A Study of the Sampling Error in Satellite Rainfall Estimates Using Optimal Averaging of Data and a Stochastic Model

THOMAS L. BELL[†] AND PRASUN K. KUNDU*

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

(Manuscript received 28 April 1995, in final form 20 September 1995)

ABSTRACT

A method of combining satellite estimates of rainfall into gridded monthly averages suitable for climatological studies is examined. Weighted averages of the satellite estimates are derived that minimize the mean squared error of the grid-box averages. A spectral model with nonlocal, scaling, diffusive behavior at small distances, tuned to tropical Atlantic (GATE) statistics, is developed to study the optimal weighting method. Using it, the effect of optimal weighting for averaging data similar to what will be provided by the Tropical Rainfall Measuring Mission (TRMM) satellite is examined. The improvement in the accuracy of the averages is found to be small except for higher-latitude grid boxes near the edges of the satellite coverage. The averages of data from a combination of TRMM and a polar orbiting instrument such as SSM/I, however, are substantially improved using the method. A simple formula for estimating sampling error for each grid box is proposed, requiring only the local rain rate and a measure of the sample volume provided by the satellite.

1. Introduction

Climate models are unable to represent all of the physical laws governing the behavior of the atmosphere, oceans, and land surfaces at the level of detail necessary to describe them exactly. Many physical processes must be approximated using simplified, parameterized equations. The values of the parameters can sometimes be determined from the basic physical laws or estimated from analysis of field experiments, but not infrequently they must be adjusted based on comparisons of the climate model behavior with climatological statistics. [See Peixoto and Oort (1992), for example.]

Some of the most important and yet most difficult of processes to represent involve precipitation. They are important because much of the heating of the atmosphere occurs through latent heat release accompanying precipitation, but their complexity requires severe simplifications of the equations used to describe them in the models. For this reason climate modelers are particularly interested in comparing the performance of their models with observed precipitation.

The Tropical Rainfall Measuring Mission (TRMM) satellite is planned for launch in 1997 as a joint venture of the United States and Japan (Simpson et al. 1988) with the participation of many other countries. Because it will include the first satellite-borne meteorological radar, its rainfall estimates are expected to be more accurate than estimates obtained from earlier satellite systems. Its orbit will enable it to provide coverage from latitudes 38°S to 38°N.

The problem of converting the satellite observations into a gridded product convenient for comparison with climate models is not straightforward because the satellite views any one grid box only occasionally and, in general, partially. These fragmentary observations must somehow be combined into as accurate an estimate of the time-averaged rain rate over the grid box as is possible.

A method of doing this will be explored here that is a step beyond simply adding up the observations and dividing by their number. It attempts to take advantage of the spatial and temporal correlations of rain in order to extrapolate from observed areas and times into the unobserved areas and times. The method is closely related to Kriging (e.g., Cressie 1993) or optimal interpolation (e.g., Gandin 1963) adapted to the satellite estimation problem. Its framework can in principle accommodate the corrections needed when there is a diurnal cycle and the problem of combining estimates from several instruments on one satellite or from several satellites. The addition of estimates from geosynchronous satellites is particularly attractive since the high sampling rate of the geosynchronous satellites reduces sampling error considerably (Atlas and Bell 1992; Adler et al. 1993). Geosynchronous satellite estimates are, however, in general much less accurate

[†] Visiting Professor, University of Tokyo, Center for Climate System Research, Tokyo, Japan.

* Applied Research Corporation, Landover, Maryland.

Corresponding author address: Dr. Thomas L. Bell, Laboratory for Atmospheres, Code 913, Goddard Space Flight Center, Greenbelt, MD 20771.

than the TRMM estimates are expected to be, and an optimal strategy for combining the two will require knowledge of the statistics of the errors of the two systems.

In the following section the method of combining the satellite estimates will be developed, allowing for the possibility of correcting the gridded results for the presence of a diurnal cycle and for combining estimates from different instruments or satellites with different error characteristics. In section 3 some simplifying statistical assumptions for the study will be described. Section 4 develops a stochastic model spectrum to represent the GATE (Global Atmospheric Research Program Atlantic Tropical Experiment) covariance statistics. The model is used to study the optimal weighting approach. Section 5 obtains the parameter values of the spectral model that fits the GATE data well, and an empirical fit to the covariance behavior of the model useful for the numerical calculations. In section 6, the optimal weighting for gridding TRMM data at various latitudes is obtained using the GATE statistics. Since rain statistics vary somewhat from place to place, the true optimal weights for various locations may differ from those suggested by the GATE statistics, and so the results here should be considered as exploratory in nature. To improve on the weights obtained here, one will have to use the local statistical properties as they become available.

The usefulness of optimal weighting methods for TRMM data is discussed in section 7, and the consequences of neglecting longer-timescale fluctuations in the data are considered. The impact of adding data from a second satellite instrument like the polar orbiting SSM/I is briefly examined, showing increased effectiveness of optimal weighting methods. In section 8, a simple method of estimating sampling error is developed, and a generalization of the method to estimate sampling errors easily in areas with different statistics is proposed. Section 9 offers some concluding remarks. Two appendices supply some algebraic and numerical details.

2. Optimal weighting method

The problem posed in this section is to produce a gridded map of rainfall on a monthly timescale using satellite observations. If $R(\mathbf{x}, t)$ is the rain rate at point $\mathbf{x}$ at time t and A is the area of a particular grid box, then the monthly averaged rain rate can be written

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

where $T = 1$ month. In the examples studied here the size of the grid box will be taken to be of order 500 km on a side. Note that the rain volume (per unit area) that falls in the month is just $\bar{R}T$.

During the course of one month, the grid box A will be visited by satellites at times t_i, $i = 1, \cdots, n$. The TRMM microwave instrument (TMI), for example, will see grid boxes centered at 0° and at 30° latitudes, with a pattern of observing times like those shown in Fig. 1, assuming a 760-km swath width for the instrument and a circular orbit at 350-km altitude and 35° inclination and using an approximate orbital calculation suggested by Brooks (1977). [A description of the orbital approximation may also be found in Shin and North (1988).] Grid boxes near the equator are visited at fairly regularly spaced intervals, whereas grid boxes at 30° are seen four or five times in a row at intervals of about 1.5 h, followed by an 18-h gap.

The swath of the TMI generally covers only part of the area A, as in Fig. 2, for example. Denote by A_i the area of the portion of the box observed by the satellite instrument at time t_i. Figure 3 shows the histogram of the fractions A_i/A for a box at 30° seen by the TMI during one month. Partial observations are common. Geosynchronous satellites, in contrast, can supply observations of a grid box every half hour over its entire area—but with much less accuracy.

Denote by R_i the true area-averaged rain rate over the observed portion of the grid box; that is,

$$R_i = \frac{1}{A_i}\int_{A_i} d^2\mathbf{x} R(\mathbf{x}, t_i). \qquad (2.2)$$

Estimates $\hat{R}_i$ of this average will be generated from the satellite measurements, with unpredictable random errors ϵ_i expressed as

$$\hat{R}_i = R_i + \epsilon_i. \qquad (2.3)$$

It will be assumed that the estimates have been corrected for biases, so that

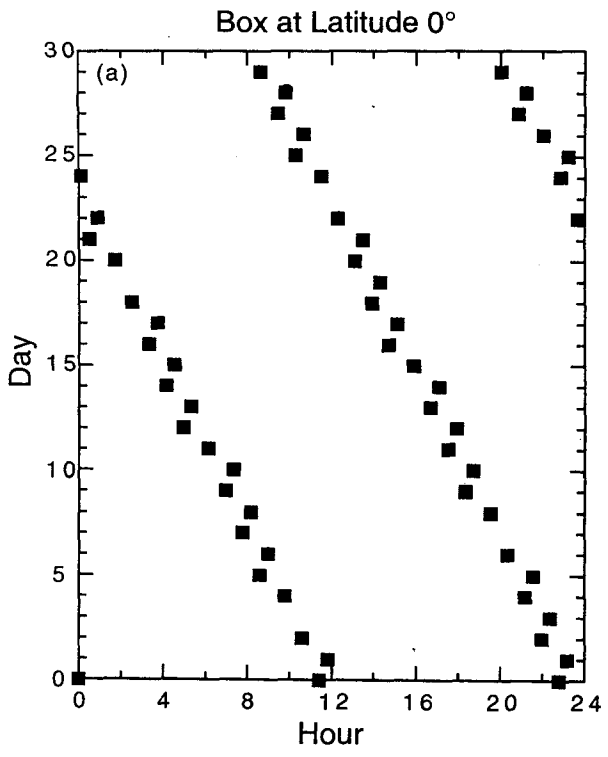
$$\langle \epsilon_i \rangle = 0. \qquad (2.4)$$

Here the angular brackets indicate the expected value over the climatological distribution of occurrences of rain in the area of the grid box and also over the distribution of instrument noise. The distribution of errors ϵ_i may depend on the fraction of the area observed and the time t_i (if, for instance, there is a diurnal cycle); it could depend on how much and what types of rain are occurring at the time (e.g., stratiform, convective, etc.); it could also depend on the estimation method and the satellite used.

A method of combining the estimates $\hat{R}_i$ to produce the best possible estimate of $\bar{R}$ [Eq. (2.1)] is needed. It is proposed here to estimate $\bar{R}$ with a linear combination of the individual estimates $\hat{R}_i$, that is, with

$$\hat{\bar{R}} = \frac{1}{n}\sum_{i=1}^{n} w_i \hat{R}_i, \qquad (2.5)$$

where the weights w_i are chosen to optimize the estimate $\hat{\bar{R}}$ and the factor $1/n$ has been introduced so that the weights w_i are of order 1 in size. The weights are constrained to produce an unbiased estimate, so that



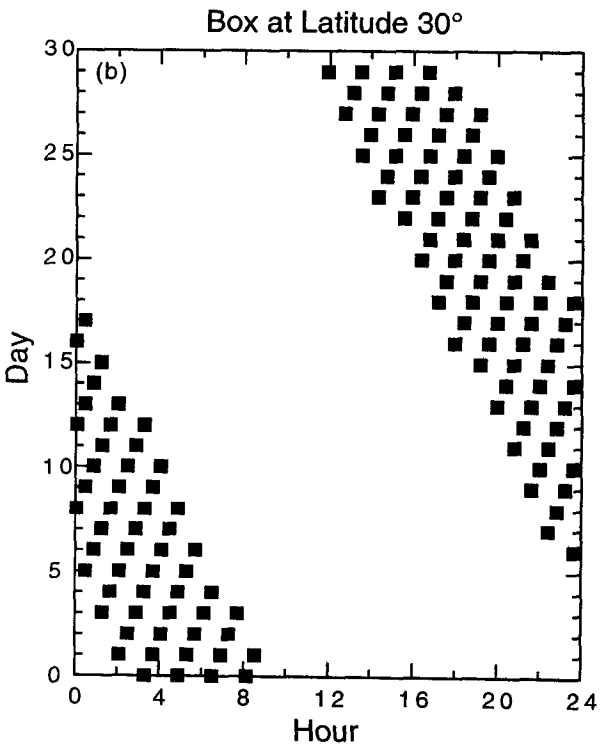


FIG. 1. Visiting times during one month of 512 km × 512 km box centered at latitudes (a) 0° and (b) 30° for TRMM microwave instrument swath width 760 km. Satellite in 350-km altitude circular orbit with 35° inclination.

