg_2^{(i)} \cdots dg_{N_i}^{(i)} dg_1^{(j)} dg_2^{(j)} \cdots dg_{N_j}^{(j)} \times B_{ij}[\{g_n^{(i)}\}_{n=1}^{N_i}, \{g_n^{(j)}\}_{n=1}^{N_j}; f] P_i[\{g_n^{(i)}\}_{n=1}^{N_i}] P_j[\{g_n^{(j)}\}_{n=1}^{N_j}], \quad \text{(A24)}$$

where

$$B_{ij}[\{g_n^{(i)}\}_{n=1}^{N_i}, \{g_n^{(j)}\}_{n=1}^{N_j}; f] = \{1 - \Delta t^{(i)} \exp(-\pi i f T)/TG(fT) \sum_{n=1}^{N_i} \frac{g_n^{(i)}}{\mu^{(i)}} \exp[2\pi i f t_n^{(i)}]\} \times \{1 - \Delta t^{(j)} \exp(\pi i f T)/TG(fT) \sum_{n=1}^{N_i} \frac{g_n^{(j)}}{\mu^{(j)}} \exp[-2\pi i f t_n^{(j)}]\}. \quad \text{(A25)}$$

Since $\mu_n^{(i)} = \mu^{(i)}$ and $\gamma_n^{(j)} = \gamma^{(j)}$ for every n, (24) can be written into

$$E_{ij} = 2\tau_0\sigma_A^2 \int df \frac{G^2(fT)}{(G(f\Delta t^{(i)})G(f\Delta t_j))} \times \frac{(G(f\Delta t^{(i)}) - 1)(G(f\Delta t_j) - 1)}{(1 + 4\pi^2\tau_0^2 f^2)}. \quad (A26)$$

Next examine the integration kernel (also called the filter)

$$\frac{G^2(fT)\{G[f\Delta t^{(i)}] - 1\}\{G[f\Delta t^{(j)}] - 1\}}{G[f\Delta t^{(i)}]G[f\Delta t^{(j)}]}, \quad (A27)$$

when N_i, N_j, and T are sufficiently large. As mentioned before, when it is stated that N_i, N_j, and T are sufficiently large, typically it is required that $N_i, N_j \geqslant 30$, and $T \geqslant 15$ days. Without loss of generality, assume $\Delta t^{(j)} \geqslant \Delta t^{(i)}$. Let $\Delta_{ij} = \Delta t^{(j)}/\Delta t^{(i)}$. So $\Delta_{ij} \geqslant 1$. Consider the case that

$$\Delta_{ij} = k_{ij} + r_{ij}, \quad (A28)$$

k_{ij} positive integer and r_{ij} rational in (0, 1).

One can easily show (see Fig. 3) that

$$\frac{\sin(N\pi\Delta t f)}{N\sin(\pi\Delta t f)} \sim \frac{1}{N\Delta t}\sum(-1)^{(N-1)n}\delta(f - n/\Delta t) \quad \text{when } N \to \infty. \quad (A29)$$

Notice that

$$\frac{G(fT)}{G(f\Delta t^{(i)})} = \frac{\sin[N_i\pi\Delta t^{(i)}f]}{N_i\sin[\pi\Delta t^{(i)}f]} \quad (A30)$$

and that $N_i \geqslant N_j$ since $\Delta_{ij} = \Delta t^{(j)}/\Delta t^{(i)} \geqslant 1$ and $N_i\Delta t^{(i)} = N_j\Delta t^{(j)} \approx T$. Therefore, $G(fT)/G[f\Delta t^{(i)}]$ is a faster delta-convergent sequence than $G(fT)/G[f\Delta t^{(j)}]$. Further, notice that

$$G(0 \pm 0) = 1, \quad \text{and} \quad G(n) \sim 0 \quad \text{and} \quad G(n\Delta_{ij}) \sim 0 \quad \text{when} \quad n > 0. \quad (A31)$$

