WATER RESOURCES RESEARCH, VOL. 21, NO. 5, PAGES 764–768, MAY 1985

Synthesis of Radar Rainfall Data

WITOLD F. KRAJEWSKI AND KONSTANTINE P. GEORGAKAKOS

Hydrologic Research Laboratory, National Weather Service, National Oceanic and Atmospheric Administration, Silver Spring, Maryland

A method of generating synthetic radar-rainfall data is described. The data are generated by imposing random noise on a given, high-quality radar-rainfall field. Certain conditions are imposed on the resultant rainfall field so that the noise parameters are prespecified. The conditions pertain to the second order statistics of the generated rainfall fields: the mean, the variance, the correlation, and the variance of the logarithmic ratio of the resultant field to the original field. Accuracy of the generation method is evaluated from implementing a test case using Global Atmospheric Research Program Atlantic Tropical Experiment radar data. The method can be used in a number of different, mainly hydrologic, applications. These include validation of radar and rain gage data merging procedures, testing of various methods for computation of mean areal precipitation, and sensitivity analysis of rainfall-runoff models.

1. Introduction

In recent years, radar sensors have found wide applicability in the measurement of rainfall fields. This is mainly due to their ability to map the spatial characteristics of rainfall. It is a fact, however, that in some cases, large observation errors occur [*Harrold et al.*, 1973; *Wilson and Brandes*, 1979; *Collier et al.*, 1983]. Since standard rain gages offer a much more accurate way to measure point values of rainfall, procedures are being developed to merge radar and rain gage observations [*Brandes*, 1975; *Crawford*, 1979; *Eddy*, 1979]. The purpose is to obtain the best estimate of the rainfall field, taking advantage of the spatial detail that the radar gives and of the high point accuracy of the gages.

One of the generic problems faced while developing merging procedures is validation. Reliable synthesis of radar-rainfall fields can be very helpful as an alternative to costly field experiments. By using rainfall synthesis one can control the ensemble statistics of the generated fields so that the instrument-observation errors are simulated in the validation process. Thus it becomes possible to evaluate various merging procedures for statistically different error fields.

Since the statistical characteristics of radar-rainfall error fields are basically unknown, it would be difficult to use direct methods of random field generation such as those described by *Mejia and Rodriguez-Iturbe* [1974] or *Mantoglou and Wilson* [1981]. In this work we propose a methodology to derive error-field statistics, and then using direct methods, to generate radar-rainfall fields.

Our methodology avoids the explicit specification of the rainfall field statistics by acting on the point values of the observed fields. Thus the original field which could be, for example, a high-quality radar rainfall field, is taken as known. A random noise, which is Gaussian, isotropic, and has predefined second-order statistics, is imposed at each point of the original field. The noise level varies from point to point based on the local original field characteristics such as magnitude and gradient.

In addition to its use in the validation of merging procedures the proposed methodology can be used in the design of rainfall observation systems and in the testing of mean areal precipitation estimators or rainfall-runoff models, to mention only a few applications.

Paper number 5W0115.
0148-0227/85/005W-0115$02.00

Other attempts to use existing high-quality radar fields for the generation of synthetic spatial radar-observed rainfall were those by D. R. Greene et al. (unpublished manuscript, 1980). The method presented below provides improvements over those by Green et al., which resorted to trial and error for the specification of the statistical parameters of the noise field in order to obtain an ensemble of fields with specified spatial properties.

In the next sections the proposed methodology is described, followed by a discussion of an example implementation.

Method of Generation

The basic idea of the procedure is to generate fields from an existing high-quality radar field by imposing a noise field of known statistics such that the ensemble of the resultant realizations meets certain conditions. The conditions pertain to the spatially averaged second-order statistics of the generated fields.

It is due to the imposed conditions that the second-order statistics of the noise field are obtained. Such a procedure was made necessary by the lack of knowledge of the radar-noise field statistics.

