Uncertainty Quantification of Mean-Areal Radar-Rainfall Estimates

EMMANOUIL N. ANAGNOSTOU* AND WITOLD F. KRAJEWSKI

Iowa Institute of Hydraulic Research, University of Iowa, Iowa City, Iowa

JAMES SMITH

Department of Civil Engineering and Operations Research, Princeton University, Princeton, New Jersey

(Manuscript received 30 July 1997, in final form 3 April 1998)

ABSTRACT

The most common rainfall measuring sensor for validation of radar-rainfall products is the rain gauge. However, the difference between area-rainfall and rain gauge point-rainfall estimates imposes additional noise in the radar–rain gauge difference statistics, which should not be interpreted as radar error. A methodology is proposed to quantify the radar-rainfall error variance by separating the variance of the rain gauge area-point rainfall difference from the variance of radar–rain gauge ratio. The error in this research is defined as the ratio of the ''true'' rainfall to the estimated mean-areal rainfall by radar and rain gauge. Both radar and rain gauge multiplicative errors are assumed to be stochastic variables, lognormally distributed, with zero covariance. The rain gauge area-point difference variance is quantified based on the areal-rainfall variance reduction factor evaluated in the logarithmic domain. The statistical method described here has two distinct characteristics: first, it proposes a range-dependent formulation for the error variance, and second, the error variance estimates are relative to the mean rainfall at the radar product grids. Two months of radar and rain gauge data from the Melbourne, Florida, WSR-88D are used to illustrate the proposed method. The study concentrates on hourly rainfall accumulations at 2- and 4-km grid resolutions. Results show that the area-point difference in rain gauge rainfall contributes up to 60% of the variance observed in radar–rain gauge differences, depending on the radar grid size, the location of the sampling point in the grid, and the distance from the radar.

1. Introduction

Because weather radar does not measure rainfall directly, algorithms are used to estimate rainfall from radar observations. It is widely recognized that radar-rainfall algorithms estimate rainfall with a high degree of uncertainty (e.g., Zawadzki 1984; Austin 1987; Joss and Waldvogel 1990). Some of the uncertainty sources include the storm-to-storm and within-the-same-storm variability of drop size distribution, the variations of reflectivity with height, the temporal and spatial radar sampling associated with nonlinear averaging of highly variable precipitation fields, the radar signal attenuation, and the radar hardware miscalibration and noise. Therefore, use of radar-derived rainfall products in hydrometeorological applications requires proper characterization of the estimation error.

Recent studies on radar-rainfall uncertainty concerned the mean-field bias (Anagnostou et al. 1997) and range-dependent, systematic, and random errors (Kitchen and Jackson 1993; Smith et al. 1996). The objective of this study is a quantification of the variance of radar-rainfall estimation error with focus on 4 km $\times$ 4 km maps of hourly accumulations. This has implications for operational flood forecasting and operation of water resource systems, optimal merging of rainfall products from different sensors, validation of quantitative precipitation forecasts using radar data, and assimilation of radar-rainfall estimates into numerical weather prediction models.

A common way of assessing the accuracy of radar-rainfall estimates is through comparisons with observations from rain gauge networks. Although rain gauge observations are considered to accurately represent rainfall at a point, their accuracy in representing mean-areal rainfall of the size of a radar pixel is low (Ciach and Krajewski 1999). This difference between mean-areal rainfall and point-rainfall rate is attributed to two combined effects of the small-scale variability of rainfall (Seed and Austin 1990) and the large difference between

* Current affiliation: Department of Civil and Environmental Engineering, University of Connecticut, Storrs, Connecticut.

Corresponding author address: Dr. Emmanouil N. Anagnostou, University of Connecticut, 261 Glenbrook Road, U-37, Storrs, CT 06269.
E-mail: manos@agnes.gsfc.nasa.gov

a radar and rain gauge sampling area (approximately eight orders of magnitude). The effect of spatial rainfall variability on the rain gauge sampling—and as a result on the radar-to-rain gauge comparisons—has long been recognized (e.g., Hendrick and Comer 1970; Zawadzki 1973, 1975; Krajewski 1987). Recent studies in the literature provide additional evidence that rain gauge rainfall can significantly deviate from the true mean-areal rainfall. Kitchen and Blackall (1992) show that the hourly area–point (AP) root-mean-square (rms) difference can be up to 150% with respect to the mean rainfall for convective storms in England. Ciach and Krajewski (1999) show that in the Tropics, the contribution of the AP rms difference to the radar-to-rain gauge (RG) rms difference is over 50% for hourly accumulations, and it can be significant even at the daily time scale ($\sim$20%). Therefore, one can easily realize that RG rms difference is not an actual representation of the difference between radar and true mean-areal rainfall (RT). To retrieve the variance of the desired error variable RT, the variance of the AP difference should be separated from the variance of the RG difference.

The concept of error separation in radar-rainfall estimation has been discussed in an empirical study by Barnston (1991), who applied it to the storm totals. Ciach and Krajewski (1999) developed a rigorous mathematical apparatus for the radar-rainfall error variance separation and provided an illustration of its application for tropical rainfall over a wide range of temporal scales (from 10 min to 48 h). Herein, we follow the same conceptual framework with several notable differences.

We define the RG and AP differences in a multiplicative sense, hereafter called G/R and A/P ratios. The advantage of our multiplicative (i.e., ratios) approach is that the error variance is directly related to the magnitude of rainfall (unlike in the case of additive errors, where the error variance applies uniformly to all rainfall values). Both G/R and A/P ratios are assumed to be random variables with stationary lognormal distributions. To analyze the variances of G/R and A/P ratios, natural logarithmic (log) transformation of the variables is applied. The assumption is made that over a long period the mean of the log-transformed variables is zero; this means that the radar-rainfall estimator and rain gauge measurements are unbiased. In our radar-rainfall uncertainty parameterization the $\log(G/R)$ variance is allowed to change with distance from the radar, while the $\log(A/P)$ variance is independent of radar range. This requires the assumption of statistical homogeneity of rainfall with range. A three-parameter power-law function is used to model the range-dependent $\log(G/R)$ variance. A least squares nonlinear regression, applied on long-term rain gauge and radar-rainfall data, is used to estimate the function's parameters. We assume spatial independence of the radar-rainfall estimation error. For the variance partitioning to apply, the assumption is made that $\log(G/R)$ and $\log(A/P)$ variables have zero covariance (Johnson and Wichern 1992). These assumptions are further discussed in subsequent sections and the conclusions.