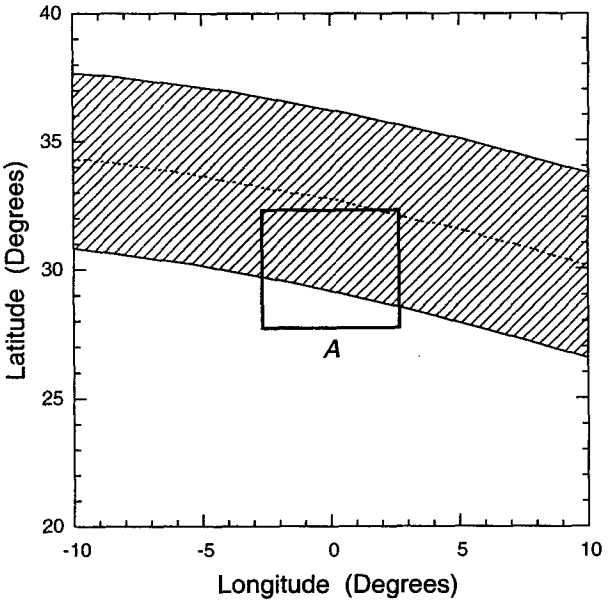


FIG. 2. Example of partial overlap of TMI swath with box A. Box dimensions are 512 km × 512 km with sides lying along meridian and latitude lines.

$$\langle \hat{\bar{R}} \rangle = \langle \bar{R} \rangle. \tag{2.6}$$

This is by no means the only approach possible. Satellite estimates can be made at the resolution of the footprints of the instruments (~10 km). The estimate (2.5) makes very little use of this higher resolution information since it starts from simple area averages $\hat{R}_i$ of the high-resolution retrieved rain rates. Optimal strategies could in principle be devised that use weighted averages of the individual footprint estimates, which amounts in effect to optimal interpolation (or extrapolation) of the rain rates where they are measured to the unmeasured portions of the space–time volume

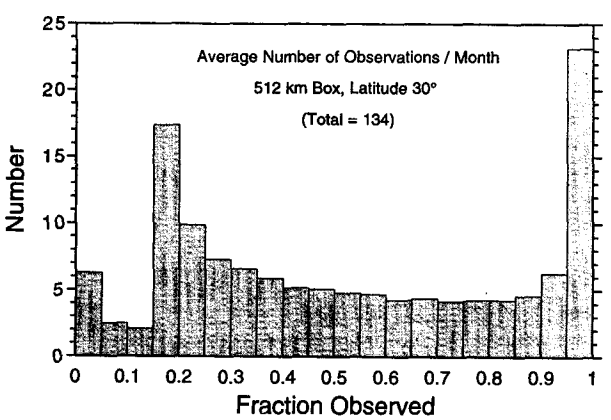


FIG. 3. Average number of occurrences during one month of fractional observations by TRMM of a 512-km box located at 30° latitude.

for which the average (2.1) is desired. This is computationally much more difficult and depends sensitively on the space and time correlations of point rain rates, which are generally only imperfectly known. The approach taken here is comfortably intermediate between doing "nothing at all," that is, using straightforward averages, and optimal interpolation in its most detailed application. It differs from the optimal weighting approach explored by North et al. (1993) for combining estimates from multiple satellites in that here the weights given to the observations from each overflight of a satellite are individually optimized, whereas North et al. (1993) obtain monthly averages for each satellite with straightforward averaging and then combine the various satellite monthly averages with optimized weights. The weighting method (2.5) includes the North et al. (1993) weighting as a special case but is more complicated to implement.

The estimate (2.5) is special in that it does not include a constant ($\hat{R}_i$ independent) term in the estimate, which is commonly used in optimal interpolation strategies. That is, we have not written

$$\hat{\bar{R}} = \frac{1}{n}\sum_{i=1}^{n} w_i \hat{R}_i + b.$$

If such a constant term b were included, and no rain were observed by the satellite during the month, such an estimate would in effect supply the climatological mean instead of zero as its estimate of the grid-box average. This is felt to be undesirable because it is easy to imagine long-term shifts in the large-scale rain patterns, such as those that accompany El Niño–Southern Oscillation (ENSO) events (e.g., Rasmusson and Wallace 1983; Lau and Chan 1986), and a weighting strategy that permits the grid-box averages to drop to zero when the rain systems move away from the grid box seems preferable to one that only drops to the prior climatological mean.

The weights w_i in (2.5) will be chosen to minimize the mean squared error

$$\sigma_E^2 = \langle(\hat{\bar{R}} - \bar{R})^2\rangle, \tag{2.7}$$

subject to the normalization constraint (2.6). Because of the constraint (2.6), expression (2.7) can be rewritten as

$$\sigma_E^2 = \langle(\hat{\bar{R}}' - \bar{R}')^2\rangle, \tag{2.8}$$

where primes on random variables denote deviations from the ensemble mean. Note that for each month and grid box, the weighted averages (2.5) will not in every case be closer to the true mean $\bar{R}$ than a simple average of the observations. They are only more accurate in the average sense that they will have a lower mean-squared error σ_E^2 than the simple-average estimates.

The TRMM satellite will not sample all times of the day equally often during a month (see Fig. 1, for example). Monthly averages of the satellite observations will therefore favor some parts of the day over others, leading to a possible bias in $\hat{\bar{R}}$ if there is a significant diurnal cycle. The weights in (2.5) can reduce this potential bias if enough data has been collected so that the climatological mean rain rate as a function of the time of day is known. If this mean is denoted by $\mu(t)$, so that $\langle R_i\rangle = \mu(t_i)$, then the constraint (2.6) can be written

$$\frac{1}{n}\sum_{i=1}^{n} w_i \mu(t_i) = \frac{1}{T}\int_0^T dt\, \mu(t). \tag{2.9}$$

If there is no appreciable diurnal cycle [$\mu(t) = \mu$], the constraint (2.9) reduces to the simpler form

$$\sum_{i=1}^{n} w_i = n, \quad [\mu(t) = \mu]. \tag{2.10}$$

The weights w_i are obtained by minimizing (2.8) subject to the normalization constraint (2.9). This can be accomplished by minimizing (2.8) with a Lagrange multiplier times the constraint added, that is, by minimizing

$$\left\langle\left(\frac{1}{n}\sum_i w_i \hat{R}_i' - \bar{R}'\right)^2\right\rangle - 2\lambda\sum_i w_i \mu(t_i), \tag{2.11}$$

where λ is the Lagrange multiplier. We will omit specifying the range of the sum (1 to n) from now on. The factor of 2 has been introduced for convenience.

The weights that minimize (2.11) are obtained by setting its derivatives with respect to each w_i equal to zero. The linear equations for the weights that result are

$$n^{-1}\sum_j \langle(R_i' + \epsilon_i)(R_j' + \epsilon_j)\rangle w_j = \langle(R_i' + \epsilon_i)\bar{R}'\rangle + \lambda\mu(t_i). \tag{2.12}$$

These equations and the constraint (2.9) fix the weights w_i and the multiplier λ. They are easy to solve once the coefficients are known since they are linear in the w_i. They provide a general framework for combining the partial observations of the satellite into a best estimate of the grid-box averages.

The results derived here apply equally well to the problem of combining data from several instruments or satellites. In that case the index i in the weighted average (2.5) is used to label both the time of observation and from which instrument or satellite the estimate is obtained. The equations are unchanged. What is changed is the level of difficulty in obtaining the optimal weights since a reasonable statistical model of the errors ϵ_i of the different satellite estimates must be developed in order to evaluate the terms in (2.12).

3. Statistical assumptions for calculations

In order to gain experience with the optimal-weighting approach, a simplified version of the problem is

examined here. It is based on realistic sampling characteristics of the TRMM satellite; that is, the observation times and coverage of the grid boxes are representative of what the TRMM satellite is expected to provide. The statistical behavior of the rain observed by the satellite is taken to be similar to what was observed in GATE.

We shall neglect the effect of retrieval errors in the TRMM estimates here. This is a reasonably good approximation if the root-mean-square (rms) error $\langle \epsilon_i^2 \rangle^{1/2}$ of the estimates is less than perhaps 30% of the rms area-averaged rain rate $\langle (R_i')^2 \rangle^{1/2}$. This assumption is probably valid (except for cases where the area A_i observed contains only a few satellite footprints) if the estimation errors for each footprint are independent, as discussed by Wilheit (1988) and Bell et al. (1990).

We shall also neglect the effect of the diurnal cycle. This is not an unreasonable approximation in cases where the amplitude of the diurnal cycle is relatively weak compared with $\langle (R_i')^2 \rangle^{1/2}$, as in GATE (see Bell and Reid 1993, for example). With these two assumptions, (2.12) for the optimal weights simplifies to

$$n^{-1} \sum_j \langle R_i' R_j' \rangle w_j = \langle R_i' \bar{R}' \rangle + \lambda, \qquad (3.1)$$

with $\Sigma_i w_i = n$.

The covariances $\langle R_i' R_j' \rangle$ and $\langle R_i' \bar{R}' \rangle$ must be calculated from a statistical model of the rain-rate fields encountered by the satellite. The left-hand side covariance in (3.1), when written out explicitly in terms of the rain-rate field $R(\mathbf{x}, t)$, becomes

$$\langle R_i' R_j' \rangle = \frac{1}{A_i A_j} \int_{A_i} d^2\mathbf{x} \int_{A_j} d^2\mathbf{y}\, c(\mathbf{x}, t_i; \mathbf{y}, t_j), \qquad (3.2)$$

with the point covariance function defined as

$$c(\mathbf{x}, t; \mathbf{y}, u) \equiv \langle R'(\mathbf{x}, t) R'(\mathbf{y}, u) \rangle. \qquad (3.3)$$

The covariance on the right-hand side of (3.1) is

$$\langle R_i' \bar{R}' \rangle = \frac{1}{A_i A T} \int_{A_i} d^2\mathbf{x} \int_{A} d^2\mathbf{y} \int_0^T dt\, c(\mathbf{x}, t_i; \mathbf{y}, t). \qquad (3.4)$$

In order to proceed, a representation of the covariance $c(\mathbf{x}, t; \mathbf{y}, u)$ is needed. The model used here is described in the following section. Readers who wish to skip the details of representing the space–time covariance may proceed directly to section 6.

4. Model of rain rate covariance

As was emphasized in the sampling error calculations of Bell et al. (1990), rain rate fields (and many other geophysical fields as well) have spatial and temporal correlations with the property that spatial averages have longer timescales as the averaging area increases. Such behavior is commonly encountered in turbulent fluids. It is a simple consequence of the fact that the variability on larger scales is governed by motions that require more time to span the larger distances and thus occur more slowly. See Batchelor (1953), for example. We wish to construct a model of the space–time correlations of rain with this property. This phenomenon is relatively easy to represent in terms of the Fourier spectrum of the field, but calculations of the covariances in (3.2) and (3.4) are awkward to carry out using Fourier expansions. We have therefore elected to deal with this problem in two stages. In the first stage, a simple Fourier spectral representation with as few parameters as possible is adjusted until it reproduces the essential aspects of GATE space–time correlations. In the second stage, a more complicated empirical fit to the space–time correlations implied by the spectral model is used for the actual calculations of (3.2) and (3.4). These two stages will now be described in more detail.

The rain field statistics are taken to be homogeneous and isotropic. That is, within the grid-box area A the statistics are assumed not to depend on location or orientation. This is certainly not true in reality, but because the only quantities that matter to the optimal-weighting procedure are the statistics of (generally large) area averages, this approximation is probably adequate for our purposes. Moreover, because the optimal weights minimize the mean-squared error, the latter is insensitive to small deviations of the weights from their optimal values. It is expected that the weights obtained by minimizing an approximate expression for the error are still close enough to being optimal with respect to the exact mean-squared error as long as the departure from homogeneity and isotropy is not too large.

