WATER RESOURCES RESEARCH, VOL. 29, NO. 5, PAGES 1489–1504, MAY 1993

Real-Time Rainfall Prediction at Small Space-Time Scales Using a Two-Dimensional Stochastic Advection-Diffusion Model

KENJI JINNO, AKIRA KAWAMURA, AND RONNY BERNDTSSON[1]

Department of Civil Engineering, Kyushu University, Fukuoka, Japan

MAGNUS LARSON AND JANUSZ NIEMCZYNOWICZ

Department of Water Resources Engineering, University of Lund, Lund, Sweden

A model based on the two-dimensional stochastic advection-diffusion equation is developed to forecast properties of individual rain cells in urban areas such as speed and spatial rainfall intensity. Two different modeling approaches are employed, and examples of the results are given. The first approach involves a Gaussian distribution as an analytic solution to the advection-diffusion equation, whereas the second one entails a double Fourier series expansion of the rainfall intensity field. Both modeling approaches are used to predict the rainfall intensity field over a small 12-gage urban catchment in southern Sweden. The model parameters are continuously updated by extended Kalman filtering. The Fourier series approach is shown to be the most flexible for practical applications and to give the most accurate forecasts. This model approach gives acceptable forecasts for a lead time of 1–5 min. It gives consistently smaller prediction errors compared to both the Gaussian solution and simple extrapolation calculations. The effect of system noise level on the forecast accuracy and model performance is discussed. The model can be used not only to predict in real time the spatial rainfall, but also to parameterize the variability pattern of small-scale spatial rainfall into a set of physically based parameters, thus separating the effects of advective velocity, turbulent diffusion, and development/decay.

1. INTRODUCTION

Real-time prediction of the space-time rainfall field in urban areas is motivated for on-line decision making to optimize the operation of urban hydrological systems, thus minimizing pollutant discharge from drainage systems, avoiding flooding, and making maximum use of the available storage volume in sewers. Prediction of the space-time rainfall field and its dynamic behavior is required as input to runoff simulation models. The requirements of the temporal and spatial resolution of the rainfall input are especially high for urban catchments [e.g., *Berndtsson and Niemczynowicz*, 1988]. Small catchment areas, dense building structure, and thus a high degree of impermeable areas with a resulting rapid runoff, imply that the smallest spatial units of the rainfall field (individual cells) are of primary importance.

Most of the existing models for real-time control of hydrological systems use a stochastic description of the rainfall field [*Georgakakos and Hudlow*, 1984; *Georgakakos and Bras*, 1984; *Georgakakos*, 1986]. Models for short-term rainfall forecasting or nowcasting have in most cases used radar data as input [*Browning and Collier*, 1989; *Bellon and Austin*, 1984; *Einfalt and Denoeux*, 1987]. However, most urban catchments are still not equipped with radar; instead urban rainfall is usually observed by rain gages. Thus in order to extend the use of measurements from rain gages there is a need to develop methods for rainfall forecasting with respect to the specific requirements of urban hydrological management. If radar data also are at hand, the radar can provide information on a larger scale embedding the rain gage system (e.g., by specifying boundary conditions for the forecasting area).

Real-time prediction of urban-scale rainfall relies on two fundamentals: (1) an understanding of the small-scale space-time characteristics of the rainfall field, and (2) an effective technique to determine the rainfall intensities in time and space. The former has been elaborated on in a related paper (R. Berndtsson et al., Some Eulerian and Lagrangian statistical properties of rainfall at small space-time scales, submitted to *Journal of Hydrology*, 1992; hereinafter Berndtsson et al., submitted manuscript, 1992), and the latter will be dealt with in this paper.

There are several features that a useful forecasting model must exhibit due to the special characteristics of small-scale rainfall variability and the requirements of the urban hydrological management system. First, the model should be capable of reconstructing the irregular shape of the small-scale rainfall field from a set of physically meaningful mathematical parameters. Second, the model should be flexible enough to allow for a dynamic behavior of the rainfall field in both time and space. Third, the model should be able to predict the rainfall field trend in time. Fourth, the model should permit a high degree of automation allowing for use in an interactive urban management system [e.g., *Browning and Collier*, 1989].

Kumar and Foufoula-Georgiou [1990] described whole contour methods as one of the most promising future techniques for extrapolative short-time forecasting (see also

[1]On leave from Department of Water Resources Engineering, University of Lund, Lund, Sweden.

Paper number 92WR02849.
0043-1397/93/92WR-02849$05.00

Bohne [1988]). One type of whole contour method is Fourier domain shape representation, which is computationally intensive but amenable for use in automatic systems [e.g., *Kawamura et al.*, 1991]. Consequently, one of the model approaches evaluated in this paper involves a Fourier domain shape representation of the rainfall field. Since the rainfall intensity field may change quickly both in time and space a model is needed that continuously tracks the behavior of the rainfall field in real time and recursively updates the parameters of the model. Thus, using the latest estimate of the model parameters a prediction is made for the next time period. In the present study the theory of Kalman filtering has been employed for the recursive estimation procedure [*Kalman*, 1960; *Kalman and Bucy*, 1961; *Athans et al.*, 1968]. Finally, a model is also needed for a convenient mathematical description of the physical characteristics of individual rain cells. Here, a stochastic advection-diffusion equation has been chosen to describe the temporal evolution of a rain cell in two spatial dimensions at ground level. The primary factor limiting the period of valid extrapolation is the development or decay of the rainfall [*Browning and Collier*, 1989]. Therefore the ability of the model to consider these phenomena is crucial for reliable forecasting. In the present model the development or decay of rainfall intensity was considered by adding a source-sink term to the advection-diffusion equation. To summarize, the physical properties of the rain cell are described by an advection-diffusion equation. This equation is solved by two methods, namely, analytically by a Gaussian-shaped solution and by a double Fourier series expansion. The parameter estimation technique relies on extended Kalman filtering.

2. Methodology and Model Formulation

This section describes the two approaches to modeling the temporal and spatial distribution of rainfall intensities. The basic equations of the two approaches for characterizing an individual rain cell are described in section 2.1. Section 2.2 gives an overview of the Kalman filter formulation. Section 2.2.1 gives the formulation of the first model approach, the analytic solution involving a Gaussian distribution to the basic equation, and section 2.2.2 gives the formulation of the second model approach, i.e., the Fourier series solution.

2.1. *Basic Model Formulation*

Rainfall from individual cells as observed at ground level is described in the model by a two-dimensional stochastic advection-diffusion equation. Several simplifying assumptions are necessary in order to employ such a description of the rain cell. The main concern in this study is rainfall at small space-time scales. It is therefore assumed that during short periods (minutes), the main processes acting on the raindrops within an individual rain cell are advection and diffusion (Berndtsson et al., submitted manuscript, 1992). Thus the raindrops are considered a passive constituent advected by a turbulent flow field. Furthermore, it is assumed that the change in the vertical direction of the three-dimensional field of raindrops is negligible, i.e., it is assumed that the two-dimensional Gaussian-shaped rainfall intensity field on a horizontal plane in space is reflected as a two-dimensional Gaussian-shaped intensity field at ground level. Under this assumption we may regard a two-dimensional advection-diffusion equation in the horizontal plane as an analogy of the flux of rainfall at ground level. The aptness of this assumption is indicated by the often found Gaussian-shaped spatial intensity pattern of individual rain cells [*Sharon*, 1972; *Zawadzki*, 1973; *Marshall*, 1980]. The Gaussian function constitutes a particular solution to the advection-diffusion equation.

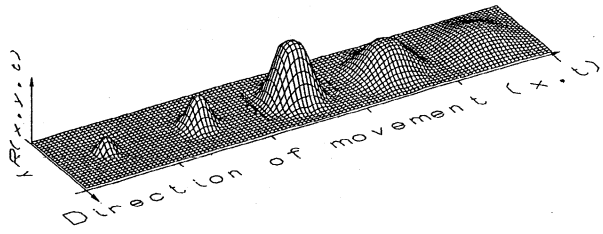

Fig. 1. Schematic of a space-time rain cell evolution as predicted by the advection-diffusion equation used in the modeling approach.

If the x axis is taken along the main direction of movement of the advective rain cell (note the difference between our definition of advection and meteorological convection), the two-dimensional stochastic advection-diffusion equation may be expressed as (assuming spatial homogeneity regarding rainfall speed, diffusion, and development/decay):

$$\frac{\partial R(x, y, t)}{\partial t} + u \frac{\partial R(x, y, t)}{\partial x} = D_x \frac{\partial^2 R(x, y, t)}{\partial x^2} + D_y \frac{\partial^2 R(x, y, t)}{\partial y^2} - \gamma R(x, y, t) + \varepsilon(x, y, t) \quad (1)$$

where

$R(x, y, t)$ rainfall intensity in space (x, y) and time (t), m/min;
u advective speed of the rainfall cell in the direction of movement, m/min;
D_x, D_y diffusion coefficients in the x and y direction, respectively, m^2/min;
γ development/decay coefficient of the rainfall intensity, min^{-1};
$\varepsilon(x, y, t)$ stochastic component with zero-mean Gaussian white noise in time and space, m/min^2.

Equation (1) has frequently been used to describe turbulent diffusion of a conservative property in a fluid medium [e.g., *Calder*, 1965]. It should be stressed, however, that the use of the advection-diffusion equation in this paper is based on the validity of the assumptions made above and should not be considered as an attempt to physically describe the properties of the turbulent and thermodynamic behavior of convective air resulting in individual rain cells. Instead, (1) is used to conveniently describe in mathematical terms the speed and spatial intensity of a rain cell as observed at ground level. Accepting this, (1) embodies the foremost properties of an individual rain cell as observed at ground level and a set of functional parameters with a physical interpretation in terms of the overall cell behavior (Figure 1). The diffusion coefficients D_x and D_y describe the rate at which the cell expands horizontally. The development/decay coefficient γ describes the development (cumulus stage) and decay (dissipating stage) of the cell in terms of change in

strength. Negative values of γ represent a situation when the cell develops or expands (cumulus stage), i.e., the rainfall intensity increases. Similarly, positive values of γ represent a decaying situation (dissipating stage) when the rainfall intensity decreases. The last term in (1) is a stochastic component that is added to consider a certain degree of error and uncertainty inherent in modeling of the rainfall. In the following this term is assumed to be completely random in both time and space. However, at locations where a trend may be present, such as the case of a marked topographical gradient inducing orography, this term can be separated into a deterministic and a stochastic component.

The x axis in (1) is taken along the direction of movement of the cell. This is done in order to eliminate advection in the y direction and off-diagonal diffusion and consequently to reduce the number of parameters to estimate. In a related paper (Berndtsson et al., submitted manuscript, 1992) it was shown that the speed of individual cells is important to consider in small catchments. The initial values of advective speed and direction of movement, however, may be difficult to estimate in a real-time forecasting situation. If information on wind speed and direction is known beforehand [e.g., *Niemczynowicz*, 1984], these data may be used as initial values, which significantly reduces the time for parameter identification in the model. Radar observations and cell-tracking algorithms can also provide velocity and direction estimates. This kind of information may be provided by nearby meteorological stations.

