Journal of Hydrology, 44 (1979) 97—123

[1]

REAL-TIME ESTIMATION OF VELOCITY AND COVARIANCE STRUCTURE OF RAINFALL EVENTS USING TELEMETERED RAINGAGE DATA — A COMPARISON OF METHODS

EDWARD R. JOHNSON and RAFAEL L. BRAS*

National Weather Service, Hydrologic Research Laboratory, Silver Spring, MD 20910 (U.S.A.)
Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139 (U.S.A.)

(Received November 2, 1978; revised and accepted March 9, 1979)

ABSTRACT

Johnson, E.R. and Bras, R.L., 1979. Real-time estimation of velocity and covariance structure of rainfall events using telemetered raingage data — a comparison of methods. J. Hydrol., 44: 97—123.

Short-term rainfall prediction in time and space requires parameter estimation in real time. This paper discusses several methods of using telemetric raingage data to estimate the non-stationary mean and variance of a rainfall event. Similarly, real-time estimation of residual covariance structure and storm velocity is discussed. The performance of the suggested techniques are illustrated with examples. The parameter estimation corresponds to a forecasting model developed by Johnson and Bras. The goal of this article is to inform other researchers of the advantages and disadvantages of different methodologies for real-time estimation of rainfall statistical characteristics.

INTRODUCTION

To implement a rainfall prediction scheme of the type presented by Johnson and Bras (1978), it is necessary to describe the structure of the event. For a stochastic model, a minimal description requires the mean function and the covariance function.

Throughout this paper, it is assumed that a real-time forecasting model of the following general form is being used:

$$\mathbf{i}(t) = \mathbf{m}(t) + \mathbf{r}(t) \tag{1}$$

where $\mathbf{i}(t)$ = vector of N rainfall rates at time t; $\mathbf{m}(t)$ = vector of mean values at time step t; $\mathbf{r}(t)$ = vector of residuals at time t; and N = number of raingages in the prediction scheme.

There are two basic statistical approaches to estimating rainfall structure. A multi-realization approach assumes that the current rainfall event is a member of a class of rainfall events all with the same mean and variance function.

*To whom correspondence should be addressed.

Taking a multi-realization approach requires a scheme for classifying events and identifying in real time the class of each particular event.

A single-realization approach assumes that each event is unique. As a result, no scheme is necessary to classify events and no error is introduced by possible misclassification in real time. All of the techniques investigated in this paper are based on a single-realization assumption.

Different sources of data may make different estimation procedures appropriate. This work assumes that information is collected from a network of raingages which are interrogated at regular intervals.

ESTIMATION OF MEAN AND VARIANCE

Rainfall is a complex process; rates vary both in time and space. The problem addressed here is to estimate the expected value of the rainfall rate at particular locations at particular times, and the variance about that expected value. During rainfall prediction it will be necessary to estimate future statistics of the rainfall event. Throughout this paper the term "estimation" will be used for past, present and future statistics.

It is apparent that some external information or assumption is required to allow any estimation. If the mean value is allowed to vary arbitrarily over space and time, then only one sample is available at any (x,y,t) and no estimation is possible. In this section, several alternative assumptions for the structure of the mean and variance function are examined. In general, the discussion will focus on estimation of the mean, although the final form of the estimation procedure will be used to estimate both the mean and variance.

The simplest assumption would be that the mean and variance are stationary for the duration of the rainfall event, i.e.:

$$m(x,y,t) = m \qquad \text{and} \qquad \sigma(x,y,t) = \sigma \tag{2), (3}$$

Assuming stationarity greatly simplifies the problem of estimating the mean and variance functions; every sampled value for the current storm can be used in the estimation procedure. If rainfall is stationary in the mean for the duration of the event, then the rainfall rate has a constant expected value throughout the storm.

Unfortunately, evidence suggests that rainfall is not stationary in time. It is generally accepted that storms of a certain type (e.g., thunderstorms) tend to have the heaviest rainfall in a particular part of the storm (e.g., the first third). This behavior is clearly evident on average percentage mass curves for the particular storm type (see Eagleson, 1970). If rainfall rates were stationary in the mean, then the average percentage mass curve should be a straight line, not the strongly curved forms usually observed.

Using percentage mass curves to estimate the mean function of a storm in space and time will require many assumptions about storm structure and type (see Bras and Rodríguez-Iturbe, 1976). It would require estimation of total event depth and duration, a storm classification scheme, and specification of the spatial structure of the event as well.

For real-time rainfall prediction, it seems reasonable to let the actual event carry as much weight as possible by imposing minimal assumptions regarding the spatial or temporal structure. As a rule, only those parameters which are required should be estimated, and these should be estimated in a simple and direct fashion.

The time axis of the mass curve at a point is measured relative to the local start of the storm at that point. This suggests that the temporal structure of the mean at several points be examined relative to the local start of the storm at each point, not on an absolute "wall clock" time scale.

To a large extent, the spatial structure of the storm is specified by the pattern of storm arrival at various points. In real-time estimation the storm itself defines the arrival at each gaged location, so it may be possible to avoid imposing any strong assumption about the spatial structure of the event. The technique developed to estimate the mean and variance vectors draws heavily on the above ideas.

The technique used to estimate the non-stationary mean and variance of rainfall is called the storm counter method. It depends on one fundamental assumption: relative to the time that a storm arrives at a location, the time history of the mean and variance is identical for all locations. It is not assumed that the storm will arrive at all locations, e.g., convective cells might completely miss some points.

Data from raingages are recorded in discrete time steps. The storm counter counts the number of time steps since the start of rainfall at each raingage. The mean and variance can be estimated for each storm counter using the common statistical expression for the sample mean and variance:

$$\hat{m}(s) = \frac{1}{n(s)} \sum_{\substack{i=1 \\ t(i,s)\neq 0}}^{N} q_i[t(i,s)] \tag{4}$$

where

$\hat{m}(s)$ = estimate of mean rainfall rate for storm counter s

$n(s)$ = number of gages with storm counter at least as high as s, i.e., sample size for storm counter s

$t(i,s)$ = indexing function giving time step for storm counter s at gage i; if storm counter s at gage i is not available $t(i,s) = 0$

$q_i(t)$ = measured rainfall rate at gage i at time step t

N = number of gages in raingage network

$$\hat{\sigma}^2(s) = \frac{1}{n(s)-1} \sum_{\substack{i=1 \\ t(i,s)\neq 0}}^{N} [q_i\{t(i,s)\} - \hat{m}(s)]^2 \tag{5}$$

where

$\hat{\sigma}^2(s)$ = estimated variance of rainfall rate for storm counter s

The sample size $n(s)$ may vary considerably from one storm counter to another.

Certain effects are ignored by the storm counter method. For example, the storm counter method does not consider orographic effects or storm aging (i.e., variations of the storm in real time). The basic idea in using a non-stationary mean is that the mean function will capture some of the structure of the event. Ignoring some structural details places a heavier burden on the stochastic component. By ignoring orographic effects or storm aging, the storm counter method produces higher estimates of the non-stationary variance than a model which accurately includes these effects. However, an inaccurate orographic model might have higher variance than the simple storm counter method, likewise for an inaccurate model of any type of storm structure.

Usually, the sample size will vary from one storm counter to another. (The exception is the case where rainfall starts simultaneously at all locations.) In all cases, there will be one or more gages where rainfall first occurs, and these gages will have the highest storm counters. Storm counters greater than the highest recorded storm counter will have no sample at all. Other storm counters might have a very small sample. Nevertheless, an estimate of the mean and variance for a storm counter with no sample may be required.

What should be used to estimate the mean and variance for a storm counter with a small (or zero) sample? The only reasonable approach is to include the sample from nearby storm counters until an acceptable sample size is reached. What is an acceptable sample size? The problem is to set a lower limit on the sample size.

If a storm counter (n) has a sample size smaller than the minimum allowed, the sample values for the next higher and next lower storm counters are included ($n \pm 1$). The process is repeated ($n \pm 2$, $n \pm 3, \ldots$) until the minimum sample size is achieved.

A high value for the minimum sample forces all rainfall values to be included in the mean and variance estimate. This is equivalent to assuming stationarity because the mean and variance estimates would be the same for every storm counter.

In order to estimate the mean or variance at any point using the storm counter method, it is necessary to know the value of the storm counter for that point. If it has begun to rain, the storm counter is simply the number of time steps since it started to rain. If it has not started to rain, it is necessary to estimate the time when it will begin to rain. In the storm counter framework, estimation of storm arrival is equivalent to estimation of the value of negative storm counters.

Several techniques were investigated by the authors (Johnson and Bras, 1978) to estimate storm arrival. They all have in common a fairly high uncertainty in the estimated arrival time. The methods fall into two broad categories: storm velocity methods and regression methods.