The rain statistics will also be assumed not to vary during the course of a day or of a month; that is, they will be assumed to be stationary. Homogeneity and stationarity imply that the lagged covariance (3.3) of the rain-rate field at points $\mathbf{x}$ and $\mathbf{y}$ at the instants t and u has the simple form

$$c(\mathbf{x}, t; \mathbf{y}, u) = c(\boldsymbol{\rho}, \tau), \qquad (4.1)$$

with $\boldsymbol{\rho} = \mathbf{y} - \mathbf{x}$ and $\tau = u - t$. Isotropy implies that $c(\boldsymbol{\rho}, \tau) = c(\rho, \tau)$, where $\rho = |\boldsymbol{\rho}|$. We can express $c(\rho, \tau)$ in terms of the Fourier integral

$$c(\rho, \tau) = (2\pi)^{-3/2} \int_{-\infty}^{\infty} d\omega \int d^2\mathbf{k} e^{i(\mathbf{k}\cdot\boldsymbol{\rho} - \omega\tau)} \tilde{c}(\mathbf{k}, \omega), \qquad (4.2)$$

where integration is over the entire wavenumber $\mathbf{k}$ plane and the Fourier transform $\tilde{c}(\mathbf{k}, \omega)$ represents the (unnormalized) power spectrum of the random field. As a consequence of isotropy, $\tilde{c}(\mathbf{k}, \omega) = \tilde{c}(k, \omega)$, where $k = |\mathbf{k}|$.

A simple model of the power spectrum will be used here that is suggested by the discrete version employed

by Bell et al. (1990) in constructing a gridded rainfall model. The rain rate field $R(\mathbf{x}, t)$ is written as a sum of spatial Fourier modes with amplitudes $a(\mathbf{k}, t)$. Each amplitude is assumed to be described by a linear stochastic equation of the form

$$\frac{da(\mathbf{k}, t)}{dt} = -\frac{1}{\tau_k} a(\mathbf{k}, t) + f_{\mathbf{k}}(t), \qquad (4.3)$$

where $f_{\mathbf{k}}$ is a white-noise forcing and the timescale $\tau_{\mathbf{k}}$ of the process is taken to have the form

$$\tau_k = \frac{\tau_0}{(1 + k^2 L_0^2)^{1+\nu}}, \qquad (4.4)$$

where τ_0 and L_0 are characteristic time and length scales. The form (4.4) assumed here is suggested by (18) of Bell et al. (1990). In effect, the timescale τ_k decreases at small spatial scales (large k) as a power law but at large spatial scales (small k) approaches τ_0. The forced diffusion model considered by North and Nakamoto (1989) corresponds to the special case $\nu = 0$. With $\nu \neq 0$, (4.3) represents a randomly forced field with a kind of nonlocal diffusion; the diffusion scales at small distances differently from ordinary diffusion. The timescales associated with various length scales also vary with areal size differently from the way they do with ordinary diffusion. A nonzero value of ν was found to be particularly helpful in fitting the spatial correlations of the model to the GATE correlations. Although (4.3) is written as a full dynamical model for rainfall, it should be emphasized that only its predictions for the second-moment statistics (the spectrum) are actually used here.

Using standard techniques of time series analysis (Jenkins and Watts 1968, for example), the power spectrum implied by (4.3) is found to be

$$\tilde{c}(\mathbf{k}, \omega) = \frac{F_0}{\omega^2 + 1/\tau_k^2}, \qquad (4.5)$$

where F_0 is a constant and, upon substituting (4.4),

$$\tilde{c}(\mathbf{k}, \omega) = \frac{F_0 \tau_0^2}{\tau_0^2 \omega^2 + (1 + k^2 L_0^2)^{2+2\nu}}. \qquad (4.6)$$

The model spectrum thus has four parameters, F_0, τ_0, L_0, and ν, that must be determined from the data.

5. Fitting the spectral model to GATE data

In fitting a parameterized statistical model to data, one's first choice might be to determine the values of the parameters using maximum likelihood methods. In the case of a spectral model, this would normally lead to fitting the model spectrum to the space–time spectrum of the data. There is, however, a difficulty with this approach. The area covered by the GATE data is a square 280 km on a side, and it is required that the spectral model reproduce the statistics of area averages on this scale (and extrapolate to scales nearly twice that) as faithfully as possible. It is notoriously difficult to estimate the spectral power of a series at scales similar in magnitude to the length of the series, the reason being that ordinary spectral analysis implicitly introduces spurious periodicity in the series on these scales. [See Jenkins and Watts (1968) for discussion of some of these problems.] Although there are techniques for attempting to reduce the distortions in the spectral estimates at large scales, it was decided to determine the parameters of the model by fitting the model behavior directly to the statistics of area averages of the data on the scales of most interest to us. In effect, the fitting process is skewed so that the statistics of the largest scales are reproduced best.

The GATE dataset consists of a series of radar-derived rain-rate maps on a 4-km grid. The maps are 400-km diameter circles centered at 8°30′N latitude and 23°30′W longitude in the Atlantic Ocean about 1000 km west of the African coast in the intertropical convergence zone (ITCZ). The maps are generally spaced 15 min apart, with occasional large gaps. The radar data were collected during the summer of 1974. The methods used in producing this dataset are described by Hudlow and Patterson (1979). Only the data from a 280-km square box inscribed within the 400-km circle are used in the statistical analysis. Spatial correlations of gridpoint rain rates, variances of area-averaged rain rates for a range of areal sizes, and lagged autocorrelations of area-averaged rain rate are obtained from the dataset.

Let us first list the steps we shall follow in obtaining the model parameters before going into the details.

1) Since the GATE data is gridded on a 4-km grid, the spatial correlation of 4-km box averages implied by the model spectrum is calculated, and the spectrum parameters L_0 and ν are adjusted until the calculated model correlations fit the GATE spatial correlations well.

2) The overall normalization F_0 of the spectrum is obtained by fitting the model variances of area-averaged rain rate to the same variances determined from the data.

3) Finally, the correlation timescale τ_0 is obtained from fitting the lagged autocorrelation of area-averaged rain rate.

It is interesting, before doing this, to examine the spatial covariance $c(\rho, \tau)$ of point rain rate implied by the spectral model. This is obtained from (4.2). The integration over ω is easily done and gives

$$c(\rho, \tau) = (2\pi)^{-1} \int d^2\mathbf{k} e^{i\mathbf{k}\cdot\boldsymbol{\rho}} c(\mathbf{k}, \tau), \qquad (5.1)$$

with

$$c(\mathbf{k}, \tau) \equiv \sqrt{\frac{\pi}{2}} F_0 \tau_k e^{-|\tau|/\tau_k}. \qquad (5.2)$$

Carrying out the angular integration in the **k** plane then yields

$$c(\rho, \tau) = \int_0^\infty dk k J_0(k\rho) c(k, \tau), \qquad (5.3)$$

where $J_0(x)$ is the Bessel function of the first kind of order zero. At zero lag ($\tau = 0$), the k integral can be done (Abramowitz and Stegun 1970, section 11.4.44), with the result

$$c(\rho, 0) = \gamma_0 C_\nu\left(\frac{\rho}{L_0}\right), \qquad (5.4)$$

where γ_0 is a constant related to F_0 in the spectrum (4.6) by

$$F_0 = \sqrt{\frac{2}{\pi}}\,\Gamma(1+\nu)\left(\frac{L_0^2}{\tau_0}\right)\gamma_0 \qquad (5.5)$$

and where we have introduced the function

$$C_\nu(z) \equiv \left(\frac{z}{2}\right)^\nu K_\nu(z), \qquad (5.6)$$

with $K_\nu(z)$ denoting the modified Bessel function of order ν (Abramowitz and Stegun 1970, section 9.6). From the behavior of $K_\nu(z)$ near the origin it follows that for $\nu > 0$, $C_\nu(z) \to \Gamma(\nu)/2$ as $z \to 0$, and in this case the point variance $c(0, 0)$ is finite and well defined. However, for $\nu < 0$, which fits the GATE data best,

$$c(\rho, 0) \sim \left(\frac{\rho}{L_0}\right)^{-2|\nu|} \qquad (5.7)$$

as $\rho \to 0$, and the point variance is infinite. This is also the case for $\nu = 0$, for which $c(\rho, 0)$ diverges logarithmically, as mentioned by North and Nakamoto (1989). It has sometimes been hypothesized that the statistics of instantaneous point rain rates may be nearly singular as a consequence of scaling behavior (Lovejoy and Mandelbrot 1985, for instance). Averages over any nonzero area have finite variance, however, and so this phenomenon causes no problem with fitting the model to the data, which consist only of area-averaged measurements.

The behavior of $c(\rho, \tau)$ for $\rho = 0$ can likewise be obtained by returning to (5.1). After doing the angular integral in **k** space and making a few judicious integration variable changes, an expression for $c(0, \tau)$ can be obtained with limiting behavior

$$c(0, \tau) \to \frac{\Gamma(1+\nu)}{2(1+\nu)}\gamma_0 \times \begin{cases} \Gamma(-q)(|\tau|/\tau_0)^q, & \tau \to 0, \\ \dfrac{\exp(-|\tau|/\tau_0)}{|\tau|/\tau_0}, & \tau \to \infty, \end{cases} \qquad (5.8)$$

where $q = \nu/(1+\nu) < 0$.

a. Obtaining the spectral model parameters

Let us turn now to the calculation of the model spatial correlation between two areal averages of the rain-rate field. Denote by $\mathcal{A}$, $\mathcal{B}$ two boxes each of area L^2 with their centers separated by **s**, as depicted in Fig. 4. The covariance of these two box averages is

$$\mathcal{C}_{\mathcal{AB}}(\mathbf{s}, \tau) = \langle R'_{\mathcal{A}}(t) R'_{\mathcal{B}}(t+\tau)\rangle, \qquad (5.9)$$

where $R_{\mathcal{A}}$ and $R_{\mathcal{B}}$ are the area-averaged rain rates over the respective boxes $\mathcal{A}$ and $\mathcal{B}$. The covariance can be written in terms of the point covariance (3.3) as

$$\mathcal{C}_{\mathcal{AB}}(\mathbf{s}, \tau) = \frac{1}{L^4}\int_{\mathcal{A}} d^2\mathbf{x} \int_{\mathcal{B}} d^2\mathbf{y}\, c(\mathbf{y} - \mathbf{x}, \tau). \qquad (5.10)$$

The correlation of box averages is just the ratio of the covariance to the variance of the box averages,

$$\Phi_{\mathcal{AB}}(\mathbf{s}, \tau) = \frac{\mathcal{C}_{\mathcal{AB}}(\mathbf{s}, \tau)}{\sigma^2_{\mathcal{A}}}, \qquad (5.11)$$

$$\sigma^2_{\mathcal{A}} \equiv \mathcal{C}_{\mathcal{AA}}(0, 0). \qquad (5.12)$$

Model spatial correlations $\Phi_{\mathcal{AB}}(\mathbf{s}, 0)$ with $L = 4$ km were computed for various choices of ν and L_0. Algebraic and numerical details are described in appendix A. Using trial and error, values of ν and L_0 were found for which the model spatial correlations agree well with the spatial correlation of the GATE Phase I data (covering the 18-day period 28 June to 16 July 1974). Figure 5 shows the GATE spatial correlations (dots) for various spatial separations, as reported by Bell et al. (1990), and the model correlation (smooth curve) for $\nu = -0.11$ and $L_0 = 104$ km. Bell et al. (1990) estimate that the 95% confidence limits for the GATE spatial correlations at separations >80 km are approximately ±0.1. Considering the estimated level of sampling error, the agreement is perhaps better than would have been expected.

Next, the variances of area-averaged rain rate $\sigma^2_{\mathcal{A}}$ are used to set the overall normalization of the spectrum, γ_0, which is related to F_0 by (5.5). Figure 6 shows the variances computed from GATE Phase I data (A.

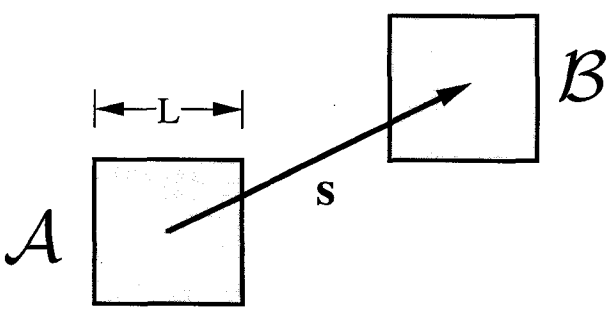


FIG. 4. Diagram of boxes $\mathcal{A}$ and $\mathcal{B}$ in (5.9).

