Evaluation of Sampling Errors of Precipitation from Spaceborne and Ground Sensors

CHARLES E. GRAVES*

Climate System Research Program, Texas A&M University, College Station, Texas

JUAN B. VALDÉS

Department of Civil Engineering and Climate System Research Program, Texas A&M University, College Station, Texas

SAMUEL S. P. SHEN

Department of Mathematics, University of Alberta, Edmonton, Alberta, Canada

GERALD R. NORTH

Climate System Research Program, Texas A&M University, College Station, Texas

(Manuscript received 30 August 1991, in final form 15 January 1992)

ABSTRACT

The spatial and temporal characteristics of rainfall over Oklahoma and Kansas are analyzed in this paper using the raingage data collected during the Preliminary Regional Experiment for STORM-Central (PRE-STORM). The autocorrelation function and the spectrum are obtained directly from both processing the raingage data and using a theoretical stochastic model of space–time precipitation. This theoretical model serves as an intermediate step to obtain more information from the raingage records. The spectra obtained are then compared with those obtained from oceanic precipitation in the GARP (Global Atmospheric Research Program) Atlantic Tropical Experiment (GATE) and with that obtained from analyzing raingage records in east Texas. Finally, the spectra are used to evaluate the sampling errors that are due to the spatial gaps in measurements. The sampling error is expressed as an integral over the product of the spectral density of the stochastic rain field and a filter function. This filter function solely depends on the space–time configuration of the measurement instruments. The values of the analytical and numerical results on the sampling error are obtained for ground, spaceborne, and combined sensors of precipitation for several aggregation levels in space and time and alternative spacing and visiting times. It was found that sampling errors of land precipitation are higher than those reported for ocean precipitation.

1. Introduction

Precipitation measurements on a global basis have a high priority in several national and international meteorological space plans. Spaceborne sensors are the main instruments that can provide this information (cf. Simpson et al. 1988). Our present program includes evaluation of the errors due to spatial and/or temporal gaps in measurements from space sensors and associated ground networks for the estimation of area-averaged time-lumped precipitation. The simplest approach to the quantification of these sampling errors appears to be that of North and Nakamoto (1989). Their approach requires the mathematical representation of the precipitation process or at least its space–time spectral density. An alternative approach to estimate these sampling errors is to use Monte Carlo methods (e.g., Bell et al. 1990). Other studies where the sampling errors of precipitation measurements for raingage networks are Huff (1970), Zawadski (1973), Rodríguez-Iturbe and Mejía (1975), Bras and Rodríguez-Iturbe (1976), Silverman et al. (1981), and Gabriel (1981). A summary of raingage-network design approaches is found in Bras and Rodríguez-Iturbe (1985).

North and Nakamoto (1989) established an analytical technique to estimate the mean-square sampling error of measurements of the rainfall process due to intermittence either in space (raingages) or in time (spaceborne sensors). In their work, the mean-square sampling error ϵ^2 is expressed as an integral over fre-

* Currently at Department of Earth and Atmospheric Sciences, St. Louis University, St. Louis, MO.

Corresponding author address: Juan B. Valdés, Dept. of Civil Engineering, Texas A&M University, College Station, TX 77843-3136.

quency f and two-dimensional wavenumber (ν_x, ν_y) with an integrand consisting of two factors: a design-dependent filter $|H(f, \nu_x, \nu_y)|^2$ and the frequency–wavenumber spectral density of the precipitation field $S(f, \nu_x, \nu_y)$. Namely,

$$\epsilon^2 = \iiint |H(f, \nu_x, \nu_y)|^2 S(f, \nu_x, \nu_y) df d\nu_x d\nu_y. \quad (1)$$

The appealing aspect of (1) is that the information about the sampling design and the properties of the rain field are clearly factored. Thus, the sampling error from a given measurement design (e.g., raingage network and satellite) is solely governed by the spectral characteristics of rainfall. It is interesting that this figure of merit, ϵ^2, only depends on the second moments of the rain field.

Our present analysis of rainfall spectra is motivated by its presence in (1). Likewise, our discussion of the spectra is directed toward the resultant sampling errors. This present study examines spectra from data in two distinct manners. On one hand, we examine the spectra directly from observational records. On the other hand, we examine them by way of a stochastic rainfall model tuned by the observational data. For our study there are only a few cases [like the GARP (Global Atmospheric Research Program) Atlantic Tropical Experiment (GATE)] where rainfall datasets exist in a convenient form with large areal coverage and high temporal and spatial resolutions. In these cases, standard spectral estimation techniques can be applied. To help compensate, we include the additional tool of the stochastic rainfall model developed by Waymire et al. (1984; hereafter the WGR model). It can extrapolate the missing information and provides an alternative way to estimate the spectrum. It should be acknowledged that the assumptions of the model limit the shape of the spectrum.

In this paper, the observational spectra are obtained from the GATE and from the PRE-STORM (Preliminary Regional Experiment for STORM-Central) experiments. In addition, the study examines the spectra from the WGR model tuned to both the GATE and PRE-STORM experiments and to raingage data from eastern Texas. Where possible, the observational and model-based spectra will be compared. One purpose for using both GATE and PRE-STORM experiments is to compare and contrast the differences in midcontinental precipitation to oceanic precipitation.

After a discussion of the data and spectral estimates, our attention turns to the stochastic rainfall model. An analytical expression for the spectrum in terms of the model parameters will be used along with values for the various parameters estimated from the data. Finally we conclude with a discussion of the sampling errors and their consequences for estimating large space–time averages of rainfall like those proposed for the Tropical Rainfall Measuring Mission (TRMM; Simpson et al. 1988).

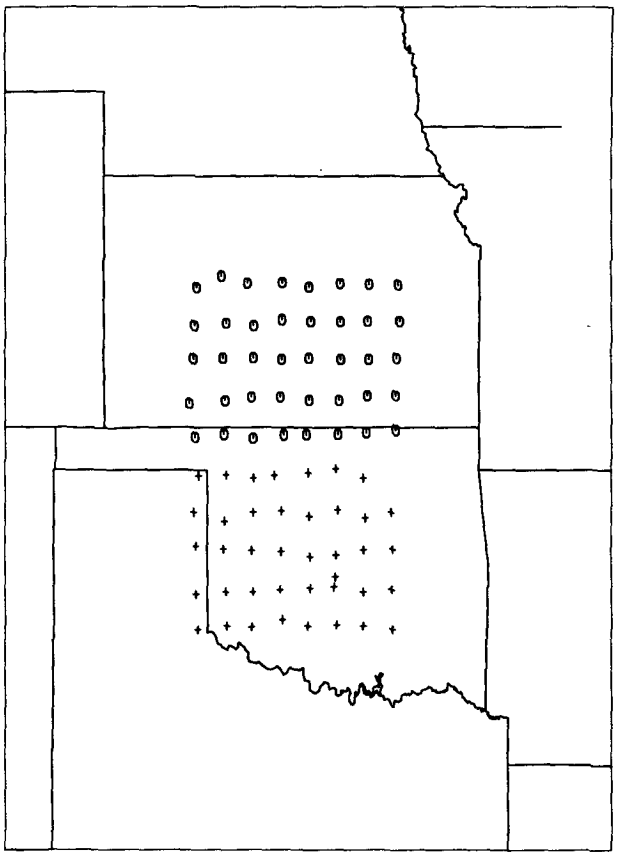

FIG. 1. The PRE-STORM area over the central United States. The circles indicate PAM locations and the crosses indicate the SAM sites.