Storm velocity estimation is discussed later, nevertheless storm velocity defines an "upwind" direction from the point where storm arrival is to be estimated. Even in a very dense raingage network it is very unlikely that another gage will be precisely upwind of another, so that a path of some finite width must be searched upwind for a gage where the storm has arrived. Such a scheme was tested for a variety of rainfall events and the path width varied to maximize the accuracy of the storm arrival estimate. For these tests, no path width could be found that gave accurate estimates. Furthermore, no estimate of the uncertainty of storm arrival is provided by such a scheme.

Regression techniques do provide a measure of the uncertainty of storm arrival estimates, and primarily for this reason a regression method was selected as the methodology to be used in estimating storm arrival. In addition, for the test cases examined, the chosen regression was more accurate than the storm velocity method. If radar measurements were available, the performance of the storm velocity method should be greatly improved.

The regression methodology follows. Since it has rained at some locations, there will be some points with positive storm counters. These points can be used to estimate the parameters of an equation:

$$L_i = a_0 + a_1 x_i + a_2 y_i + \epsilon_i \tag{6}$$

where L_i = storm counter at the ith gage; x_i, y_i = location of the ith gage; and ϵ_i = error at the ith gage.

The estimated parameters can be used to estimate the storm counter at gages where it has not yet rained. Presumably, the predicted storm counters will not be positive at locations where it has not yet rained. It should be emphasized that the regression method requires an extrapolation of the function chosen, i.e., parameters will be estimated from gages with positive storm counters and extrapolated to locations with negative storm counters.

Naturally, eq. 6 will describe storm arrival better for some types of events (e.g., frontal storms) than for others (e.g., convective cells). In essence, it is difficult to predict the arrival of a "rough" or cell-type storm, and this difficulty will be reflected in the goodness of fit of eq. 6. The model accounts for the uncertainty in the predicted storm arrival in the following way.

The mean value (and variance) at a point where it has not rained is found as a weighted sum of the mean values (variance) from a number of storm counters. Eq. 6 provides an estimate, L_i, of the storm counter value at the ith gage. Also, the standard error of estimate of eq. 6 can be found, $E[\epsilon_i^2]^{\frac{1}{2}}$. Rather than estimate the mean and variance of the ith gage from the single storm counter nearest to L_i, all storm counters within two standard errors of L_i are used. The means and variances of these various storm counters are not given equal weight, the weights are proportional to a normal density centered on L_i with a variance of $E[\epsilon_i^2]$. It is known that it has not begun to rain at the ith gage, therefore positive values of the storm counter at the ith gage are not used in the weighting scheme.

Occasionally, it may happen that all the values within two standard errors of eq. 6 are positive, i.e., it should already be raining. In these cases, the storm is considered to have "missed" the point; the storm counter is considered to be $-\infty$, i.e., the storm will never arrive.

The effect of using a weighted mean and variance instead of the estimates from a single storm counter is to "hedge" the uncertainty in storm arrival. When eq. 6 fits poorly, more storm counter values will be included in the weighted mean.

Unless it begins to rain simultaneously at all gages in the network, the gages will have a variety of storm counter values. Therefore, each gage could have a mean and variance at each time step that is different from (all) other gages.

It is difficult to formulate a direct test of the procedure to estimate the non-stationary mean and variance. Instead, it is proposed to examine the hypothesis that the normalized residuals have a mean value of zero at each time step. If forced to reject this hypothesis, some doubt is thrown on the storm counter method in the sense that it does not capture the structure of the storm. On the other hand, accepting this hypothesis does not prove the optimality of the storm counter method — another technique may exist that produces a zero mean field with lower variance. In short, the testing method is not definitive, it is simply a test of whether the storm counter method is reasonable. The mathematics are developed below.

Define:

$$\hat{\epsilon}_{jt} = (q_{jt} - \hat{m}_{jt})/\hat{\sigma}_{jt} \tag{7}$$

where

$\hat{m}_{jt}$ = estimate of mean at gage j at time step t
$\hat{\sigma}_{jt}$ = estimate of standard deviation at gage j at time step t
q_{jt} = rainfall rate (measured) at gage j at time step t
$\hat{\epsilon}_{jt}$ = estimate of normalized residual at gage j at time step t

There is a problem in defining $\hat{\epsilon}_{jt}$ using eq. 7, because $\hat{\sigma}_{jt}$ might be zero. For now, $\hat{\epsilon}_{jt}$ will be defined equal to zero in this case. Let:

$$\hat{M}_t = \frac{1}{N} \sum_{j=1}^{N} \hat{\epsilon}_{jt} \tag{8}$$

where $\hat{M}_t$ = estimate of spatial mean of normalized residuals at time step t; and N = number of raingages in network.

The hypothesis to be examined is $M_t = 0$ where M_t is the true value of the spatial mean of the normalized residual at time t. Therefore, the variance of $\hat{M}_t$ is needed:

$$\operatorname{var} \hat{M}_t = E(\hat{M}_t - M_t)^2 = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} E[\hat{\epsilon}_{it}\hat{\epsilon}_{jt}] \tag{9}$$

The terms $E[\hat{\epsilon}_{it}\hat{\epsilon}_{jt}]$ are residual covariances. They can be computed using the residual covariance function which is described below.

The mean of the normalized residuals can be estimated at every time step. These estimates are not independent. It is possible to compute the lag-one covariance of the mean estimate:

$$\operatorname{cov}(\hat{M}_t\hat{M}_{t+1}) = E[(\hat{M}_t - M_t)(\hat{M}_{t+1} - M_{t+1})] = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} E[\hat{\epsilon}_{it}\hat{\epsilon}_{jt+1}] \quad (10)$$

One of the assumptions to be made in estimating the residual covariance function is stationarity. As a result of this assumption, var M_t and cov(M_tM_{t+1}) are constant with time as long as N is held constant. Therefore, a lag-one correlation coefficient of the estimates of the mean of the normalized residuals is defined:

$$\rho_M = \operatorname{cov}(\hat{M}_t\hat{M}_{t+1})/\operatorname{var}\hat{M}_t \quad (11)$$

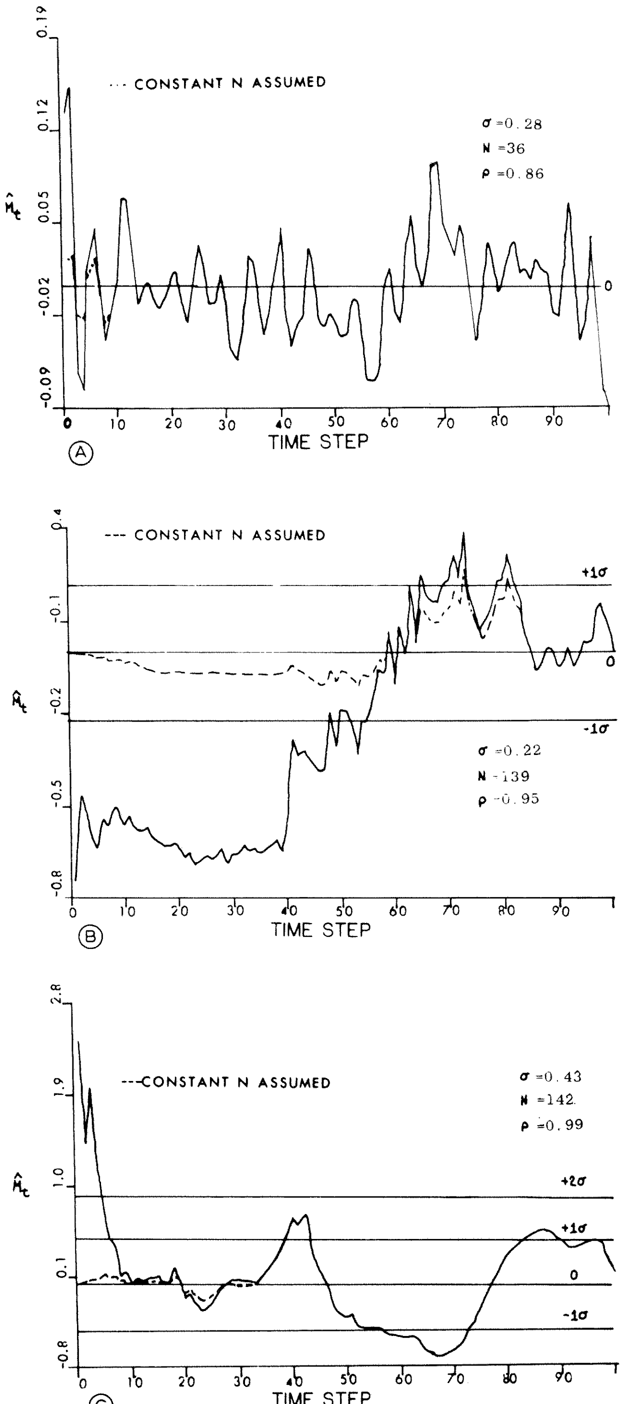
Returning to eq. 7, there is a problem in defining $\hat{\epsilon}_{jt}$ for time steps before the storm arrives at gage j. The correct approach would be to consider these values undefined and reduce the sample size N to include only points where the storm has begun. This would give a sample size that varies from time step to time step and therefore the variance of the sample mean, var $\hat{M}_t$, and its lag-one covariance, cov$(\hat{M}_t\hat{M}_{t+1})$, would be time-varying. To simplify computation of var $\hat{M}_t$ and ρ_M, the sample size N is held constant in eqs. 9 and 10. Every sample point used to estimate $\hat{M}_t$ has the same variance and no sample value ($\hat{\epsilon}_{jt}$) is perfectly correlated with any other sample value. As a result, every sample value contributes to reducing var $\hat{M}_t$ and the variance shown is a lower bound on var $\hat{M}_t$ at each time step.