If $G(x, y)$ is the generated field and $O(x, y)$ is the original, high-quality radar field, then the error field $\Delta(x, y)$ is commonly [*Hudlow et al.*, 1979] expressed as

$$\Delta(x, y) = \log_{10}\left(\frac{G(x, y)}{O(x, y)}\right) \qquad (x, y) \in A \tag{1}$$

where x, y are the field-point coordinates, and A is the field domain.

For the purposes of this study we take $\Delta(x, y)$ to be the product of a random field and a deterministic component according to

$$\Delta(x, y) = \varepsilon(x, y)S(x, y) \tag{2}$$

In (2), $\varepsilon(x, y)$ is a stationary and, in general, anisotropic, Gaussian random field of mean μ, variance σ^2, and correlation function $\rho(\tau_1, \tau_2)$, with τ_1, τ_2 denoting spatial lags in the two directions x and y. $S(x, y)$ is a deterministic function which makes $\Delta(x, y)$ a nonstationary random field. We adopt the form of $S(x, y)$ given by D. R. Greene et al. (unpublished manuscript, 1980)

$$S(x, y) = \langle|\nabla O(x, y)|\rangle O_{\max}(x, y) + O(x, y)\langle|\nabla O(x, y)|_{\max}\rangle \cdot (2\langle|\nabla O(x, y)|_{\max}\rangle O_{\max}(x, y))^{-1} \tag{3}$$

where

$\langle|\nabla O(x, y)|\rangle$ average absolute value of the gradient computed in four directions around the point (x, y) in the original field;
$\langle|\nabla O(x, y)|_{\max}\rangle$ maximum absolute value of the gradient in the original field;
$O(x, y)$ original field value at the point (x, y);
$O_{\max}(x, y)$ maximum value in the original field.

The form of function $S(x, y)$ in (3) and (2) implies high errors where high gradients and high magnitudes occur.

We note at this point that the development of the methodology is independent of the particular form of $S(x, y)$ in (3), so that any deterministic, real function of (x, y) can be used.

Eliminating $\Delta(x, y)$ from (1) and (2) yields

$$G(x, y) = O(x, y)10^{\varepsilon(x,y)S(x,y)} \tag{4}$$

It should be noted here that the use of (4) for the generation of $G(x, y)$ leaves the zero-rainfall areas of $O(x, y)$ unaltered. The changes in the nonzero areas of $O(x, y)$ are the ones that produce fields with the desirable statistics.

If one has a mechanism for generating the random component field $\varepsilon(x, y)$, then, using (4) and the original high-quality radar field $O(x, y)$, one can produce a realization of $G(x, y)$.

There are several methods for the generation of $\varepsilon(x, y)$, given its statistical parameters μ, σ^2, and $\rho(\tau_1, \tau_2)$. The turning bands method (TBM) presented by *Mantoglou and Wilson* [1982] is an efficient one in terms of accuracy and cost.

The TBM gives us a way to generate the field $\varepsilon(x, y)$ in (4) if its second-order statistics are known. We obtain these statistics by imposing certain conditions on the generated fields.

Due to the fact that $G(x, y)$ is a nonstationary random field, we need to specify operational measures of its statistical properties. Thus we define these measures as follows.

The spatial mean R of the field is

$$R = \frac{1}{|A|}\int_A E\{G(x, y)\}\,dx\,dy \tag{5}$$

where A is the generation domain with area $|A|$, and $E\{\ \}$ denotes expectation of the value of G at the point (x, y).

The spatially averaged variance P of the field $G(x, y)$ is

$$P = \frac{1}{|A|}\int_A E\{[G(x, y) - E\{G(x, y)\}]^2\}\,dx\,dy \tag{6}$$

The spatially averaged correlation $\rho_x(\tau_1)$ of the field $G(x, y)$ in direction x is

$$\begin{aligned}\rho_x(\tau_1) = \frac{1}{|A|}\int_A & E\{[G(x, y) - E\{G(x, y)\}]\\ &\cdot [G(x+\tau_1, y) - E\{G(x+\tau_1, y)\}]\}\\ &\cdot (E\{[G(x, y) - E\{G(x, y)\}]^2\}\\ &\cdot E\{[G(x+\tau_1, y) - E\{G(x+\tau_1, y)\}]^2\})^{-1/2}\,dx\,dy\end{aligned} \tag{7}$$

with τ_1 denoting the spatial lag in the x direction.