The variance of the $\log(A/P)$ variable is estimated based on the mean-areal variance reduction factor proposed by Rodriguez-Iturbe and Mejia (1974). Another approach that is based on the variogram function derived from short lagged-distance rain gauge data is described in Kitchen and Blackall (1992). Other studies, related to rain gauge sampling variations due to differences between point measurements and areal rainfall have been reported by Huff (1970), Silverman et al. (1981), and Seed and Austin (1990), among others. Most of the past studies concerned large spatial (10–100 km) and temporal (storm totals to monthly accumulations) scales. In this research, we are interested in spatiotemporal scales that define the real-time radar-rainfall products (i.e., hourly rainfall accumulations and 2- to 4-km radar product grid sizes).

To apply variance partitioning our procedure requires the following: 1) log transformation of variables (ratios) and 2) the additional assumption that the expectation of log-transformed variables is similar to the logarithm of the variables' expectation. We will provide evidence that for hourly rainfall greater than 0.5 mm h^{-1} the effect of log transformation is insignificant. Due to the above assumptions and constraints, caution should be used in applying the results of this study to other spatiotemporal scales from the scales discussed in this paper.

An extensive dataset from the Melbourne, Florida, WSR-88D is used to illustrate our methodology. It spans a period of 59 days (3 August–30 September 1995) and consists of volume scan reflectivity observations and 1-min resolution rain gauge measurements, aggregated to hourly accumulations. The dataset includes several isolated convective storms and two tropical systems (Hurricane Erin and Tropical Storm Jerry). It was compiled by the National Aeronautic and Space Administration's (NASA's) Tropical Rainfall Measuring Mission Ground Validation Team for a special experiment organized by NASA. Information regarding the experiment, the full dataset statistics, and the data quality can be found in Krajewski et al. (1996).

The rain gauges used in this study are divided into two clusters according to their geographic location. The clusters were selected such that the rain gauges in each cluster have homogeneous long-term statistics (Court and Griffiths 1986). The first cluster is organized along the eastern coast of the Florida peninsula (St. John's rain gauge network), while the second cluster consists of rain gauges concentrated at the Cape Canaveral peninsula (Kennedy Flight Center network). Figure 1 shows a view of the Florida map with the radar location and the two rain gauge clusters. Rainfall statistics for the clusters are presented in Table 1.

Figure 2 shows the conditional ($>$0.5 mm h^{-1}) mean, variance, and probability of detection (POD) for hourly rain gauge accumulations. POD is the probability (in percent) that rain gauge–accumulated rainfall is greater

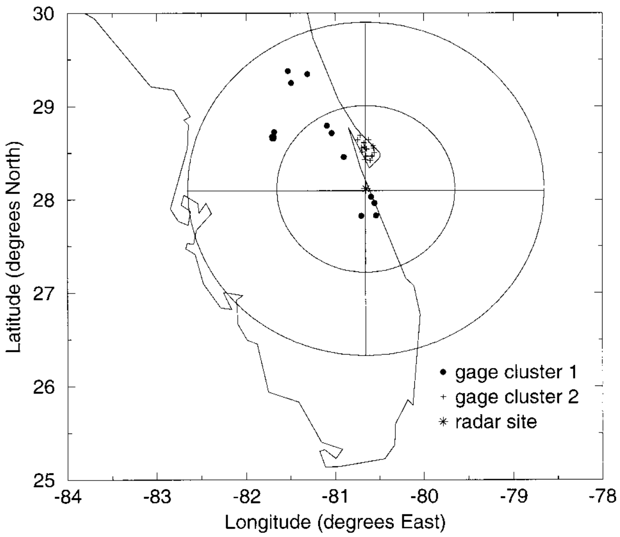

FIG. 1. Map of southern Florida showing the radar site and the locations of the rain gauge clusters. The inner and outer circles correspond to 100 and 200 km radar ranges, respectively.

TABLE 1. Rain gauge data sample statistics.

Rain gauge network	Cluster 1	Cluster 2
Number of gauges	17	15
Wet G/R pairs	2701	1967
Conditional mean (mm h^{-1})	3.02	3.90
Quantiles (0.1, 0.9) (mm h^{-1})	0.25, 8.37	0.25, 9.66

than a rainfall threshold given that the corresponding gridded radar-rainfall estimate is greater than this threshold. Rain gauge sample statistics are presented for each cluster separately. Each point in the plots represents a rain gauge. The average values of the clusters' statistics are shown with straight lines. One can notice systematic differences between the mean statistics of the two clusters. For example, the POD and mean of the first cluster is approximately 20% less with respect to the second cluster. The variances, though, seem to be similar in magnitude. Within the same cluster there are no obvious systematic changes in the above statistics. To avoid inhomogeneity effects in the rain gauge statistics, we decided to use only data from the first cluster of rain gauges. This cluster was chosen because it captures a larger range of radar subgrid-size intergauge distances (0.6–4 km) and radar ranges (10–170 km). The second cluster was used for an independent evaluation of the error variance model predictions.

The gridded radar-rainfall data are generated based on a real-time radar-rainfall estimation algorithm (Anagnostou and Krajewski 1999a). In this study the algorithm was used to convert volume scan reflectivity data to gridded hourly rainfall accumulations. The spatial resolution of the radar products are either the 2 km × 2 km Cartesian grid or the Hydrologic Rainfall Analysis Project (HRAP) grid plane (~4 km × 4 km). The algorithm uses a climatologically parameterized range correction formula to reduce the range effect in radar-rainfall products. The mean-field radar–rain gauge bias was evaluated every hour from corresponding radar–rain gauge observations, based on a statistical procedure. Details of the algorithm and evaluation of its performance can be found in Anagnostou and Krajewski (1999a,b).

This paper is organized as follows. In section 2 we present the definition of radar error and the concept of variance partitioning in the logarithmic domain. In sections 3 and 4 we discuss procedures for estimating the $\log(G/R)$ and $\log(A/P)$ variances, respectively. A case study based on two months of hourly radar and rain gauge–rainfall data is presented in section 5, and we close with conclusions and recommendations for further research in section 6.