2. The spectrum of PRE-STORM

PRE-STORM was a field experiment in 1985 designed to record meteorological information from convective activity over Kansas and Oklahoma. Figure 1 shows the PRE-STORM area and the mesonet sites that collected the rainfall data used in this study. These sites were placed nearly 50 km apart on an 10×8 grid. The circles in Fig. 1 denote the locations of the Portable Automated Mesonet (PAM) stations, while the crosses denote the locations of the Stationary Automated Mesonet (SAM) stations. Further details on the PRE-STORM experiment are discussed by Meitín and Cunning (1985).

Rainfall in 5-min aggregations were reported at each of the mesonet sites from 1 May 1985 through 27 June 1985. In this study, the records at each site were examined for missing and obviously spurious data. These missing and spurious data were replaced by linearly interpolated values. In nearly all cases, this resulted in setting these flagged points to zero. At most sites, less than 5% of the data were interpolated, although at eight locations about 15% of the data were interpolated.

For our study, the data were analyzed as a continuous space–time field. The space–time spectral density

was estimated separately for May and June from non-overlapping 15-min aggregated samples. The spectral densities were obtained by squaring the Fourier coefficients and smoothing in wavenumber and frequency space (see, for example, Bartlett 1978; Chatfield 1980). With a raingage spacing of 50 km, much of the spatial information of individual storms was lost and the spatial spectra represent only the largest spatial scales of rainfall. Additionally, some of the power at large wavenumbers was most likely aliased into the lower wavenumbers. As a result, these spectral densities can only provide a lower limit on the sampling errors of raingage networks because the power at large wavenumbers is a major contributor to these errors. However, not all space–time information was lost. As seen in the PRE-STORM spectra in Figs. 2 and 3, there is a concentration of power along a line indicating eastward propagation in both May and June. The eastward-propagating power is spread about a 25 km h^{-1} phase speed denoted in the figures by dashed lines. This phase speed is representative of the large-scale motion of the rainfall field and will be used in later sections involving the stochastic rainfall model (this propagation is also observed at the other north–south wavenumbers as well). In contrast, there is no apparent propagation in the north–south direction. By experience, the eastward drift is not unexpected since storm systems typically move from west to east across Kansas and Oklahoma.

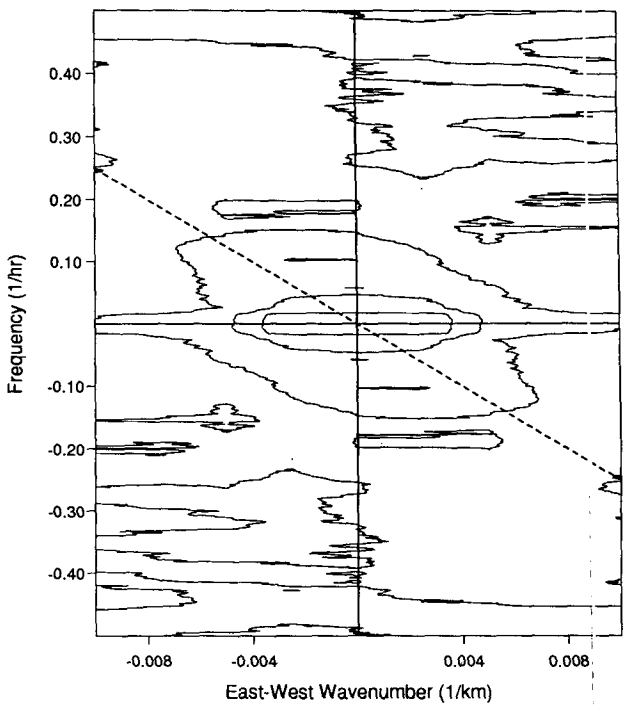

FIG. 3. Frequency–wavenumber spectrum of PRE-STORM rainfall for June. This portion of the spectrum is for the north–south space wavenumber zero (i.e., $\nu_y = 0$). The contour lines follow a logarithmic scale. The ratio between adjacent contour lines is $e^{-0.5}$.

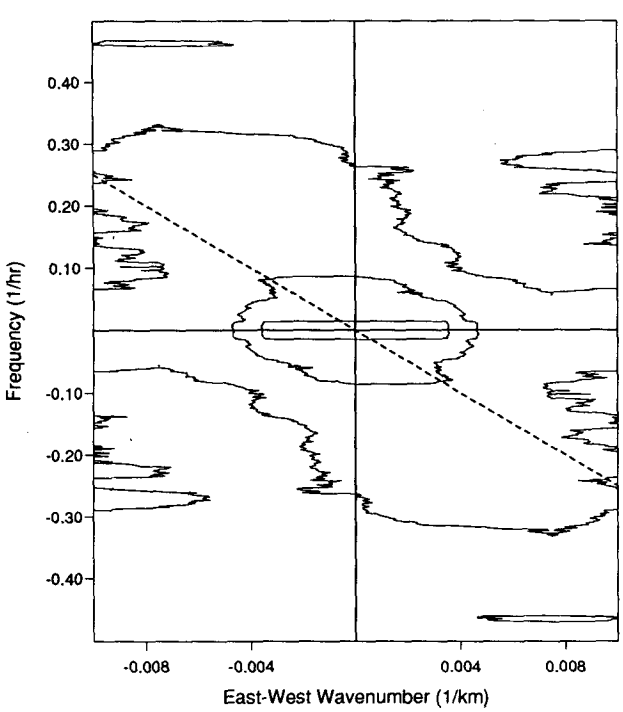

FIG. 2. Frequency–wavenumber spectrum of PRE-STORM rainfall for May. This portion of the spectrum is for the north–south space wavenumber zero (i.e., $\nu_y = 0$). The contour lines follow a logarithmic scale. The ratio between adjacent contour lines is $e^{-0.5}$.

Also observed in Figs. 2 and 3 is the general character of both spectral densities. Besides the eastward-propagating power, the spectra are red, decreasing power with increasing wavenumber and frequency. Furthermore, the spectral densities are quite noisy, a characteristic of the strong spatial and temporal intermittency of rainfall. The temporal spectral density with $\nu_x = 0$ and $\nu_y = 0$ (hereafter the temporal spectral density) for the PRE-STORM data is shown in Figs. 4 and 5. This plot contains the data along the $\nu_x = 0$ axis shown in Figs. 2 and 3. These figures reveal the rapid decrease in power out to time scales near 4 h. At higher frequencies, the power is more uniform, tending toward a white spectrum. The flat nature of the spectra at these frequencies may reflect the sampling characteristics of the sparse raingage network, where individual storms often appear as δ functions within the measurement network.