Fig. 1 shows $\hat{M}_t$ vs. time for five storms. Two values of $\hat{M}_t$ are shown at each time step: (1) the value that keeps a constant sample size by assuming that unknown $\hat{\epsilon}$ are all zero (a biased estimate); (2) the value with a time-varying sample size. The two values become identical when it has rained at every gage. Of course, the biased estimate is always closer to zero.

All figures show the number of gages which eventually have rain (N) and the standard deviation of $\hat{M}_t$ (σ). The lag-one correlation of $\hat{M}_t$ is also reported in some cases. The computed correlations are quite high, indicating that a high value of $\hat{M}_t$ is very likely to be followed by another high value.

The storms are taken from three data sources. The synthetic storm of Fig. 1A was created by a rainfall generator (Bras and Rodríguez-Iturbe, 1976). The other four events are true rainfall data collected from two dense raingage networks. The time steps in all cases are of 5-min durations.

The METROMEX network is the data source for the storms in Fig. 1D and E. There are over 200 gages in this network on a grid of approximately 3 km spacing. The gages are standard weighing-bucket gages — nearly all with 24-h charts. The Chickasha network has over 150 weighing-bucket gages on a grid of approximately 5 km spacing, and it is the source of storms of Fig. 1B and C. (The six digit number for the storm represents the date, i.e., MM081575 is the storm of August 15, 1975.)


··· CONSTANT N ASSUMED
σ =0.28
N =36
ρ =0.86
$\hat{M}_t$
0.19
0.12
0.05
-0.02
-0.09
0
0
10
20
30
40
50
60
70
80
90
TIME STEP
A
--- CONSTANT N ASSUMED
+1σ
0
-1σ
σ =0.22
N =139
ρ =0.95
$\hat{M}_t$
0.4
-0.1
-0.2
-0.5
-0.8
0
10
20
30
40
50
60
70
80
90
TIME STEP
B
---CONSTANT N ASSUMED
σ =0.43
N =142
ρ =0.99
+2σ
+1σ
0
-1σ
$\hat{M}_t$
2.8
1.9
1.0
0.1
-0.8
0
10
20
30
40
50
60
70
80
90
TIME STEP
C

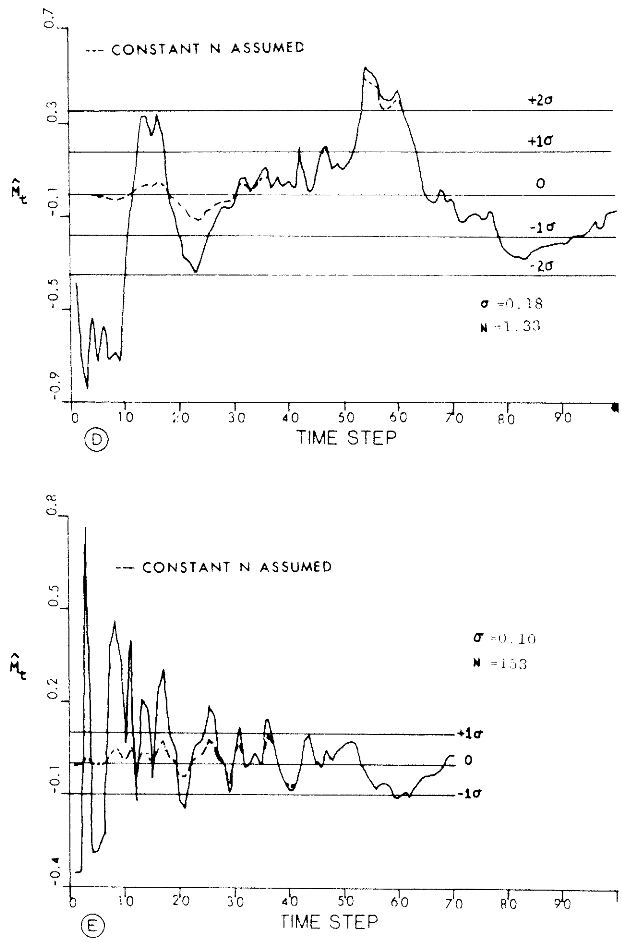


Fig. 1. Mean of residuals: (A) synthetic storm; (B) storm CK060162; (C) storm CK041167; (D) storm MM080175; and (E) storm MM082575.

As expected, the storm counter method performs better for some cases than for others. Several of the storms seem to have a period of about 1-h duration when rainfall rates are higher over the entire network than the value expected from the storm counter method — notably storm MM080175 (Fig. 1D). Early in the storm, the sample size N can be quite small. Storm CK060162 (Fig. 1B) has many time steps of very light rain at a small number of gages preceding a period of heavier more general precipitation.

The storm counter method is not always completely successful in capturing the storm structure as indicated by those storms where $\hat{M}_t$ is significantly greater than 0.0 for some period of time. On the other hand, normalizing by the mean and variance produced by the storm counter method does remove a great deal of the structure of the rainfall event. Again, this technique of testing cannot show the optimality of the storm counter method, but it does show that the technique is consistent and does identify some of the structure of the event.

RESIDUAL COVARIANCE FUNCTION

The mean and variance of a rainfall event can be estimated in real time by the storm counter method. To complete the definition of rainfall structure, it is necessary to specify the residual covariance structure. In an off-line posterior analysis, the residual covariance between two raingage locations can be statistically estimated, using the $\hat{\epsilon}$ of eq. 7. However, a real-time analysis for prediction will certainly require estimates of the residual covariance when no sample values ($\hat{\epsilon}$) are known. As a result, it is necessary to assume a functional form for the residual covariance.

An attempt was made to avoid assuming a particular covariance function. An estimation procedure based on smoothing sample covariances was used. A moving average over a range of storm distances was implemented. The problem with this approach is that it may violate the required positive-definiteness property of the covariance matrix. Some effort at explicitly including a positive-definiteness constraint was made, but no feasible estimation procedure could be developed.

It is reasonable to assume that for any time lag τ, the covariance will have a maximum value at some particular offsets $[\Delta x_{max}(\tau), \Delta y_{max}(\tau)]$. For lag zero, $\Delta x_{max} = \Delta y_{max} = 0$, i.e., the covariance function has a maximum at the origin. If the storm is moving, it should be less variable in the coordinate system that moves with the storm than in a coordinate system fixed to the ground. This argues that relative maxima will all be in a straight line, i.e.:

$$\Delta x_{max}(\tau) = U_x \tau \qquad \text{and} \qquad \Delta y_{max}(\tau) = U_y \tau \tag{12}$$

where U_x = x-direction component of storm velocity; and U_y = y-direction component of storm velocity.

There is evidence of this behavior in rainfall (Zawadski, 1973; Marshall, 1975).

Eq. 12 assumes that the storm velocity components are constant. If the storm velocity varies with time, the effect on the covariance is quite complex. A variable storm velocity does more than simply relocate the covariance maxima at each lag so that they are no longer colinear. A variable storm velocity implies that the location of covariance maxima are a function of absolute time, not just time lag. In short, variable storm velocity implies non-stationarity of the covariance which would make covariance estimation difficult. Furthermore, for a rainfall prediction scheme, it would be necessary to either propose the form of future variation in the storm velocity or assume that it will remain constant at the value of the last estimate.

If the covariance decreases uniformly in all directions from the maxima defined by eq. 12, it is isotropic in a coordinate system that moves with the storm. It is important to realize that isotropy in the moving coordinates implies anisotropy in a fixed coordinate system with the storm's movement as the source of this anisotropy.

The distance from the maxima of eq. 12 is measured by the storm distance d. Define:

$$\begin{aligned} d(x_i,y_i,t_1;x_j,y_j,t_2) &= [\{(x_i - U_x t_1) - (x_j - U_x t_2)\}^2 \\ &\quad + \{(y_i - U_y t_1) - (y_j - U_y t_2)\}^2]^{\frac{1}{2}} \\ &= [\{\Delta x_{ij} - U_x \tau_{12}\}^2 + \{\Delta y_{ij} - U_y \tau_{12}\}^2]^{\frac{1}{2}} = d_{ij\tau_{12}} \end{aligned} \qquad (13)$$

where

$\Delta x_{ij} = x_i - x_j$; $\quad \Delta y_{ij} = y_i - y_j$; $\quad$ and $\quad \tau_{12} = t_1 - t_2$

In essence, the use of the storm distance reduces the dimensionality of the estimation. The covariance function used for this study is the exponential.

$$E[\epsilon(x_i, y_i, t+\tau)\ \epsilon(x_j,y_j,t)] = \alpha_\tau \exp(-\beta_\tau\, d_{ij\tau}) \qquad (14)$$

At each lead τ, two parameters must be estimated, α_τ and β_τ. Also required are estimates of U_x and U_y, the storm velocity components.