Two modeling approaches are exemplified and discussed in the paper within the framework of the general methodology. The first approach involves an analytic solution to (1) under certain initial and boundary conditions and by excluding the stochastic component, whereas the second approach encompasses an expansion of the rainfall field $R(x, y, t)$ and the stochastic component $\varepsilon(x, y, t)$ in terms of Fourier series. In the first case, a Gaussian-shaped solution is employed to predict the rainfall intensity distribution in time and space. In the second case, the Fourier series expansion is used to predict the temporal and spatial rainfall field. In both cases, simulated and measured rain gage data are used to study the model performance. Also in both cases, we compare with simple extrapolation calculations in order to interpret the practical value of the models.

The two modeling approaches represent two sophistication levels of the rain cell description. In the first case, individual cells are assumed to be described by a Gaussian-shaped intensity pattern in space. In the second case, the spatial intensity pattern is allowed to vary freely as described by the Fourier series expansion. Only as initial conditions, a Gaussian-shaped distribution is assumed. Consequently, the results of the modeling will show the sophistication level for the cell description that is necessary for realistic real-time forecasting of rainfall in small urban catchments.

There are several reasons why a Fourier series representation was chosen instead of a numerical scheme, e.g., a finite-difference approach. Many studies during the recent years have shown that a finite-difference approach gives large numerical error, especially if the advective term in the advection-diffusion equation dominates the transport [e.g., *Jinno et al.*, 1989]. A highly dominant advection term may be expected for small-scale advective rainfall (e.g., Berndtsson et al., submitted manuscript, 1992). A major disadvantage of using a numerical scheme is also that the grid points need to match the rain gage locations. Since rain gages are usually not installed at regular grid points a rather detailed and complicated grid scheme would have been required. Besides this, every time a rain gage location is changed or when a rain gage is malfunctioning, the grid scheme has to be changed. Another problem is how to treat the boundaries along the rain gage network. Data for the boundaries are usually not at hand or trivially estimated for rain gage networks. Also, the computational load is much larger for a numerical scheme as compared to a Fourier series approach. The Fourier series approach assumes a uniform velocity field and uniform parameter values. In summary, a Fourier series approach is superior in many ways compared to a numerical scheme for the special characteristics of the advection-diffusion equation and the model domain which a rain gage network represents.

Excluding the stochastic term and assuming instantaneous occurrence of rainfall at time t_0 at location (x_0, y_0) on the infinite plane, the analytic solution of (1) is given by a Gaussian distribution as [e.g., *Brutsaert*, 1974]

$$R(x, y, t) = \frac{I}{4\pi(D_x D_y)^{1/2}(t - t_0)} \exp\left\{-\frac{[x - x_0 - u(t - t_0)]^2}{4D_x(t - t_0)} - \frac{(y - y_0)^2}{4D_y(t - t_0)} - \gamma(t - t_0)\right\} \quad (2)$$

where I is the strength of the cell (cubic meters per minute), t_0 is the occurrence time of the cell (minutes), and x_0, y_0 are the coordinates for the initial location of the cell (meters). Hereafter, we refer to (2) subject to the above initial and boundary conditions as the Gaussian solution. The rainfall intensity may be predicted at any time and location once the parameters I, u, D_x, D_y, γ, x_0, y_0, and t_0 have been identified. By use of a filtering technique these unknown parameters are estimated and updated at each time step. In this case the solution of (2) is nonlinear, and the extended Kalman filter needs to be applied [e.g., *Bras and Rodriguez-Iturbe*, 1985; *Jinno et al.*, 1989, 1990].

In the second modeling approach, the second-order partial differential equation (equation (1)) is transformed into a set of ordinary differential equations by applying a double Fourier series expansion. The rainfall field $R(x, y, t)$ and the coupled Gaussian white noise $\varepsilon(x, y, t)$ are expanded as

$$R(x, y, t) = A_{0,0}(t)/2 + \sum_{m=1}^{M}\sum_{n=0}^{N}[A_{m,n}(t)\cos F_1(x, y, m, n) + B_{m,n}(t)\sin F_1(x, y, m, n)] + \sum_{m=0}^{M}\sum_{n=1}^{N}[C_{m,n}(t)\cos F_2(x, y, m, n) + D_{m,n}(t)\sin F_2(x, y, m, n)] \quad (3)$$

$$\varepsilon(x, y, t) = \sum_{m=1}^{M} \sum_{n=1}^{N} [E_{m,n}(t) \cos F_1(x, y, m, n) + F_{m,n}(t) \sin F_1(x, y, m, n) + G_{m,n}(t) \cos F_2(x, y, m, n) + H_{m,n}(t) \sin F_2(x, y, m, n)] \quad (4)$$

where

$$F_1(x, y, m, n) = 2\pi mx/\lambda_x + 2\pi ny/\lambda_y \quad (5)$$

$$F_2(x, y, m, n) = 2\pi mx/\lambda_x - 2\pi ny/\lambda_y \quad (6)$$

In (3)–(6), $A_{m,n}(t)$, $B_{m,n}(t)$, $C_{m,n}(t)$, $D_{m,n}(t)$, $E_{m,n}(t)$, $F_{m,n}(t)$, $G_{m,n}(t)$, and $H_{m,n}(t)$ are Fourier coefficients, M and N are the number of terms in the expansion procedure, and λ_x and λ_y are wavelengths for the x and y directions, respectively, that cover the entire area where rain gages are installed. Substitution of (3)–(6) into (1) yields a set of ordinary differential homogeneous equations with respect to the wave numbers m and n according to

$$\begin{bmatrix} dA_{m,n}(t)/dt \\ dB_{m,n}(t)/dt \\ dC_{m,n}(t)/dt \\ dD_{m,n}(t)/dt \end{bmatrix} = \begin{bmatrix} -P_{m,n} & -Q_m & 0 & 0 \\ Q_m & -P_{m,n} & 0 & 0 \\ 0 & 0 & -P_{m,n} & -Q_m \\ 0 & 0 & Q_m & -P_{m,n} \end{bmatrix} \begin{bmatrix} A_{m,n}(t) \\ B_{m,n}(t) \\ C_{m,n}(t) \\ D_{m,n}(t) \end{bmatrix} + \begin{bmatrix} E_{m,n}(t) \\ F_{m,n}(t) \\ G_{m,n}(t) \\ H_{m,n}(t) \end{bmatrix} \quad (7)$$

where

$$P_{m,n} = D_x(2\pi m/\lambda_x)^2 + D_y(2\pi n/\lambda_y)^2 + \gamma \quad (8)$$

$$Q_m = u(2\pi m/\lambda_x) \quad (9)$$

2.2. *Kalman Filter Formulation*

The system and the observation equation of the Kalman filter are expressed in discrete form as

$$\mathbf{X}(k + 1) = \mathbf{\Phi}(k)\mathbf{X}(k) + \boldsymbol{\alpha}(k) + \mathbf{v}(k) \quad (10)$$

$$\mathbf{Y}(k + i) = \mathbf{\Gamma}(k + i)\mathbf{X}(k + i) + \boldsymbol{\beta}(k + i) + \mathbf{w}(k + i) \quad (11)$$

where

$\mathbf{X}(k + 1)$ system vector to be estimated;
$\mathbf{\Phi}(k)$ known state transition matrix;
$\boldsymbol{\alpha}(k)$ known constant vector;
$\mathbf{v}(k)$ white Gaussian system noise vector;
$\mathbf{Y}(k + i)$ observation vector;
$\mathbf{\Gamma}(k + i)$ known observation matrix;
$\boldsymbol{\beta}(k + i)$ known constant vector;
$\mathbf{w}(k + i)$ white Gaussian observation noise vector;
k time step;
i time interval of observations.

As seen above in (10) and (11), the model errors are considered by treating the advection and diffusion parameters as random variables. The estimate of the state vector at time step $k + i$, calculated using the observation obtained at time step k, is denoted $\hat{\mathbf{X}}(k + i|k)$ and at time step $k + i$, $\hat{\mathbf{X}}(k + i|k + i)$. If $\hat{\mathbf{X}}(k|k)$ is known after the observation is obtained at time step k, $\hat{\mathbf{X}}(k + i|k)$ and $\hat{\mathbf{X}}(k + i|k + i)$ are calculated at time step $k + i$ as

$$\hat{\mathbf{X}}(k + i|k) = \mathbf{\Phi}(k + i - 1)\hat{\mathbf{X}}(k + i - 1|k) + \boldsymbol{\alpha}(k + i - 1) \quad (12)$$

$$\hat{\mathbf{X}}(k + i|k + i) = \hat{\mathbf{X}}(k + i|k) + \mathbf{K}(k + i)\boldsymbol{\nu}(k + i) \quad (13)$$

where

$$\mathbf{K}(k + i) = \mathbf{P}(k + i|k)\mathbf{\Gamma}^T(k + i) \cdot [\mathbf{\Gamma}(k + i)\mathbf{P}(k + i|k)\mathbf{\Gamma}^T(k + i) + \mathbf{W}(k + i)]^{-1} \quad (14)$$

$$\boldsymbol{\nu}(k + i) = \mathbf{Y}(k + i) - \hat{\mathbf{Y}}(k + i|k) \quad (15)$$

$$\hat{\mathbf{Y}}(k + i|k) = \mathbf{\Gamma}(k + i)\hat{\mathbf{X}}(k + i|k) + \boldsymbol{\beta}(k + i) \quad (16)$$

in which $\mathbf{W}(k + i)$ is the covariance matrix of the observation noise $\mathbf{w}(k + i)$ and T denotes transposed.

Similarly, the state estimate error covariance matrix at time step $k + i$, calculated using the observation obtained at time step k, is denoted $\mathbf{P}(k + i|k)$ and at time step $k + i$, $\mathbf{P}(k + i|k + i)$. They are calculated as

$$\mathbf{P}(k + i|k) = \mathbf{\Phi}(k + i - 1)\mathbf{P}(k + i - 1|k)\mathbf{\Phi}^T(k + i - 1) + \mathbf{V}(k + i - 1) \quad (17)$$

$$\mathbf{P}(k + i|k + i) = [\mathbf{I}_U - \mathbf{K}(k + i)\mathbf{\Gamma}(k + i)]\mathbf{P}(k + i|k) \quad (18)$$

where $\mathbf{V}(k + i - 1)$ is the covariance matrix of the system noise $\mathbf{v}(k + i - 1)$ and $\mathbf{I}_U$ is a unit matrix. Details on this procedure can be found in the works by *Athans et al.* [1968], *Bras and Rodriguez-Iturbe* [1985], or *Kawamura et al.* [1984, 1989].

2.2.1. *Formulation of the Gaussian solution.* For the Gaussian solution, (2), the eight-element system vector $\mathbf{X}_G(t)$ to be estimated is

$$\mathbf{X}_G(t) = [I, u, D_x, D_y, \gamma, x_0, y_0, t_0]^T \quad (19)$$

The true values of the parameters I, u, D_x, D_y, γ, x_0, y_0, and t_0 are assumed constant in time (and in space) for the Kalman filter procedure so that

$$dI(t)/dt = 0 \quad (20)$$

$$du(t)/dt = 0 \quad (21)$$

$$dD_x(t)/dt = 0 \quad (22)$$

$$dD_y(t)/dt = 0 \quad (23)$$

$$d\gamma(t)/dt = 0 \quad (24)$$

$$dx_0(t)/dt = 0 \quad (25)$$

$$dy_0(t)/dt = 0 \quad (26)$$

$$dt_0(t)/dt = 0 \tag{27}$$

However, note that this does not mean that the estimated parameter values do not change in time throughout the process of the extended Kalman filter [e.g., *Kawamura et al.*, 1986].