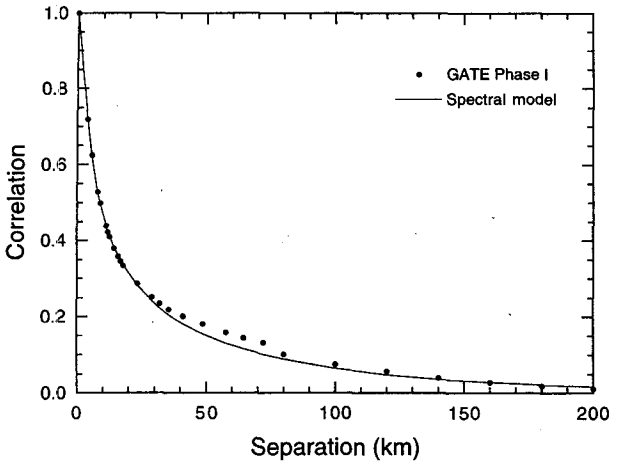


FIG. 5. Comparison of the spatial correlations of 4-km gridded rain rate found from Phase I of GATE and for the spectral model with $\nu = -0.11$ and $L_0 = 104$ km.

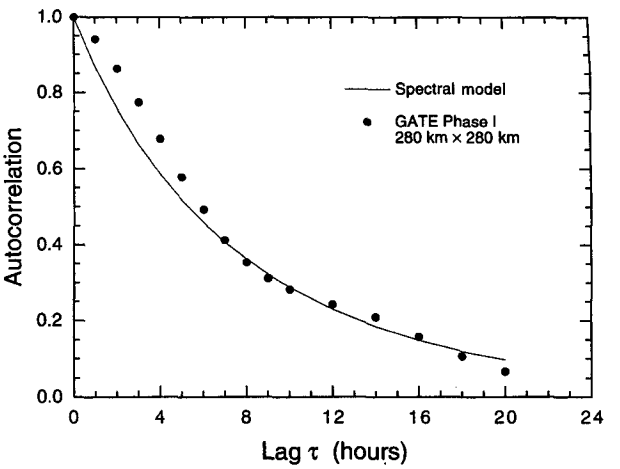


FIG. 7. Lagged autocorrelation $\Phi_{\mathcal{A}\mathcal{A}}(0, \tau)$ of area-averaged rain rate for $L = 280$ km, comparing GATE data and spectral model behavior.

McConnell and L. S. Chiu, personal communication, 1987) and $\sigma^2_{\mathcal{A}}$ from the spectral model, with $\gamma_0 = 1.0$ mm^2 h^{-2}, for various box sizes $L = \mathcal{A}^{1/2}$. (The data point at $L = 354$ km was obtained using the full 400-km diameter area covered by the GATE data.) The 95% confidence limits on the point at $L = 280$ km are based on the assumption that $R_{\mathcal{A}}$ is normally distributed with an 8-h correlation time and may be underestimated.

Finally, the parameter τ_0 was determined by fitting the autocorrelation of box-averaged rain rate for $L = 280$ km, obtained by setting $s = 0$ in (5.10), to the GATE autocorrelations for the same sized box. Details of the algebra and numerical techniques are given in appendix A. The quality of the fit is shown by Fig. 7. Although the spectral model captures the overall decay of correlations, and beyond about lag 7 h agrees quite well, it underestimates the correlation at smaller lags. To improve the fit would require going to an autoregressive model higher in order than the first-order model used in (4.3). Bell and Reid (1993) were, in fact, able to fit this behavior to a second-order model. The underestimate of lagged correlation for small lags causes the spectral model to underestimate slightly the effectiveness of the optimal weighting and to overestimate sampling error when compared with what a better-fitting model would give. The numerical difficulties in dealing with a higher-order model would be much greater, however.

The parameter values for the spectral model are summarized in Table 1.

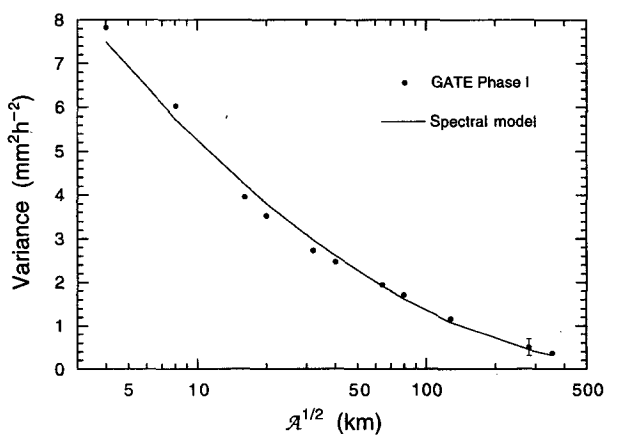


FIG. 6. Variance of area-averaged rain rate versus the square root of the area, ranging from the grid box size 4 km to the effective size of the GATE area. The error bar at $L = 280$ km is explained in the text.

b. *Some characteristics of the spectral model fit*

As one check of the quality of the fit, the lagged correlations of box averages for $L = 64$ km are compared in Fig. 8 to the GATE Phase I statistics for four boxes in the GATE area. The NE-quadrant box's autocorrelation is stronger than the other three boxes' and is evidence for some inhomogeneity in the statistics in the area. We have been unable to discover a physical explanation for the anomaly. The spectral model agrees reasonably well with the autocorrelations in three of

TABLE 1. Parameters of spectral model fit to GATE Phase I data. Note that F_0 is related to these parameters by (5.5).

Parameter	Value	
γ_0	1.0	mm^2 h^{-2}
ν	−0.11	—
L_0	104	km
τ_0	13	h

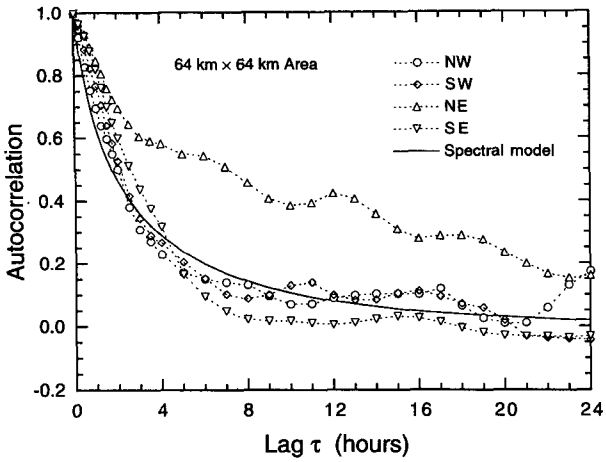


FIG. 8. Lagged autocorrelations of area-averaged rain rate for L = 64 km, comparing the behavior of the spectral model and of GATE data averaged over four boxes centered 100 km northwest, southwest, northeast, and southeast of the center of the GATE area.

the quadrants except that, as in the 280-km box case, it slightly underestimates autocorrelations for small lags.

Note the presence for both sized areas of a long "tail" in the model autocorrelations that would be missed by simple exponential fits to the autocorrelations. The model's lagged autocorrelation for 4-km box averages drops to $1/e$ at $\tau \approx 0.2$ h, rather smaller than the 0.9-h correlation time estimated for a box at the center of the GATE area (from Laughlin 1981). However, the integral correlation time, defined as

$$\tau_{\text{int}}(L) \equiv \int_0^\infty \Phi_{AA}(0, \tau)d\tau, \qquad (5.13)$$

which would be identical to the lag at which the autocorrelation drops to $1/e$ if the autocorrelation were a simple exponential $\exp(-\tau/\tau_{\text{int}})$, ranges from 1.5 h at L = 4 km to 10 h at L = 280 km, reflecting the substantial nonexponential tails in the autocorrelation. The integral correlation time generally exceeds the GATE correlation times (estimated from the lag at which the autocorrelation has dropped to $1/e$) by a factor of about 1.7. It is well described by the power-law form $\tau_{\text{int}}(L) \approx 0.67\, L^{0.49}$ h.

Some asymptotic properties of the spectral model are derived in appendix A.

c. *Empirical space–time correlation fit to spectral model*

In order to carry out the calculation of the covariances in (3.2) and (3.4), required in order to get the optimal weights, the covariance $c(\mathbf{x}, t; \mathbf{y}, u)$ is needed. It is computationally too expensive to carry out the spectral transform (4.2) for each separation $(\mathbf{s}, \tau)$ needed. An empirical interpolation formula to represent the covariance has therefore been developed.

It was assumed that the spatial integrals in (3.2) and (3.4) would be done numerically using an 8-km grid, twice as coarse as the GATE data grid. The spectral model was used to calculate the space–time covariances of L = 8-km boxes from (5.10) and the methods described in appendix A. These covariances were fitted to a form

$$\mathcal{C}_{ab}(\mathbf{s}, \tau) \approx \sigma_a^2 \Phi_0(s) \exp\{-[|\tau|/\tau(s)]^{\mu(s)}\}, \qquad (5.14)$$

with factorizable time dependence, for a set of values of s ranging between 0 and 400 km. The exponential dependence on τ was suggested by (5.2). Algebraic tractability of time integrals over the covariance was also a factor in the choice of the time dependence. In (5.14) the zero-lag spatial correlation $\Phi_0(s)$ and the characteristic timescales and exponents $\tau(s)$ and $\mu(s)$ depend on the separation s and, implicitly, on the size L of the grid boxes a and b. The forms of the functions were taken to be

$$\Phi_0(s) = \begin{cases} (a_1 s + a_2)^{-a_3} e^{-s/a_4}, & s \geq L, \\ 1, & s = 0, \end{cases} \qquad (5.15)$$

which is the same as was used successfully in Bell et al. (1990) and

$$\tau(s) = \begin{cases} b_1 s^{b_2} + b_3, & s \geq L, \\ \tau(0), & s = 0; \end{cases} \qquad (5.16)$$

$$\mu(s) = \begin{cases} c_1 s^{c_2} + c_3, & s \geq L, \\ \mu(0), & s = 0. \end{cases} \qquad (5.17)$$

The coefficients $a_1, \ldots, a_4, b_1, b_2, b_3, c_1, c_2, c_3$ were determined by standard nonlinear least squares fitting procedures. The specific forms in (5.15)–(5.17) are chosen primarily to allow efficient machine computation without appreciable loss of accuracy. It should be emphasized that although a large number of parameters are needed in order to represent the model covariance function $\mathcal{C}_{ab}(\mathbf{s}, \tau)$ to sufficient accuracy, they are not all independent: the covariance $\mathcal{C}_{ab}(\mathbf{s}, \tau)$ is completely determined by the box size L and by the model spectrum, which depends on only the four independent parameters listed in Table 1.