Clearly, the ideal approach would be to simultaneously estimate all the covariance parameters, U_x, U_y, α_0, β_0, α_1, and β_1 for a rainfall prediction scheme with only one prediction lead value. In a multi-lead prediction format, other height and decay parameters must be added to the list. The 6-parameter (or 8- or 10- or more) non-linear estimation problem is formidable, especially when it must be solved in real time.

The problem can be greatly simplified if independent estimates of U_x and U_y are available. The estimation of storm velocity is discussed below. With U_x and U_y assumed known, the covariance parameters (α_τ and β_τ) can be estimated separately for each lag τ. The following constraints must be imposed on α_τ and β_τ to insure the validity of eq. 14, $\alpha_\tau \geqslant 0$; $\alpha_\tau \leqslant \alpha_0$; $\beta_\tau > 0$.

The next section describes the procedure used to estimate the covariance parameters, α_τ and β_τ.

Covariance estimation

The first step in estimating the covariance parameters is to create the sample values. These sample values are estimated by the following equation:

$$\hat{c}_{ij\tau} = \frac{1}{T_1(i,j)} \sum_{t=T_2(i,j,\tau)}^{T_3(i,j,\tau)} \hat{\epsilon}_{it+\tau}\, \hat{\epsilon}_{jt} \qquad (15)$$

where

$\hat{c}_{ij\tau}$ = estimates of covariance between gage i and gage j at lag τ

$\hat{\epsilon}_{it}$ = estimated normalized residual at gage i at time step t (see eq. 7)

$T_1(i,j)$ = length of shortest record at gage i or j, i.e., maximum number of data values included in $\hat{c}_{ij\tau}$ at any lag τ

$T_2(i,j,\tau)$, $T_3(i,j,\tau)$ = start and end time steps so that all $\hat{\epsilon}_{it+\tau}$ and $\hat{\epsilon}_{jt}$ in eq. 15 exist

The quantity $T_3 - T_2$ gives the number of data values included in a particular estimate $\hat{c}_{ij\tau}$. Define:

$$n_{ij\tau} = T_3(i,j,\tau) - T_2(i,j,\tau) \tag{16}$$

Eq. 15 is the usual sample covariance estimate (recognizing that $\hat{\epsilon}_{it}$ has a zero mean). Using the exponential covariance function gives

$$\hat{c}_{ij\tau} = \hat{\alpha}_\tau \exp(-\hat{\beta}_\tau d_{ij\tau}) + \eta_{ij\tau} \tag{17}$$

where $\eta_{ij\tau}$ = error in estimating $\hat{c}_{ij\tau}$ using eq. 15 plus effects of errors in estimation of $\hat{\alpha}_\tau$ and $\hat{\beta}_\tau$; $\hat{\alpha}_\tau$ = estimated value of α_τ; and $\hat{\beta}_\tau$ = estimated value of β_τ.

The problem is to estimate α_τ and β_τ so that the error terms $\eta_{ij\tau}$ are as small as possible. Presumably, the values of $\hat{c}_{ij\tau}$ corresponding to large $n_{ij\tau}$ are more accurate on the average, so they should receive a higher weight in the procedure used to estimate α_τ and β_τ. A weighted sum of squares is formed:

$$\phi(\hat{\alpha}_\tau, \hat{\beta}_\tau) = \sum_{\substack{\text{all } i,\\ \text{all } j,\\ \pm\tau}} n_{ij\tau}[\hat{c}_{ij\tau} - \hat{\alpha}_\tau \exp(-\hat{\beta}_\tau d_{ij\tau})]^2 \tag{18}$$

The values of $\hat{\alpha}_\tau$ and $\hat{\beta}_\tau$ are defined by:

$$\hat{\alpha}_\tau, \hat{\beta}_\tau = \underset{\hat{\alpha}_\tau\ \hat{\beta}_\tau}{\text{Min}}\ \phi(\hat{\alpha}_\tau, \hat{\beta}_\tau) \tag{19}$$

The solution of eq. 19 is a straightforward problem in nonlinear optimization which can be solved with a Newton—Raphson method.

In order to test the covariance estimation, a process with known covariance is needed. A rainfall synthesis model (Bras and Rodríguez-Iturbe, 1976) was used to generate a rainfall event. The event has a non-stationary mean and variance. It moves at constant velocity. It has an exponential covariance function of the residuals with true lag-zero parameters $\alpha_0 = 1.0$ and $\beta_0 = 0.5$. The event was sampled at 5-min time step at 36 "gages" in a 6×6 uniform grid with grid size 2.0 mi. The storm counter method described above was used to estimate the means and variances. The estimated covariance parameters of the residuals $\hat{\alpha}_0$ and $\hat{\beta}_0$ for each time step are shown in Fig. 2.

Estimation of α is very stable, as expected. No estimate of β is possible for the first ten steps because the sample size is too small, so the default value

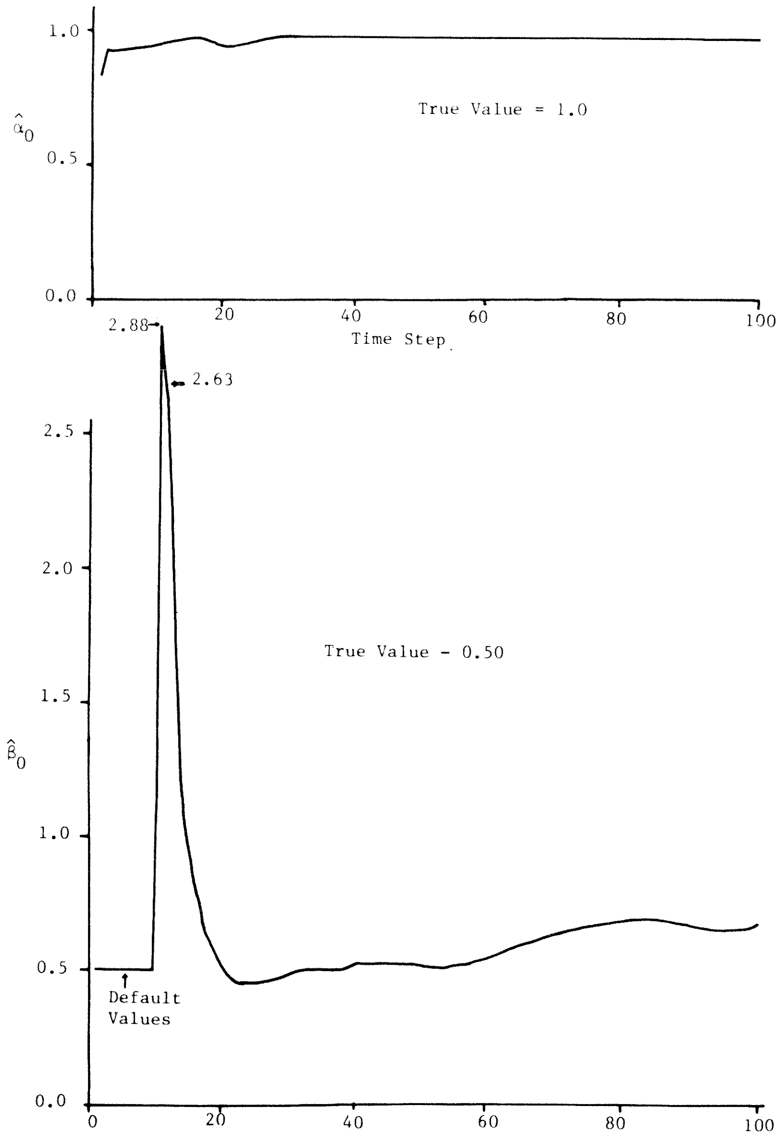


Fig. 2. Covariance estimation test.

(0.5) is shown. The first four estimates of β are far off, but the value stabilizes quickly.

It should be emphasized that $\hat{\alpha}_0$ and $\hat{\beta}_0$ are dependent on the procedure used to estimate means and variance. Covariance parameters at higher lags are also related to the storm velocity.

DEFINING STORM VELOCITY

It is not simple to define the storm velocity. Rainfall rates form a complex pattern that both moves and changes with time. Due to changes in the storm, the movement of one function of the pattern (e.g., the point of maximum rainfall rate) may not be the same as the movement of some other function of the pattern (e.g., the centroid). The storm velocity is not a physical parameter. It has to be defined as some function of rainfall rates, and different mathematical definitions of the storm velocity produce different "storm velocities."