Similarly, we can define the correlation $\rho_y(\tau_2)$ in direction y with τ_2 denoting the spatial lag in the y direction.

Equations (5), (6), and (7) describe the spatially averaged field-expected value, field variance, and field correlation of $G(x, y)$.

Another measure of variance used often in the radar literature [*Hudlow et al.*, 1979] in place of (6) is the variance V of the logarithmic ratio of (1) defined as

$$\begin{aligned}V = \frac{1}{|A|}\int_A & E\left\{\log_{10}{}^2\left(\frac{G(x, y)}{O(x, y)}\right)\right\}dx\,dy\\ &- \frac{1}{|A|}\int_A E^2\left\{\log_{10}\left(\frac{G(x, y)}{O(x, y)}\right)\right\}dx\,dy\end{aligned} \tag{8}$$

By setting the expressions in (5), (6) or (8), and (7) to prespecified values, one can, in principle, obtain expressions for μ, σ^2, and $\rho(\tau_1, \tau_2)$. Then, one can generate the $\varepsilon(x, y)$ field using the TBM and subsequently generate the $G(x, y)$ field from (4).

The first step is to obtain expressions for the expectations in (5)–(8). Because of the exponential form of (4) we used the well-known relationships between the moments of log normal and normal random variables [e.g., *Vanmarcke*, 1983]. The relationships are applied at each point (x, y) in the field. Use of these relationships gives the desired expression for the expectations. *Krajewski and Georgakakos* [1985] give the details of the derivations.

If one specifies design values R_0, V_0, or P_0, $\rho_{x0}(\tau_1)$ in (5), (8), or (6) and (7), respectively, and substitutes the derived expectations, one obtains

$$\frac{1}{|A|}\int_A O(x, y)\exp\left\{\frac{1}{2}(\ln 10)^2 S^2(x, y)\sigma^2 + (\ln 10)S(x, y)\mu\right\}dx\,dy = R_0 \tag{9}$$

$$\frac{\sigma^2}{|A|}\int_A S^2(x, y)\,dx\,dy = V_0 \tag{10}$$

$$\begin{aligned}\frac{1}{|A|}\int_A & O^2(x, y)[\exp\{2(\ln 10)^2 S^2(x, y)\sigma^2\\ &+ 2(\ln 10)S(x, y)\mu\} - \exp\{(\ln 10)^2 S^2(x, y)\sigma^2\\ &+ 2(\ln 10)S(x, y)\mu\}]\,dx\,dy = P_0\end{aligned} \tag{11}$$

$$\begin{aligned}\frac{1}{|A|}\int_A & \{\exp\{(\ln 10)^2 S^2(x, y)\sigma^2\rho(\tau_1, 0)\} - 1\}\\ &\cdot \{\exp\{(\ln 10)^2 S^2(x, y)\sigma^2\} - 1\}^{-1}\,dx\,dy = \rho_{x0}(\tau_1)\end{aligned} \tag{12}$$

Again, similar expression can be derived for $\rho_{y0}(\tau_2)$. Solving simultaneously the equations (9), (10), or (11), and (12), and its corresponding form for y direction, one can obtain values for μ, σ^2, $\rho(\tau_1, 0)$, and $\rho(0, \tau_2)$.

Note that in (12) we used relationships of lognormal variables that are strictly true when the field is stationary. Therefore the expression in (12) is approximate, since the stationary field $\varepsilon(x, y)$ is multiplied by the function $S(x, y)$, yielding a nonstationary (in general) product field. The approximation is better for smaller lags τ_1 and smoother functions $S(x, y)$.

Assuming, for example, an exponential, anisotropic correlation function for $\varepsilon(x, y)$, of the type

$$\rho(\tau_1, \tau_2) = \exp\{-(h_1{}^2\tau_1{}^2 + h_2{}^2\tau_2{}^2)^{1/2}\} \tag{13}$$

knowledge of $\rho(\tau_1, 0)$ and $\rho(0, \tau_2)$ gives estimates of h_1 and h_2. Therefore $\rho(\tau_1, \tau_2)$ can be defined.