The quality of the fit of $\Phi_0(s)$ to the spatial correlation is as good as or better than that of the spectral-model spatial correlation to the GATE data shown in Fig. 5 and is not shown. The fits of $\mathcal{C}_{ab}(\mathbf{s}, \tau)$ to (5.14) as a function of lag τ are shown in Fig. 9 for separations s = 0, 129, and 400 km. (Fits at several dozen other separations were examined but are not shown here.) The quality of the fits is excellent for $s \leqslant 100$ km and $s \geqslant 300$ km but decreases somewhat in the range 100 km $\leqslant s \leqslant$ 300 km. The movement of the inflection point to larger τ with increasing spatial separation s is

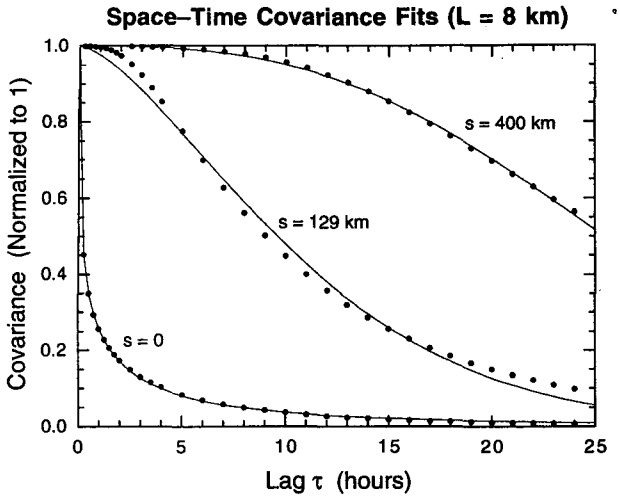


FIG. 9. Fits of the empirical interpolation formula (5.14) to spectral model covariances for $L = 8$ km boxes at separations $s = 0$, 129, and 400 km. Plots are normalized to one at $\tau = 0$ for each separation. Spectral model results are shown as dots; empirical fits as solid lines.

an interesting phenomenon but might be difficult to establish with radar data because of the low level of correlations at these separations and lags.

The parameters $\mu(s)$ and $\tau(s)$ obtained from the above fits at many separations s are then themselves fitted to the forms (5.16) and (5.17) using a nonlinear least squares procedure. These latter fits are shown in Fig. 10. Because these calculations are done for a rectilinear grid, there is a possibility for anisotropic effects to enter, especially for small separations s. However, there was little evidence of this in the cases checked, and the isotropic forms independent of the direction of s used here appeared to work well.

The parameters obtained following the procedures outlined above are shown in Table 2 for two box sizes,

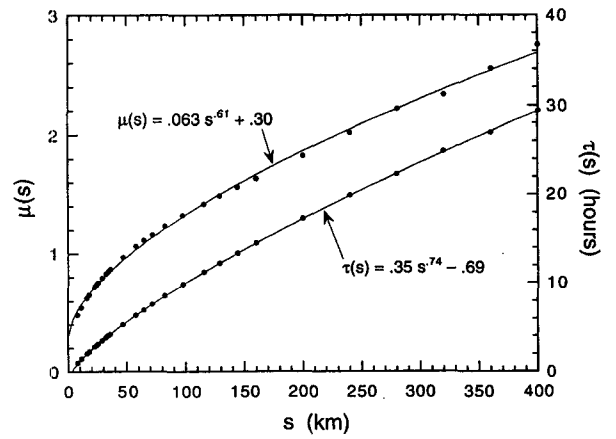


FIG. 10. Fits of (5.16) and (5.17) to the characteristic times and exponents obtained from fits like those shown in Fig. 9. The values at $s = 0$ are excluded from the fit. See Table 2 for these values.

TABLE 2. Parameter values used in the empirical interpolation formulas (5.14)–(5.17) for two grid-box sizes with areas $a = L^2$. Assumes separation s in kilometers. The original values used in formula (5.15) by Bell et al. (1990) are shown for comparison.

		Spectral model		Bell et al. (1990)
		$L = 4$ km	$L = 8$ km	$L = 4$ km
σ_a^2	[mm^2 h^{-2}]	7.5	5.7	7.0
a_1		0.8244	0.6968	1.193
a_2		−1.0208	−3.0495	−2.165
a_3		0.3146	0.2611	0.2902
a_4		70.12	71.40	78.11
b_1		0.2548	0.3476	
b_2		0.8043	0.7446	
b_3		−0.2724	−0.6877	
τ (0)	[h]	0.2285	0.4543	
c_1		0.1313	0.0629	
c_2		0.4814	0.6070	
c_3		0.1345	0.2994	
μ (0)		0.3307	0.3840	

$L = 4$ km and $L = 8$ km, as well as the parameters used for fitting the spatial correlation form (5.15) directly to the GATE data, which are given in Bell et al. (1990). Note that only the empirical fits for $L = 8$ km were used in the optimal weighting calculations. For the larger areas that matter most to the calculations here, the model generally reproduces GATE correlations quite well for lags 0–24 hours, though at some intermediate lags it underestimates correlations. These underestimates would tend to overestimate sampling error and to underestimate improvements possible with optimal weighting when compared with what would be calculated using a more accurate model. Given the level of sampling uncertainty in the statistics due to the size of the GATE dataset itself, however, it is difficult to justify a more elaborate modeling effort.

6. Optimal weights

With a model of the rain rate covariance in hand, we are now ready to calculate the optimal weights from the linear equations (3.1), which have the simple form

$$n^{-1} \sum_j P_{ij} w_j = Q_i + \lambda, \tag{6.1}$$

where the coefficient matrix

$$P_{ij} \equiv \langle R_i' R_j' \rangle \tag{6.2}$$

and the vector

$$Q_i \equiv \langle R_i' \bar{R}' \rangle \tag{6.3}$$

are given by the integrals (3.2) and (3.4) over the point covariance $c(\mathbf{x}, t; \mathbf{y}, u)$. The spatial integrals are done numerically on an 8-km grid using the approximation for the covariance (5.14) with the parameter values in Table 2 for $L = 8$ km. Algebraic and numerical details

may be found in appendix B. Once the matrix P_{ij} and vector Q_i have been computed, numerical solution of (6.1) for the weights w_i is straightforward.

As a point of reference to aid in understanding the optimal weights, it is helpful to consider a limiting situation in which rain rates at slightly different times and places are uncorrelated with each other. In this situation, the matrix $\langle R_i' R_j' \rangle$ becomes diagonal, and after a small amount of algebra, (3.1) simplifies to

$$n^{-1}(\sigma^2/A_i)w_i^{(0)} = \left(\frac{\sigma^2}{A}\right) + \lambda, \qquad (6.4)$$

where σ^2 is the "point" variance of the rain-rate field in this limiting situation. These equations can be easily solved for $w_i^{(0)}$ and λ with the help of the constraint $\Sigma_i w_i^{(0)} = n$, and one obtains

$$w_i^{(0)} = \alpha \frac{A_i}{A} \quad (\text{"simple weights"}), \qquad (6.5)$$

where A_i/A is the fraction of the grid-box area A viewed by the satellite instrument at time t_i, and $\alpha = n/(\Sigma_i A_i/A)$ is a normalization factor.

We refer to these as the simple weights since they are a plausible choice for averaging the satellite data in the absence of any information about the statistics of the rain within the grid box. It is easy to show that if the satellite instrument footprints are uniformly distributed within the swath and the monthly average of the satellite data is estimated using a simple arithmetic average of all the footprint estimates that fall within the grid box during the month, the weighting (6.5) is exactly equivalent to such an average. This kind of averaging of satellite data is commonly used in gridding satellite estimates of rainfall, although satellite footprints are sometimes nonuniformly distributed in the swath.

Another possible approach to the problem of adding up the satellite data to obtain gridded monthly averages might be to treat the satellite observations as a time sequence of estimates and to attempt to approximate the time integral of rain rate (2.1) as a kind of numerical integration using the time sequence of observations $\hat{R}_i$. A possible candidate might be

$$\hat{\bar{R}} = \frac{1}{T}\sum_i \frac{\Delta t_{i-1} + \Delta t_i}{2}\hat{R}_i,$$

with $\Delta t_i = t_{i+1} - t_i$, $\Delta t_0 = \Delta t_n = 0$. From this perspective, observations that are close together (small Δt_i) would get less weighting than widely spaced observations (large Δt_i). This approach would be plausible when the satellite observations are frequent enough that rain rates change very little from one observation to the next and if most of the grid-box area is observed each time. This is unfortunately not the case for low earth-orbiting satellites carrying the more accurate microwave instruments, and so weightings that attempt to implement a kind of "time integration" will generally perform much worse than the simple weighting.

To investigate the behavior of the optimal weighting scheme, optimal weights were calculated for 512 km × 512 km grid-box monthly averages using the GATE statistics. The satellite observation pattern is what is expected for TRMM as discussed in section 2, and the swath width is assumed to be 760 km (side to side), corresponding to the TRMM microwave instrument (TMI). Optimal weights for boxes centered at latitudes 0° and 30° are shown in Figs. 11 for one month of observations. Also shown is the straight line the weights would fall on if the simple weighting (6.5) were optimal (i.e., if there were no correlations).

These plots give the impression that there is a unique optimal weighting associated with each box fraction

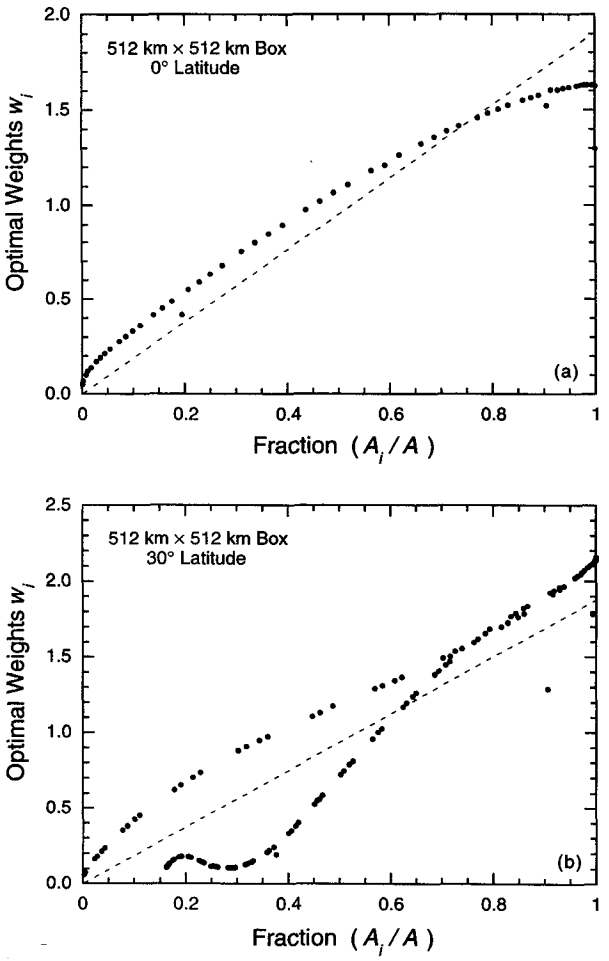


FIG. 11. Optimal weights calculated for 512-km grid box centered at latitudes (a) 0° and (b) 30° for one month of observations by the TRMM microwave instrument. There are 58 observations in (a) and 134 observations in (b). Here A_i/A is the fraction of the area observed by the satellite. The optimal weights would lie along the straight lines if there were no correlations.

seen by the satellite. There is a good reason for this. The optimal weighting assigned to each observation depends on the overall sequence of observations in which it is embedded. Since correlations are quite small for time separations greater than 15 or 20 hours, most of the weights are determined almost entirely by the character of nearby observations and not by their placement relative to the beginning and end of the month. The orbit is highly predictable, and if the longitude at which the satellite crosses the equator (moving northward) is known, the fraction of the grid box that will be seen during the orbit can be predicted from the orbit of the satellite and the instrument scanning pattern. Moreover, the equatorial crossing points of the satellite are themselves highly predictable; they move westward 23.3° with each orbit in the orbital approximation used here. It can thus be seen that each fraction observed is associated with a unique, predictable pattern of observations before and after it. The TRMM satellite samples areas with a "typical sequence" occurring many times during a month. This is the reason for the simple pattern emerging in the plots of the weights. The more complex pattern occurring for the 30° latitude box can be "unwrapped" by plotting the weights as a function of the equatorial crossing longitude associated with each observation (not shown).