No matter how the "true" storm velocity is defined mathematically, it is necessary to develop a storm velocity estimator that relies on the available rainfall data.

Tracking techniques

The most intuitive approach to storm velocity estimation is to identify the location of some feature of the storm pattern at a specified time and see how far it moves in the next time step. In short, the idea is to track the movement of some storm feature.

The common problem shared by all tracking techniques is that the data must be adequate to identify accurately the feature being tracked. Certainly any source of rainfall data will have some amount of inaccuracy, but there are two additional fundamental deficiencies of rainfall data that affect the accuracy of tracking techniques — lack of spatial coverage and resolution.

The data collection system, whether composed of radar, raingages, or both, covers only a limited area. As a result, some part of the data collected at each time step is from a "new" part of the storm as it moves into the data collection area. Certain tracking techniques are susceptible to mismatching the new part of the storm pattern. Basically, some techniques assume or imply that the unobserved parts of the storm pattern all have zero rainfall.

To illustrate the effect of lack of coverage on various tracking techniques, a one-dimensional example storm is created. The example storm is shown in Fig. 3. The rainfall rate at time step 1 is described by $q_1(x)$. The pattern moves 0.2 km between time step 1 and 2, i.e., $q_1(x) = q_2(x + 0.2)$ and the true storm velocity is 0.2 km per time step. The example storm is measured from $x = -0.1$ to $x = +0.9$; the measured storm is shown in Fig. 3b. Future sections will use the above storm to illustrate problems associated with several sampling techniques.

As discussed previously, the only data source considered in this paper is raingages. This places a serious restriction on the spatial resolution of the storm pattern. Some tracking techniques that may work well for radar data will fail using raingage data, even from dense raingage networks.

A variety of tracking techniques are briefly discussed below. Certain of these methods were tested on one or more synthesized rainfall events. The

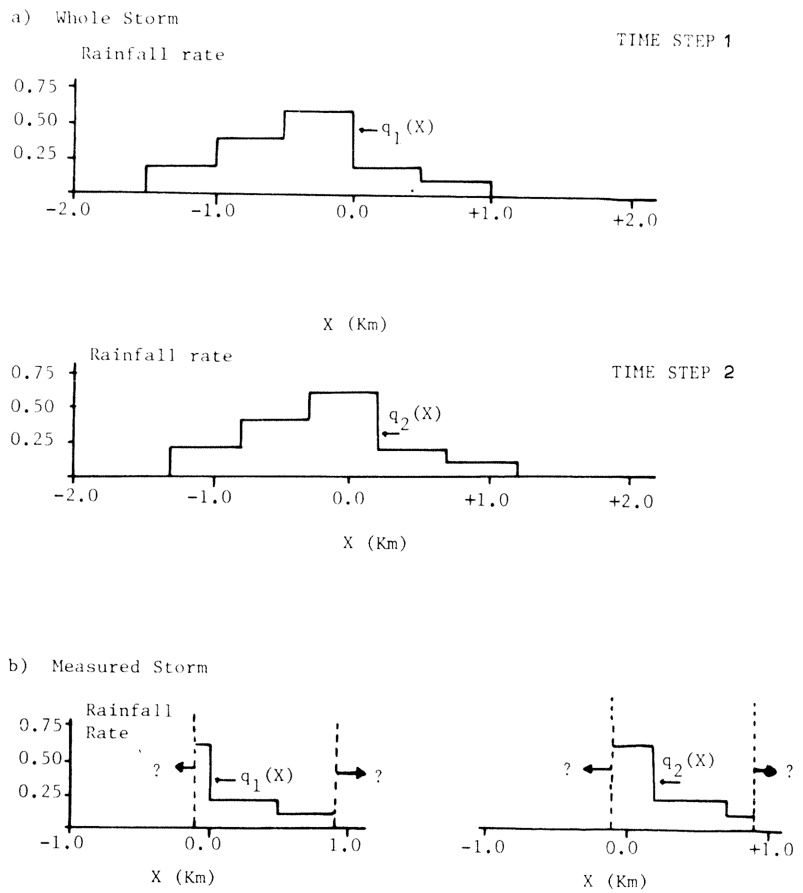


Fig. 3. One-dimensional example storm.

rainfall generator (Bras and Rodríguez—Iturbe, 1976) used to synthesize the various events creates a frozen rainfall pattern that can be moved at any speed and direction over a raingage network. Only the most promising techniques were tested extensively.

The lack of resolution implicit in raingage data makes it impossible to track a single feature of the raingage pattern such as the point of heaviest rainfall or the center of a cell (however defined). Even with high-resolution data, it is better to track the whole storm pattern in some way and thus use all available information.

Centroid. In this approach, the storm movement is defined by the movement of the centroid:

$$X(t) = \iint_A xq(x,y,t)\,\mathrm{d}x\mathrm{d}y \Big/ \iint_A q(x,y,t)\,\mathrm{d}x\mathrm{d}y \qquad (20)$$

where $q(x,y,t)$ = rainfall rate; $X(t)$ = location of centroid in x-direction (E—W) at time t; and A = data collection area.

The velocity is defined by:

$$U_x = [X(t_2) - X(t_1)]/(t_2 - t_1) \tag{21}$$

where U_x = storm velocity component in x-direction. The y-component of storm velocity is similarly defined.

Unfortunately, it is possible for the movement of the estimated centroid to be the opposite of the movement of the true centroid. A new part of the storm enters the data collection network at each time step. If the new part of the storm has a heavy rainfall rate, it can move the centroid estimate in a direction contrary to the storm movement.

Applying the centroid tracking technique to the one-dimensional storm (Fig. 3) shows exactly this behavior. The centroids of the measured storm pattern at time steps 1 and 2 are at x = 0.250 km and x = 0.233 km respectively, giving a storm velocity estimate of −0.017 km per time step. The true velocity is +0.200 km per time step. The centroid of the entire storm moves from −0.383 km to −0.183 km, giving a correct velocity estimate.

Composite matching functions. Three functions have been investigated to define how well rainfall patterns match at two time steps:

$$C_1(\alpha,\beta) = \frac{1}{S} \iint_S q(x,y,t_1)q(x+\alpha,\ y+\beta,\ t_2)\,\mathrm{d}x\mathrm{d}y \tag{22}$$

$$C_2(\alpha,\beta) = \frac{1}{S} \iint_A [q(x,y,t_1) - q(x+\alpha,\ y+\beta,\ t_2)]^2\,\mathrm{d}x\mathrm{d}y \tag{23}$$

$$C_3(\alpha,\beta) = \frac{1}{S} \iint_S \left| q(x,y,t_1) - q(x+\alpha,\ y+\beta,\ t_2)\right|\,\mathrm{d}x\mathrm{d}y \tag{24}$$

where $q(x,y,t)$ = rainfall rate at time t at location (x,y); and S = area compared.

The basic procedure is to optimize the match defined by C_1, C_2 or C_3 to find (α^*,β^*), the x- and y-components of storm movement. In other words:

$$\alpha^*, \beta^* = \underset{\alpha,\beta}{\mathrm{Max}}\ C_1(\alpha,\beta), \quad \text{or} \quad \underset{\alpha,\beta}{\mathrm{Min}}\ C_2(\alpha,\beta), \quad \text{or} \quad \underset{\alpha,\beta}{\mathrm{Min}}\ C_3(\alpha,\beta)$$

The match point (α^*,β^*) defines the storm velocity components:

$$U_x = \alpha^*/(t_2 - t_1) \quad \text{and} \quad U_y = \beta^*/(t_2 - t_1) \tag{25}$$

Rainfall data are collected at two points in time, t_1 and t_2, and compared using eqs. 22—24. Since the size of the data collection system does not change, the area compared (S) would be a function of the displacements α and β as depicted in Fig. 4A. There is no guarantee that the matching function is comparable at different values of α and β, even when normalized by S. This difficulty can be avoided.