2. Formulation of radar-rainfall error model

The uncertainty in radar-rainfall estimates is modeled based on the following formulation:

$$R_A(s, u) = [R_R(s, u)][\varepsilon_R(s, u)], \quad \text{if} \quad R_R(s, u) > r, \quad (1)$$

where R_A (mm) is the "true" mean-areal rainfall and R_R (mm) is the gridded radar-rainfall accumulation. The random variable $\varepsilon_R(s, u)$ represents the multiplicative radar-rainfall error. Variables u and s are indices of the

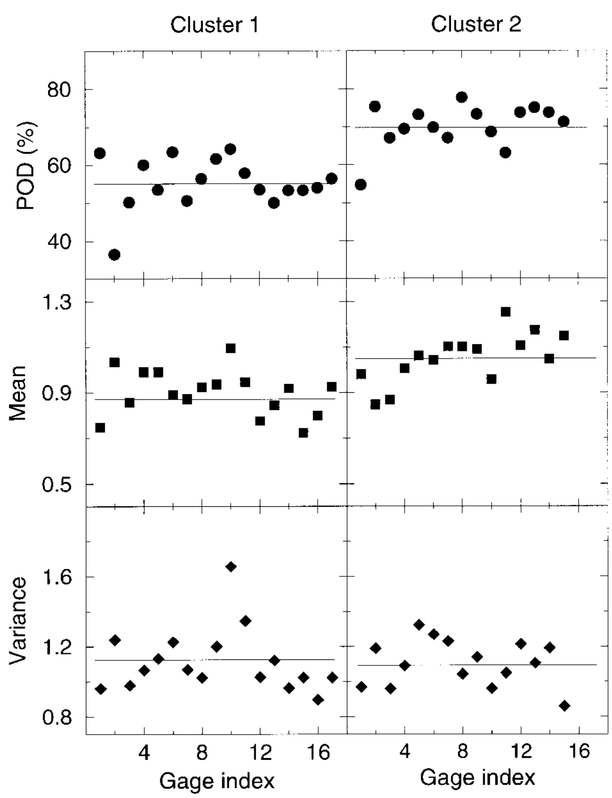

FIG. 2. Long-term hourly rain gauge log-rainfall statistics.

radar grid location and time, respectively, and variable r (mm) is the lower rainfall threshold to which (1) applies. The random variable ε_R is assumed to have a lognormal distribution with median 1 (mean of log-transformed variable is zero) and range-dependent variance, Σ_R. Spatial independence is assumed between ε_R variables with lagged distances greater than the radar pixel size (2–4 km).

Due to lack of accurate area-rainfall (R_A) data, evaluation of Σ_R should rely solely on spatially sparse rain gauge (R_G) reports. Hence, replacing variable R_A with variable R_G (1) becomes

$$R_G(s, j) = [R_R(s, j)][\varepsilon_{\mathrm{GR}}(j)], \quad \text{if} \quad R_R(s, j) > r, \quad (2)$$

where j is the rain gauge location index, $\varepsilon_{\mathrm{GR}}$ is the radar–rain gauge (G/R) rainfall ratio, and R_G (mm h^{-1}) is the rain gauge rainfall. We assume that $\varepsilon_{\mathrm{GR}}$ is a random variable having a lognormal distribution with median 1 and range-dependent variance. The assumption of median 1 for $\varepsilon_{\mathrm{GR}}$ implies that rain gauge measurements are unbiased. Applying log transformations to the variables and rearranging terms, the second-order moment statistics of (2) are

$$\mathrm{var}[\log\varepsilon_{\mathrm{GR}}] = \mathrm{var}[\log R_G - \log R_R]. \quad (3)$$

If we assume that the radar grid area A is filled with rain gauges, that the rain gauge–rainfall measurements R_G accurately represent rainfall at a point, and that rainfall is greater than a lower threshold, r, somewhere in A, then the conditional areal-rainfall R_A for an accumulation period s can be represented by the following integral:

$$R_A = \frac{1}{A}\int_A (R_G(s, u)\,|\,R_G(s, u) > r)\, du, \quad (4)$$

where u denotes the rain gauge location. At this point let us introduce the approximation

$$\log R_A \approx \overline{\log R_G} + C = \frac{1}{A}\int_A [\log R_G(s, u) + C\,|\,R_G(s, u) > r]\, du. \quad (5)$$

Equation (5) can be interpreted as follows. For a given minimum rainfall threshold the difference between the log of average rainfall is different by a constant from the average of the log of rainfall computed conditionally on that threshold. Figure 3 demonstrates this assumption for a number of hourly rain gauge rainfall thresholds. In these calculations, the area-integral of Eq. (5) was replaced by a time summation over a period of two months of rain gauge data from Melbourne, Florida. It is clear from Fig. 3 that 0.5 mm h^{-1} is a cutoff threshold where the variability of $\log R_A - \log R_G$ difference becomes negligible (<0.04). This means that $\log R_A - \overline{\log R_G}$ difference can be represented by a constant factor, C. Based on the above discussion, (3) is rewritten as

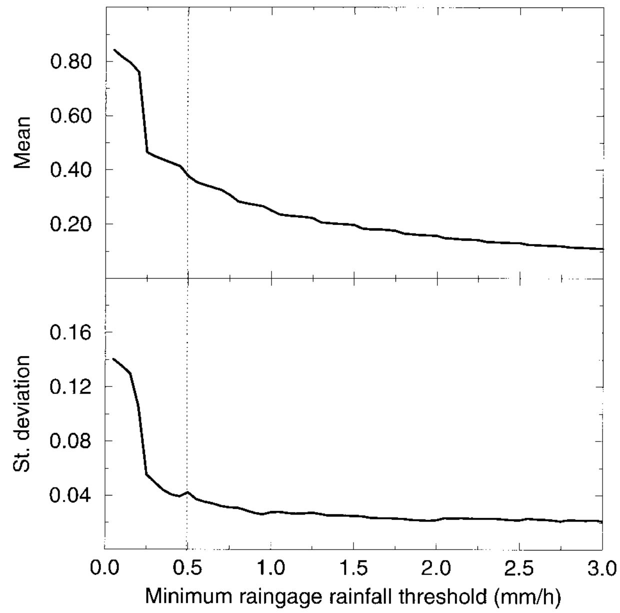

FIG. 3. Normalized difference between the log-transformed mean rain gauge rainfall and the mean of the log-transformed rain gauge rainfall values vs minimum rainfall rate threshold. The statistics are computed based on 32 rain gauges.