There is no strong peak in these spectral densities at the diurnal cycle. Two possible reasons may contribute to a lack of a spectral peak at this frequency. First of all, there is a phase progression across the central United States in the diurnal cycle (Wallace 1975). Thus, the spatial average contained in the temporal spectral densities (i.e., zero wavenumber component) would act to smear the peak at the diurnal cycle. Second, only a record of rainfall 57 days long is being analyzed, a relatively short time series considering the

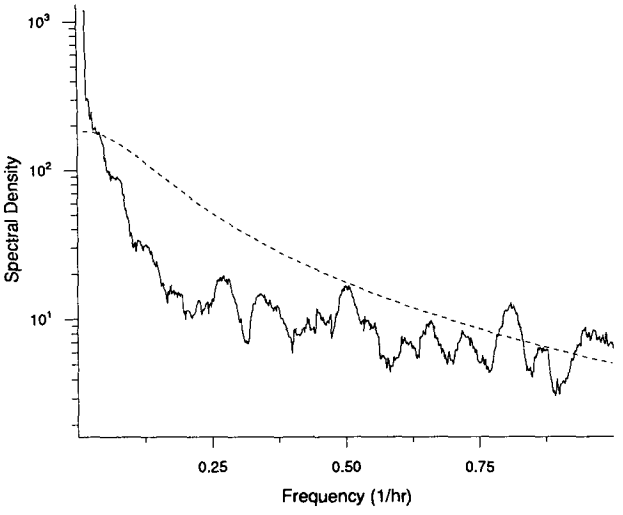

FIG. 4. Temporal spectrum of PRE-STORM rainfall for May. The dashed line indicates the spectra derived from the WGR model. Both spectra are normalized by their variance.

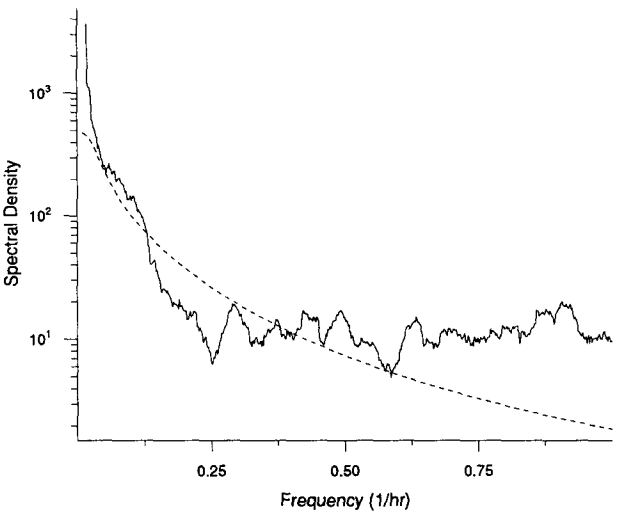

FIG. 5. Temporal spectrum of PRE-STORM rainfall for June. The dashed line indicates the spectra derived from the WGR model. Both spectra are normalized by their variance.

large fraction of time it does not rain. [It is suggested that 90 days are needed to extract a diurnal cycle (T. Bell, personal communication).]

Another spectrum derived directly from observations is from the GATE experiment and was first reported by Nakamoto et al. (1990) from the radar records of Arkell and Hudlow (1977). The temporal spectra of GATE are shown in Fig. 6 along with a portion of the PRE-STORM spectrum. While there are differences in the resolution and sampling characteristics between GATE and PRE-STORM, a comparison of the spectra provides a "quick look" into the possible satellite sampling-error differences between oceanic- and land-based rainfall. All the spectra in Fig. 6 have been normalized following Nakamoto et al. (1990) to highlight the differences in the shapes of the spectral density functions. Besides the peak near the diurnal cycle in one of the GATE spectra, both the GATE and PRE-STORM spectra follow the same general shape down to time scales near 1 h. Below this frequency, the GATE spectra drop off at a faster rate than the PRE-STORM spectra. The results of Nakamoto et al. (1990) indicated that at these higher frequencies the GATE spectra followed an f^{-4} dependence, although the PRE-STORM spectral densities from period of 2 days down to 4 h appear to follow an $f^{-5/3}$ dependence. This slope is consistent with the slope reported by Crane (1990) for similar frequencies obtained from radar measurements. The spatial smoothing of the GATE data (4-km averages) may account for some of the reduced variability at higher frequencies. In contrast, the sampling error of the sparse network of gauges in PRE-STORM may enhance the "white" character observed in the higher frequencies in Figs. 4 and 5 and may have also accounted for the absence of a diurnal cycle. The consequence of the sharper decay of the temporal spectra of GATE implies a smaller satellite sampling error (demonstrated in later sections).

3. Parameters and spectrum of WGR model

The WGR stochastic rainfall model incorporates many observed features of rainfall. A couple of features specifically included into the model are "rainbands" and the clustering of rainfall in both space and time. The model represents rain in a hierarchical approach, with rain cells embedded in cluster potential centers that are in turn embedded in rainbands. [Readers are referred to the original paper by Waymire et al. (1984) for further details.]

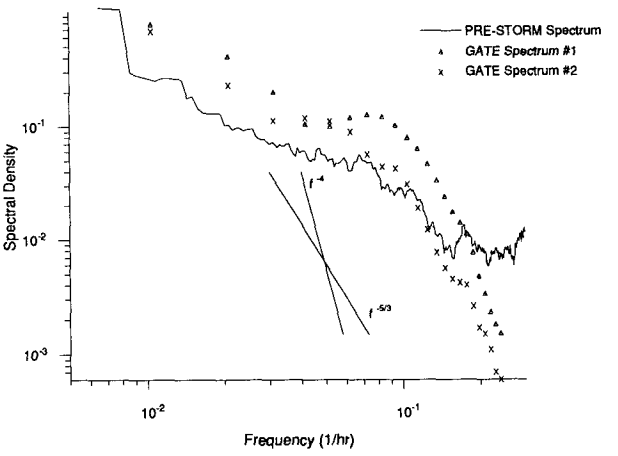

FIG. 6. A comparison of GATE and the PRE-STORM frequency spectra over the entire record. The GATE spectra has been adopted from Nakamoto et al. (1990). All spectra have been normalized following Nakamoto et al. The straight lines indicate possible slopes observed in these spectra.

TABLE 1. WGR-estimated parameters from PRE-STORM for May. The SSQ is a measure of the fit, μ is the mean, σ^2 is the variance, ρ_1 is the temporal lag-1 autocorrelation, and ρ_x is the simultaneous cross correlation of sites 50 km to the east.