By examining the observational sequence in detail, it is possible to understand some of the qualitative behavior seen for the weights. There are a few "outliers" visible in these plots, for example. These correspond to the observations at the beginning and end of the month. Each observation $\hat{R}_i$ can be considered to have a certain information content concerning the monthly total rainfall. Observations at the beginning and end of the month contain unneeded information about what happened before and after the month (due to time correlation) and, possessing this "excess" information, tend to be weighted less than observations in the interior of the month. The reason for the two "branches" in Fig. 11b for the 30° latitude box can likewise be understood. Referring to Fig. 1b, it can be seen that satellite observations at this latitude occur in "bursts" of four or five at slightly less than 24-h intervals. It is found that the lower branch corresponds to the weightings of the two or three interior observations in the bursts. The upper branch with higher weights corresponds to the two outermost observations in the bursts, whose information is less redundant, because they are bordered by 18-h gaps.

The optimal weights calculated for the cases in which only a tiny fraction A_i/A of the area is seen are always substantially larger than the simple weights. This is because of the spatial correlation. If rain is seen in a small corner of the area A, it can be inferred that there is probably rain occurring in nearby regions. The optimal weighting takes advantage of this by giving these observations extra weight. The effect, proportionally speaking, is largest for the very smallest observed areas.

How much difference does the optimal weighting make? The weights for the 0° latitude box (Fig. 11a), for example, are not very different from the simple weights, except for small A_i/A, which do not contribute much to the averages. As one measure of the impact of the choice of weights, the sampling errors computed from (2.8) for the optimal and simple weighting can be compared. Figure 12 shows the sampling errors, expressed as the ratio of the rms error to the mean rain rate $\langle R \rangle = 0.445$ mm h^{-1} for GATE Phase I. Generally speaking, the degree of improvement of optimal weighting over the simple weighting is not large. The simple weighting is near to being optimal for TRMM data. Only for boxes near latitude 30° does the optimal weighting reduce sampling error significantly, amounting to a 15% reduction in the error variance σ_E^2.

There is a perhaps surprising amount of variation in the sampling error from one latitude to another. This is due to real changes in the satellite sampling with latitude. The sampling errors shown in Fig. 12 are somewhat larger than were estimated in Bell et al. (1990), largely due to the reduced swath width of the microwave instrument presently planned for the satellite.

7. Discussion of sampling error results

Except near latitude 30°, the amount of improvement in the TRMM grid-box estimates using optimal weighting of the data is small and (except near 30°) would not justify the effort of obtaining the optimal weights, in our opinion. The reason for the smallness of the reduction in sampling error is that the TRMM observations are generally too regularly spaced and/or are not very correlated. It is only when the observations be-

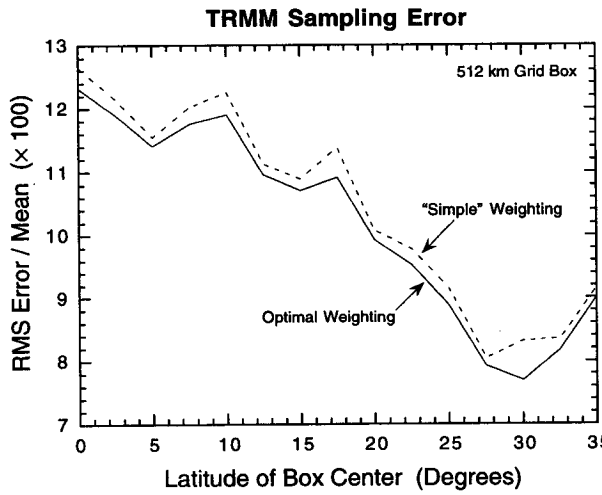


FIG. 12. Percent sampling error computed from modeled GATE statistics for optimal and simple averaging of the TRMM satellite observations.

come significantly correlated that the optimal weighting begins to differ from the simple weighting and acts to reduce the error.

It is anticipated that the TRMM measurements will eventually be used to calibrate the rain estimates from microwave instruments on other satellites and that the observations of several satellites can be combined to reduce sampling error in the gridded rain products. Such a scenario can be explored using the methods described here. As an example, the calculations in the previous section were done assuming that microwave estimates are available from two satellites, one TRMM-like and the other in a sun-synchronous orbit with the characteristics of the SSM/I (Special Sensor Microwave/Imager) on the DMSP (Defense Meteorological Satellite Program) *F11* satellite. The latter satellite was assumed to be in a circular orbit at 833-km altitude and 98.7° inclination, with average observation times of 0530 A.M. and 1730 P.M. local time and a swath width of 1400 km. The ratios $\sigma_E/\langle R\rangle$ for each satellite alone and for the combined data are given in Table 3.

Not surprisingly, the addition of the SSM/I observations reduces the sampling error. Optimal weighting is more effective in reducing sampling error for the combined data than for either satellite alone and lowers the error variance by 15% when applied to the combined data. As mentioned above, optimal weighting methods become more effective as the satellite observations become more correlated with each other and when the sampling pattern is irregular, as is the case here with the combination of two satellite observational patterns. The optimal weights are no longer simply related to the fraction of the grid box seen, as they were in Fig. 11. Since the pattern of time intervals between the observations of one satellite and those of the other satellite change during the course of the month, the optimal weights vary accordingly.

This optimal weighting calculation ignores the effect of the retrieval error terms in (2.12). When data from several satellites are combined and the sampling error becomes smaller, the justification for neglecting the retrieval error terms becomes weaker, especially when different instruments have different levels of retrieval error. When the observations of two instruments overlap, the more accurate instrument's estimates should be favored. The optimal weighting computed from (2.12), including retrieval error statistics, would accomplish this. However, the estimation of retrieval error statistics is not easy, especially if correlations in the retrieval errors become important, as discussed in Atlas and Bell (1992). Sampling error can be largely eliminated by addition of data from geosynchronous satellites, but the issue of adjusting the weights according to the retrieval errors and diurnal biases for the various satellite estimates becomes even more pressing.

The model calculations carried out here have simplified the statistical behavior of the observed rain by assuming that it is homogeneous and isotropic over the grid-box area. The temporal behavior of the model has also been simplified in that the presence of a diurnal cycle has been ignored, as well as lower-frequency phenomena such as is associated with the easterly waves with periods of 2 to 5 days that are strong in this part of the Tropics (see, for example, Thompson et al. 1979). The easterly wave activity can be clearly seen in the lagged correlations in the GATE data at lags of 2 to 5 days (not shown). How much do the results here depend on having ignored this longer-timescale variability? Since the satellite sampling is relatively good on the timescales of the easterly waves, it seems likely that the neglect of variability at these frequencies should not be so important, and in fact it is possible to show that the contributions from these frequencies to the optimal weighting and to sampling error tend to be much reduced compared to the impact of the higher frequencies, using arguments along the lines of North and Nakamoto (1989) and expressing the optimal weights as a filter acting on the rain-rate power spectrum.

We have experimented with introducing a 5-day wave somewhat artificially into the calculations. The optimal weights are changed very little, and the rms sampling error estimates increase by a factor of order 1.1. The increase can be understood as resulting from the negative correlations at lags of several days: sampling error is smallest when the rain at the sampled times is representative of the unsampled rain. Negative correlations imply that the sampled rain is in fact very different from the unsampled rain at certain unsampled times, and so introduction of wave activity at longer timescales while keeping the shorter-timescale correlations relatively unchanged leads to an increase in sampling error. Our sampling error estimates are thus slightly smaller than they would be if the longer-timescale statistics of the rain were better represented, but probably not by a significant amount, based on these calculations.

TABLE 3. SSM/I data combined with TMI data. Sampling errors for monthly averages over a 512-km box at the equator. Columns 2 and 3 give the rms sampling errors divided by the mean rain rate, 0.445 mm h^{-1}, for simple and optimal averaging.

	Percent error	
Satellite instruments	Simple average	Optimal average
TMI	12.5	12.2
SSM/I	10.8	10.7
TMI + SSM/I	8.3	7.6

8. A simple estimate of the sampling error

Because of the lack of correlation in rain from one TRMM observation to the next (at lower box latitudes), the sampling errors can be rather well estimated

by assuming that they decrease as the inverse of the square root of the sample size. This is also true of the error dependence on areal coverage since areas that are sufficiently well separated are uncorrelated, as is evidenced by the model behavior in which the variance of area-averaged rain rate decreases inversely with area [(A.14)] when the areas are large. As a measure of the sample size for a month of satellite observations, consider the quantity

$$S = \sum_i A_i/A. \tag{8.1}$$

It is the sum of the box fractions observed by the satellite during the month. If rain events in a large area and over a long time span are statistically somewhat similar and uncorrelated, we would expect the rms sampling error σ_E defined in (2.7) to depend on S as

$$\sigma_E \sim S^{-1/2}. \tag{8.2}$$

Figure 13 shows a plot versus S of the rms sampling error (with the simple weighting) for monthly averages of TMI data for a number of box locations and, in addition, the sampling errors calculated for the anticipated swath width (220 km) of the precipitation radar (PR) on the TRMM satellite for various box latitudes. A fit to an inverse square root behavior is superimposed. It can be seen that by simply calculating the satellite instrument's sampling volume S (8.1) one can estimate the sampling error to within a percentage point or so using (8.2), provided one has representative sampling-error calculations to extrapolate from. It should be noted that the coefficient of the fit indicated in Fig. 13 depends on the size of the grid box and to some extent on the TRMM orbit. It depends a great deal on the rainfall statistics. The possibility that this dependence can be made explicit will be discussed next.

The error estimates here have all been derived assuming the observed rain is statistically similar to the rain measured in GATE. To estimate sampling error for each grid box in a gridded global dataset, account will have to be taken of the changes in the statistics with geographical location and season. In Bell et al. (1990), a formula for sampling error was derived for regularly spaced observations at intervals Δt sufficiently large that the observations are uncorrelated, which, in the notation of this paper, can be written

$$\frac{\sigma_E}{\langle R\rangle} \approx \frac{1}{p^{1/2}} \frac{\sigma_a(r>0)}{\langle R_a(r>0)\rangle} \left(\frac{L_{\text{corr}}^2}{A}\right)^{1/2} \left(\frac{\Delta t}{T}\right)^{1/2}. \tag{8.3}$$

Here p is the probability of rain occurring in the instrument-footprint area a; $\sigma_a(r>0)$ is the rms rain rate in area a, where the average is over events with nonzero rain; $\langle R_a(r>0)\rangle$ is likewise the average rain rate in the area a conditional on nonzero rain; and L_{corr} may be thought of as being the typical spatial dimension of a rainy event. It is a kind of correlation length associated with the spatial correlation function

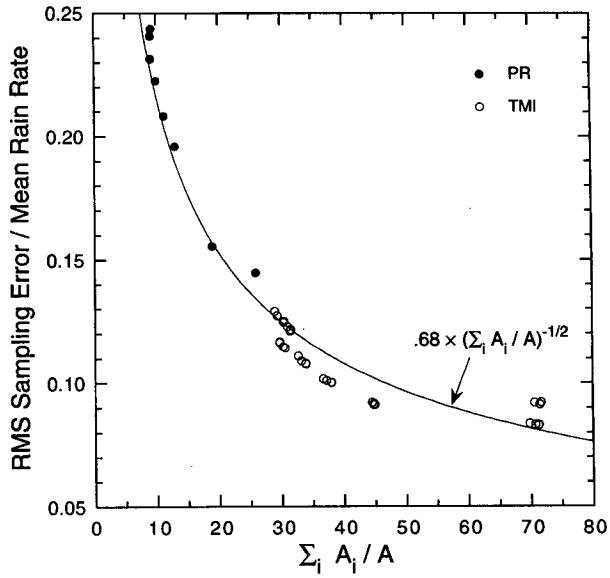


FIG. 13. Sampling error for 512-km boxes at latitudes 0°, 5°, . . . , 35°, and at longitudes 0°, 30°, 60°, and 90°, viewed by the TRMM microwave instrument (TMI), and boxes at the same latitudes and 0° longitude viewed by the TRMM radar (PR), plotted versus the sum of the box fractions observed by the instruments during the month. Simple averaging is assumed. The smooth curve shows a least squares fit to an inverse square root dependence of the error on sample size.