$$\mathrm{var}[\log\varepsilon_{\mathrm{GR}}] \approx \mathrm{var}[\log R_G + (\log R_A - \overline{\log R_G} - C) - \log R_R] = \mathrm{var}[(\log R_G - \overline{\log R_G} - C) - (\log R_R - \log R_A)]. \quad (6)$$

The first and second terms of the variance at the right-hand side of (6) are the rain gauge area-point difference error ε_G and the radar-rainfall error ε_R, respectively. This is shown in the following. Applying logarithmic transformations to the variables of (1) and rearranging terms we get the following expression for the radar-rainfall error:

$$\log\varepsilon_R = \log R_A - \log R_R. \quad (7)$$

The rain gauge error is defined as

$$R_A = R_G\varepsilon'_G, \quad (8)$$

where ε'_G is the rain gauge multiplicative error. For notational convenience, we define $\varepsilon_G = 1/\varepsilon'_G$. From the assumption of (5), after log transformations and rearrangement of terms, we get

$$\log\varepsilon_G \approx \log R_G - \overline{\log R_G} - C. \quad (9)$$

From (7) and (9), Eq. (6) becomes

$$\mathrm{var}[\log\varepsilon_{\mathrm{GR}}] \approx \mathrm{var}[\log\varepsilon_G - \log\varepsilon_R] = \mathrm{var}[\log\varepsilon_G] + \mathrm{var}[\log\varepsilon_R] - 2\,\mathrm{Cov}[\log\varepsilon_G, \log\varepsilon_R]. \quad (10)$$

Assuming now that $\mathrm{cov}[\log\varepsilon_G, \log\varepsilon_R]$ is zero (or at least close to zero) and rearranging terms, we get the following expression:

$$\text{var}[\log\varepsilon_R] = \text{var}[\log\varepsilon_{GR}] - \text{var}[\log\varepsilon_G]. \quad (11)$$

A physical interpretation of zero covariance is that the factors affecting radar measurements are weakly related to the rain gauge sampling error, which is due to the high spatial rainfall variability. Some of the main radar error factors are 1) temporal variations in the *Z*–*R* relationship; 2) strong updrafts and downdrafts in convective clouds; 3) sampling of hail, snow, and mixed-phase hydrometeors; 4) returns due to ground clutter and anomalous propagation of the radar beam; and 5) signal noise and nonlinear beam integration effects.

The next two sections present procedures for evaluating the variances of $\log\varepsilon_{GR}$ and $\log\varepsilon_G$. Assuming an unbiased estimator (i.e., zero mean of $\log\varepsilon_R$), the variance of $\log\varepsilon_R$ is related to the variable of interest (Σ_R) as

$$\Sigma_R = \exp[2\ \text{var}(\log\varepsilon_R)] - \exp[\text{var}(\log\varepsilon_R)]. \quad (12)$$

Under certain conditions, the variance of the multiplicative radar error Σ_R is related to the conditional mean-square additive error of radar-rainfall estimates as

$$W(s, u) = \Sigma_R R_R(s, u)^2, \quad (13)$$

where $W(s, u)$ (mm^2 h^{-2}) is the mean-square error at radar grid location u for hour s and other variables have been previously defined. For (13) to hold the estimates (R_R) must be unbiased and rainfall within the radar grid must be greater than a minimum threshold r ($\sim$0.5 mm h^{-1}). The appendix shows how (13) is derived based on the above conditions.

3. Radar–rain gauge ratio variance

A power-law model is used to represent the range dependence in the variance of $\log\varepsilon_{GR}$. The model has the following form (Smith et al. 1989):

$$\text{var}[\log\varepsilon_{GR}(S)] = \phi + \delta(S/S_0)^\gamma, \quad (14)$$

where ϕ, δ, and γ are the model coefficients. Variable S (km) is the distance from the radar, and S_0 is equal to 200 km, which is the maximum range for estimating the radar-rainfall error. The estimation of the model's coefficients is discussed next.

For a given minimum rainfall-rate threshold (r) and long-term radar and rain gauge rainfall data, the sample variance of $\log R_G - \log R_R$ difference is evaluated for each radar–rain gauge location u as

$$\upsilon_u = \frac{1}{N_u}\sum_{s=1}^{N_u}\{[\log R_G(s,u) - \log R_R(s,u)]^2 \,|\, R_G(s,u) > r\}, \quad (15)$$

where υ is the sample variance, u is the rain gauge location, and N_u is the u radar–rain gauge location's sample size. Radar–rain gauge locations with small sample sizes (e.g., $N_u < 30$) are excluded from the analysis. This constraint eliminated three rain gauge locations in cluster 1, while for the remaining locations, N_u was between 60 and 100 pairs. Substituting the sample variances to (14) we get

$$\upsilon_j = \phi + \delta(S_j/S_0)^\gamma, \quad (16)$$

where j is the rain gauge index. Applying a nonlinear least squares regression at the (υ_j, S_j) pairs, the coefficients of (14) are estimated.

4. Rain gauge area-point ratio variance

The formulations discussed here provide expressions for evaluating the variance of the ratio between radar grid areal rainfall and its estimate at a single point within the averaging domain. We follow a method that relies on the mean-areal rainfall variance reduction factor proposed by Rodriguez-Iturbe and Mejia (1974). The method makes the assumption of second-order stationarity and requires calculations of the sample mean, variance, and correlation function. Kitchen and Blackall (1992) proposed a method that relates the area-point ratio variance to the areal integral of the rain gauge log-rainfall point variogram. The methods are conceptually similar, given that correlograms and variograms are related for stationary random fields; however, the variogram method does not require calculations of the sample mean and variance. In the following we discuss how the variance reduction factor applies in estimating the area-point ratio variance. For discussion on the variogram method, refer to Anagnostou (1997).