By adding system noise $\mathbf{v}(t)$, (20)–(27) lead to the following system equation:

$$d\mathbf{X}_G(t)/dt = \mathbf{0} + \mathbf{v}(t) \tag{28}$$

Discretization of (28) by difference leads to [e.g., *Bras and Rodriguez-Iturbe*, 1985]

$$\mathbf{X}_G(t + \Delta t) = \mathbf{X}_G(t) + \mathbf{v}(t)\Delta t \tag{29}$$

where Δt is the discrete time interval. In (10) (compare (29) and (10)) $t = k$, $t + \Delta t = k + 1$, Φ is a unit matrix, $\alpha = \mathbf{0}$, and $\mathbf{v}(t)\Delta t = \mathbf{v}(k)$.

The observation vector $\mathbf{Y}(k)$ of the extended Kalman filter is the rainfall intensity distribution observed at the 12 rain gage locations in Lund with a sampling interval of 1 min. With unknown initial and boundary conditions, the rainfall distribution is predicted using the extended Kalman filter on the basis of the information given by the rain gages.

For the Gaussian solution each element of the observation vector $\mathbf{Y}(k)$ corresponds to the observed rainfall at each rain gage located at point (x, y) and is derived for time step k from (2). By adding the observation noise $w(x, y, t)$ it becomes

$$\mathbf{Y}(k) = \mathbf{R}[\mathbf{X}_G(k)] + \mathbf{w}(k) \tag{30}$$

where $\mathbf{R}[\mathbf{X}_G(k)]$ is the vector function whose elements equal the right-hand side of (2) and $\mathbf{w}(k)$ is the observation noise vector whose elements are composed of $w(x, y, t)$.

In this case $\mathbf{R}[\mathbf{X}_G(k)]$ is a nonlinear vector function in $\mathbf{X}_G(k)$, and a Taylor series expansion is applied with $\mathbf{X}^*_G(k)$ as the estimated parameter yielding

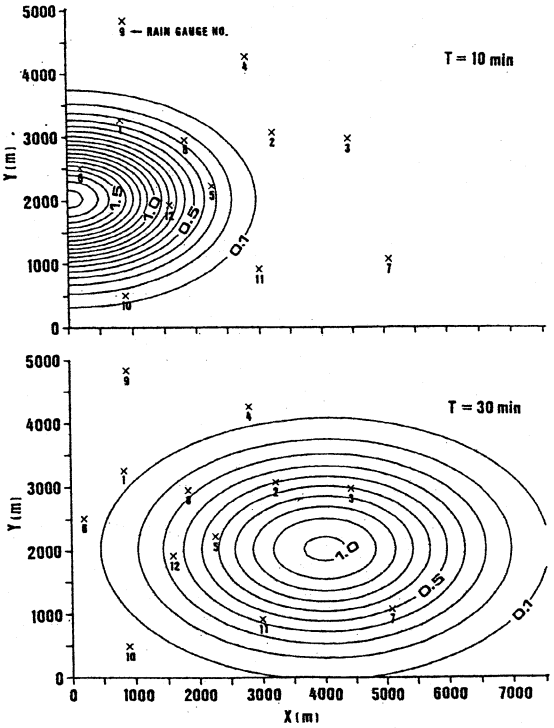

Fig. 3. Synthetic rain cell used for model evaluation moving over the rain gage network in Lund.

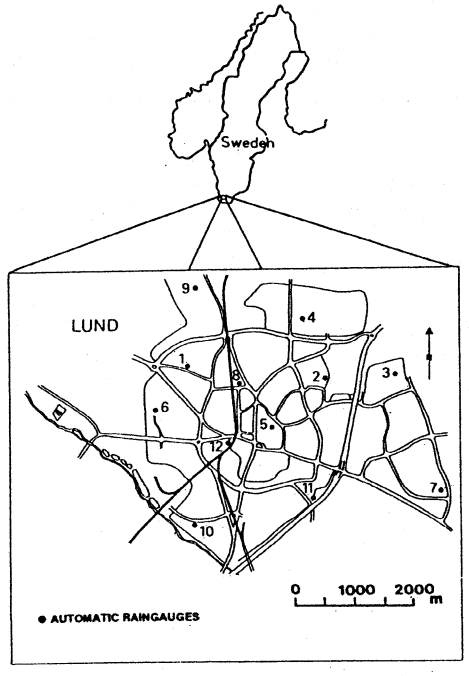

Fig. 2. The experimental catchment in Lund [*Niemczynowicz*, 1984].

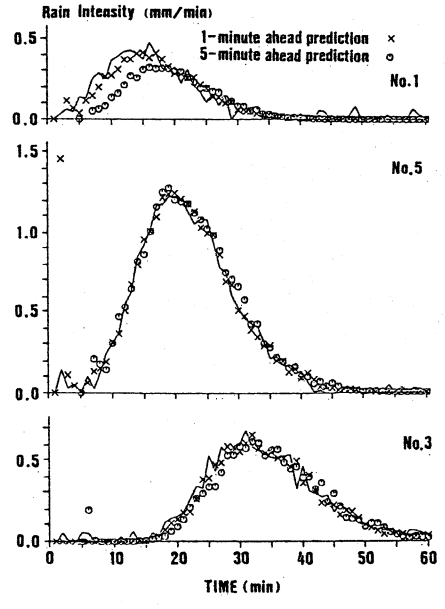

Fig. 4. Resulting 1- and 5-min-ahead real-time predictions of rainfall intensity for three gages (the solid line shows the synthetic rainfall with 0.05 mm/min observation noise).

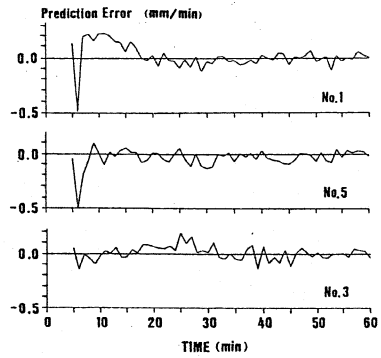

Fig. 5. The 5-min-ahead prediction error for three gages.

$$\mathbf{Y}(k) = \mathbf{J}[\mathbf{X}^*_G(k)]\mathbf{X}_G(k) + \mathbf{R}[\mathbf{X}^*_G(k)] - \mathbf{J}[\mathbf{X}^*_G(k)]\mathbf{X}^*_G(k) + \mathbf{w}(k) \quad (31)$$

where $\mathbf{J}[\mathbf{X}^*_G]$ is the Jacobian matrix for the vector function $\mathbf{R}(\mathbf{X}_G)$ in terms of the vector $\mathbf{X}_G$ defined as

$$\mathbf{J}[\mathbf{X}^*_G] = \left[\frac{\partial \mathbf{R}}{\partial I}, \frac{\partial \mathbf{R}}{\partial u}, \frac{\partial \mathbf{R}}{\partial D_x}, \frac{\partial \mathbf{R}}{\partial D_y}, \frac{\partial \mathbf{R}}{\partial \gamma}, \frac{\partial \mathbf{R}}{\partial x_0}, \frac{\partial \mathbf{R}}{\partial y_0}, \frac{\partial \mathbf{R}}{\partial t_0}\right]_{\mathbf{X}_G = \mathbf{X}^*_G} \quad (32)$$

Thus the observation matrix $\mathbf{\Gamma}$ and the constant vector $\boldsymbol{\beta}$ in the observation equation (11) for the Gaussian solution are (compare (11) and (31))

$$\mathbf{\Gamma}(k) = \mathbf{J}[\mathbf{X}^*_G(k)] \quad (33)$$

$$\boldsymbol{\beta}(k) = \mathbf{R}[\mathbf{X}^*_G(k)] - \mathbf{J}[\mathbf{X}^*_G(k)]\mathbf{X}^*_G(k) \quad (34)$$

The term $\mathbf{X}^*_G(k)$ is replaced by $\hat{\mathbf{X}}_G(k|k-j)$, which is the estimate of the state vector $\mathbf{X}_G$ at time step k based on the observation at time step $k-j$.

2.2.2. *Formulation of the Fourier series expansion.* When employing the Fourier series expansion ((3)–(6)), the system vector $\mathbf{X}_F(t)$ to be estimated by the extended Kalman filter is

$$\mathbf{X}_F(t) = [u, D_x, D_y, \gamma, A_{0,0}(t), \cdots, A_{m,n}(t), B_{m,n}(t), C_{m,n}(t), D_{m,n}(t), \cdots]^T \quad (35)$$

The number of elements in $\mathbf{X}_F(t)$ is $4MN + 2M + 2N + 5$. The true values of the parameters u, D_x, D_y, and γ are assumed to be constant in time and space in the extended Kalman filter so that

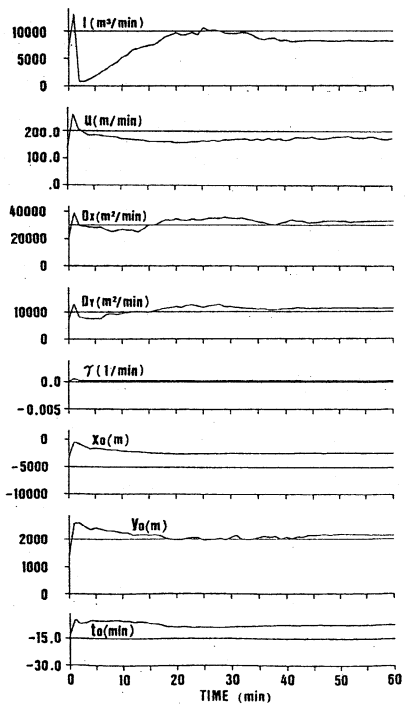

Fig. 6. Results of the identification procedure for estimating unknown model parameter values (the horizontal lines depict the true values).

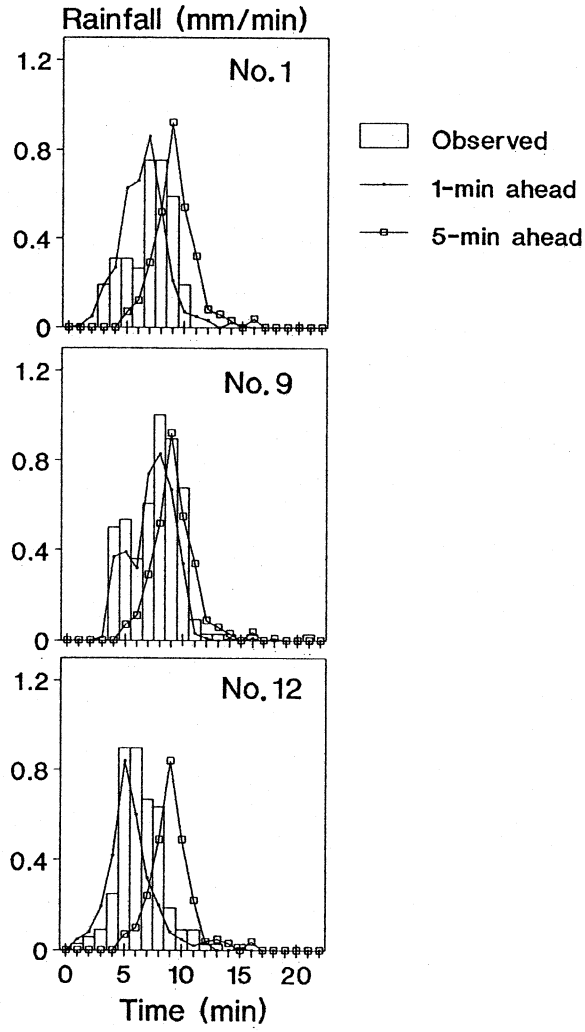

Fig. 7. Time series of 1- and 5-min-ahead real-time predictions for the event of June 25, 1979, 0101–0122 using simple extrapolation calculations.