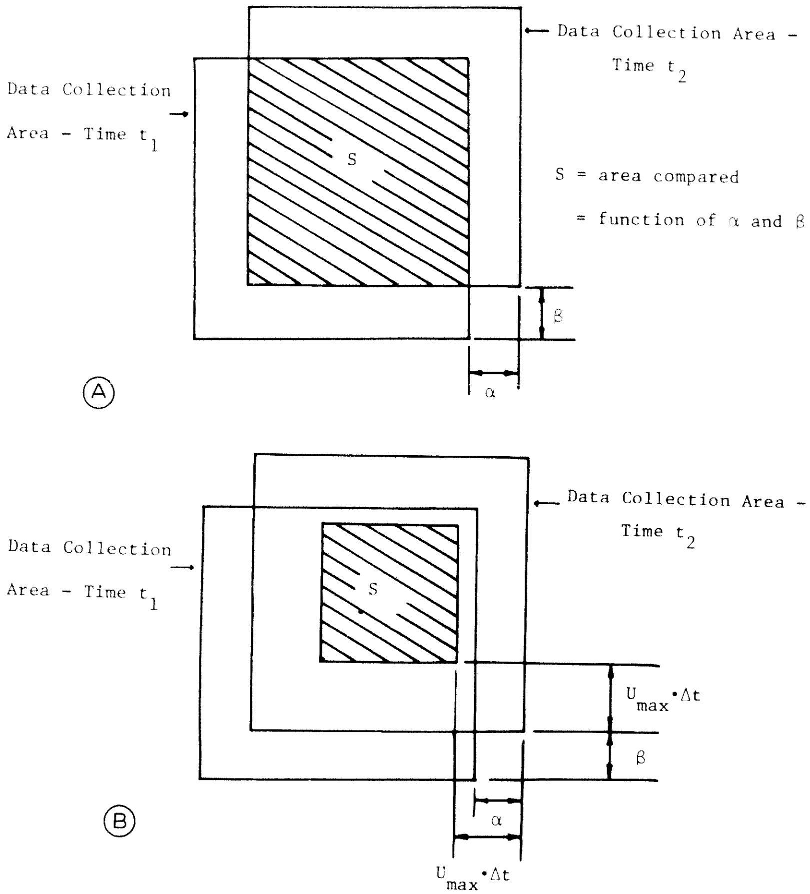


Fig. 4. Matching area.

It should be possible to specify a priori a maximum possible storm velocity U_{max}. This implies a minimum possible area of comparison S. The area compared in eqs. 22—24 is limited to this minimum value of S. The procedure is depicted in Fig. 4B. Limiting the matching area in this way does disregard some data, but it guarantees that the value of the matching function can be compared for various displacements.

The maximum product form (C_1 in eq. 22) has been used to estimate storm velocity by at least one researcher (Zawadski, 1973) using radar data. The maximum product form requires that the whole storm be within the data collection system, i.e., that the rainfall rate be zero elsewhere. If not, a new part of the storm pattern with heavy rainfall can enter the data collection network and bias the results in a manner similar to the centroid movement. Zawadski used the maximum product function (C_1) to track a group of convective cells, and the problem did not arise.

Applying the maximum product form to the one-dimensional sample storm shows the difficulty clearly. For this case:

$$C_1(\alpha) = \frac{1}{0.6}\int_{0.1}^{0.7} q_1(x)q_2(x+\alpha)\,dx$$

where

$$q_1(x) = q_2(x+0.2)$$

$$q_1(x) = \begin{cases} 0.6, & \text{for} \quad -0.1 \leqslant x < 0.0 \\ 0.2, & \text{for} \quad 0.0 \leqslant x < 0.5 \\ 0.1, & \text{for} \quad 0.5 \leqslant x < 0.9 \\ \text{unknown}, & \text{for} \quad x < -0.1 \quad \text{or} \quad x \geqslant 0.9 \end{cases}$$

The limits of integration are chosen to allow a maximum storm velocity of ± 0.2 km per time step. The computed values of $C_1(\alpha)$ are shown in Table I. Notice that $C_1(\alpha)$ is a maximum for $\alpha = -0.2$, but the value of that correctly identifies the storm velocity is $\alpha = +0.2$

TABLE I

$C_1(\alpha)$ for example storm

$C_1(\alpha)$	α	Notes
0.0733	−0.2	maximum value of $C_1(\alpha)$
0.0600	−0.1	
0.0467	0.0	
0.0317	+0.1	
0.0300	+0.2	true value of α

The absolute deviation function (C_3) has computational advantages over the squared deviation function (C_2). Since both gave identical results in tests, the absolute deviation (C_3) is preferred.

The three matching criteria are defined in integral form, but the rainfall rate is not known as a spatially continuous function. It is necessary to approximate the double integral form of C_1, C_2 and C_3.

As a first step, the rainfall generator (Bras and Rodríguez-Iturbe, 1976) was used to create a known event sampled on a uniformly-spaced grid of 50×50 points. The grid spacing was very small (~0.25 mi.). Since the grid spacing of sampling points (i.e., raingages) was uniform, the double integral form of the matching functions could be approximated by a double summation at discrete values of α and β corresponding to multiples of the grid spacing. This test gave perfect results for both C_2 and C_3 — the true values of U_x and U_y were identified in every case (to the accuracy of the grid spacing).

A network of 2500 raingages (50×50) is totally unrealistic. Also, no real network is uniformly spaced. To be useful, some approximation of $C_2(\alpha, \beta)$ or $C_3(\alpha, \beta)$ must be found that uses realistic data.

The next approach tried was to fit a continuous surface to the observed rainfall rates and use this fitted surface in the expression for $C_2(\alpha, \beta)$ or $C_3(\alpha, \beta)$. Hopefully a fitting function could be found that would allow eq. 23 or 24 to be integrated analytically as a function of the parameters of the fitting function and the values α and β. Neither a multiquadradic (Hardy, 1971; Shaw and Lynn, 1972), Fourier (Davis, 1973), nor polynomial (Davis, 1973) fit yields a useful form. Other interpolation schemes such as bi-cubic spines (Shaw and Lynn, 1972) or kriging (Davis, 1973; Guarascio et al., 1976) calculate weights as a function of location and therefore do not have a universal parameter set; these methods are even more difficult to integrate analytically. In short, this approach does not work.

Instead of using the fitted function directly in the expression for $C_3(\alpha, \beta)$, it can be used to find interpolated rainfall rates at points forming a uniform grid. If the interpolated grid has the same number of points as there are raingages, then the interpolation scheme is merely "correcting". the true locations to form a uniform grid. If there are more points in the interpolated grid, the data have been "extended" (by introducing the structure of the fitting function). This approach is by far the most promising tracking technique, and it has been tested for a number of synthetic events. The results of these tests are presented below where the accuracy of the tracking method is compared to the regression approach described in the following section.

Regression

It is difficult to estimate the storm velocity from point data. Trying to find the cross-covariance maxima leads to a high-order non-linear parameter estimation problem. Raingage data lack the spatial resolution to accurately track the storm pattern. What is required is a linear estimation technique that relies on the temporal (not spatial) detail of raingage records.

It is possible to estimate the storm velocity via a linear regression. The approach was first used by Marshall (1975), and has been applied by at least one other investigator (Shearman, 1977).

Referring to Fig. 5, consider two raingages, i and j, a distance d apart. A storm from direction θ moving at speed V will take t_{ij} to travel between i and j in the direction of storm movement where:

$$t_{ij} = d\cos(\alpha - \theta)/V \tag{26}$$

Expanding $\cos(\alpha - \theta)$ gives:

$$\begin{aligned} t_{ij} &= (\cos\theta/V)\, d\cos\alpha + (\sin\theta/V) d\sin\alpha = (\cos\theta/V)\, y_{ij} + (\sin\theta/V) x_{ij} \\ &= b_1 y_{ij} + b_2 x_{ij} \end{aligned} \tag{27}$$

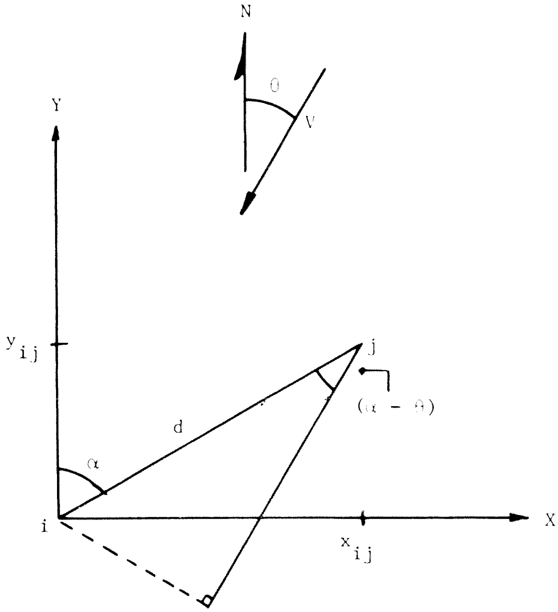


Fig.5. Storm passing over gages i and j (after Marshall, 1975).

where $x_{ij} = (x_j - x_i)$ = x-direction distance from gage i to gage j; $y_{ij} = (y_j - y_i)$; V = magnitude of storm velocity; and θ = direction of storm movement (from North axis).

If t_{ij} can be estimated from the raingage data, then the parameters of eq. 27 can be estimated as a linear regression. From the estimated parameters, $\hat{b}_1$ and $\hat{b}_2$, estimates of V and θ are found:

$$\hat{V} = (\hat{b}_1{}^2 + \hat{b}_2{}^2)^{-\frac{1}{2}} \qquad \text{and} \qquad \hat{\theta} = \tan^{-1}(\hat{b}_2/\hat{b}_1) \qquad (28), (29)$$

Finally, the two storm velocity components can be estimated:

$$\hat{U}_x = -\hat{V}\sin\hat{\theta} \qquad \text{and} \qquad \hat{U}_y = \hat{V}\cos\hat{\theta} \qquad (30), (31)$$

How should t_{ij} be estimated? The basic idea is to find the best match between the rainfall records at gages i and j. The time lag giving the best match is used to estimate t_{ij}. Marshall (1975) uses the cross-correlation function as a matching criterion. There is a slight computational advantage to a different matching criterion:

$$\hat{t}'_{ij} = \underset{\hat{t}'_{ij}}{\text{Min}} \frac{1}{T - \hat{t}'_{ij}} \sum_{t=1}^{T-\hat{t}'_{ij}} \left| q_i(t + \hat{t}'_{ij}) - q_j(t) \right| \qquad (32)$$

where t_{ij} = time lag of match point measured in time steps; $q_i(t)$ = measured rainfall rate at gage i at time step t; and T = number of time steps in storm (up to present).