With μ, σ^2, $\rho(\tau_1, \tau_2)$ known, one can use the TBM to generate realizations of $\varepsilon(x, y)$.

In the particular case of an isotropic $\varepsilon(x, y)$ field with correlation function

$$\rho(\tau) = \exp\{-h\tau\} \tag{14}$$

the design equation (12) for $\rho_{x0}(\tau_1)$ and corresponding ex-

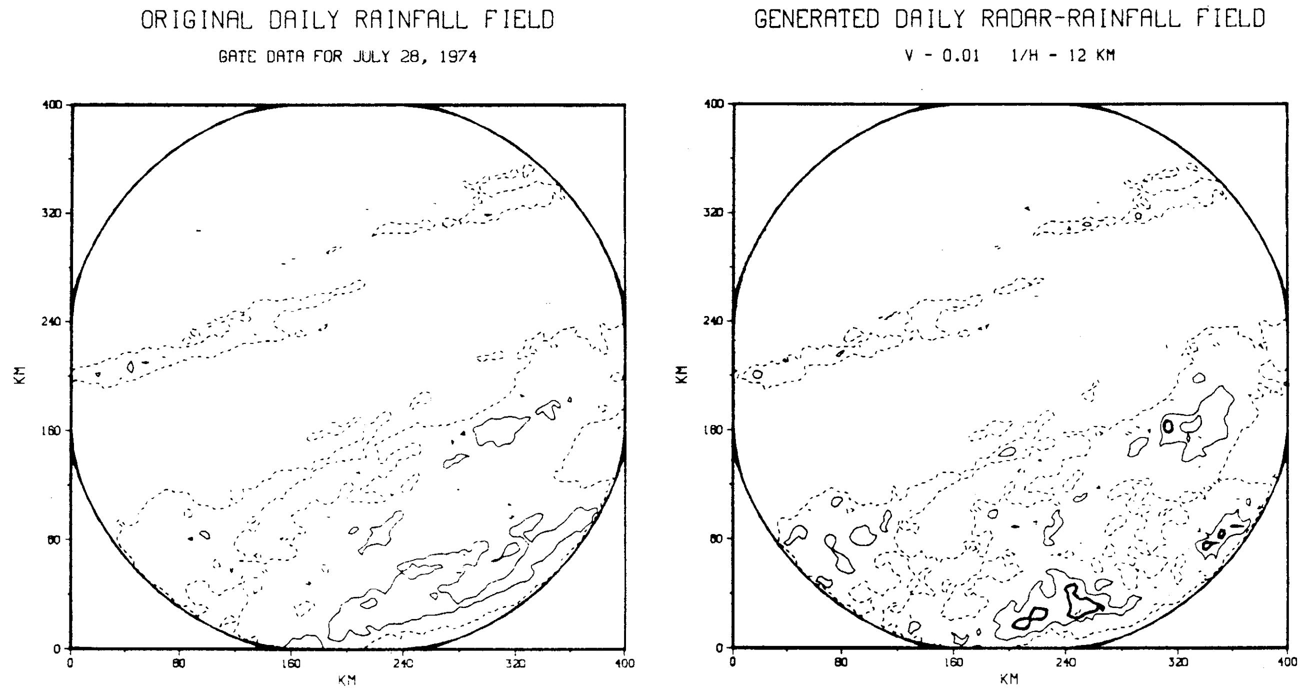

Fig. 1. Original daily rainfall field from GATE data for July 28, 1974 (left), and generated daily radar-rainfall field for $V = 0.01$ and $1/h = 12$ km (right). Dashed curves correspond to contours at the 1.5 mm/h level, solid curves correspond to contours at the 4 mm/h level, and thick solid curves correspond to contours at the 11.5 mm/h level.

pression for $\rho_{y_0}(\tau_2)$ consolidate to the following:

$$\frac{1}{|A|}\int_A \{\exp\{(\ln 10)^2 S^2(x, y)\sigma^2\rho(\tau)\} - 1\} \cdot \{\exp\{(\ln 10)^2 S^2(x, y)\sigma^2\} - 1\}^{-1}\, dx\, dy = \rho_{G_0}(\tau) \qquad (15)$$

where $\rho_{G_0}(\tau)$ is the design value of the correlation for the $G(x, y)$ field.