From the definition of mean-areal rainfall given by Eq. (4) and based on the approximation shown in (5), the area-point ratio variance can be evaluated as

$$\text{var}[\log R_G - \log R_A] \approx E\{[\log R_G(s, u_0) - \overline{\log R_G} - C]^2 \,|\, R_G(s, u_0) > r\}, \quad (17)$$

where u_0 is the location of the sampling point (rain gauge) within the averaging domain (radar products grid). Assuming homogeneous first- and second-order statistics within the averaging domain, which implies that mean (μ_G) and variance (σ_G^2) of log rainfall are the same at any point in the product grid, one can expand (17) to

$$\text{var}(\log R_G - \log R_A) \approx \text{var}(\log R_G) + \text{var}[(\overline{\log R_G} - C)] - 2\ \text{cov}[\log R_G, (\overline{\log R_G} - C)], \quad (18)$$

where the first term, $\text{var}[\log R_G]$, is equal to σ_G^2. The second and third terms of (18) are expressed with respect to the point variance and correlation function of log rainfall as (Rodriguez-Iturbe and Mejia 1974)

$$\mathrm{var}[(\overline{\log R_G} - C)] = E[(\overline{\log R_G} - C - \mu_G)^2] \approx \frac{\sigma_G^2}{A^2} \int\int_A \rho(u, v)\, du\, dv \qquad (19)$$

$$\mathrm{cov}[\log R_G, (\overline{\log R_G} - C)] = E[(\log R_G - \mu_G)(\overline{\log R_G} - C - \mu_G)] \approx \frac{\sigma_G^2}{A} \int_A \rho(u_0, u)\, du, \qquad (20)$$

where $\rho(\ )$ is the spatial correlation of log rainfall. From (18), (19), and (20) we obtain the area-point difference variance:

$$\mathrm{var}[\log R_G - \log R_A] \approx \sigma_G^2 \left(1 - \frac{2}{A} \int_A \rho(u_0, u)\, du + \frac{1}{A^2} \int\int_{A^2} \rho(u, v)\, du\, dv\right), \qquad (21)$$

where the expression in the parenthesis is the mean-areal variance reduction factor (Rodriguez-Iturbe and Mejia 1974).

The point log-rainfall variance and spatial correlation can be estimated from long-term rain gauge data. Based on the assumption of stationarity, the spatial correlation function is described in terms of separation distance only:

$$\rho(u, v) = f\{|u - v|, [R_G(s, u) > r, R_G(s, v) > r]\}, \qquad (22)$$

where u and v are locations in space. The most commonly used correlation functions are the exponential, with or without a ''nugget constant'' (Kitanidis 1986). The nugget constant is used to model the rapid decorrelation of rainfall at very small lag distances (of the order of meters). Ciach and Krajewski (1999) argue that at the scales of radar products the ''nugget effect'' is the dominant source of spatial decorrelation of rainfall. The correlation functions considered in our study are

$$\rho(|u - v|) = \exp\{|u - v|/\lambda_e\} \quad \text{and} \qquad (23)$$

$$\rho(|u - v|) = \rho_0 \exp\{|u - v|/\lambda_e\}, \qquad (24)$$

where λ_e (km) is the spatial correlation distance and ρ_0 is the nugget constant.

The above parameters are evaluated based on the sample lagged-distance correlation values $\hat{\rho}(d_{ij})$ derived from rain gauge pairs with separation distances ($d_{ij} = |u_i - u_j|$) less than the products' scale (2–4 km). The sample lagged-distance correlation is evaluated as

$$\hat{\rho}(d_{ij}) = \frac{1}{\hat{\sigma}_G^2}\left(\frac{1}{n_{ij}} \sum_{s=1}^{N_s} \{\log R_G(s, i)\, \log R_G(s, j)\,|\,[R_G(s, i) > r, R_G(s, j) > r]\} - \hat{\mu}_G^2\right), \qquad (25)$$

where N_s is the total number of accumulation periods s, and n_{ij} is the number of samples used in the (i, j) rain gauge pair. The rain gauge log-rainfall sample mean ($\hat{\mu}_G$) and variance ($\hat{\sigma}_G^2$) are computed as

$$\hat{\mu}_G = \frac{1}{n} \sum_{s=1}^{N_s} \sum_{i=1}^{N_G} [\log R_G(s, i)\,|\,R_G(s, i) > r] \qquad (26)$$

and

$$\hat{\sigma}_G^2 = \frac{1}{n} \sum_{s=1}^{N_s} \sum_{i=1}^{N_G} [\log R_G(s, i)^2\,|\,R_G(s, i) > r] - \hat{\mu}_G, \qquad (27)$$

where N_G is the total number of rain gauges, and n is the number of samples used in calculating the rain gauge log-rainfall statistics.

5. Case study

In this section we provide an illustration of the expressions derived above. The scales of interest are hourly rainfall accumulations at 2 km × 2 km square grids and the HRAP (about 4 km × 4 km) grid. The statistics presented hereafter are conditioned on rainfall rates exceeding 0.5 mm h^{-1}. This minimum rainfall threshold is applied first to minimize the effect of small rainfall rates in log transformations and second to satisfy the approximation shown in Eq. (5).

As a first step, we evaluated the range-dependent variance of $\log\varepsilon_{GR}$ [Eq. (14)]. Figure 4 shows the radar–rain gauge mean-square log-rainfall difference, evaluated for each rain gauge location u [Eq. (15)] versus distance from the radar. For the 2 km × 2 km grid, the variance of logarithmic G/R ratios range from 0.35 close to the radar to approximately 1.0 at a 180-km range. These numbers translate into G/R ratio standard deviations of approximately 0.7 and 2.2, respectively. Figure 4 also shows the regression line and 90% confidence intervals of the power-law function [Eq. (16)], as fitted on the first cluster data (solid circles). Notice that the confidence bounds of the 4-km averaging domain is almost twice as large as the 2-km domain's bounds. This is attributed solely to the rain gauge–rainfall uncertainty, which increases as the radar grid area increases from 4 to 16 km^2. It is clear that all calibration and most of