Then:

$$\hat{t}_{ij} = \hat{t}'_{ij}\Delta t \tag{33}$$

where Δt = time step length.

Eqs. 32 and 33 can be used to find $\hat{t}_{ij}$ for each distinct pair of raingages i and j. The need for computational efficiency is apparent, since there are $N \times (N-1)/2$ distinct pairs in an N gage network, and the number of terms in eq. 32 grows with time.

Certainly eq. 32 will not always locate the true match point. Since the sample size tends to be large and the computational effort high, it is reasonable to avoid estimates which are likely to be in error. Several criteria were established to try to improve the quality of the sample used in eq. 27.

If either gage i or gage j has fewer than four rainfall values since the storm started at that point, $\hat{t}_{ij}$ is not computed. Clearly, some minimum number of data values should be required, although the choice of four as the minimum is arbitrary.

Large values of $\hat{t}_{ij}$ indicate that the storm takes a long time to travel from gage i to gage j. While this will surely be true for gages which are far apart, large values of $\hat{t}_{ij}$ will tend to be less accurate due to changes within the storm. Computed values of $\hat{t}_{ij}$ which are greater than one hour are discarded. Again, the choice of one hour is arbitrary.

If the start of rainfall at gage i is more than one hour different from the start at gage j, then $\hat{t}_{ij}$ is not computed. If a computed $\hat{t}_{ij}$ is more than 1 h different from the difference of rainfall start times at i and j, then it is not used.

The effect of these rules is to remove sample values of $\hat{t}_{ij}$ which are most likely to be inaccurate.

If all goes well, a large number of values of $\hat{t}_{ij}$ will be available to estimate the parameters of eq. 27, and a standard linear least-squares regression procedure will give an excellent fit. Sometimes the regression approach will not work because the sample size is too small or the regression lacks significance. In these cases, the centroid is used to estimate storm velocity.

The regression can fail in two ways: insufficient sample or lack of significance. If there are fewer than ten sample values of $\hat{t}_{ij}$, the regression is not attempted.

The significance of the regression is tested; the hypothesis $b_1 = b_2 = 0$ must be rejected with 99% confidence using the F-statistic of the regression (Draper and Smith, 1966).

The regression estimate of storm velocity performed well for a number of synthetic test storms, including cases where storm tracking techniques failed. However, there were a few test cases where the storm velocity estimates diverged from the true values for a number of time steps. Examining these problem test cases showed that a few values of $\hat{t}_{ij}$ were significantly different from the true values, but the majority of values were reasonably accurate. Therefore, a robust regression procedure was implemented to automatically remove outliers.

TABLE II

Comparison of the regression approach and storm tracking techniques

Tracking	Regression
(1) relies more on the description of the spatial variation of rainfall intensities than the time history of the event	(1) relies more on the description of the time variation
(2) conceptually based on a storm velocity that may change with time	(2) conceptually based on a constant storm velocity
(3) gives no indication when the estimate is poor	(3) gives an indication that the estimate is poor

If the standard error of the regression is less than 0.15 h, the robust regression procedure is not needed. If not, then the outliers are identified and removed from the sample values of $\hat{t}_{ij}$. Any $\hat{t}_{ij}$ more than two standard errors away from the latest regression is considered an outlier.

Once the outliers are identified and removed, a new regression is performed on the pruned sample. This new regression is also subject to the standard error test. If it fails, the entire sample is pruned again and another iteration of robust regression performed. A limit of ten passes of robust regression is imposed.

In most of the test cases, the robust regression was not required. When used, the robust regression usually converged to a velocity estimate close to the true value in two or three passes. The other cases converged, but more slowly.

The regression approach to storm velocity estimation produces reasonably accurate estimates. When it fails, it indicates the poor estimate by lack of fit, and therefore allows for alternative estimates to be used when it fails. The contrast between the regression approach and storm tracking techniques is shown in Table II. On conceptual grounds, the regression scheme seems to be more appropriate when the data source is a raingage network alone. Radar data which have much greater spatial detail are better suited to a tracking approach.

Testing of velocity estimates

It is impossible to test storm velocity estimation methods on real storms because there is no way to find the true storm velocity. Instead, various synthetic events moving at different known storm velocities were created using a rainfall generation model (Bras and Rodríguez-Iturbe, 1976). These test storms were used to compare the accuracy of the matching technique with the regression technique.

For each test case, a generated rainfall event was sampled on a 6 × 6 grid (36 points) at a spacing of about 2 mi. The multiquadratic equation (Hardy, 1971) was used to extend the sample to a 50 × 50 uniform grid at a spacing of 0.25 mi. Using the 50 × 50 uniform grid, $C_3(\alpha,\beta)$ (see eq. 24) was approximated for discrete values of α and β to find α^* and β^*, the optimal values. Complete enumeration of all possible values of α and β is computationally burdensome, so a gradient search procedure was used to find α^* and β^* to minimize $C_3(\alpha,\beta)$. The same generated values (on the 6 × 6 grid) were used as input to the regression method.

The results of the first test case are shown in Fig. 6A. Both techniques perform reasonably well. The regression method does not have a large enough sample until time step 4. The value of $\hat{U}_x$ produced by the regression is far off at time step 4, but thereafter both methods perform fairly well. The

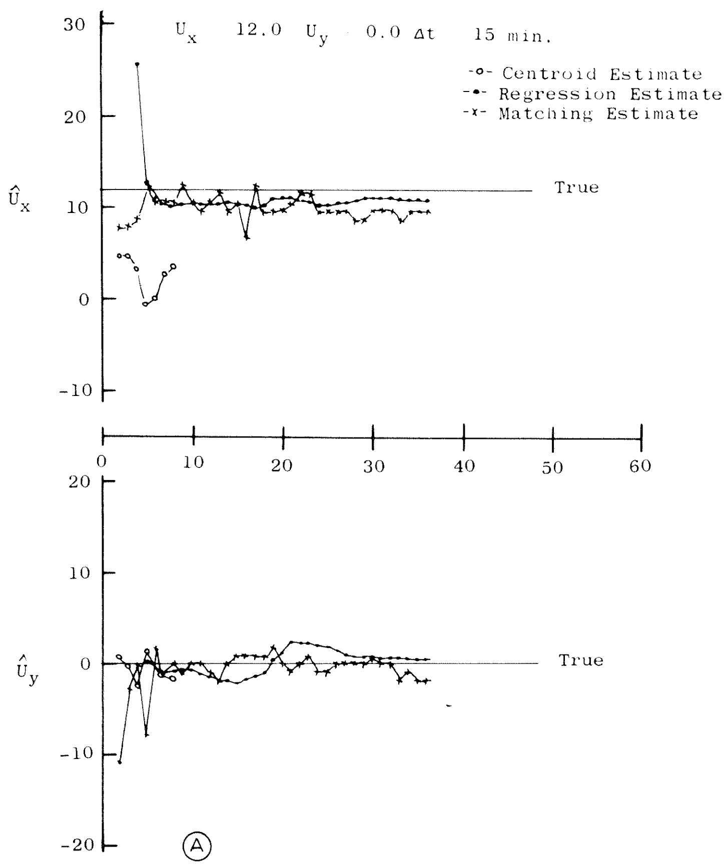
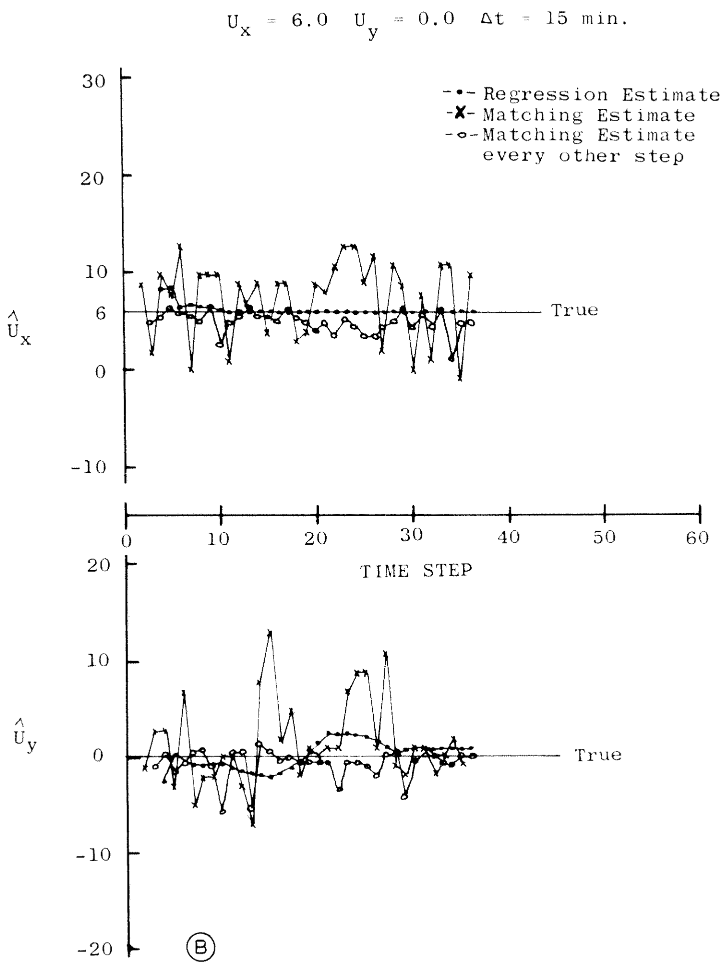