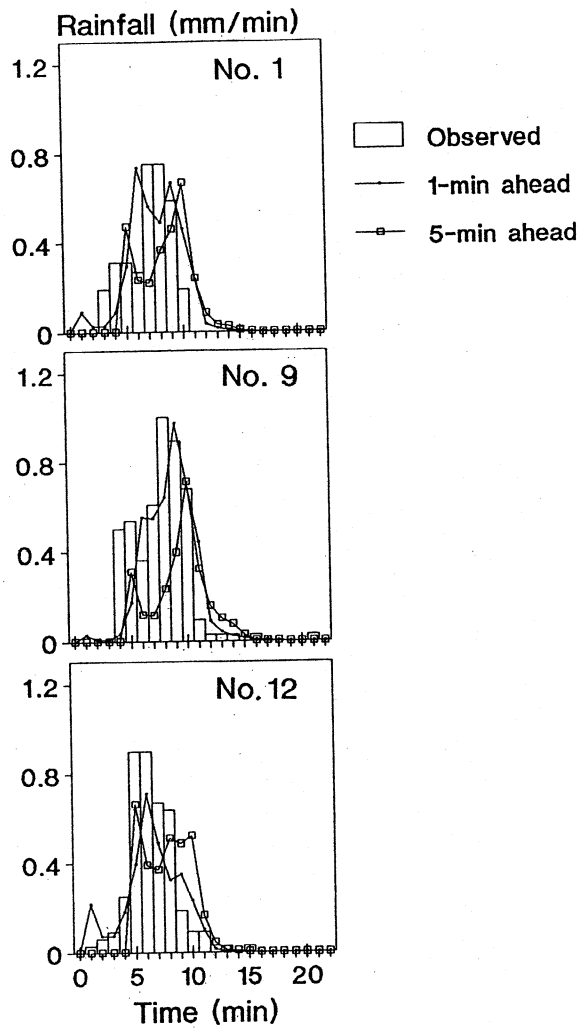

Fig. 8. Time series of 1- and 5-min-ahead real-time predictions for the event of June 25, 1979, 0101–0122 using the Gaussian solution.

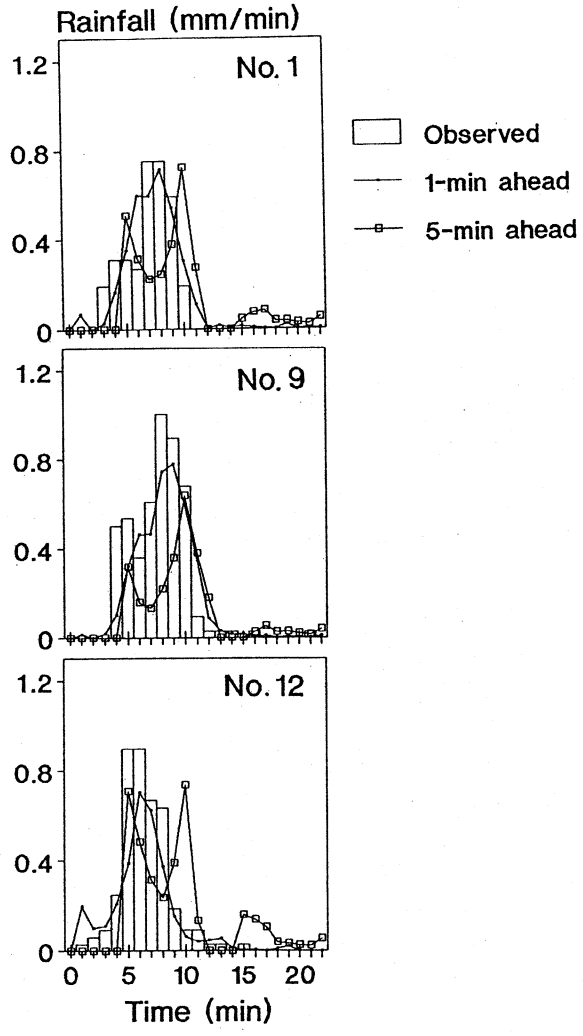

Fig. 9. Time series of 1- and 5-min-ahead real-time predictions for the event of June 25, 1979, 0101–0122 using the Fourier series approach.

TABLE 1. Comparison of Model Performance for a Single-Cell Event Using Simple Extrapolation, the Gaussian Solution, and the Fourier Series Approach

System Noise Level (v), %	Identified Final Parameters				Prediction Error, mm/min	
	u, m/min	D_x, $\times 10^5$ m^2/min	D_y, $\times 10^5$ m^2/min	γ, $\times 10^{-3}$ min^{-1}	$J_{1\text{-min}}$	$J_{5\text{-min}}$
			Simple Extrapolation			
...	...	...	...	...	0.18	0.38
			Gaussian Solution			
0	1030	5.0	2.5	−5.8	0.23	0.26
2	1333	4.2	2.5	1.6	0.14	0.20
5	993	−5.1	4.9	2.3	0.31	0.43
			Fourier Series Expansion			
0	1155	5.6	2.8	1.1	0.12	0.22
2	1694	6.7	3.6	0.3	0.13	0.26
5	1522	11.7	5.0	0.0	0.13	0.27

Data are for the event of June 25, 1979, 0101–0122 (Figures 7–9).

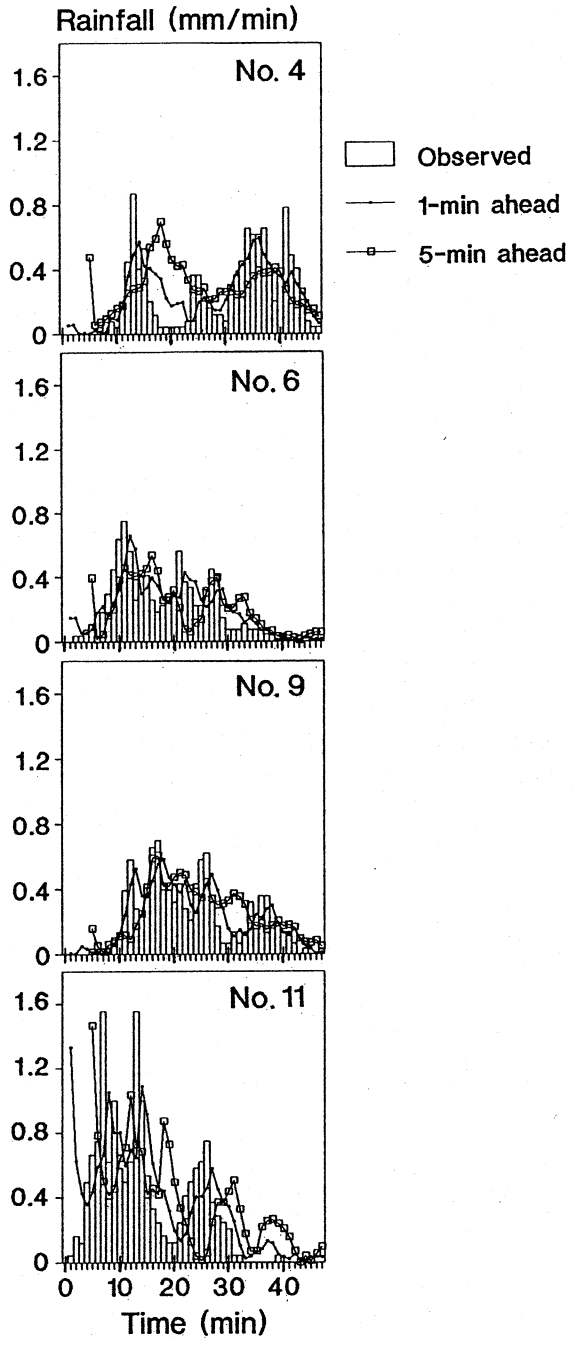

Fig. 10. Time series of 1- and 5-min-ahead real-time predictions for the event of July 13, 1980, 1047–1133 using the Fourier series approach.

$$du(t)/dt = 0 \tag{36}$$

$$dD_x(t)/dt = 0 \tag{37}$$

$$dD_y(t)/dt = 0 \tag{38}$$

$$d\gamma(t)/dt = 0 \tag{39}$$

As pointed out before, however, this does not mean that the parameter values do not change during the estimation procedure of the extended Kalman filter.

Combining (7) and (36)–(39), and adding a system noise $\mathbf{v}(t)$, the system equation is obtained as

$$d\mathbf{X}_F(t)/dt = \mathbf{f}[\mathbf{X}_F(t)] + \mathbf{v}(t) \tag{40}$$

where $\mathbf{f}[\mathbf{X}_F(t)]$ is the vector function which elements are equal to the right-hand side of (36)–(39) and the first term of the right-hand side of (7). The term $\mathbf{f}[\mathbf{X}_F(t)]$ in (40) is a set of nonlinear functions of $u(t)$, $D_x(t)$, $D_y(t)$, $\gamma(t)$, $A_{0,0}(t)$, $A_{m,n}(t)$, $B_{m,n}(t)$, $C_{m,n}(t)$, and $D_{m,n}(t)$. Here, a Taylor series expansion of $\mathbf{f}[\mathbf{X}_F(t)]$ is carried out, and second-order terms or higher are neglected in order to linearize for the extended Kalman filter to yield

$$d\mathbf{X}_F(t)/dt = \mathbf{J}[\mathbf{X}^*_F(t)]\mathbf{X}_F(t) + \mathbf{f}[\mathbf{X}^*_F(t)] - \mathbf{J}[\mathbf{X}^*_F(t)]\mathbf{X}^*_F(t) + \mathbf{v}(t) \tag{41}$$

where $\mathbf{J}[\mathbf{X}^*_F(t)]$ is the Jacobian matrix for the vector function $\mathbf{f}[\mathbf{X}_F(t)]$ taken at $\mathbf{X}_F(t) = \mathbf{X}^*_F(t)$. The ijth component of $\mathbf{J}[\mathbf{X}^*_F(t)]$ is defined as

$$\{\mathbf{J}[\mathbf{X}^*_F(t)]\}_{ij} = \left.\frac{\partial\{\mathbf{f}[\mathbf{X}_F(t)]\}_i}{\partial\{\mathbf{X}_F(t)\}_j}\right|_{\mathbf{X}_F(t)=\mathbf{X}^*_F(t)} \tag{42}$$

Discretization of (41) by difference gives

$$\mathbf{X}_F(t+\Delta t) = \{\mathbf{J}[\mathbf{X}^*_F(t)]\Delta t + \mathbf{I}_U\}\mathbf{X}_F(t) + \{\mathbf{f}[\mathbf{X}^*_F(t)] - \mathbf{J}[\mathbf{X}^*_F(t)]\mathbf{X}^*_F(t)\}\Delta t + \mathbf{v}(t)\Delta t \tag{43}$$

Comparing (43) and (10), and letting $t = k$, $t + \Delta t = k + 1$, one finds that

$$\Phi(k) = \{\mathbf{J}[\mathbf{X}^*_F(t)]\Delta t + \mathbf{I}_U\} \tag{44}$$

$$\alpha(k) = \{\mathbf{f}[\mathbf{X}^*_F(t)] - \mathbf{J}[\mathbf{X}^*_F(t)]\mathbf{X}^*_F(t)\}\Delta t \tag{45}$$

$$\mathbf{v}(k) = \mathbf{v}(t)\Delta t \tag{46}$$

The vector $\mathbf{X}^*_F(k)$ is replaced by $\hat{\mathbf{X}}_F(k|k-j)$, which is the estimate of the state vector $\mathbf{X}_F(k)$ at time step k based on the observation at time step $k - j$.