The next section exemplifies the application of the proposed methodology.

Numerical Implementation

This section presents an example of the generation of rainfall fields from an original, high-quality radar field. Accuracy of preservation of the design statistics in the generated ensemble of rainfall fields is studied in connection with the number of generated fields and the magnitude of the statistics themselves.

The original field consisted of daily radar data from the international Global Atlantic Tropical Experiment (GATE) conducted in 1974. A detailed description of the GATE data is given by *Hudlow and Patterson* [1979]. The original radar field corresponds to spatially averaged daily accumulations for July 28, 1974. Spatial averages were computed in 4 × 4 km domains.

For the purposes of this example, the design equations (9), (10), (11), and (15) were studied.

During the example runs we generated rainfall fields from the original radar field with mean equal to the spatial average of the original field (1.16 mm/h) and with prespecified values for the logarithmic-ratio variance V_0 (0.005, 0.01, and 0.03). The correlation condition consisted of specifying values for the correlation distance 1/h of the $\varepsilon(x, y)$ field (4, 12, and 20 km).

For illustration purposes, Figure 1 presents an example of original and generated fields.

In a true generation process, one specifies the value of the correlation of the $G(x, y)$ field and then, using (15), one obtains the value of the correlation of the $\varepsilon(x, y)$ field. Since our purpose was to study the capabilities of the method for a range of correlation values, we specified several values of the correlation of $\varepsilon(x, y)$ by specifying the correlation distance 1/h and then we used (15) to compute the correlation ρ_G of the $G(x, y)$ field. Given that (15) is strictly true for stationary fields $S(x, y)\varepsilon(x, y)$ we computed ρ_G for the smallest possible lag $\tau = 4$ km so that $S(x, y) \approx S(x + \tau, y) \approx S(x, y + \tau)$.

For all the combinations of R_0, V_0, and $1/h$ we generated three ensembles of rainfall fields with the number of fields per ensemble, NS, equal to 10, 25, and 50. We then computed the statistics R, V, P, $\rho_G(\tau = 4$ km) for each ensemble and we compared them with their theoretical values obtained from (9), (10), (11), and (15), respectively. We also computed μ and σ^2 from each ensemble and compared them to the values obtained by solving (9) and (10). Thus we were able to evaluate the accuracy of the TBM generator.

Because the specification of high V and P statistics will sometimes yield physically unacceptable values of the precipitation rates, we monitored the number of values exceeding an arbitrarily chosen rate, set to 50 mm/h, which is close to the observed world record value of 55 mm/h for daily data [*Chow*, 1964]. Consequently, we give guidelines on the specification of V, which is a normalized measure, so that generation of realistic rainfall fields results.

Table 1 contains the results obtained for all cases. The results form nine sets that cover the nine combinations of V and $1/h$ values specified. In Table 1 the sets are arranged in three rows (V fixed) and three columns ($1/h$ fixed). For all cases, the value of R remained equal to 1.16 mm/h.

The table displays the prespecified values of μ, σ^2, V, P, ρ_G ($\tau = 4$ km), as well as the percent errors ((prespecified value) − (computed value))/(prespecified value) × 100 that

TABLE 1. Percent Errors in the Generation Process for Nine Sets of Specified Values