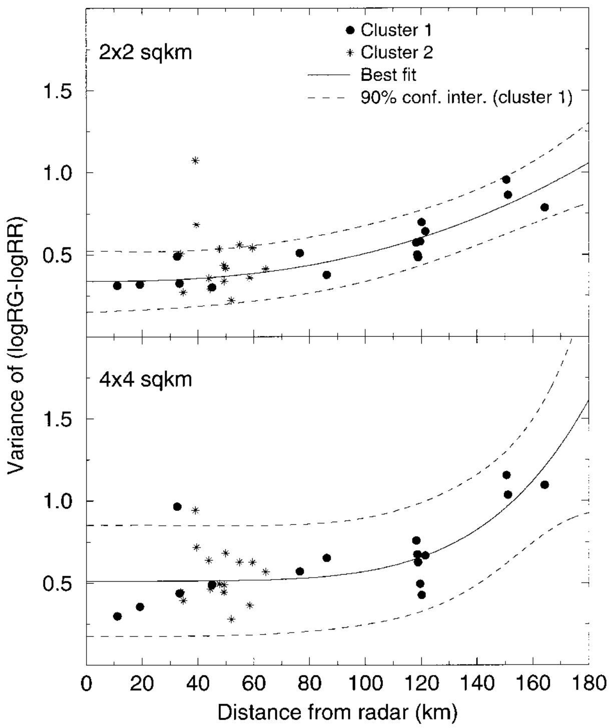

FIG. 4. Mean-square radar–rain gauge log-rainfall difference for 2 km (upper panel) and 4 km (lower panel) radar pixels. The regression line corresponds to the power-law model of Eq. (16).

the second cluster's data points are well within first cluster's 90% confidence bounds. The values of the model coefficients (ϕ, δ, γ) are shown in Table 2.

The next step includes evaluation of the rain gauge area-point ratio variance. Section 4 presented a method, which utilizes rain gauge sample estimates of the mean, variance, and spatial correlation structure of rainfall. Figure 5 shows the sample correlation between rain gauge pairs of the same cluster versus their intergauge distance. Only pairs with intergauge distances less than the 4 km (radar grid size) are presented. Because only two data points exist at intergauge distances less than 1 km, it is impossible to identify a correlation function at this small scale. Therefore, the results presented in this study are qualitative illustrations of hypothetical small-scale correlation functions.

Figure 5 shows plots of these correlation functions. The parameter values of the correlation functions are shown in Table 3. The rain gauge area-point ratio variance, var($\log\varepsilon_G$), is evaluated according to Eq. (21) based on the long-term sample rain gauge log-rainfall variance (σ_G^2), the fitted spatial correlation function (ρ), and the location of the sampling point in the averaging domain (u_0).

Figure 6 demonstrates the significance of the rain gauge location within the averaging domain with respect to the magnitude of the area-point ratio variance. The figure shows the error variance computed from Eq. (21)

TABLE 2. Gauge–radar (G/R) ratio model coefficients.

Averaging Domain	ϕ	δ	γ
2 km × 2 km	0.34	0.93	2.47
4 km × 4 km	0.51	1.87	3.02

versus normalized distances of the rain gauge locations from the domain's center. Normalization is with respect to half of the domain's size (1 and 2 km). The upper and lower panels of Fig. 6 correspond to the 2- and 4-km radar grid sizes, respectively. The vertical lines show the average distance of the first cluster's rain gauges evaluated for both the 2 km × 2 km Cartesian grid and the HRAP product grid. Table 3 summarizes the rain gauge area-point ratio variance results computed from (23) based on the two correlation functions. Notice that the correlation model with nugget constant results in an approximately 40% higher error variance for the 2 km × 2 km grid area. For the 4 km × 4 km grid area, though, the effect of the nugget constant is negligible.

Based on the fitted G/R ratio variance model and the evaluated rain gauge area-point ratio variances, the range-dependent radar rainfall error variance (Σ_R) is assessed [Eqs. (12) and (13)]. A summary of the results, for the two rain gauge error variance estimates and two discrete ranges from the radar, is presented in Tables 4 (2 km × 2 km) and 5 (4 km × 4 km). The 0.1 and 0.9 quantiles shown in these tables are with respect to uncertainty in fitting the range-dependent model [Eq. (16)] to the sample G/R ratio variances. Notice that at 2-km grid size the correlation function with nugget effect results in a lesser estimate of the normalized standard deviation of radar-rainfall error. Normalization is with respect to the mean-radar rainfall. At the 2-km grid size and short ranges (20 km) the mean ratio of rain gauge

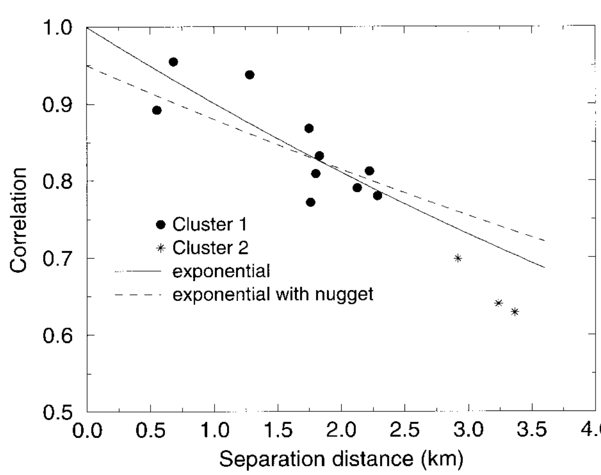

FIG. 5. Rain gauge log-rainfall correlation vs lagged distance. The solid and dashed regression lines correspond to the following correlation functions: exponential (25) and exponential with nugget effect (26).

TABLE 3. Rain gauge correlation function parameters and the corresponding area-point difference variance.

Models	Parameter values	Variance (2 × 2 km²)	Variance (4 × 4 km²)
Exponential ($1/\lambda_e$)	−0.105	0.094	0.199
Exponential and nugget (ρ_0, $1/\lambda_e$)	0.95, −0.077	0.122	0.199

to radar error variance is approximately 0.5, while at far ranges (150 km) it is 0.15. At the 4-km grid size the corresponding results are 0.6 and 0.2 at 20 and 150 km, respectively. The corresponding normalized standard deviations of the hourly radar rainfall products' errors at 2-km grid resolution are about 60% (20 km) and 140% (150 km) of the mean. At 4-km resolution the corresponding results are similar. This indicates that the increase in the variance of the *G*/*R* ratio, shown in Fig. 4, is solely due to increase of the sampling area difference between the two sensors. From Tables 4 and 5 it is clear that there is significant variability in the results. This is attributed to 1) uncertainty in fitting the *G*/*R* ratio variance model and 2) differences in the methods used to quantify the rain gauge error variance. With the currently available data, however, it is impossible to determine which method is superior. High spatial resolution (<0.1 km) measurements of rainfall are needed to quantify the ''true'' small-scale correlation structure of rainfall and assess the different assumptions made in this formulation.