The unknown parameters and the coefficients in the double Fourier series expansion are identified in the prediction process by the extended Kalman filter [*Athans et al.*, 1968; *Kawamura et al.*, 1986]. The observation vector $\mathbf{Y}(k)$ is the rainfall distribution given by the measurements at each rain gage. The observation equation for rain gage i, located at point (x_i, y_i), is derived for time step k from (3), and by adding the observation noise $w(x_i, y_i, k)$ it becomes

$$R(x_i, y_i, k) = [0, 0, 0, 0, 1/2, \cdots, \cos F_1(x_i, y_i, m, n), \sin F_1(x_i, y_i, m, n), \cos F_2(x_i, y_i, m, n), \sin F_2(x_i, y_i, m, n), \cdots] \cdot [u, D_x, D_y, \gamma, A_{0,0}(k), \cdots, A_{m,n}(k), B_{m,n}(k), C_{m,n}(k), D_{m,n}(k), \cdots]^T + w(x_i, y_i, k) \tag{47}$$

Consequently, the ith row component of the observation matrix $\Gamma(k)$ in (11) corresponds to

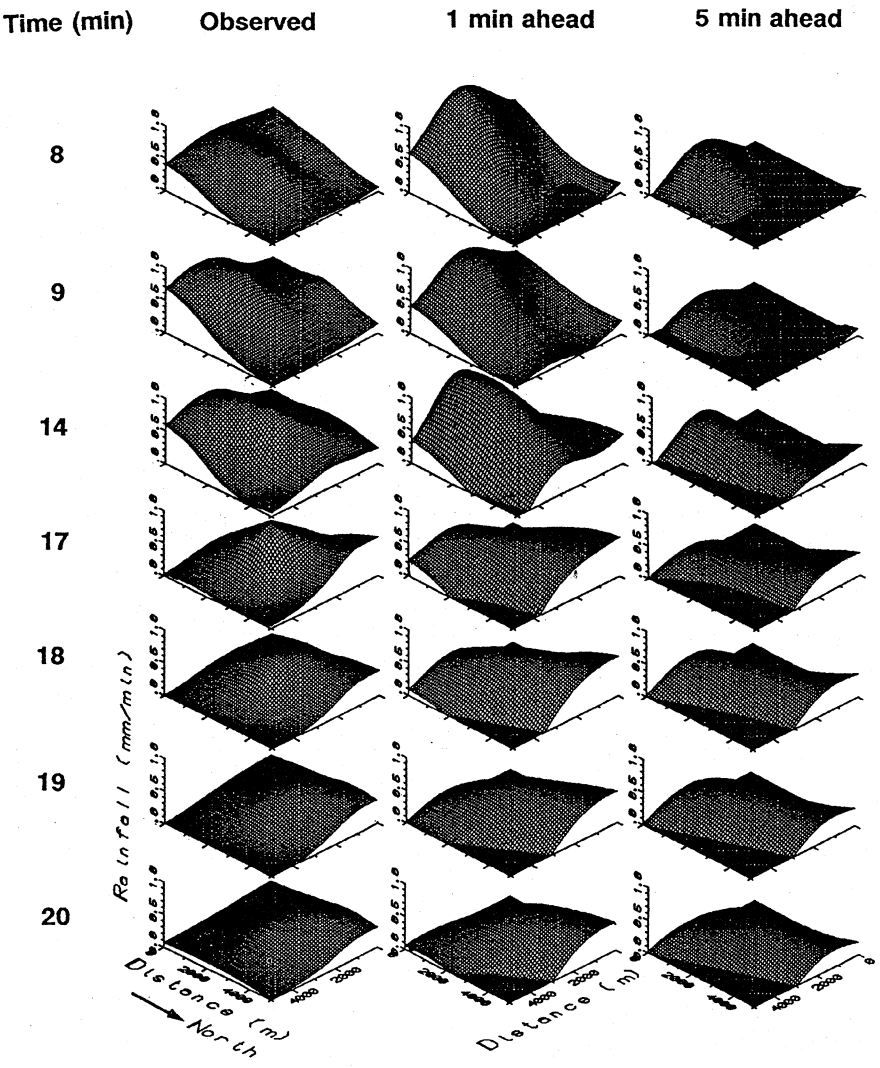

Fig. 11. Time series of the spatial rainfall distribution for the event in Figure 10.

$$\{\Gamma(k)\}_i = [0,\ 0,\ 0,\ 0,\ 1/2,\ \cdots,\ \cos F_1(x_i, y_i, m, n),$$
$$\sin F_1(x_i, y_i, m, n),\ \cos F_2(x_i, y_i, m, n),$$
$$\sin F_2(x_i, y_i, m, n),\ \cdots] \quad (48)$$

and

$$\beta(K) = 0 \quad (49)$$

Further details on applications of the extended Kalman filter can be found in the works by *Bras and Rodriguez-Iturbe* [1985], *Kawamura et al.* [1986], *Jinno et al.* [1989], and *Kawamura and Jinno* [1991].

3. Real-Time Prediction Using Generated Rainfall

The previously discussed Gaussian solution was employed for real-time prediction of rainfall from a synthetic rain cell with an elliptic shape moving over the observation network in the city of Lund (Figure 2). These calculations constituted the first test of the applicability of the proposed methodology for rainfall prediction. Figure 3 shows the rain cell moving over the rain gage network. The parameter values in (2) for the rain cell generation were set to I = 10,000 (m^3/min), u = 200 (m/min), D_x = 30,000 (m^2/min), D_y = 10,000 (m^2/min), γ = 0 (min^{-1}), x_0 = −5000 (m), y_0 = 2000 (m), and t_0 = −15 (min). The Gaussian solution was used together with the corresponding initial values for the prediction of the eight unknown model parameters I, u, D_x, D_y, γ, x_0, y_0, and t_0. These were set to 70% (the seven first parameters) and t_0 to 90% of their true values, respectively. An observation noise of 0.05 mm/min was also added to the generated rainfall intensities. The Gaussian solution proved to be rather sensitive to the initial parameter values. Consequently, the initial values had to be chosen fairly close to the true parameter values (70–90% as seen above).

Both 1- and 5-min-ahead real-time predictions of the rainfall intensity were carried out together with the corre-

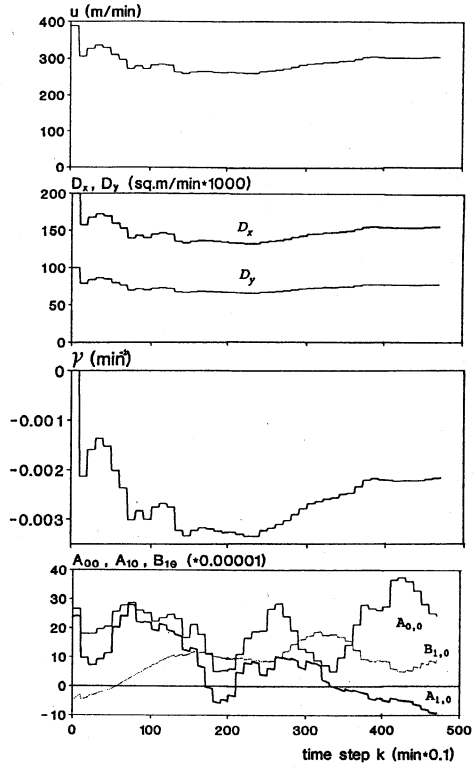

Fig. 12. Parameter identification for the event in Figures 10 and 11.

sponding estimation of the eight unknown parameters. Figure 4 displays the results of the prediction for three different gages (1, 3, and 5). It appears that the model is capable of predicting both 1- and 5-min-ahead intensities well. For gage 1, however, there is some discrepancy in the 5-min-ahead prediction during the rising stage of the hyetograph. As seen in Figure 5, which shows the corresponding 5-min-ahead prediction error for the same three gages, it takes the model about 10 min to correctly identify the parameters at gage 1.

Figure 6 illustrates the results of the parameter estimation, and the identified parameter values at the end of the calculation are very close to their true values.

Also, the Fourier series approach was applied to a synthetic rain cell. For a detailed description of the outcome of these calculations and the relation between the number of observation points, observation time interval, number of expansion terms, and prediction error, see *Kawamura et al.* [1990]. In brief, however, the number and distribution of gages appeared to be sufficient for single-cell evolutions. The number of terms in the Fourier series expansion was set to $M = N = 5$ which was found to be sufficient (cases up to $M = N = 10$ were tested). The Fourier series approach appeared more flexible and less sensitive to choice of initial values as compared to the Gaussian solution. When comparing the computational time, however, the Fourier series approach requires about 30 times the computational time than that of the Gaussian solution. This is mainly due to the larger number of parameters to estimate in the Fourier series approach. Even so, the total computational time for the Fourier series approach does not pose any serious restrictions for practical applications as shown below.

Generally, the number of terms required in the Fourier series expansion depends on the spatial rainfall variability and the density of the rain gage network (the smaller-scale features to be monitored, the more terms). This will also increase the computational load. In the present example with 12 gages, $M = N = 5$ were found to be sufficient. The CPU time required for 5-min-ahead prediction in this case (including interpolation over the model domain) was about 0.5 s using a supercomputer (Facom VP2600). Calculations on a more ordinary computer (Facom M1800) required about 2.5 s for a similar calculation. Thus the computational load does not pose any limitations for practical applications. It also needs to be pointed out that actual rain gage networks in field applications usually are less dense than in the present example. Therefore the number of terms needed in the Fourier series approach that was used in the present paper may be regarded as an upper value.