Temporal aggregation (h)					Parameter						
0.5	1	2	3	6	λ (h^{-1})	ρ_L (h^{-1})	α (h^{-1})	β (h^{-1})	i_0 (mm h^{-1})	ν	SSQ
	μ, ρ_1 σ^2, ρ_x			ρ_1, σ^2	0.0191	0.0013	2.842	0.884	151.6	5.982	0.028
		μ, ρ_1 σ^2, ρ_x		ρ_1, σ^2	0.0191	0.0011	1.031	0.761	206.1	1.796	0.038
			μ, ρ_1 σ^2, ρ_x	ρ_1, σ^2	0.0292	0.0005	1.010	0.415	187.1	2.982	0.421
μ, ρ_1 σ^2, ρ_x		ρ_1, σ^2			0.0292	0.0010	1.196	1.106	130.3	2.982	0.064
μ, ρ_1 σ^2, ρ_x		ρ_1, σ^2			0.0292	0.0012	3.788	1.106	194.2	4.648	0.112

The parameters of the WGR model are defined as follows: λ_m^{-1} is the average time (h) between storms; ρ_L is the average number of cluster potential centers in a rainband per unit area (km^{-2}); i_0 is the maximum rainfall cell intensity at the moment of birth and at the center of the cell (mm h^{-1}); σ is the cluster spread factor (km); β is the cellular birth rate (h^{-1}); $E[\nu]$ is the average number of cells per cluster potential center; $\mathbf{u} = (u_x, u_y)$ is the vector of average storm velocities (km h^{-1}); α is the attenuation coefficient in time (h^{-1}); and D is the attenuation coefficient in space (km). The original formulation assumed no particular distribution for ν or i_0. In this version, ν is chosen to be a Poisson random variable, while i_0 is chosen to be a constant.

The parameters were estimated using a technique that is similar to the one described by Islam et al. (1988) by using temporally aggregated records. This numerical method finds the minimum of the sum of the squares of the normalized deviations, that is,

$$\min(\mathrm{SSQ}) = \sum_{n=1}^{M} \left[\frac{f_n(\mathbf{p})}{\theta_n} - 1 \right]^2,$$

where $f_n(\mathbf{p})$ is the WGR-derived nth statistic for the given set of the parameter $\mathbf{p}$, and θ_n represents the nth sample statistic obtained from the data. Typically the mean, variance, lag-1 autocorrelation, and simultaneous cross correlation at one temporal aggregation level are used along with the variance and lag-1 autocorrelation at another aggregation level. Only six of the nine parameters are typically obtained in this manner (i.e., $M = 6$); three parameters—$\mathbf{u}$, D, and σ—are determined from the meteorological conditions. As will be discussed later, prescribing D and σ will restrict our ability to examine the sampling error of raingage networks.

In the PRE-STORM study, D and σ were chosen to be 2.0 and 3.5 km, respectively, while the velocity was $u_x = 25$ km h^{-1} and $u_y = 0$ km h^{-1} (based on the spectra shown in Figs. 2 and 3). Tables 1 and 2 show the estimated WGR parameters for the PRE-STORM analysis. The smaller the value of SSQ, the better the fit of the parameters to the sample statistics. These parameters indicate a storm arrival time near 60 h (λ^{-1}) and a storm duration of 1–4 h (β^{-1}). Individual cells typically lasted for 20–30 min, while the initial intensity at the center of the cell was quite strong near 150 mm h^{-1}. The analysis in later sections of the PRE-

TABLE 2. WGR-estimated parameters from PRE-STORM for June. The SSQ is a measure of the fit, μ is the mean, σ^2 is the variance, ρ_1 is the temporal lag-1 autocorrelation, and ρ_x is the simultaneous cross correlation of sites 50 km to the east.

Temporal aggregation (h)					Parameter						
0.5	1	2	3	6	λ (h^{-1})	ρ_L (h^{-1})	α (h^{-1})	β (h^{-1})	i_0 (mm h^{-1})	ν	SSQ
	μ, ρ_1 σ^2, ρ_x			ρ_1, σ^2	0.0153	0.0015	2.409	0.263	146.9	10.007	0.091
		μ, ρ_1 σ^2, ρ_x		ρ_1, σ^2	0.0114	0.0015	0.992	0.229	149.2	4.987	0.036
			μ, ρ_1 σ^2, ρ_x	ρ_1, σ^2	0.01379	0.0015	0.875	0.209	146.9	3.963	0.0141
μ, ρ_1 σ^2, ρ_x		ρ_1, σ^2			0.0433	0.0019	2.677	2.106	154.0	2.722	0.296
μ, ρ_1 σ^2, ρ_x		ρ_1, σ^2			0.0375	0.0025	2.677	1.106	154.0	2.648	0.349

STORM data using the WGR model incorporates the values for the lowest SSQ for both months.

The analysis of the WGR model parameters for the GATE experiment was previously done by Valdés et al. (1990). Their estimation procedure relied on spatial statistics rather than temporal statistics as done here. Several examples of their estimated parameters are provided in Table 3. The comparison between the PRE-STORM and GATE parameters reveals the differences between rainfall over the tropical oceans and that over land as viewed by the WGR model. The major differences are found in parameters ρ_L and i_0, which indicate that the rain over land is more sparse, yet the intensity of each cell is greater than that over the tropical ocean. As will be discussed later, this result suggests that rainfall over land will have a larger sampling error mainly due to its larger spatial variability, but the normalized error will be only slightly larger.

Valdés et al. (1990) derived the analytical form of the frequency–wavenumber spectrum of the WGR model. This spectrum depends on all nine parameters of the model parameters, and its analytical expression is as follows:

$$S(f, \nu_x, \nu_y) = \theta_1 \frac{\alpha F(D, 0)}{\alpha^2 + \Theta^2} + \theta_2 \frac{2\alpha\beta(\beta^2 - \alpha^2)}{(\alpha^2 + 4\pi^2 f^2)(\beta^2 + 4\pi^2 f^2)} \delta(\nu_x)\delta(\nu_y) + \theta_3 \frac{\alpha\beta(\beta^2 - \alpha^2)}{(\alpha^2 + \Theta^2)(\beta^2 + \Theta^2)} \frac{F(D, \sigma)}{4\pi(D^2 + \sigma^2)}, \quad (2)$$

where

$$F(D, \sigma) = 8\pi(D^2 + \sigma^2) \times \exp[-4\pi^2(D^2 + \sigma^2)(\nu_x^2 + \nu_y^2)]$$

$$\Theta = 2\pi(\nu_x u_x + \nu_y u_y + f)$$

$$\theta_1 = \frac{\lambda_m E[\nu]\rho_L \pi D^2 i_0^2}{2\alpha}$$

$$\theta_2 = \frac{2\lambda_m \beta E[\nu]^2 \rho_L^2 \pi^2 D^4 i_0^2}{\alpha(\beta^2 - \alpha^2)}$$

$$\theta_3 = \frac{2\lambda_m \beta E[\nu(\nu - 1)]\rho_L \pi^2 D^4 i_0^2}{\alpha(\beta^2 - \alpha^2)},$$

and where $\delta(x)$ is the Dirac delta function, ν_x and ν_y are the spatial wavenumbers, and f is the temporal frequency. Here the spectrum is normalized such that

$$\sigma_R^2 = \iiint S(f, \nu_x, \nu_y) df d\nu_x d\nu_y = \theta_1 + (\beta - \alpha)\left[\theta_2 + \frac{\theta_3}{4\pi(D^2 + \sigma^2)}\right],$$

where σ_R^2 is the variance of the rainfall field at a point in space. Since the rainfall field is assumed to be homogeneous and stationary, this variance is a constant.