Finally, the following approximation is given:

$$\frac{G^2(fT)\{G[f\Delta t^{(i)}] - 1\}\{G[f\Delta t^{(j)}] - 1\}}{G[f\Delta t^{(i)}]G[f\Delta t^{(j)}]} \sim \frac{1}{T}\sum_{n\neq 0}\frac{(-1)^{(k+1)n}\sin(N_j n r_{ij}\pi)}{N_j\sin(n r_{ij}\pi)} \times \delta\left[f - \frac{n}{\Delta t^{(i)}}\right]. \quad (A32)$$

Here when nr_{ij} becomes an integer, the expression $\sin(N_j n r_{ij}\pi)/[N_j\sin(n r_{ij}\pi)]$ is defined by $(-1)^{N_j}$. This approximation and by (A26) immediately gives

$$E_{ij} = \frac{2\tau_0\sigma_A^2}{T}\sum_{n\neq 0}(-1)^{(k_{ij}+1)n}\frac{\sin(N_j n r_{ij}\pi)}{N_j\sin(n r_{ij}\pi)} \times \frac{1}{1 + 4\pi^2\tau_0^2(n/\Delta t^{(i)})^2}. \quad (A33)$$

When $r_{ij} \to 0$, one can verify that $E_{ij} \to E_i^2 = E_j^2$. If there are only two satellites (i) and (j), then $\alpha_i = \alpha_j = 0.5$. Hence, $E^2 = \alpha_i^2 E_i^2 + \alpha_j^2 E_j^2 + 2\alpha_i\alpha_j E_{ij} = E_i^2 = E_j^2$. This is the case that the satellite (i) and the satellite (j) are doing the identical sampling. This scheme certainly cannot reduce the sampling error and the second satellite is redundant.

The formula for the total sampling error can be obtained by combining (A15), (A22), and (A33):

$$E^2(\alpha) = \sum_{i=1}^{I}\alpha_i^2\left(\frac{2\tau_0\sigma_A^2}{T}\left\{\frac{\Delta t^{(i)}}{2\tau_0}\coth\left[\frac{\Delta t^{(i)}}{2\tau_0}\right] - 1\right\} + \frac{\sigma_A^2}{N_i}\left[\frac{\sigma^{(i)}}{\mu^{(i)}}\right]^2\right) + \frac{4\tau_0\sigma_A^2}{T}\sum_{i\neq j}\alpha_i\alpha_j\sum_{n\neq 0}(-1)^{(k_{ij}+1)n} \cdot \frac{\sin(N_j n r_{ij}\pi)}{N_j\sin(n r_{ij}\pi)\{1 + 4\pi^2\tau_0^2[n/\Delta t^{(i)}]^2\}}, \quad (A34)$$

where $\alpha \equiv (\alpha_1, \cdots, \alpha_I)$ is the vector of weights associated with each satellite.

REFERENCES

Baker, D. J., 1990: *Planet Earth, The View from Space.* Harvard University Press, 190 pp.

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

——, 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,** 2195–2206.

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

Laughlin, C., 1981: On the effect of temporal sampling on the observation of mean rainfall. *Precipitation Measurements from Space,* D. Atlas and O. Thiele, Eds., NASA Goddard Space Flight Center.

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

Nakamoto, S., J. B. Valdes, and G. R. North, 1990: Frequency-wavenumber spectrum for GATE Phase I rainfields. *J. Appl. Meteor.*, **29,** 842–850.

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

——, T. L. Bell, R. F. Cahalan, and F. J. Moeng, 1982: Sampling errors in the estimation of empirical orthogonal functions. *Mon. Wea. Rev.*, **110,** 699–706.

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.

——, ——, Y.-S. Ahn, and P. A. Arkin, 1990: Time scales and variability of area-averaged tropical oceanic rainfall. *Mon. Wea. Rev.*, **118,** 1507–1516.

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,** 2207–2216.

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.