4. REAL-TIME PREDICTION USING OBSERVED RAIN GAGE DATA

In order to elaborate on the choice of the two model approaches described in section 2, several modeling analyses were performed by use of 10 high-intensive rainfall events described by *Niemczynowicz and Jönsson* [1981]. In this paper selected results from the calculations are presented that have been chosen to reflect the typical features of these two approaches. These modeling approaches were also

TABLE 2. Comparison of Model Performance for Two Multiple-Cell Events (Shown in Figures 10–16) Using the Fourier Series Approach

System Noise Level (v), %	Identified Final Parameters: u, m/min	D_x, $\times 10^5$ m^2/min	D_y, $\times 10^5$ m^2/min	γ, $\times 10^{-3}$ min^{-1}	Prediction Error, mm/min: $J_{1\text{-min}}$	$J_{5\text{-min}}$
		Event of July 13, 1980, 1047–1133 Hours				
0	304	1.6	0.8	−2.2	0.20	0.26
2	368	0.4	1.3	9.5	0.21	0.43
5	404	2.4	1.2	0.2	0.22	127
		Event of August 23, 1980, 1004–1103 Hours				
0	564	6.9	3.5	3.8	0.17	0.29
2	101	0.2	2.5	9.8	0.17	0.47
5	−18	4.8	−1.9	62.2	0.17	1.06

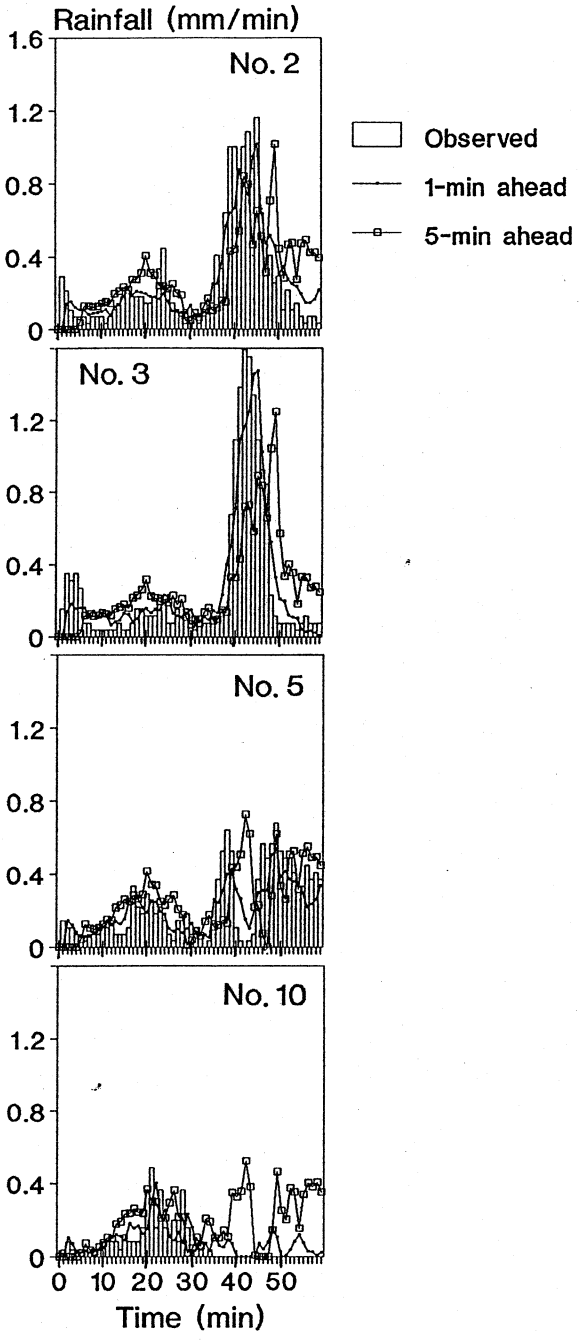

Fig. 13. Time series of 1- and 5-min-ahead real-time predictions for the event of August 23, 1980, 1004–1103 using the Fourier series approach.

compared with simple extrapolation calculations as outlined below.

In the related paper by Berndtsson et al. (submitted manuscript, 1992, Figure 6) it was shown that simple extrapolation of the rainfall field in the direction of movement for the studied space-time scales may generally be expected to behave unsatisfactorily for lead times exceeding about 2 min (correlation level of 0.9). The general performance of extrapolation calculations should be seen in relation to the space and time scales dealt with. For larger spatial scales, it is probable that good forecasts for 1 to 5 min ahead can be obtained by extrapolation of a frozen rainfall field. For spatial scales of a few square kilometers this is often not the case. In order to further evaluate the developed model, however, extrapolation calculations were done in order to compare with model results.

Figure 7 shows an example of results from an extrapolation calculation applied to an event with a rather clear single-cell evolution (event of June 25, 1979, 0101–0122). A constant rain cell speed and direction of movement were chosen based on observations at the 700-mbar altitude [see *Niemczynowicz*, 1984]. Interpolations were done by general kriging procedures, and extrapolations were performed by assuming that the rain cell will continue to travel without change in the general direction of movement.

Figure 7 shows that for 1 min ahead the predictions agree rather well with observations even if small-scale variations are not reproduced. The prediction error for a lead time of 1 min was 0.18 mm/min as an average for the 12 gages (Table 1). For 5-min-ahead predictions the agreement does not appear to be very good. This stems from the difficulty in extrapolating something outside the observation network that has not yet been observed. The prediction error for this lead time was 0.38 mm/min (Table 1). If applied to events with abrupt changes in rainfall intensity, the method may be expected to behave worse, especially for larger lead times.

4.1. *Prediction of a Single-Cell Event Using the Gaussian Solution*

The first model approach involved the Gaussian solution, and (2) was applied to the same event as for the extrapolation calculations (Figure 7). This approach proved generally to work well for events with rather low rainfall intensity and no abrupt changes in that variable. Rainfall evolutions involving high-peaked intensities during short periods were especially difficult to reproduce since the initial guess for the physical parameters becomes crucial. The rainfall event, however, shows a rather smooth rainfall evolution with no abrupt intensity changes, and consequently the prediction by the Gaussian solution agrees well with the observations (minimum prediction error averaged over all 12 gages was 0.14 mm/min; Table 1). The 5-min prediction is less accurate (minimum average prediction error of 0.20 mm/min; Table 1). Even so, it appears better than simple extrapolation calculations as seen above.

The initial parameter values are, however, difficult to estimate in the Gaussian solution. This proved to be especially true for the rather small catchment area in Lund. Initial parameter values for the direction and speed of the rain cell u were taken from meteorological information at the 700-mbar altitude. The initial location x_0 and y_0, and occurrence time of the cell t_0, proved to be especially sensitive for the model performance and difficult to assess in the Lund catchment. The parameter values for the case in Figure 8 were set to $I = 5 \times 10^4$ (m^3/min), $u = 1036$ (m/min), $D_x = 5 \times 10^5$ (m^2/min), $D_y = 2.5 \times 10^5$ (m^2/min), $\gamma = 0$ (min^{-1}), $x_0 = -15{,}000$ (m), $y_0 = -2000$ (m), and $t_0 = -10$ (min). Generally, the initial values for these parameters had to be chosen as 70–80% of their true values in order to get a reasonable model performance. This, in turn, limits the

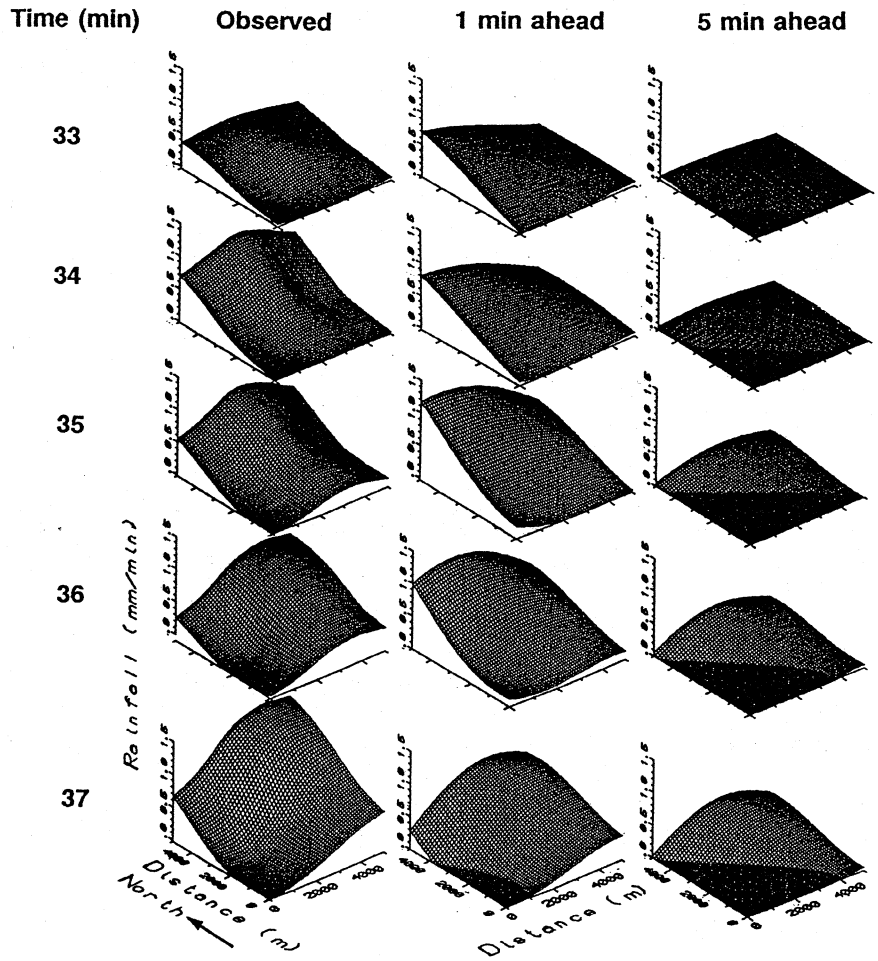

Fig. 14. Time series of the spatial rainfall distribution for the event in Figure 13.

application of the model approach in real forecasting situations, and especially in gaged catchments that are small in relation to the size of the rain cells. A further factor that limits the application of the Gaussian approach is the occurrence of multiple cells. In theory, it is possible to include multiple cells in the modeling scheme by superposition of several Gaussian solutions. In practice, however, this will become tedious, and parameter values will be complicated to estimate.

4.2. *Prediction of a Single-Cell Event Using the Fourier Series Expansion*

Figure 9 shows real-time prediction for the same event as discussed above for the extrapolation and the Gaussian solution (Figures 7 and 8), using the Fourier series approach. The discrete time interval was set to $\Delta t = 0.1$ min, and the figure shows that the 1-min-ahead predictions agree well with the observations (minimum average prediction error of 0.12 mm/min; Table 1). Deviations from a Gaussian-shaped cell structure is rapidly compensated for by the Fourier series. Because of the short duration of the rainfall, 5-min-ahead predictions are difficult to make, and the fit to the observed data shows less agreement (minimum prediction error of 0.22 mm/min; Table 1). Even so, the agreement is generally better than for simple extrapolation and the Gaussian approach.

In all modeling examples employing the Fourier series approach, a Gaussian-shaped cell was used for estimating initial parameter values for the Fourier coefficients. This Gaussian-shaped cell was generated assuming an initial cell strength I corresponding to 500 m/min over a square area of 100 m^2. The initial location for this structure was chosen according to observed speed and wind direction and the initial values given above for the Gaussian solution ($t_0 = -10$ min). Consequently, the initial values for the physical parameters u, D_x, D_y, and γ were set to be the same for both the Gaussian and the Fourier series approach (Figures 8 and 9). The choice of initial parameter values proved to be much less sensitive for the model performance in the Fourier series approach compared to the Gaussian solution. Consequently, the Fourier series approach appears to be more flexible and more suitable for practical forecasting applications.