In (2), the temporal shape of the spectrum has both an f^{-2} and an f^{-4} dependence. This is a consequence of the exponential descriptions for the two temporal time scales in the WGR model. The spatial character of the spectrum is nearly Gaussian, although it is modified by an advection process (the advection term $\nu_x u_x + \nu_y u_y$ is found in Θ). These spectral shapes are relatively simple and can only represent the gross spectral character of rainfall. In Figs. 4 and 5, both the observed and WGR temporal spectra are shown for May and June 1985. In the figures these spectra were normalized so that $\int S(f, \nu_x, \nu_y)/\sigma_R^2 df d\nu_x \nu_y = 1.0$. The WGR spectra captures the general shape of the observed spectra, although it does highlight the sharp decay at low frequencies and weak decay at high frequencies in the observed spectra.

An explanation for these separate regions may simply be that individual storms are the source of high-frequency power that appears to be nearly random, and hence almost "white," while the background environment that supports the rainfall is the main source of the low-frequency power. The WGR model tries to capture this characteristic with two temporal scales α^{-1}

TABLE 3. WGR-estimated parameters for GATE (taken from Valdés et al. 1990). The SSQ is a measure of the fit, μ is the mean, σ^2 is the variance, ρ_1 is the temporal lag-1 autocorrelation, and ρ_x is the simultaneous cross correlation of neighboring grid points.

Spatial aggregation (km)				Parameter						
4 × 4	8 × 8	16 × 16	32 × 32	λ (h^{-1})	ρ_l (h^{-1})	α (h^{-1})	β (h^{-1})	i_0 (mm h^{-1})	ν	SSQ
μ, σ^2 ρ_1, ρ_x	σ^2, ρ_1			0.0128	0.0038	1.73	0.355	55.1	3.82	0.021
μ, σ^2 ρ_1, ρ_x		σ^2, ρ_1		0.0146	0.0024	1.32	0.185	77.8	9.44	0.063
μ, σ^2 ρ_1, ρ_x		σ^2, ρ_1		0.0131	0.0029	1.32	0.228	78.8	8.66	0.179
	μ, σ^2 ρ_1, ρ_x	σ^2, ρ_1		0.0235	0.0038	1.33	0.469	55.1	5.38	0.023
	μ, σ^2 ρ_1, ρ_x		σ^2, ρ_1	0.0235	0.0060	1.32	0.639	57.0	3.30	0.116

and β^{-1}. The inability of the WGR model to capture this characteristic may possibly be a consequence of the simple exponential temporal dependence for the storm attenuation and birth rate of cells.

4. Sampling errors for the WGR model

As previously mentioned, North and Nakamoto (1989) provide a formalism for evaluating the sampling error from a variety of measuring designs. The two components to the sampling errors are the rainfall spectrum, $S(f, \nu_x, \nu_y)$, and the design-specific filter $H(f, \nu_x, \nu_y)$, as shown in (1). Specifically, for the case of a satellite sampling every Δt and viewing a complete box of area $L \times L$ over time T, the design-dependent filter is

$$H(f, \nu_x, \nu_y) = G(\nu_x L)G(\nu_y L)G(fT)\left[1 - \frac{1}{G(f\Delta t)}\right], \quad (3)$$

where

$$G(x) = \frac{\sin(\pi x)}{\pi x}. \quad (4)$$

Using the formulas in (2) and (3), the approximate mean-square sampling error for a satellite-based sensor can be shown to be (see the Appendix for details)

$$\epsilon_s^2 = \frac{1}{L^2 T}\left\langle \frac{8\pi\theta_1 D^2}{\alpha}\left[\frac{\alpha\Delta t}{2}\coth\left(\frac{\alpha\Delta t}{2}\right) - 1\right] + 2(\theta_2 L^2 + \theta_3)\left\{\frac{\Delta t}{2}\left[\beta\coth\left(\frac{\alpha\Delta t}{2}\right) - \alpha\coth\left(\frac{\beta\Delta t}{2}\right)\right] + \frac{\alpha^2 - \beta^2}{\alpha\beta}\right\}\right\rangle, \quad (5)$$

where $N\Delta t = T$. The approximations require large sampling areas and long times; the larger the sampling area and the longer the sampling time, the more accurate the approximations. For this study, the sampling area is sufficiently large and the sampling time is sufficiently long so that (5) is an acceptable approximation based upon the accuracy test of (5) given in the Appendix.

The design-dependent filter of a network of raingages equally spaced in both axes at a distance Δl for the same space–time volume is

$$H(f, \nu_x, \nu_y) = G(fT)G(\nu_x L)G(\nu_y L) \times \left[1 - \frac{1}{G(\nu_x\Delta l)G(\nu_y\Delta l)}\right]. \quad (6)$$

In the limit of a large number of gauges ($L/\Delta l \gg 1$), the sampling error in the WGR model can be written as (again see the Appendix for details)

$$\epsilon_g^2 = \frac{4}{TL^2}\left\{8\pi\theta_1 D^2\alpha\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{\exp[-4\pi^2 D^2(n^2 + m^2)/\Delta l^2]}{[\alpha^2 + 4\pi^2(u_x n + u_y m)]/\Delta l^2} + 2\theta_3\alpha\beta(\beta^2 - \alpha^2)\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{\exp[-4\pi^2(D^2 + \sigma^2)(n^2 + m^2)/\Delta l^2]}{[\alpha^2 + 4\pi^2(u_x n + u_y m)^2/\Delta l^2][\beta^2 + 4\pi^2(u_x n + u_y m)^2/\Delta l^2]}\right\}. \quad (7)$$

In the above formula, there is no term involving the quantity θ_2. This is because the θ_2 term represents the zero wavenumber component to the autocorrelation of the rainfall field. However, there is a dependence associated with terms accompanying θ_2, thus θ_2 is found in the sampling error for satellites. This is explicitly shown in (2). Hence there is no contribution from this term in the raingage sampling error.

There are several cases in the current study in which the above formulation is not appropriate. The results in those situations are obtained by numerical integration. However, the above expression reveals the strongest dependence of the sampling error on the WGR parameters and will be helpful in later discussions.