$\Phi_{ab}(s, 0)$ defined in (5.11). Formula (8.3) suggests an approximate method of extrapolating the GATE-based sampling error estimates to other climatologies.

The quantity $T/\Delta t$ in (8.3) is the number of samples taken during the period T (one month, in our case). It can be generalized to the quantity S in (8.1). Short et al. (1993) suggest that the ratio $\sigma_a(r>0)/\langle R_a(r>0)\rangle$ is relatively constant over a range of areas a, averaging times, types of data (rain gauge or radar), and climates. Variations in area-averaged rain rate $\langle R\rangle$ can be largely accounted for by changes in the probability of rain p, as evidenced by the success of predicting rain amount in an area from the fraction of area covered by rain, which has been improved on by the introduction of threshold methods (e.g., Kedem et al. 1990; Atlas et al. 1990). Taking advantage of this fact and using $\langle R\rangle = p\langle R_a(r>0)\rangle$, we can recast (8.3) in the form

$$\frac{\sigma_E}{\langle R\rangle} \propto (\langle R\rangle AS)^{-1/2} \tag{8.4}$$

since the variation in the terms in (8.3) that are omitted in (8.4) are less important than the ones kept. Based on Fig. 13, we can extract the proportionality constant in (8.4) to write

$$\frac{\sigma_E}{\langle R\rangle} \approx 0.68\left[\left(\frac{\langle R\rangle}{0.445 \text{ mm h}^{-1}}\right)\left(\frac{A}{(512 \text{ km})^2}\right)S\right]^{-1/2}, \tag{8.5}$$

where the numerical values 0.445 mm h^{-1} and (512 km)^2 are the mean rain rate and grid-box area for the points in Fig. 13. The numerical coefficient in (8.5) can vary slightly from one satellite to another because the sampling volume S is only an approximate measure of the effective coverage by the satellite. The coefficient changes to 0.66 for sampling by a single SSM/I instrument, for example.

The inverse square root dependence on the local mean rain rate in (8.4) is particularly striking. It can in fact be used to interpret the TRMM sampling error results of Oki and Sumi (1994). They sampled hourly rain-rate fields from radar data calibrated with the AMeDAS (Automatic Meteorological Data Acquisition System) rain gauge network over southern Japan using the predicted orbit of the TRMM satellite and assuming instrumental coverage over a 700-km swath. Monthly averages of rain over 5° × 5° grid boxes centered near latitude 33.5°N were obtained with the simulated satellite sampling and compared with the actual monthly averages to obtain estimates of the mean squared error of the sample estimates. Figure 14 shows the ratio $\sigma_E/\langle R\rangle$ for each of 12 months given in Table 3 of Oki and Sumi (1994). Each point is obtained from three to four years of data, five (overlapping) grid boxes at similar latitudes, and four staggered sampling patterns. The error bars represent 95% confidence limits based on the chi-squared distribution and an estimate of the number of independent experiments entering into the error calculation. The $\langle R\rangle^{-1/2}$ behavior, shown as a least-squares fit to the points, is consistent with this, although the size of the error bars does not place very stringent demands on the fit. A point corresponding to what GATE statistics at these latitudes would give (from Fig. 12) is shown on the same plot, scaled to take into account the differing swath width and grid-box size in their study.

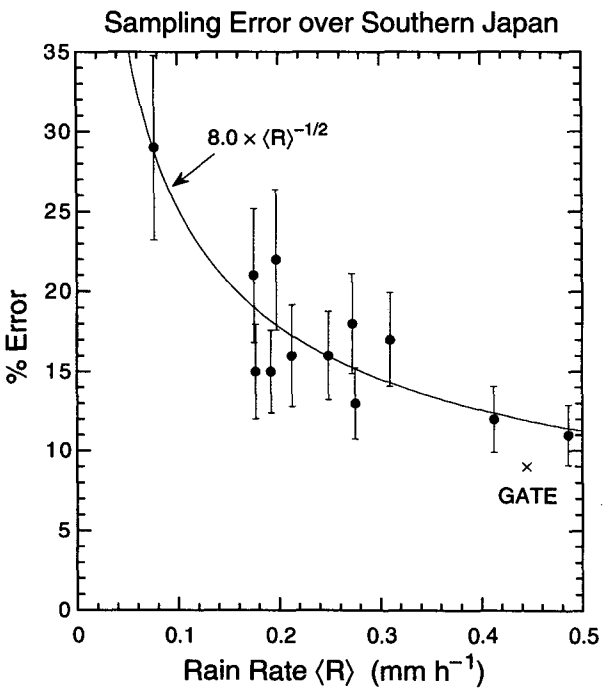


FIG. 14. Relative sampling error $\sigma_E/\langle R\rangle$, in percent, for a TRMM-like satellite viewing 5° × 5° grid boxes centered near latitude 33.5°N over southern Japan for each of 12 months (from Table 3 of Oki and Sumi 1994). Error bars are estimated 95% confidence limits. Also shown (×) is the sampling error calculated using GATE statistics (from Fig. 12), appropriately scaled.

Equation (8.4) can also be used to understand the sampling error results of Seed and Austin (1990). It appears to explain at least qualitatively the error distribution seen in the study by Weng et al. (1994) and seems to be remarkably successful in describing sampling errors inferred for gridded SSM/I results. This topic will be examined in greater detail in a subsequent paper.

9. Conclusions

The optimal weighting method investigated here offers a framework for combining satellite rainfall estimates into gridded products suitable for climatological studies. It can take advantage of prior knowledge of diurnal and seasonal changes in the statistics to correct for biases that might enter into averages of satellite data, and it provides a way to combine overlapping datasets taking into account their different levels of accuracy. To make use of the method, however, requires knowledge of the statistical characteristics of estimation errors for the various satellite instruments.

When data from a single satellite such as TRMM or SSM/I are averaged, the reduction in errors for the gridded product may not be big enough to warrant the extra labor involved in determining the optimal weights, except perhaps for grid boxes near the turnaround latitude of the TRMM satellite. The benefits of using the method become substantial when TRMM data are combined with one or more polar-orbiting satellites, however.

It is suggested that a good estimate of mean squared sampling error can be rather simply obtained by assuming that it is proportional to the average rain rate and inversely proportional to a measure of the sample size obtained by the satellite during the averaging period. This offers the chance that the gridded satellite product can provide the user with sampling error estimates for each grid point adjusted to the local sampling characteristics and climate.

Acknowledgments. One of the authors (TLB) would like to express his gratitude to the Center for Climate System Research of the University of Tokyo for the warm hospitality extended to him during his stay there. The helpful comments of two reviewers are gratefully acknowledged.

APPENDIX A

Details of Spectral Model Covariance Calculations

The variances and spatial and temporal correlations implied by the spectral model (4.6) will be derived here for box-averaged rain rates. Equation (5.10) for the covariance of the average rain rate in boxes $\mathcal{A}$ and $\mathcal{B}$, with a translation of the integration variables, can be written

$$\mathcal{C}_{\mathcal{AB}}(\mathbf{s}, \tau) = \frac{1}{L^4}\int_0^L dx_1 \int_0^L dx_2 \int_0^L dy_1 \int_0^L dy_2 c(\mathbf{s} + \mathbf{y} - \mathbf{x}, \tau). \quad \text{(A.1)}$$

Setting $\mathbf{u} = \mathbf{y} - \mathbf{x}$ and making use of the identity

$$\int_0^L dx \int_0^L dx' f(x - x') = \int_{-L}^L du(L - |u|)f(u), \quad \text{(A.2)}$$

expression (A.1) for $\mathcal{C}_{\mathcal{AB}}$ becomes

$$\mathcal{C}_{\mathcal{AB}}(\mathbf{s}, \tau) = \frac{1}{L^4}\int_{-L}^L du_1 \int_{-L}^L du_2 (L - |u_1|) \times (L - |u_2|)c(\mathbf{s} + \mathbf{u}, \tau). \quad \text{(A.3)}$$

For $\tau = 0$, substituting the explicit expression (5.4) into (A.3) and changing variables to $\boldsymbol{\xi} = \mathbf{u}/L$, one obtains

$$\mathcal{C}_{\mathcal{AB}}(\mathbf{s}, 0) = \gamma_0 \int_{-1}^1 d\xi_1 \int_{-1}^1 d\xi_2 (1 - |\xi_1|)(1 - |\xi_2|) \times C_\nu\left(|\boldsymbol{\xi} + \mathbf{s}/L| \frac{L}{L_0}\right). \quad \text{(A.4)}$$

These integrals were evaluated using the numerical integration package in *Mathematica* by Wolfram Research, Inc. (1988) for various separations $\mathbf{s}$.

For $\tau \neq 0$, it proved easier to return to the Fourier integral representation of the correlation [(5.1)]. Upon substituting this into (A.1) and carrying out the spatial integrals, one obtains

$$\mathcal{C}_{\mathcal{AB}}(\mathbf{s}, \tau) = (2\pi)^{-1} \int d^2\mathbf{k} e^{i\mathbf{k}\cdot\mathbf{s}} G^2\left(\frac{k_1 L}{2}\right) \times G^2\left(\frac{k_2 L}{2}\right) c(\mathbf{k}, \tau), \quad \text{(A.5)}$$

where $G^2(x) = (\sin x/x)^2$ is the Bartlett filter function familiar from time series analysis and $c(\mathbf{k}, \tau)$ is given in (5.2). In the isotropic case, (A.5) becomes

$$\mathcal{C}_{\mathcal{AB}}(s, \tau) = \frac{2}{\pi}\int_0^\infty \int_0^\infty dk_1 dk_2 \cos(k_1 s_1) \cos(k_2 s_2) \times G^2\left(\frac{k_1 L}{2}\right) G^2\left(\frac{k_2 L}{2}\right) c(k, \tau). \quad \text{(A.6)}$$

The lagged covariance of area-averaged rain rate ($s = 0$) has the still simpler form

$$\mathcal{C}_{\mathcal{AA}}(0, \tau) = \frac{2}{\pi}\int_0^\infty \int_0^\infty dk_1 dk_2 G^2\left(\frac{k_1 L}{2}\right) \times G^2\left(\frac{k_2 L}{2}\right) c(k, \tau) \quad \text{(A.7)}$$

$$= \sqrt{\frac{2}{\pi}} F_0 \int_0^\infty \int_0^\infty dk_1 dk_2 G^2\left(\frac{k_1 L}{2}\right) \times G^2\left(\frac{k_2 L}{2}\right) \tau_k e^{-|\tau|/\tau_k}. \quad \text{(A.8)}$$

Numerical integration was done using the IMSL (1991) routine QAND on a Cray Y-MP computer.

The integral autocorrelation time introduced in (5.13) can be calculated more directly from the inverse Fourier transform of (A.7), which can be written

$$\tilde{\mathcal{C}}_{\mathcal{A}}(\omega) = (2\pi)^{-1/2} \int_{-\infty}^\infty d\tau \mathcal{C}_{\mathcal{AA}}(0, \tau) e^{i\omega\tau} \quad \text{(A.9)}$$

$$= \frac{2}{\pi}\int_0^\infty dk_1 \int_0^\infty dk_2 G^2 \frac{k_1 L}{2} G^2\left(\frac{k_2 L}{2}\right) \times \tilde{c}(\mathbf{k}, \omega). \quad \text{(A.10)}$$

From the definition of $\tau_{\text{int}}(L)$ in (5.13) and (A.9) for $\omega = 0$, it follows that

$$\tau_{\text{int}}(L) = \sqrt{\frac{\pi}{2}} \frac{\tilde{\mathcal{C}}_{\mathcal{A}}(\omega = 0)}{\sigma_{\mathcal{A}}^2}. \quad \text{(A.11)}$$

We conclude with a few remarks about some asymptotic properties of the spectral model. As is suggested by Fig. 6, the variance of area-averaged rain rate $\sigma_{\mathcal{A}}^2$ in the spectral model diverges as the area size decreases. The divergence is dictated by the singular behavior of (5.7) noted for the point covariance. Substituting the first two terms in the expansion of (5.4) for small ρ into (A.4) gives the behavior (using numerical integration), for $L \to 0$,

$$\sigma_{\mathcal{A}}^2 \approx 16.80\, L^{-0.22} - 4.89 \text{ mm}^2 \text{ h}^{-2}, \quad \text{(A.12)}$$

which is already quite accurate at $L = 4$ km.