Prior to application of the Fourier series approach to real rain gage data, test runs were made to check the sensitivity of the choice of wavelengths λ_x, λ_y, and the number of

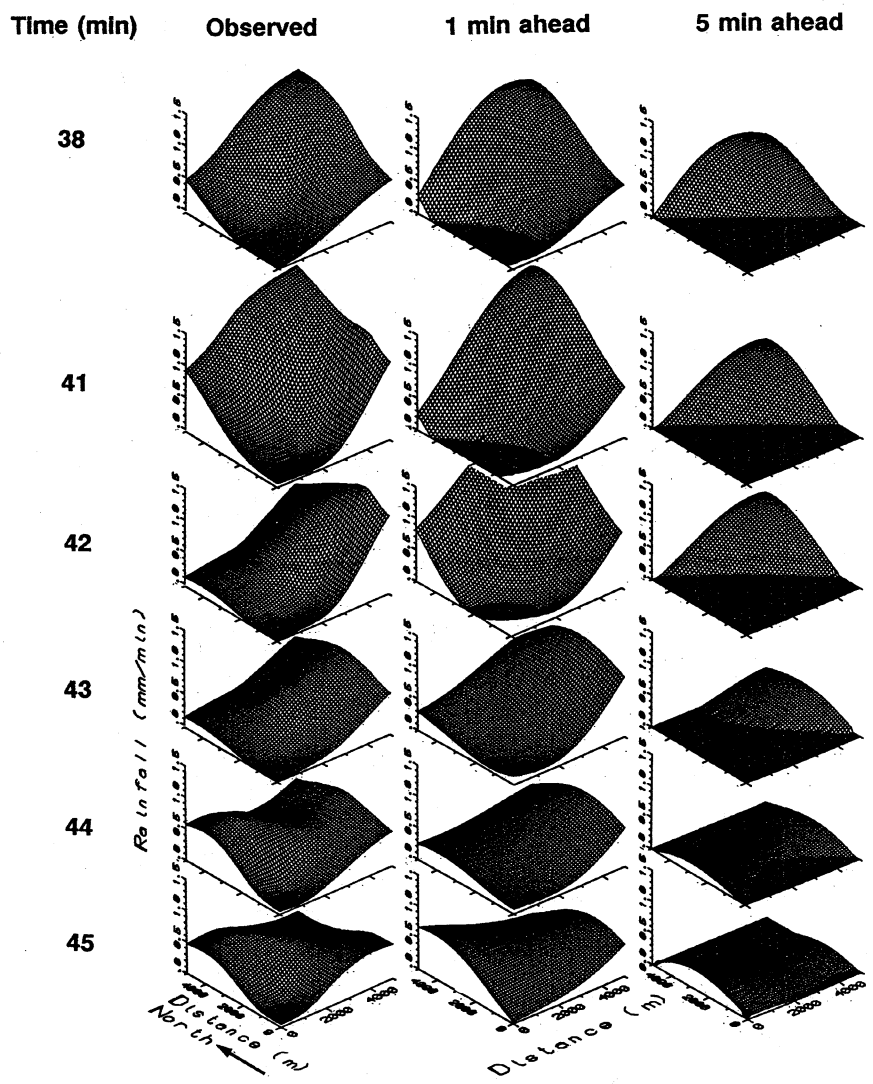

Fig. 14. (continued)

terms in the Fourier series expansion. It was found that the prediction performance is not significantly affected by the choice of wavelengths. This is because the Fourier coefficients compensate each other and the variation in the physical parameters tends to be more dependent on the system noise level. In this case the wavelengths λ_x and λ_y were set to 30,000 and 10,000 m, respectively. The number of terms necessary in the Fourier series expansion can be tested by monitoring the magnitude of higher-order Fourier coefficients. If the magnitude is negligibly small, then higher-order terms are not necessary. Thus in the present modeling examples, both M and N were set to 5.

Table 1 shows a comparison between the extrapolation technique, the Gaussian solution, and the Fourier series approach, respectively, for the event displayed in Figures 7–9. The table shows the identified final parameter values as a function of system noise level for the Gaussian solution and the Fourier series approach, and the corresponding prediction errors (the prediction error was calculated as the root-mean-square error for all gages for the model period). The system noise level is expressed as percent of the initial value for the physical parameters (i.e., the eight parameters in the Gaussian solution and the four physical parameters in the Fourier series approach). The system noise level for all Fourier coefficients (except $A_{0,0}(t)$, $A_{m,0}(t)$, $B_{m,0}(t)$, $C_{0,n}(t)$, and $D_{0,n}(t)$) in the Fourier series approach was set to a constant value of 5% of the initial power spectrum for the most dominant wave numbers $((A_{1,1}(0)^2 + B_{1,1}(0)^2 + C_{1,1}(0)^2 + D_{1,1}(0)^2)/4)$ and applied uniformly over the power spectrum for the system noise.

It is seen that the Fourier series approach in general gives the lowest prediction errors. An increase in the system noise level, however, appears generally to induce a higher prediction error for the Fourier series approach. A larger system noise level implies a larger fluctuation of the physical parameters. In some cases this may induce unreasonable estimates

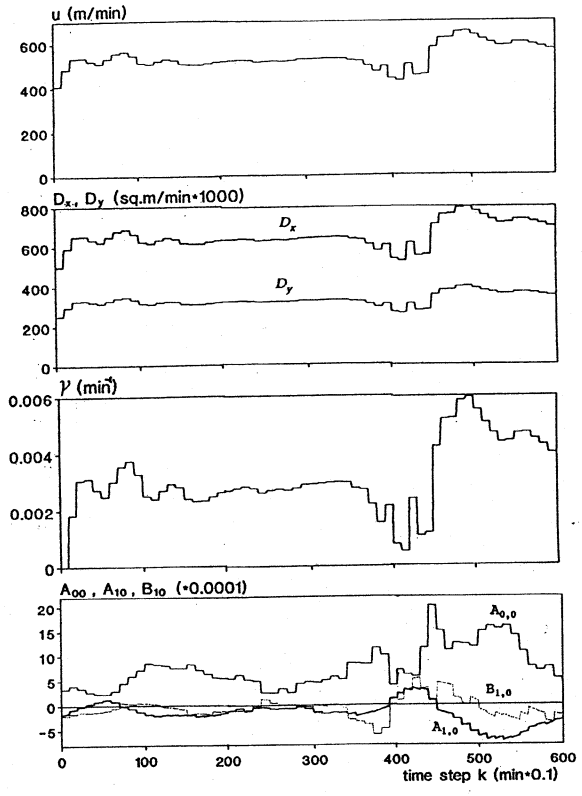

Fig. 15. Parameter identification for the event in Figures 13 and 14.

of the parameters even for the Gaussian solution. This is seen in Table 1 for a system noise level of 5% for the Gaussian solution. In this case the parameter D_x attained a negative final value. When the system noise level is set to zero or a small value, however, the parameters tend to converge to physically rational values. For the Gaussian solution, on the other hand, this means that the initial guess of parameter values has to be very close to the true values. When the system noise level is set to zero for the physical parameters in the Fourier series approach, the Fourier coefficients tend to adjust the changing behavior of the rainfall cell.

Because of the higher flexibility of the Fourier series approach it was decided to further investigate the model behavior for this approach using observed multiple-cell events as input. The results of this analysis is described in section 4.3.

4.3. *Prediction of Multiple-Cell Events Using the Fourier Series Expansion*

It should be noted that the Fourier series approach is not limited to single-cell events only. Any spatial shape in the rainfall intensity pattern may be represented by the Fourier series. The parameters of the model, however, are assumed uniform in space, which means that all cells in the modeling domain are assumed to behave similarly. This, however, may not be justified if a larger area is considered.

Figure 10 shows the Fourier series approach applied to a multiple-cell event (July 13, 1980, 1047–1133). The 1-min-ahead predictions in most cases give reasonable estimates. Figure 11 shows the corresponding time series of the spatial rainfall distribution during time 8–20 min (the time in Figure 11 corresponds to the time given in Figure 10). The observed values in the figure were obtained by general kriging procedures, implying that the values have been smoothed in space. Even though this procedure makes individual extreme peaks smooth, the general spatial cell evolution is clearly displayed in Figure 11. Thus Figure 11 shows a single cell as it moves across the catchment from southeast to northwest. The general spatial cell behavior for 1-min-ahead predictions is reproduced well by the model, whereas the 5-min-ahead predictions are less accurate. However, although the model underestimates the cell intensity peaks for 5-min-ahead predictions, it is seen that the general spatial characteristics agree rather well. The same initial model conditions as above (Figure 9) for the Fourier series coefficients were used assuming a Gaussian-shaped cell structure, which is seen in the first two diagrams for the 5-min-ahead predictions in Figure 11. It takes the model about 10 min to change from this cell structure to the observed shape (time 8–14 min in Figure 11).

The results of the parameter identification for the event in Figures 10 and 11 are illustrated in Figure 12. The major changes in the estimated parameter values occur during the passage of cell peaks (time 1–15, 20–30, and 35–45 min). The parameters u, D_x, and D_y, however, attain rather constant values compared to γ and the Fourier series coefficients. In the above example, the system noise level was set to zero, which proved to give reasonable model results and rational physical parameter values (see Table 2).

Figure 13 shows the observed and predicted time series for a second multiple-cell event (August 23, 1980, 1004–1103). Here the model also seems to be able to reasonably well predict 1- and 5-min-ahead rainfall intensities. A time series of the spatial distribution of the rainfall event in Figure 13, showing how the model is able to adapt to a multiple-peaked cell, is displayed in Figure 14. At least two cells seem to occur almost simultaneously over the catchment area (time 40–45 min). The first cell moves over the catchment from northwest to southeast (time 33–43 min). This cell is followed by a smaller one moving in the same direction (time 44–45 min). For 1-min-ahead predictions the model manages to reproduce these spatial changes. The 5-min-ahead predictions underestimate the intensities but appear to be able to reproduce the general spatial characteristics. The underestimation is to a major extent caused by the rapid changes which occur at time scales less than 5 min, which are difficult to consider in the 5-min-ahead predictions.

Figure 15 illustrates the corresponding result of the parameter identification for the case displayed in Figures 13 and 14. In this case the system noise was also set to zero, which made the smallest prediction errors concur with rational parameter values (Table 2).

Table 2 shows a summary of the two above multiple-cell events considering the system noise levels, identified final parameter values, and the prediction error. The system noise is defined as above for Table 1. As in Table 1 it can be seen that a high system noise level (5%) may give large prediction error and unreasonable physical parameter values. If the system noise level is kept at a modest level of 0–1%, rather small prediction errors are achieved.

Figure 16 shows the autocorrelation of the residuals be-

tween observed and predicted values at representative rain gage stations for the event given in Figures 10 and 11. The higher correlations for some stations indicate that improvements could be made regarding the state estimator. However, since the number of rain gages used in the modeling is small, it is difficult to calibrate the model efficiently in space (for multiple-cell events). The high autocorrelations for the 5-min values indicate the presence of correlated residuals for some gages that are mainly due to the lag between predictions and observations. The autocorrelation indicates that the rainfall area changes its properties rapidly, which is difficult for the model to consider.