It is beneficial to evaluate the sampling error when both sensors, spaceborne and ground, are available. With a combination of sensors, the estimate of the space–time-averaged rainfall may be chosen such to minimize the sampling error. Using the same sampling schemes adopted here, North et al. (1991) derived an analytical form for the minimum sampling error for a combined network of raingages and spaceborne sensors. They found that the combined sampling error is approximately

$$\epsilon^2 \approx \frac{\epsilon_s^2\epsilon_g^2}{\epsilon_s^2 + \epsilon_g^2}. \quad (8)$$

5. Evaluation of sampling errors

The sampling errors were evaluated from both the observed and WGR-derived spectra for the PRE-STORM experiment. When appropriate, the analytical expressions from the previous section were used, otherwise the results were obtained by numerical integration. The sampling errors from GATE and eastern Texas raingage data were obtained from the corresponding WGR-derived spectra. In all cases, $L = 500$ km and $T = 30$ days, while for the satellite sensor, $\Delta t = 12$ h. These parameters are consistent with TRMM-like products. The sampling errors are also presented in percent of ϵ/σ_A, where σ_A^2 is the variance of area averages. Over the tropical oceans, this ratio is about

1.25, however, for PRE-STORM the ratio is larger, about 2.0.

a. Sampling errors for PRE-STORM

The sampling errors estimated from the PRE-STORM observational spectra are shown in Table 4. In the table, the results for May and June have been averaged together. The maximum number of raingages considered is only 25 (a 5 × 5 grid) because of the limited spatial resolution. With larger raingage networks, the design-dependent filter begins to vary significantly over the spectrum's bandwidth. This phenomenon affects the accuracy of the numerical integration. As mentioned before, the raingage sampling errors in Table 4 should only be considered as a lower bound, since all the variance at wavelength shorter than 100 km was missed (or aliased into longer wavelengths). To illustrate, a simple extrapolation of the spectra of the form $\exp[-(\nu_x + \nu_y)/\nu_0]$, with $\nu_0 = 0.05$ km^{-1}, was carried out. The resulting sampling errors were a factor of 2 larger than those reported in Table 4. Therefore, to obtain a more accurate estimate of the sampling errors requires a larger range of wavenumbers than that available in the PRE-STORM experiment.

The results in Table 4 also suggest that for our particular sampling scheme, the satellite sampling errors are near 9%. Due to the high temporal resolution of the PRE-STORM data, this estimate is considered to be much more accurate than the raingage error estimates. Temporal variability is the major factor contributing to the satellite sampling errors. This result is surprisingly small and only slightly larger than the 5% error found in a similar analysis with GATE data (Nakamoto et al. 1990). Another way to examine this sampling error is to look at the area-averaged autocorrelation function shown in Fig. 7. The figure reveals that the autocorrelation function has a long memory and the exponential decay of an autoregressive model is not an appropriate representation. Consequently, the time scale obtained from an autoregressive estimate is not representative. As an example, Fig. 7 also shows an autocorrelation function based on an autoregressive model with a 3-h time scale. A more appropriate time scale in this situation is the integral time scale determined as the integral of the autocorrelation function, namely,

$$\tau_0 = \int \rho(\tau) d\tau.$$

In this case, the integral time scale was found to be 5 h. This time scale is approximately one-half of that found in GATE (Nakamoto et al. 1990). This is consistent with the slightly larger error in PRE-STORM than found in GATE. This larger sampling error is still smaller than those reported by Seed and Austin (1990), who found sampling errors near 22% for the Patrick Air Force Base radar in Florida. The area considered by Seed and Austin, however, was 50% smaller than the PRE-STORM area. Part of this discrepancy can be attributed to the fact that with the increase in area comes a longer autocorrelation time and hence a smaller sampling error.

The sampling errors obtained from the WGR analysis are presented in Table 5, and they generally agree with the observational results. With satellite sampling error near 12%, the differences in the temporal spectral densities observed in Figs. 4 and 5 result in a slight increase in the sampling error. The smaller sampling error in the observations is most likely attributed to the long tail of the autocorrelation function that is not captured by the WGR model. This was also noticed in the WGR model when comparing to GATE data (Valdés et al. 1990).

The spatial spectrum in the WGR model depends mainly on D and σ [see (7)], thus, the accuracy of the raingage sampling errors is dependent on the representativeness of the chosen parameters. Since the sampling errors are consistent with the observational results, the

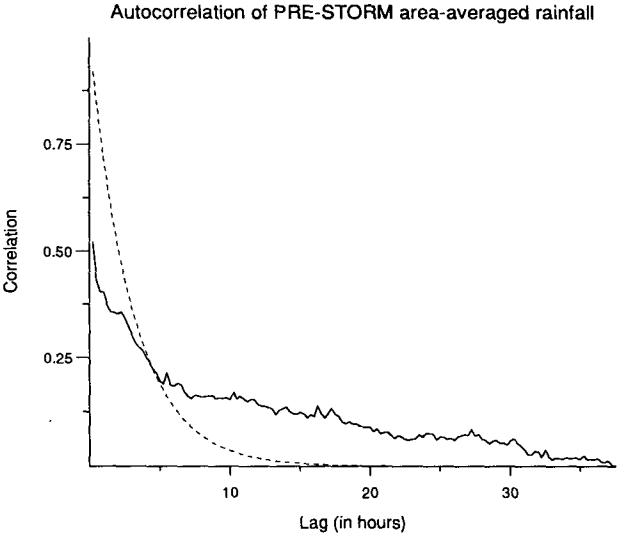

FIG. 7. Autocorrelation function of PRE-STORM rainfall over the entire record. The dashed line indicates an exponential autocorrelation function with a time scale of 3 h.

TABLE 4. Sampling errors for spaceborne, ground, and combined sensors for PRE-STORM data. The results for May and June have been averaged together.

Number of gauges	Δl (km)	Satellite (%)	Raingages (%)	Combined (%)
25	100	8.5	10.5	6.6
16	125	8.5	11.1	6.7
9	167	8.5	16.6	7.5
4	250	8.5	24.6	8.0
1	500	8.5	57.6	8.4

TABLE 5. Sampling errors for spaceborne, ground, and combined sensors based on the WGR model for the PRE-STORM data. Both May and June results have been averaged together.

Number of gauges	Δl (km)	Satellite (%)	Raingages (%)	Combined (%)
2500	10	11.7	0.1	0.1
625	20	11.7	1.1	1.0
277	30	11.7	2.3	2.3
100	50	11.7	5.5	5.0
44	75	11.7	10.1	7.7
25	100	11.7	15.1	9.2
21	108	11.7	16.7	9.6
4	250	11.7	47.2	11.3

WGR-derived errors appear to also be near the lower bound.

b. *Sampling errors for GATE*

The sampling errors from the WGR model for GATE are shown in Table 6. From the table it may be seen that raingages give better results than the satellite only for very dense networks, the latter being highly unlikely to be implemented in the tropics. For networks with a spacing of approximately 100 km or more between gauges, the satellite-based sensor will have smaller estimation errors. These results are higher than those obtained by North and Nakamoto, but similar to PRE-STORM, the WGR model was unable to capture the long autocorrelation times observed in the GATE data (see Valdés et al. 1988, 1990). It is of particular interest that even when the WGR model has shorter autocorrelation times than those found directly from GATE data, the satellite still does a better job in estimating the climatic monthly mean value of rainfall, and in that sense they are consistent with the findings of North and Nakamoto.

c. *Sampling errors for eastern Texas area*

The Brazos River valley in the vicinity of College Station, Texas, provides the setting for an another application described in this paper. Precipitation in this area has different characteristics than those observed in oceanic rainfall, but still no significant orographic effects were detected in the study by Koepsell and Valdés (1991).