Fig. 6A. For caption see p. 121.

centroid estimate is also shown (for time steps 2—8) in Fig.6A. The discrete values of α and β correspond to approximately 1 mi./h for a 15-min time step.

Test storm one seems to indicate that either method might be used. However, further testing showed that the good performance of the matching technique during the first test was fortuitous. The storm velocity of the first test is close to an even multiple of the grid spacing of the 6 × 6 grid.

The second test storm has one-half the velocity of storm *1*. The results are shown on Fig. 6B. The regression estimate performs much better than the matching technique. An attempt was made to improve the matching technique by comparing every other time step rather than every time step, i.e., giving the storm more time to move. The improvement is apparent, but there is no indication of when the every-other-step procedure should be used. In a test case, the storm velocity is known to be slow, so the every-other-step procedure



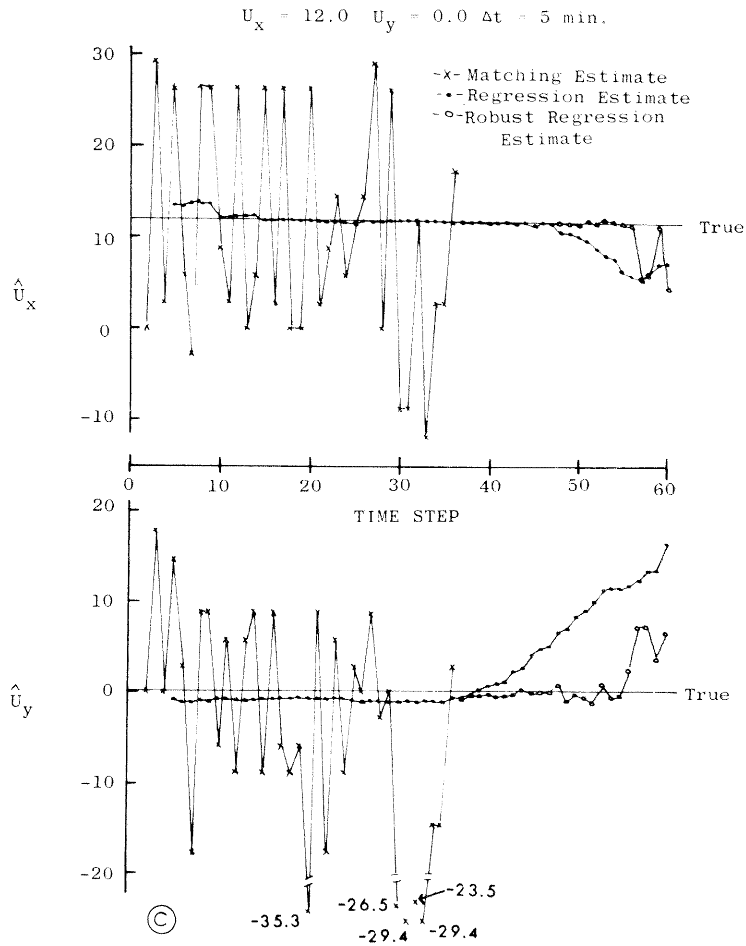


Fig.6. Velocity estimation tests: (A) storm *1*; (B) storm *2*; and (C) storm *3*.

is indicated. In the real world, the unreliable estimate obtained by the matching technique does not give an indication that an every-other-step procedure would be more reliable.

The third test storm is shown in Fig.6C. The storm velocity is identical to storm *1* (U_x = 12.0) but the sampling interval is reduced (from 15 to 5 min). This placed an unbearable strain on the matching technique to discern a smaller amount of movement over each time step. The regression estimate performs well up to time step 40 or so. At that time, a number of erroneous sample values of $\hat{t}_{ij}$ (see eq. 32) begin to be produced. The robust regression scheme reduced the sensitivity of the estimated U_x and U_y to these erroneous sample values and produces more accurate $\hat{U}_x$ and $\hat{U}_y$. The most divergent values (time steps 57, 58 and 60) are produced when the robust regression was terminated due to the number of iterations criterion.

As a result of the three test storms shown, as well as other test results not shown, the authors recommend the regression approach to estimate storm velocity from raingage data.

CONCLUSIONS

This paper shows real-time parameter estimation for a rainfall forecasting model. Hopefully, future researchers will benefit from the illustrations of the several methodologies presented.

The storm counter technique is easy to implement in real time and capable of representing significant features of a rainfall event. Although testing of the procedure can only be described as a qualified success, it is obvious that the non-stationarity in the mean and variance of the rainfall process is captured by the storm counter method. The method permits handling non-stationarities without major, inflexible, structural assumptions.

Estimating the covariance structure in real time proved to be a challenging problem. Functional forms are required to preserve positive definiteness of covariance. The optimization methodology utilized to estimate parameters worked well and efficiently. It was necessary, though, to estimate storm velocities outside of the covariance identification procedure.

Testing of several velocity estimators ranging from storm feature tracking to covariance maximization clearly converged in a regression procedure based on time histories of raingages. The selected estimation procedure proved to be stable and convergent under many conditions.

All the parameters estimated and discharged in this paper can be used in real-time forecasting of rainfall. Interested readers are referred to Johnson and Bras (1978) for forecasting applications.

ACKNOWLEDGEMENTS

This work was sponsored by the National Science Foundation under Grant ENG 76-118-17 to the Massachusetts Institute of Technology, Department of Civil Engineering, Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics. Publication support was provided by the National Weather Service, Hydrologic Research Laboratory.

REFERENCES

Bras, R.L. and Rodríguez-Iturbe, I., 1976. Rainfall generation: a nonstationary time-varying multidimensional model. Water Resour. Res., 12(3): 450—456.

Davis, J.C., 1973. Statistics and Data Analysis in Geology. Wiley, New York, N.Y., 504 pp.

Draper, N.R. and Smith, H., 1966. Applied Regression Analysis. Wiley, New York, N.Y., 407 pp.

Eagleson, P.S., 1970. Dynamic Hydrology. McGraw-Hill, New York, N.Y., 448 pp.

Guarascio, M., David, M. and Huijbregts, C. (Editors), 1976. Advanced Geostatistics in the Mining Industry. D. Reidel, Boston, Mass., 461 pp.

Hardy, R.L., 1971. Multiquadric equations of topography and other irregular surfaces. J. Geophys. Res., 76(8): 1905—1915.

Johnson, E.R. and Bras, R.L., 1978. Short term rainfall prediction: a non-stationary multivariate stochastic model. Ralph M. Parsons Lab. Water Resour. Hydrodyn., Rep. No. 233, M.I.T., Cambridge, Mass., 247 pp.

Marshall, R.J., 1975. A stochastic model to simulate moving storms. In: Precipitation Analysis for Hydrologic Modeling, Nat. Symp. Am. Geophys. Union Hydrol. Sect., Jun. 26—28, 1975, Univ. Calif., Davis, Calif., pp. 259—300.

Mejía, J.M. and Rodríguez-Iturbe, I., 1974. On the synthesis of random field sampling from the spectrum: an application to the generation of hydrologic spatial processes. Water Resour. Res., 10(4): 705—711.

Shaw, E.M. and Lynn, P.P., 1972. Areal rainfall evaluation using two surface fitting techniques. Hydrol. Sci. Bull., 17(4): 419—433.

Shearman, J.R., 1977. The speed and direction of movement of storm rainfall patterns with reference to urban storm sewer design. Hydrol. Sci. Bull., 22(3): 421—431.

Zawadski, I.I., 1973. Statistical properties of precipitation patterns. J. Appl. Meteorol., 12: 459—472.

DATA SOURCES

Project METROMEX, Illinois State Water Survey, Box 232, Urbana, Ill.
The Southern Great Plains Watershed Research Center, Chickasha, Okla.