Overall, the Fourier series approach demonstrated a high degree of flexibility to rapidly varying cell shapes. The less accurate 5-min-ahead predictions should be considered in relation to the choice of the rainfall events. The rainfall events used for the model testing were the most extreme observations during the 3-year measurement period. They proved to be exceptionally variable with high peak intensities during a few minutes. If rainfall events with longer duration and lower intensities had been chosen, the 5-min-ahead predictions would have been more in agreement with measurements. One of the main objectives of this paper was, however, to investigate if predictions can be made for short-term and extreme rainfall that are of major concern in urban hydrology. The paper also shows the general applicability of a Fourier domain shape method as a technique in urban hydrology and urban management.

5. SUMMARY AND DISCUSSION

It was shown that the two-dimensional stochastic advection-diffusion equation may be used together with a Fourier domain shape method and the extended Kalman filter to forecast small-scale space-time rainfall in real time. It is argued that urban hydrological problems impose special requirements regarding rainfall input to drainage and management models. In most urban areas, radar data are still not available on a sufficiently small scale to model individual cells. However, in areas where radar is installed these data may provide valuable information on larger-scale movement and thus be used as, e.g., boundary conditions for the proposed methodology [e.g., *Cluckie and Collier*, 1991].

Simple extrapolation calculations proved to agree well with observations for a lead time of 1 min. If the lead time exceeds about 2 min, however, the approach using one of the models presented may be expected generally to display better agreement between prediction results and measurements. The reason for this is that both model approaches include a physically based description of major cell features. A major difficulty, however, when using the Gaussian solution is the estimation of initial model parameters. Also, in some cases the rapidly changing cell structures are inadequately described by the Gaussian solution. The Fourier series approach, on the other hand, represents a flexible tool to predict rapidly changing cell structures in time and space. The choice of initial conditions is less sensitive for the model performance than for the Gaussian solution. Also, the Fourier series approach is not limited to single-cell events only.

It should be stressed that the Fourier series approach to model small-scale rainfall that is presented in this paper has several advantages. The small-scale rainfall over an urban area may vary greatly depending on the speed of the rain cells, development or decay of cells, and the number of simultaneous cells. This makes initial and boundary conditions very difficult to estimate in advance (if no radar observations are at hand). In the present model an algorithm is used in which the parameter values of the advection-diffusion equation are identified simultaneously with updates of predicted rainfall. This means that under unknown initial and boundary conditions, the rainfall distribution may still be predicted during the time when the parameters are being identified. Also the Fourier series approach requires less computational time compared to a numerical scheme.

The model is especially suitable for catchments with rain gage systems but can be made more flexible if radar data are available. Since a Fourier domain shape method is used to represent the rainfall field and the Gaussian noise, the model domain does not have to be discretized. This has special advantages for irregularly spaced gaging systems. Furthermore, including or excluding individual gages in the modeling, e.g., because of periodic malfunction or testing of suitable locations for new gages, becomes trivial.

Since the model uses a recursive estimation procedure for parameter identification it is suitable for inclusion in an interactive urban management system. Further, the model can be used effectively for runoff prediction allowing for small-scale spatial variability and dynamic influence on runoff generation. If larger catchment areas can be used, the lead time for prediction can be made longer and thus more time can be spent on, e.g., managing valves, pumps, etc.

Upon request, the models presented can be received from the authors.

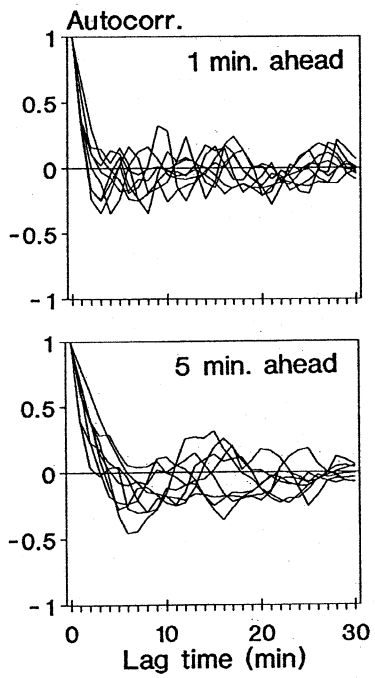

Fig. 16. Autocorrelation of the residuals between measured and predicted rainfall intensity for the event in Figures 10 and 11.

Acknowledgments. This research was partly supported by the Japan Society for the Promotion of Science and the Japanese

Ministry of Education by giving R. Berndtsson the possibility to spend a sabbatical year at the Kyushu University, and partly by the Hörjel foundation through a guest researcher grant allowing A. Kawamura to work at the University of Lund for 3 months. This support is gratefully acknowledged.

REFERENCES

Athans, M., R. P. Wishner, and A. Bertolini, Suboptimal state estimation for continuous-time nonlinear system from discrete noisy measurement, *IEEE Trans. Autom. Control, AG-13*, 504–514, 1968.

Bellon, A., and G. L. Austin, The accuracy of short-term radar rainfall forecast, *J. Hydrol., 70*, 35–49, 1984.

Berndtsson, R., and J. Niemczynowicz, Spatial and temporal scales in rainfall analysis: Some aspects and future perspectives, *J. Hydrol., 100*, 293–313, 1988.

Bohne, A. R., Extrapolative forecasting of precipitation and cloud fields, paper presented at Conference on Mesoscale Precipitation: Analysis, Simulation and Forecasting, Mass. Inst. of Technol., Cambridge, 1988.

Bras, R. L., and I. Rodriguez-Iturbe, *Random Functions and Hydrology*, 559 pp., Addison-Wesley, Reading, Mass., 1985.

Browning, K. A., and C. G. Collier, Nowcasting of precipitation systems, *Rev. Geophys., 27*, 345–370, 1989.

Brutsaert, W., More on an approximate solution for nonlinear diffusion, *Water Resour. Res., 10*, 1251–1252, 1974.

Calder, K. L., On the equation of atmospheric diffusion, *Q. J. R. Meteorol. Soc., 91*, 514–517, 1965.

Cluckie, I. D., and C. G. Collier (Eds.), *Hydrological Applications of Weather Radar*, 644 pp., Ellis Horwood, Chichester, England, 1991.

Einfalt, T., and T. Denoeux, Radar rainfall forecasting for real-time control of a sewer system, in *Urban Drainage Hydraulics and Hydrology, Proceedings, Fourth International Conference on Urban Storm Drainage, Lausanne, 31 Aug.–4 Sep., 1987*, edited by B. C. Yen, pp. 47–48, International Association for Hydraulic Research, Lausanne, Switzerland, 1987.

Georgakakos, K. P., A generalized stochastic hydrometeorological model for flood and flash flood forecasting, 1, Formulation, *Water Resour. Res., 22*, 2083–2095, 1986.

Georgakakos, K. P., and R. L. Bras, A hydrologically useful station precipitation model, 1, Formulation, *Water Resour. Res., 20*, 1585–1596, 1984.

Georgakakos, K. P., and M. D. Hudlow, Quantitative precipitation forecast techniques for use in hydrological forecasting, *Bull. Am. Meteorol. Soc., 65*, 1186–1200, 1984.

Jinno, K., A. Kawamura, T. Ueda, and H. Yoshinaga, Prediction of the concentration distribution of groundwater pollutants, in Groundwater Management, Quantity and Quality, Proceedings of the Benidorm Symposium, October, 1989, *IAHS Publ., 188*, 131–142, 1989.

Jinno, K., A. Kawamura, M. Larson, R. Berndtsson, and J. Niemczynowicz, A convective-dispersion model for real-time predictions of urban-scale rainfall, in *Planning, Control and Regional Studies, Proceedings, 5th International Conference on Urban Storm Drainage, Osaka, Japan*, vol. 3, edited by Y. Iwasa and T. Sueishi, pp. 1261–1267, Osaka University, Osaka, Japan, 1990.

Kalman, R. E., A new approach to linear filtering and prediction problems, *J. Basic Eng., 82D*, 35–45, 1960.

Kalman, R. E., and R. S. Bucy, New results in linear filtering and prediction theory, *J. Basic Eng., 83D*, 95–108, 1961.

Kawamura, A., and K. Jinno, Prediction method of concentration distribution of groundwater pollutants using a field monitoring network systems, in *Proceedings, International Symposium on Ground Water, Nashville, Tennessee*, edited by G. P. Lennon, pp. 144–149, American Society of Civil Engineers, New York, 1991.

Kawamura, A., K. Jinno, and T. Ueda, Detection of abnormality by the adaptive Kalman filtering, (in Japanese with English abstract), *Proc. Jpn. Soc. Civ. Eng., 345*(II-1), 111–121, 1984.

Kawamura, A., K. Jinno, T. Ueda, and R. R. Medina, Detection of abrupt changes in water quality time series by the adaptive Kalman filter, in Monitoring to Detect Changes in Water Quality Series, Proceedings of the Budapest Symposium, July 1986, *IAHS Publ., 157*, 285–296, 1986.

Kawamura, A., K. Jinno, T. Ueda, and H. Yoshinaga, On the on-line prediction of the concentration distribution of one-dimensional constant coefficient stochastic convective-dispersion equation based on Kalman filter, (in Japanese with English abstract), *Tech. Rep. Kyushu Univ., 62*(1), 1989.

Kawamura, A., K. Jinno, and H. Yoshinaga, On the real-time prediction of rainfall distribution by a stochastic convection-diffusion algorithm, (in Japanese), in *Proceedings of Annual Conference of Japan Society of Hydrology and Water Resources*, pp. 102–105, Japan Society of Hydrology and Water Resources, Tokyo, 1990.

Kawamura, A., K. Jinno, J. Niemczynowicz, R. Berndtsson and M. Larson, On the temporal and spatial characteristics of short-term urban-scale rainfall and its real-time prediction, (in Japanese with English abstract), *Proc. Hydraul. Eng. Jpn. Soc. Civ. Eng., 35*, 63–68, 1991.

Kumar, P., and E. Foufoula-Georgiou, Fourier domain shape analysis methods: A brief review and an illustrative application to rainfall area evolution, *Water Resour. Res., 26*, 2219–2227, 1990.

Marshall, R. J., The estimation and distribution of storm movement and storm structure, using a correlation analysis technique and raingauge data, *J. Hydrol., 48*, 19–39, 1980.

Niemczynowicz, J., An investigation of the areal and dynamic properties of short-term rainfall and its influence on runoff generating processes, *Rep. 1005*, 215 pp., Dep. of Water Resour. Eng., Lund Inst. of Sci. and Technol., Univ. of Lund, Lund, Sweden, 1984.

Niemczynowicz, J., and O. Jönsson, Extreme rainfall events in Lund 1979–80, *Nord. Hydrol., 12*, 129–142, 1981.

Sharon, D., Spatial analysis of rainfall data from dense networks, *Hydrol. Sci. Bull., 17*, 291–300, 1972.

Zawadzki, I. I., Statistical properties of precipitation patterns, *J. Appl. Meteorol., 12*, 459–472, 1973.

R. Berndtsson, M. Larson, and J. Niemczynowicz, Department of Water Resources, University of Lund, Box 118, S-22100 Lund, Sweden.

K. Jinno and A. Kawamura, Department of Civil Engineering (SUIKO), Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812, Japan.

(Received April 20, 1992;
revised November 16, 1992;
accepted December 2, 1992.)