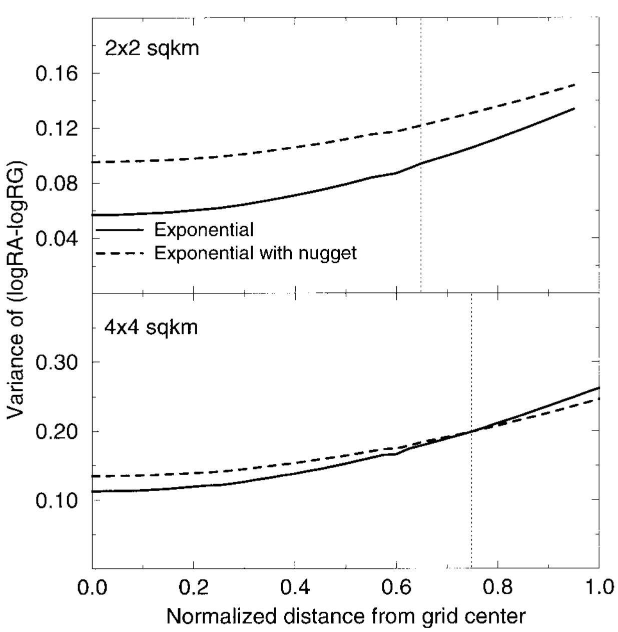

FIG. 6. Rain gauge sampling error variance in log rainfall vs distance of the rain gauge location from the center of the averaging domain. The vertical lines show the average distance of the first cluster's rain gauge locations.

6. Conclusions

A statistical method is formulated in this research in which radar error variance is derived by separating from the *G*/*R* ratio variance the rain gauge sampling (area point) error variance. Both radar and rain gauge errors are defined as the ratios of true mean-areal rainfall to that estimated by radar and rain gauges. The method fits a range-dependent model to the multiplicative radar error variances. It requires accurate information regarding the small-scale spatial structure of rainfall. Due to the different assumptions and log transformation of variables, application of the method to spatiotemporal scales other than the ones investigated in this research is not straightforward.

We demonstrate the method on a two-month period of Florida summer radar–rain gauge rainfall data. We conclude that the rain gauge rainfall uncertainty, due to subgrid rainfall variability, contributes a significant portion (up to 60%) of the differences observed in radar–rain gauge comparisons. However, the results presented in this study should be viewed with caution due to limited information on the small-scale rainfall variability. Furthermore, the presented method makes certain assumptions that should be verified experimentally before it is used to provide quantitative error variance estimates.

First, we assume that radar and rain gauge rainfall

TABLE 4. Summary of radar-rainfall error results for a 2 km × 2 km radar grid area. The results are as follows: 1) the square root of radar error variance normalized with respect to the mean radar rainfall, 2) the ratio of radar error variance to the *G*/*R* ratio variance, and 3) the ratio of rain gauge to radar error variance.

Radar + range (km)	Results number	Mean (0.1, 0.9 quantiles)	
		Exponential	Exponential and nugget
20	1	0.60 (0.38, 0.81)	0.55 (0.32, 0.77)
20	2	0.72 (0.56, 0.80)	0.64 (0.42, 0.74)
20	3	0.38 (0.80, 0.25)	0.56 (1.36, 0.35)
150	1	1.43 (1.18, 1.72)	1.37 (1.12, 1.66)
150	2	0.88 (0.86, 0.90)	0.85 (0.82, 0.87)
150	3	0.13 (0.16, 0.11)	0.18 (0.22, 0.15)

TABLE 5. The same as in Table 4, but for a 4 km × 4 km radar grid area.

Radar range (km)	Results number	Mean (0.1, 0.9 quantiles)	
		Exponential	Exponential and nugget
20	1	0.71 (0.27, 1.15)	0.71 (0.27, 1.15)
20	2	0.61 (0.25, 0.74)	0.61 (0.25, 0.74)
20	3	0.63 (2.97, 0.35)	0.63 (2.97, 0.35)
150	1	1.55 (1.04, 2.18)	1.55 (1.04, 2.18)
150	2	0.79 (0.72, 0.83)	0.79 (0.72, 0.83)
150	3	0.26 (0.39, 0.20)	0.26 (0.39, 0.20)

estimates are unbiased. This is a reasonable assumption for rain gauge rainfall, given that there are ways to quantify the systematic undercatchment of gauges due to turbulence at the gauge orifice (Nespor 1995). Kitchen and Blackall (1992) showed that the bias due to area-point difference effect in rain gauge areal-rainfall estimates is insignificant compared to the corresponding standard deviation. The bias at the radar-rainfall products can be effectively removed based on long-term radar–rain gauge comparisons (Anagnostou et al. 1997; Ciach et al. 1999; Smith and Krajewski 1991).

The assumptions that allow the variance separation are the following. The first assumption is that the log transformations of radar and rain gauge multiplicative errors are uncorrelated random variables. This assumption implies that the physical factors that affect radar measurements are weakly related to the rain gauge sampling error source, which is the small-scale spatial rainfall variability. Second, we assume homogeneous statistics within the radar products' grid domain. The third assumption is that the radar-rainfall errors are uncorrelated at lagged distances greater than the radar product grid size. Assessment of the above assumptions through simulation is subject to model specifications and constraints, which have to be verified empirically. Long-term data from rain gauges covering a few radar product grids would provide some experimental evidence regarding the validity of our method and results.

Other fundamental questions that need to be investigated are the following. How accurate are the radar and rain gauge multiplicative error statistics computed in the logarithmic scale? First, the nonbias condition, although adequate in the log scale, may not be met at the exponentiation of the log ratios. Second, we need to examine the uncertainty in deriving the multiplicative error variance from the log-transformed radar–rain gauge ratios. Finally, it is important to know how sensitive the statistics described above may be with respect to season, precipitation type, and radar site. In case of significant deviations, the parameters of the proposed statistical model should be assessed according to storm type based on physical classification criteria. These questions, to be answered, require specialized experiments involving corresponding radar data and high spatial resolution (intergauge distances from 1 to 1000 m) rain gauge networks.