Three raingages, with hourly precipitation values (Wheelock, Somerville Dam, and Washington-on-the-Brazos State Park) were used to estimate the parameters of the WGR model using time-averaged values. This area undergoes significantly different weather patterns throughout an annual cycle. During the spring, summer, and fall, maritime-tropical air masses control the climate of the region. During the winter and early spring, the climate is frequently affected by surges of cold polar Canadian air. These climate shifts result in cyclonic midlatitude rainfall events during the winter and airmass convective rainfall events during the summer. While there is a fluctuation of rainfall amounts throughout the annual cycle, there are no definitive rainy or dry seasons. This can be seen in the means, variances, and cross correlations for time-aggregated rainfall amounts. There are evidently two peak months of rainfall during the year, but each month contributes a comparable amount of rain to the annual total. Winter storms are usually long resulting in long autocorrelation times. These autocorrelation times fall off rapidly to low values during the July–August time period. In the above study, the precipitation records were divided in three periods: summer, winter, and a transition period. The parameters of the WGR model for the summer and the winter were used in our study to define the spectrum of the WGR model required to apply the North and Nakamoto technique.

From the results obtained for the two seasons (see Tables 7 and 8), it may be observed that the satellite sampling errors are around 10%. The gauges are more efficient during the winter season due to the large spatial and temporal correlations as compared with the patchy summer storms. For the evaluation of the climatic means for the summer season, raingages spaced around to 60 km or more will have larger errors than those obtained using the space sensor. For the winter season, the ground sensors are better than the space sensors for gauges spaced up to 200 km apart.

TABLE 6. Sampling errors for spaceborne, ground, and combined sensors derived from WGR parameters for GATE.

Number of gauges	Δl (km)	Satellite (%)	Raingages (%)	Combined (%)
2500	10	9.8	0.0	0.0
625	20	9.8	0.8	0.4
277	30	9.8	1.9	1.9
100	50	9.8	4.5	4.1
44	75	9.8	8.3	6.3
25	100	9.8	12.5	7.7
21	108	9.8	14.0	8.0
4	250	9.8	42.1	9.5

TABLE 7. Sampling errors for spaceborne, ground, and combined sensors for eastern Texas summer data.

Number of gauges	Δl (km)	Satellite (%)	Raingages (%)	Combined (%)
2500	10	10.9	0.2	0.2
625	20	10.9	1.5	1.4
277	30	10.9	3.3	3.1
100	50	10.9	7.6	6.2
44	75	10.9	13.6	8.5
25	100	10.9	19.9	9.5
21	108	10.9	21.9	9.7
4	250	10.9	59.7	10.7

TABLE 8. Sampling errors for spaceborne, ground, and combined sensors for eastern Texas winter data.

Number of gauges	Δl (km)	Satellite (%)	Raingages (%)	Combined (%)
2500	10	10.8	0.1	0.1
625	20	10.8	0.4	0.4
277	30	10.8	0.8	0.8
100	50	10.8	1.9	1.9
44	75	10.8	3.5	3.3
25	100	10.8	5.1	4.6
21	108	10.8	5.7	5.0
4	250	10.8	15.8	8.9

6. Discussion and final comments

The examination of the temporal spectra from PRE-STORM revealed a spectral shape that is not too dissimilar from that observed from GATE. The tail of the PRE-STORM spectra appear to have a lower slope than GATE. This result suggests that the largest difference between tropical oceanic rainfall and convective land precipitation is found at the higher frequencies. When examining the sampling errors, however, these high-frequency differences account for only a small increase in the sampling error. A more important factor in the sampling error is the larger total variability in the PRE-STORM rain field. The sampling strategy employed here requires the satellite to view the entire domain during each visit. This is not strictly the case, and Shin and North (1988) found that accounting for fractional visits increased the sampling error. Consequently in the PRE-STORM region, the satellite sampling errors appear to be large enough to warrant additional attention.

One possible way to decrease the satellite sampling error is to complement the satellite measurements with raingage measurements. The results from PRE-STORM analysis suggest that 25 gauges might be sufficient to reduce the sampling error to acceptable limits (<10%). On the other hand, the analysis of winter rainfall in eastern Texas indicates that only four raingages are required to obtain a similar sampling error, although the raingage sampling errors in our study should be considered as lower bounds.

The WGR analysis yielded satellite sampling errors that were larger than those obtained from the observations directly. This is apparently due to the long tail of the autocorrelation function that is not captured in the WGR model. A possible improvement to the WGR model is to have the ability to capture this feature. The success of the WGR model to replicate rainfall spectra is useful for estimating sampling errors over regions and periods when sufficient records do not exist for conventional techniques.

Finally, the accurate determination of the sampling errors involving a hybrid combination of designs will require more GATE-like products: large-scale, long-time datasets with high spatial and temporal resolution. Such a product maybe possible with the PRE-STORM products and we are currently starting an analysis of PRE-STORM radar records.

Acknowledgments. This research was partially funded by NASA Grants NAG-5-869 and NAGW-2427, NSERC/AES Grant 8380-1 (ARDG) of Canada, and the State of Texas Advanced Research Program Grant 160772. J. G. Meitín kindly provided the PRE-STORM data and assisted with their interpretation. All contributions are gratefully acknowledged.

APPENDIX

Derivation of Sampling Errors ϵ_s and ϵ_g

In section 4, analytical formulas for the sampling errors of satellite and raingage measurements were presented. The equations were derived from the spectrum of WGR model in section 3 and the corresponding filter functions in section (4). The formulas (5) and (7) resulted from an approximation of the integration of the product of the spectrum and the filter when L and T are large. There are many advantages in describing the sampling errors in simple analytical expressions rather than leaving them in the integral forms. First, the analytical expressions make the results very transparent. One can easily see what factors make important contributions to the sampling errors. Second, based upon the analytical expressions, the sampling errors can be computed efficiently without need of complicated numerical integrations. The details of the derivation of (5) and (7) are given below.

From (1) and (2), the error for satellite sampling can be written as

$$\epsilon_s^2 = \theta_1\alpha E_1 + \theta_2\alpha\beta(\beta^2 - \alpha^2)E_2 + \theta_3\alpha\beta\frac{(\beta^2 - \alpha^2)}{4\pi(D^2 + \sigma^2)}E_3, \quad \text{(A1)}$$

where

$$E_1 = \iiint \frac{F(D,0)}{\alpha^2 + \Theta^2} H_s^2(\nu_x, \nu_y, f)d\nu_x d\nu_y df, \quad \text{(A2)}$$

$$E_2 = \iiint \frac{\delta(\nu_x)\delta(\nu_y)}{(\alpha^2 + 4\pi^2 f^2\tau^2)(\beta^2 + 4\pi^2 f^2\tau^2)} \times H_s^2(\nu_x, \nu_y, f)d\nu_x d\nu_y df, \quad \text{(A3)}$$

$$E_3 = \iiint \frac{F(D,\sigma)}{(\alpha^2 + \Theta^2)(\beta^2 + \Theta^2)} \times H_s^2(\nu_x, \nu_y, f)d\nu_x d\nu_y df. \quad \text{(A4)}$$

In North and Nakamoto's paper (1989), it has been shown that

$$H_s^2(\nu_x, \nu_y, f) \sim \frac{1}{TL^2}\delta(\nu_x)\delta(\nu_y)\sum_{n\neq 0}\delta\left(f - \frac{n}{\Delta t}\right)$$

as $L \to \infty$ and $T \to \infty$. In our work, the asymptotic symbol "$\sim$" is replaced by the equal sign "=" since the convergence of the right-hand side was shown to be very fast. This convergence speed solely depends on the asymptotic behavior of $G(xL)$, which will be addressed below. This is the place where the approximation enters our computation. Notice that

$$\sum_{n=1}^{\infty} \frac{1}{1 + a^2 n^2} = \frac{1}{2}\left[\frac{\pi}{a}\coth\left(\frac{\pi}{a}\right) - 1\right]$$

where $a \neq 0$. With the above preparation, one can readily come out with the expression (5) for ϵ_s from (A1).