	Set 1				Set 2				Set 3			
Statistic	Specified	$NS = 10$	$NS = 25$	$NS = 50$	Specified	$NS = 10$	$NS = 25$	$NS = 50$	Specified	$NS = 10$	$NS = 25$	$NS = 50$
μ	−0.08	0.0	0.0	0.0	−0.08	−12.5	0.0	0.0	−0.08	−12.5	12.5	12.5
σ^2	0.24	12.5	8.3	4.2	0.24	12.5	8.3	4.0	0.24	12.5	8.3	4.0
V	0.005	12.0	6.0	4.0	0.005	6.0	4.0	2.0	0.005	14.0	8.0	2.0
P, mm²/h²	0.73	8.2	2.7	0.0	0.73	0.0	−2.7	−4.0	0.73	11.0	6.8	2.7
ρ_G, $\tau = 4$ km	0.32	−12.5	−15.6	−15.6	0.63	−12.9	−14.5	−14.5	0.72	−9.7	−11.1	−11.1
		Set 4				*Set 5*				*Set 6*		
μ	−0.16	0.0	0.0	0.0	−0.16	−6.3	0.0	0.0	−0.16	−6.3	12.5	12.5
σ^2	0.47	12.8	6.4	4.3	0.47	10.6	4.3	2.1	0.47	10.6	6.4	4.3
V	0.01	13.0	7.0	4.0	0.01	7.0	4.0	2.0	0.01	14.0	9.0	3.0
P, mm²/h²	1.16	4.3	0.0	−2.6	1.16	−2.6	−14.7	−10.3	1.16	14.7	11.2	4.3
ρ_G, $\tau = 4$ km	0.32	−9.4	−15.6	−15.6	0.63	−12.9	−12.9	−12.9	0.72	−9.7	−11.1	−11.1
		Set 7				*Set 8*				*Set 9*		
μ	−0.50	0.0	0.0	−2.0	−0.50	−2.0	0.0	0.0	−0.50	−4.0	6.0	6.0
σ^2	1.42	12.0	6.4	4.3	1.42	10.6	5.6	3.5	1.42	12.0	7.7	4.2
V	0.03	12.7	7.0	3.7	0.03	7.0	4.0	1.7	0.03	13.7	9.0	3.0
P, mm²/h²	4.51	55.0	46.1	38.6	4.51	49.7	42.6	37.5	4.51	51.2	48.6	39.5
ρ_G, $\tau = 4$ km	0.32	12.9	−9.4	−12.5	0.63	−11.3	−11.3	−12.9	0.72	−9.7	−9.7	−9.7

were realized during the generation process. The field mean R, not included in the table, had an error of less than one percent for all of the cases.

Inspection of the prespecified P values of Table 1 reveals that a wide range of G-field variances was included, ranging from 0.73 mm^2/h^2 up to 4.51 mm^2/h^2. Similarly, the prespecified ρ_G values suggest that a wide range of G-field correlations was studied: from 0.32 up to 0.72.

The values of the percent errors in Table 1, excluding the ones corresponding to P for sets 7, 8, and 9, are all less than 16% and, in most cases, less than 10%. In general, better accuracy is obtained as NS increases, but accuracy is very good even with $NS = 10$.

The results corresponding to statistic P, for V specified equal to 0.03 (last row of sets), show abnormal behavior compared to the rest of the results in the same row of sets and for all the rest of the variables. The cause of this phenomenon is the nonlinear relationship between V and P (see equations (6) and (9)). Because there is an exponential relationship between V and P (one cannot prespecify both V and P), small errors in approximating the specified value of V can (depending on the form of $S(x, y)$) lead to pronounced errors in the preservation of P when V and P have high-specified values. At any rate, the number of unrealistic precipitation values which resulted from the generation process, for V specified at 0.03, was unacceptable (see Table 2). Therefore the third row of sets will not normally be used in a true generation of radar fields.

TABLE 2. Number of Generated Rainfall Values that Exceed 50 mm/h

	$NS = 10$			$NS = 25$			$NS = 50$		
$1/h$, km	4	12	20	4	12	20	4	12	20
V_0									
0.005	1	1	0	3	2	0	5	4	1
0.010	1	1	0	11	7	3	27	21	15
0.030	32	41	36	92	113	99	193	232	219

SUMMARY AND CONCLUSIONS

A method for generation of radar precipitation fields was described. The method works by imposing a noise field on high-quality radar rainfall fields. The noise parameters are determined based on a set of conditions pertaining to the resultant field. In that way, nonstationary, nonergodic fields can be simulated.