For large areas, however, we expect the variance to decrease as L^{-2}, since rain rates become rapidly uncorrelated beyond separations $s \gg L_0$. This can be verified from (A.8), since for large L the factors G are highly peaked at $k = 0$ and the integrals can be done neglecting large-k behavior to obtain

$$\sigma_{\mathcal{A}}^2 \approx 2\pi\Gamma(1 + \nu)\gamma_0 \frac{L_0^2}{L^2}, \quad (L \gg L_0). \quad \text{(A.13)}$$

Finally, for large lag τ, the integrals over k_1 and k_2 in (A.8) can be done neglecting large k behavior, which is strongly damped by the exponential term, to obtain

$$\mathcal{C}_{AB}(\mathbf{s}, \tau) \approx \frac{\Gamma(1+\nu)\gamma_0}{2(1+\nu)} \frac{e^{-|\tau|/\tau_0}}{|\tau|/\tau_0} \times \exp\left[\frac{-s^2}{4(1+\nu)L_0^2 \frac{|\tau|}{\tau_0}}\right], \quad \tau \gg \tau_0, \quad \text{(A.14)}$$

which reduces to (5.8) for $s = 0$. The above formula is valid for L/L_0 and $(s^2/L_0^2)\tau_0/\tau$ less than or of order 1.

APPENDIX B

Calculation of the Optimal Weights

The values of the matrix P_{ij} and vector Q_i needed for computing the optimal weights w_i from (6.1) are calculated using numerical integration on a grid to approximate the spatial integrals in (3.2) and (3.4). The expression (3.2) for the covariance matrix $P_{ij} = \langle R_i' R_j' \rangle$ is approximated by

$$P_{ij} \approx \frac{L^4}{A_i A_j} \sum_{a \in A_i} \sum_{b \in A_j} \mathcal{C}_{ab}(\mathbf{s}, |t_j - t_i|), \quad \text{(B.1)}$$

where $\mathcal{C}_{ab}(\mathbf{s}, |\tau|)$ is the covariance (5.10) of rain rate averaged over grid boxes a and b, with areas L^2 separated by distance s and lag τ and where the notation "$a \in A_i$" means that the sum is taken over those grid boxes a whose centers lie within the satellite-observed area A_i. Here, $\mathcal{C}_{ab}(\mathbf{s}, |\tau|)$ is approximated by the empirical formula (5.14).

Likewise, the vector $Q_i = \langle R_i' \bar{R}' \rangle$ given in (3.4) is approximated by

$$Q_i \approx \frac{L^4}{A_i A T} \sum_{a \in \mathcal{A}_i} \sum_{b \in \mathcal{A}} \int_0^T dt \mathcal{C}_{ab}(\mathbf{s}, |t - t_i|). \quad \text{(B.2)}$$

Substituting the empirical formula (5.14) for $\mathcal{C}_{ab}(\mathbf{s}, \tau)$ after rewriting the time integral as

$$\int_0^T dt \mathcal{C}_{ab}(\mathbf{s}, |t - t_i|) = \left(\int_0^{T-t_i} + \int_0^{t_i}\right) d\tau \mathcal{C}_{ab}(\mathbf{s}, \tau) \quad \text{(B.3)}$$

and carrying out the time integrals, one obtains

$$Q_i \approx \frac{L^4 \sigma_a^2}{A_i A T} \sum_{a \in A_i} \sum_{b \in A} \Phi_0(s) \frac{\tau(s)}{\mu(s)} \times \left\{\gamma\left[\frac{1}{\mu(s)}, \left(\frac{T - t_i}{\tau(s)}\right)^{\mu(s)}\right] + \gamma\left[\frac{1}{\mu(s)}, \left(\frac{t_i}{\tau(s)}\right)^{\mu(s)}\right]\right\}, \quad \text{(B.4)}$$

where $\gamma(\nu, x) \equiv \int_0^x t^{\nu-1} e^{-t} dt$ is the incomplete gamma function. The functions were computed numerically using the formulas given in Press et al. (1986, section 6.2).

The calculation of the mean-squared error σ_E^2 needed to obtain the rms errors plotted in Fig. 12 requires calculating $\langle (\bar{R}')^2 \rangle$ in addition to $\langle (\hat{\bar{R}}')^2 \rangle$ and the cross term $\langle \hat{\bar{R}}' \bar{R}' \rangle$, since from (2.8)

$$\sigma_E^2 = \langle (\hat{\bar{R}}')^2 \rangle - 2\langle \hat{\bar{R}}' \bar{R}' \rangle + \langle (\bar{R}')^2 \rangle = \sum_i \sum_j P_{ij} w_i w_j - 2 \sum_i Q_i w_i + \langle (\bar{R}')^2 \rangle. \quad \text{(B.5)}$$

The last term is given by

$$\langle (\bar{R}')^2 \rangle \approx \frac{L^4}{A^2 T^2} \sum_{a \in A} \sum_{b \in A} \int_0^T dt \int_0^T dt' \mathcal{C}_{ab}(\mathbf{s}, |t - t'|). \quad \text{(B.6)}$$

The double integral can be simplified using identity (A.2) to obtain

$$\langle (\bar{R}')^2 \rangle \approx \frac{2L^4}{A^2 T} \sum_{a \in A} \sum_{b \in A} \int_0^T d\tau (1 - \tau/T) \mathcal{C}_{ab}(\mathbf{s}, \tau). \quad \text{(B.7)}$$

After substituting empirical formula (5.14) for $\mathcal{C}_{ab}(\mathbf{s}, \tau)$ and expressing the time integral in terms of the incomplete gamma function, one obtains

$$\langle (\bar{R}')^2 \rangle \approx \frac{2L^4 \sigma_a^2}{A^2 T} \sum_{a \in A} \sum_{b \in A} \Phi_0(s) \frac{\tau(s)}{\mu(s)} \times \left\{\gamma\left[\frac{1}{\mu(s)}, \left(\frac{T}{\tau(s)}\right)^{\mu(s)}\right] - \frac{\tau(s)}{T} \gamma\left[\frac{2}{\mu(s)}, \left(\frac{T}{\tau(s)}\right)^{\mu(s)}\right]\right\}. \quad \text{(B.8)}$$

However, for the case $T = 1$ mo studied here, $T \gg \tau(s)$ for all s encountered in the sums, so that the incomplete gamma functions can be well approximated by the ordinary gamma function and (B.8) becomes

$$\langle (\bar{R}')^2 \rangle \approx \frac{2L^4 \sigma_a^2}{A^2 T} \sum_{a \in A} \sum_{b \in A} \Phi_0(s) \tau(s) \times \left\{\Gamma\left[1 + \frac{1}{\mu(s)}\right] - \frac{\tau(s)}{2T} \Gamma\left[1 + \frac{2}{\mu(s)}\right]\right\}. \quad \text{(B.9)}$$

For a given optimal weighting calculation, the satellite orbit is used to generate a description of the sequence of area shapes A_i and times t_i when the satellite instrument views the specified area A. The sums in (B.1), (B.4), and (B.9) are then done, taking advantage of the fact that the summands depend only on the spatial separation s of the grid-box locations of a and b to reduce as much as possible the number of times the empirical functions have to be evaluated.

REFERENCES

Abramowitz, M., and I. A. Stegun, 1970: *Handbook of Mathematical Functions,* Applied Mathematics Series, Vol. 55 Natl. Bur. of Stand., 1964. Dover reprint, 1046 pp.

Adler, R. F., A. J. Negri, P. R. Keehn, and I. M. Hakkarinen, 1993: Estimation of monthly rainfall over Japan and surrounding waters from a combination of low-orbit microwave and geosynchronous IR data. *J. Appl. Meteor.,* **32,** 335–356.

Atlas, D., and T. L. Bell, 1992: The relation of radar to cloud area-time integrals and implications for rain measurements from space. *Mon. Wea. Rev.,* **120,** 1997–2008.

——, D. Rosenfeld, and D. A. Short, 1990: The estimation of convective rainfall by area integrals. I. The theoretical and empirical basis. *J. Geophys. Res.,* **95D,** 2153–2160.

Batchelor, G. K., 1953: *The Theory of Homogeneous Turbulence.* Cambridge University Press, 197 pp.

Bell, T. L., and N. Reid, 1993: Detecting the diurnal cycle of rainfall using satellite observations. *J. Appl. Meteor.,* **32,** 311–322.

——, 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.,* **95D,** 2195–2205.

Brooks, D. R., 1977: An introduction to orbit dynamics and its application to satellite-based earth monitoring missions. NASA Ref. Pub. 1009, National Technical Information Service, Springfield, VA, 84 pp.

Cressie, N. A. C., 1993: *Statistics for Spatial Data.* rev. ed. John Wiley & Sons, 900 pp.

Gandin, L. S., 1963: *Objective Analysis of Meteorological Fields.* (Translated by Israel Program for Scientific Translations). Gidrometeorologicheskoe Izdatel'stvo (*GIMIZ*).

Hudlow, M. D., and V. L. Patterson, 1979: *GATE Radar Rainfall Atlas.* NOAA Special Report, U.S. Govt. Printing Office, Washington, DC, 158 pp.

IMSL, Inc., 1991: *User's Manual.* IMSL Math/Library Version 2.0, International Mathematical Statistical Libraries, 1372 pp.

Jenkins, G. M., and D. G. Watts, 1968: *Spectral Analysis and Its Applications.* Holden-Day, 525 pp.

Kedem, B., L. S. Chiu, and Z. Karni, 1990: An analysis of the threshold method for measuring area-average rainfall. *J. Appl. Meteor.,* **29,** 3–20.

Lau, K.-M., and P. H. Chan, 1986: The 40–50-day oscillation and the El Niño/Southern Oscillation: A new perspective. *Bull. Amer. Meteor. Soc.,* **67,** 533–534.

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 Publications, 441 pp. [Available from Goddard Space Flight Center, Greenbelt, MD 20771.]

Lovejoy, S., and B. B. Mandelbrot, 1985: Fractal properties of rain, and a fractal model. *Tellus,* **37A,** 209–232.

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

——, 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.

Peixoto, J. P., and A. H. Oort, 1992: *Physics of Climate.* American Institute of Physics, 520 pp.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1986: *Numerical Recipes, the Art of Scientific Computing.* Cambridge University Press, 818 pp.

Rasmusson, E. M., and J. M. Wallace, 1983: Meteorological aspects of the El Niño/Southern Oscillation. *Science,* **222,** 1195–1202.

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.,* **95D,** 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 raingage data. *J. Appl. Meteor.,* **32,** 1379–1387.

Simpson, J., R. F. Adler, and G. R. North, 1988: A proposed tropical rainfall measuring mission satellite. *Bull. Amer. Meteor. Soc.,* **69,** 278–295.

Thompson, R. M., Jr., S. W. Payne, E. E. Recker, and R. J. Reed, 1979: Structure and properties of synoptic-scale wave disturbances in the intertropical convergence zone of the eastern Atlantic. *J. Atmos. Sci.,* **36,** 53–72.

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.,* **99D,** 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., Deepak, 377–385.

Wolfram Research, Inc., 1988: *Mathematica.* Version 2.0, Wolfram Research, Inc., 961 pp.