Similar approximation is applied to the case of raingages. Notice that

$$H_g^2(\nu_x, \nu_y, f) \sim \frac{1}{TL^2}\,\delta(f) \sum_{m,n\neq 0} \delta\left(\nu_x - \frac{m}{\Delta l}\right)\delta\left(\nu_y - \frac{n}{\Delta l}\right)$$

as $L \to \infty$ and $T \to \infty$. As was done for the satellite case, the asymptotic symbol "$\sim$" is replaced by the equal sign "=." With this approximation and the spectrum of the WGR model shown in section 3, the resulting expression (7) follows.

The sampling errors ϵ_s^2 and ϵ_g^2 are compared with the variance σ_A^2 of the spatial average. The quantity σ_A^2 yields the following expression:

$$\sigma_A^2 = \iiint G^2(\nu_x L) G^2(\nu_y L) P(\nu_x, \nu_y, f)\, d\nu_x d\nu_y df. \tag{A5}$$

Notice that

$$G^2(xL) \sim \frac{1}{L}\,\delta(x) \quad \text{as} \quad L \to \infty.$$

Here we would like to make a remark about the accuracy of the approximations. By the asymptotic expressions of $H_s^2(\nu_x, \nu_y, f)$ and $H_g^2(\nu_x, \nu_y, f)$ given in this section, the accuracy of the approximations (5) and (7) solely depend on the convergence speed of above asymptotic limit. To test the surprisingly fast convergence speed of the above limit, let us take $L = 50$. The following is given:

$$\int_{-0.2}^{0.2} 50G^2(50x)dx = 0.989873 \approx 1$$

$$\int_{-1.5}^{1.5} 50G^2(50x)dx = 0.998649 \approx 1$$

$$\int_{-2.0}^{2.0} 50G^2(50x)dx = 0.998987 \approx 1.$$

Thus, for any well-behaved function $f(x)$, the approximation

$$\int_{-\infty}^{\infty} 50G^2(50x)f(x)dx \approx \int_{-0.2}^{0.2} 50G^2(50x)f(x)dx \approx f(0)$$

contains a very small error. For practical sampling process, both T and L are greater than 50. Hence, we conclude that (5) and (7) are acceptable approximations.

Again using the spectrum given in section 3 immediately gives

$$\sigma_A^2 = \frac{1}{L^2}\left[4\pi D^2\theta_1 + (\beta - \alpha)\theta_3\right] + (\beta - \alpha)\theta_2. \tag{A6}$$

The expressions for the percentage errors are then as follows:

$$E_s = \left(\frac{\epsilon_s^2}{\sigma_A^2}\right)^{1/2} \times 100 \tag{A7}$$

$$E_g = \left(\frac{\epsilon_g^2}{\sigma_A^2}\right)^{1/2} \times 100. \tag{A8}$$

REFERENCES

Arkell, R., and M. Hudlow, 1977: GATE International Meteorological Radar Atlas. U.S. Department of Commerce, NOAA Publication, Washington D.C., 222 pp.

Bartlett, M. S., 1978: *An Introduction to Stochastic Processes.* Cambridge University Press, 388 pp.

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

Blackman, R. B., and J. W. Tukey, 1958: *The Measurement of Power Spectra.* Dover Publications, 190 pp.

Bras, R. L., and I. Rodríguez-Iturbe, 1976: Network design for the estimation of area mean of rainfall events. *Water Resour. Res.*, **12,** 1185–1195.

——, and ——, 1985: *Random Functions and Hydrology.* Addison-Wesley, 559 pp.

Chatfield, C., 1980: *The Analysis of Time Series: An Introduction.* Chapman and Hall, 268 pp.

Crane, R. K., 1990: Space–time structure of rain rate fields. *J. Geophys. Res.*, **95,** 2011–2020.

Gabriel, K. R., 1981: Gage density and variability of rainfall estimates: A complement to Silverman, Rogers and Dahl's study. *J. Appl. Meteor.*, **20,** 1537–1542.

Huff, F. A., 1970: Sampling errors in measurement of mean precipitation. *J. Appl. Meteor.*, **11,** 35–44.

Islam, S., R. L. Bras, and I. Rodríguez-Iturbe, 1988: Multidimensional modeling of cumulative rainfall: Parameter estimation and model adequacy through a continuum of scales. *Water Resour. Res.*, **24,** 992–995.

Koepsell, R. W., and J. B. Valdés, 1991: Multidimensional rainfall parameter estimation from sparse network. *ASCE J. Hydraulic Eng.*, **117,** 832–850.

Meitín, J., and Cunning, J., 1985: The Oklahoma–Kansas preliminary regional experiment for STORM-central. Volume I: daily operations summary. NOAA Tech. Memo. ERL ESG-20, 313 pp.

Nakamoto, S., J. B. Valdés, 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. Ocean. Technol.*, **6,** 985–992.

——, S. S. P. Shen, and R. Upson, 1991: Combining raingages with satellite measurements for optimal estimates of area-time averaged rainrates. *Water Resour. Res.*, **27,** 2785–2790.

Rodríguez-Iturbe, I., and J. M. Mejía, 1975: The design of rainfall networks in time and space. *Water Resour. Res.*, **10**(4), 713–728.

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.

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.

Silverman, B. A., L. K. Rogers, and D. Dahl, 1981: On the sampling variance of raingage networks. *J. Appl. Meteor.*, **20,** 1468–1478.

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.

Valdés, J. B., S. Nakamoto, S. S. P. Shen, and G. R. North, 1988: Satellite-based applications of the WGR model. *AGU/AMS Conf. on Mesoscale Precipitation: Analysis, Simulation and Forecasting,* Boston, Massachusetts Institute of Technology.

——, ——, ——, and ——, 1990: Estimation of multidimensional precipitation parameters by areal estimates of oceanic rainfall. *J. Geophys. Res.*, **95,** 2101–2111.

Wallace, L. M., 1975: Diurnal variations in precipitation and thunderstorm frequency over the conterminous United States. *Mon. Wea. Rev.*, **103,** 406–419.

Waymire, E., V. K. Gupta, and I. Rodríguez-Iturbe, 1984: Spectral theory of rainfall intensity at the meso-β scale. *Water Resour. Res.*, **20**(10), 1453–1465.

Zawadski, I. I., 1973: Errors and fluctuations of raingage estimates of areal rainfall. *J. Hydrol.*, **18,** 243–255.