Since the original and the "observation" (original and noise) fields are known, the method can be used in the validation procedures of various hydrologic models (radar and rain gage data merging, mean areal precipitation estimation, rainfall-runoff). The example given shows that the accuracy of the preservation of the required statistics is very good, especially for realistic values of the variance measure ($V < 0.01$), even for a relatively small number of realizations ($NS \leq 25$). The method proposed is flexible in that one can generate fields with a wide range of second-order statistics from one high-quality radar field.

When the technique is used to investigate radar and rain gage data merging, a procedure is required to synthesize the gage data. The authors are investigating techniques to generate gage values in work under preparation.

Acknowledgments. The work was performed under the National Research Council and National Oceanic and Atmospheric Administration sponsorship, while both authors were NRC-NOAA Research Associates. The comments of M. D. Hudlow and E. R. Johnson of the Hydrologic Research Laboratory of the National Weather Service (HRL/NWS), J. A. Smith of the Interstate Commission for the Potomac River Basin, R. L. Bras and A. Mantoglou of M.I.T., and L. T. Iseley of HRL/NWS are gratefully acknowledged.

REFERENCES

Brandes, E. A., Optimizing rainfall estimates with the aid of radar, *J. Appl. Meteorol., 14*, 1339–1345, 1975.

Chow, V. T. (Editor), *Handbook of Applied Hydrology*, pp. 9–47, McGraw-Hill, New York, 1964.

Collier, C. G., P. R. Larke, and B. R. May, A weather radar correction procedure for real-time estimation of surface rainfall, *Quart. J. R. Meteorol. Soc., 109*, 589–608, 1983.

Crawford, K. C., Considerations for the design of a hydrologic data network using multivariate sensors, *Water Resour. Res.*, *15*(6), 1752–1762, 1979.

Eddy, A., Objective analysis of convective scale rainfall using gages and radar, *J. Hydrol.*, *44*, 125–134, 1979.

Harrold, T. W., E. J. English, and C. A. Nicholas, The Dee weather radar project: The measurement of area precipitation using radar, *Weather*, *28*, 332–338, 1973.

Hudlow, M. D., and V. L. Patterson, GATE radar rainfall atlas, *NOAA Spec. Rep.*, p. 155, U.S. Dep. of Commer., Washington, D.C., 1979.

Hudlow, M. D., R. Arkell, V. Patterson, P. Pytlowany, and F. Richards, Calibration and intercomparison of the gate C-band radars, *NOAA Tech. Rep. EDIS 31*, p. 98, U.S. Dep. of Commer., Washington, D.C., 1979.

Krajewski, W. F., and K. P. Georgakakos, A technique for the generation of synthetic radar-rainfall data, *Tech. Note 9*, Hydrol. Res. Lab., Natl. Weather Serv., Silver Spring, Md., 1985.

Mantoglou, A., and J. L. Wilson, Simulation of random fields with the turning bands method, *Rep. 264*, Ralph M. Parsons Lab., Hydrol. and Water Resour. Syst., Dep. of Civ. Eng., MIT, Cambridge, Mass., 1981.

Mantoglou, A., and J. L. Wilson, The turning bands method for simulation of random fields using line generation by a spectral method, *Water Resour. Res.*, *18*(5), 1379–1394, 1982.

Mejia, J., and I. Rodriguez-Iturbe, On the synthesis of random fields from the spectrum: An application to the generation of hydrologic spatial processes, *Water Resour. Res.*, *10*(4), 705–711, 1974.

Vanmarcke, E., *Random Fields: Analysis and Synthesis*, MIT Press, Cambridge, Mass., 1983.

Wilson, J. W., and E. A. Brandes, Radar measurement of rainfall—A summary, *Bull. Am. Meteorol. Soc.*, *60*(9), 1048–1058, 1979.

K. P. Georgakakos and W. F. Krajewski, Hydrologic Research Laboratory, National Weather Service, National Oceanic and Atmospheric Administration, 8060 13th Street, Silver Spring, MD 20910.

(Received June 20, 1984;
revised January 29, 1985;
accepted February 6, 1985.)