Acknowledgments. This study was supported in part by NASA Grant NAG 5-2084 and by the National Oceanic and Atmospheric Administration under Cooperative Agreement NA47WH0495 between the Office of Hydrology of the National Weather Service and the University of Iowa Institute of Hydraulic Research.

APPENDIX

Mean Square Radar Rainfall Error

The conditional mean square error of radar-rainfall estimates is

$$W(s, u) = E\{[R_A(s, u) - R_R(s, u)]^2 | R_R > r\}, \quad \text{(A1)}$$

with all variables defined in section 2. For convenience the conditioning and indices (u, s) are omitted hereafter. Hence, (A1) leads to the following expressions:

$$W = E[(R_A - R_R)^2] = E[(R_A - \varepsilon_R R_A)^2]$$
$$= E[R_A^2(\varepsilon_R - 1)^2] = E[(\varepsilon_R - 1)^2]R_A^2 = \Sigma_R R_A^2. \quad \text{(A2)}$$

Replacing R_A with the unbiased estimate R_R (A2) leads to (13).

REFERENCES

Anagnostou, E. N., 1997: Real-time radar rainfall estimation. Ph.D. thesis, University of Iowa, 208 pp. [Available from Department of Civil and Environmental Engineering, University of Iowa, Iowa City, IA 52242.]

——, and W. F. Krajewski, 1999a: Real-time radar rainfall estimation. Part I: Algorithm formulation. *J. Atmos. Oceanic Technol.,* **16,** 189–197.

——, and ——, 1999b: Real-time radar rainfall estimation. Part II: Case study. *J. Atmos. Oceanic Technol.,* **16,** 198–205.

——, ——, D. J. Seo, and E. R. Johnson, 1998: Mean-field radar rainfall bias studies for WSR-88D. *Amer. Soc. Civ. Eng. J. Hydrol. Eng.,* **3,** 149–159.

Austin, P. M., 1987: Relation between measured radar reflectivity and surface rainfall. *Mon. Wea. Rev.,* **115,** 1053–1069.

Barnston, A. G., 1991: An empirical method of estimating raingage and radar rainfall measurement bias and resolution. *J. Appl. Meteor,* **30,** 282–296.

Ciach, G. J., and W. F. Krajewski, 1999: On the estimation of radar rainfall error variance. *Adv. Water Resour.,* in press.

——, ——, E. N. Anagnostou, M. L. Baeck, J. A. Smith, J. R. McCollum, and A. Kruger, 1997: Radar rainfall estimation for ground validation studies of the Tropical Rainfall Measuring Mission. *J. Appl. Meteor.,* **36,** 735–747.

Court, A., and F. Griffiths, 1986: Thunderstorm climatology. *Thunderstorm Morphology and Dynamics,* 2d ed., University of Oklahoma Press, 9–39.

Hendrick, R. L., and G. H. Comer, 1970: Space variations of pre-

cipitation and its implications for rain gauge network design. *J. Hydrol.,* **10,** 151–163.
Huff, F. A., 1970: Sampling errors in measurements of mean precipitation. *J. Appl. Meteor.,* **9,** 35–44.
Johnson, R. A., and D. W. Wichern, 1992: *Applied Multivariate Statistical Analysis.* Prentice Hall, 642 pp.
Joss, J., and A. Waldvogel, 1990: Precipitation measurement and hydrology. *Radar in Meteorology,* D. Atlas, Ed., Amer. Meteor. Soc., 577–606.
Kitanidis, P. K., 1986: Parameter uncertainty in estimation of spatial funtions: Bayesian analysis. *Water Resour. Res.,* **22,** 499–507.
Kitchen, M., and R. M. Blackall, 1992: Representativeness errors in comparisons between radar and gauge measurements of rainfall. *J. Hydrol.,* **134,** 13–33.
——, and P. M. Jackson, 1993: Weather radar performance at long range—Simulated and observed. *J. Appl. Meteor.,* **32,** 975–985.
Krajewski, W. F., 1987: Co-kriging of radar-rainfall and rain gauge data, *J. Geophys. Res.,* **92** (D8), 9571–9580.
——, and Coauthors, 1996: Radar-rainfall estimation studies for TRMM ground validation. IIHR Tech. Rep. 379, Iowa Institute of Hydraulic Research, University of Iowa, 205 pp. [Available from IIHR, University of Iowa, Iowa City, IA 52242.]
Nespor, V., 1995: Investigation of wind-induced error of precipitation measurements using a three-dimensional numerical simulation. Ph.D. dissertation, Swiss Federal Institute of Technology, Zurich, Switzerland, 109 pp. [Available from Swiss Federation Institute of Technology Zurich, ETH Zentrum, CH-8092 Zurich, Switzerland.]
Rodriguez-Iturbe, I., and J. M. Mejia, 1974: The design of rainfall network in time and space. *Water Resour. Res.,* **10,** 713–728.
Seed, A., and G. L. Austin, 1990: Variability of summer Florida rainfall and its significance for the estimation of rainfall by gauges, radar and satellites. *J. Geophys. Res.,* **95,** 2207–2215.
Silverman, B. A., L. K. Rogers, and D. Dahl, 1981: On the sampling variance of rain gauge networks. *J. Appl. Meteor.,* **20,** 1468–1478.
Smith, J. A., and W. F. Krajewski, 1991: Estimation of the mean field bias of radar rainfall estimates. *J. Appl. Meteor.,* **29,** 2505–2514.
——, R. C. Shedd, and M. L. Walton, 1989: Parameter estimation for NEXRAD hydrology sequence. Preprints, *24th Conf. on Radar Meteorology,* Tallahassee, FL, Amer. Meteor. Soc., 259–263.
——, D.-J. Seo, M. L. Baeck, and M. D. Hudlow, 1996: An intercomparison study of NEXRAD precipitation estimates. *Water Resour. Res.,* **32,** 2035–2045.
Zawadzki, I., 1973: Errors and fluctuations of rain gauge estimates of areal rainfall. *J. Hydrol.,* **18,** 243–255.
——, 1975: On radar–rain gauge comparison. *J. Appl. Meteor.,* **14,** 1430–1436.
——, 1984: Factors affecting the precision of radar measurements of rain. Preprints, *22d Conf. on Radar Meteorology,* Zurich, Switzerland, Amer. Meteor. Soc., 251–256.