Hydrological Sciences - Journal - des Sciences Hydrologiques, **35**, 3, 6/1990

A dynamic model for short-scale rainfall disaggregation

DEMETRIS KOUTSOYIANNIS & THEMISTOCLE XANTHOPOULOS
National Technical University of Athens, Department of Civil Engineering, Division of Water Resources, 5 Iroon Polytechniou Str., 157 73 Zografou, Greece

Abstract The single-site dynamic disaggregation model developed and presented in this paper is a generalized step-by-step approach to stochastic disaggregation problems. The forms studied concern low-level variables with Markovian structure and normal or gamma marginal distributions. Combined with a rainfall model, the disaggregation scheme gives a rainfall generator transforming monthly rainfall into events and hourly amounts. A particular application of the generator, based on historical data, is used to illustrate and test the model.

Un modèle dynamique de désagrégation de pluies à courte échelle

Résumé Le modèle de désagrégation dynamique mono-site presenté ici, est une approche généralisée pas à pas des problèmes de désagrégation stochastique. Les formes étudiées concernent des variables de niveau inférieur avec une structure Markoviènne et une distribution marginale normale ou gamma. Combiné avec un modèle de pluie, le schéma de désagrégation a fourni un générateur de pluie, transformant les pluies mensuelles en événements et pluies horaires. Une application de ce générateur, basée sur les données historiques, est donnée afin d'illustrer et tester le modèle.

INTRODUCTION

The linear disaggregation model developed by Valencia & Schaake (1972, 1973), along with the contributions of Mejia & Rousselle (1976), Tao & Delleur (1976), Hoshi & Burges (1979), Todini (1980) and Stedinger & Vogel (1984) is the most important scheme for stochastic disaggregation problems, and has been widely used in hydrological applications. However, as first pointed out by Valencia & Schaake (1972), their model is not suitable for the disaggregation of rainfall in time scales shorter than monthly (e.g. daily); this is due to the skewed distributions and the intermittent aspect of the rainfall process in short time scales. Other disaggregation models have been proposed and used, particularly for the disaggregation of rainfall, but they do not exhibit the generality of the Valencia-Schaake linear scheme. Grace & Eagleson (1966) proposed an urn model for the disaggregation of storm depth

Open for discussion until 1 December 1990

into shorter durations. Schaake *et al.* (1972) developed a Markov chain model for the disaggregation of monthly rainfall into daily amounts. Woolhiser & Osborn (1985) presented a scheme for the disaggregation of an individual storm depth into fractional depths, each corresponding to one tenth of the storm's duration; their scheme was based on a non-dimensionalized Markov process resulting from successive transformations of the real rainfall process. The problem of the simulation of the internal time distribution of a storm was also studied by Marien & Vandewiele (1986), where the disaggregation scheme developed applies to properly defined fractional variables with gamma distributions.

The single-site dynamic disaggregation model developed and presented in this paper is a generalized step-by-step approach to stochastic disaggregation problems. The model development was intended for application to short-scale rainfall disaggregation problems. Combined with a rainfall model, the disaggregation scheme gives a rainfall generator disaggregating monthly rainfall into events and hourly amounts.

THE DYNAMIC DISAGGREGATION MODEL

The essential elements of the dynamic disaggregation model, described in detail by Koutsoyiannis (1988), are the following:

(a) The disaggregation of a high-level variable, Z, into its k components (low-level variables, X_i, $i = 1, ..., k$), is performed in $k - 1$ sequential steps.

(b) At the beginning of the ith step, the amount-still-to-go, S_i, is known, and X_i is generated. The remaining quantity $S_{i+1} = S_i - X_i$ is transferred to the next step.

(c) In each step the distribution function of (X_i, S_i), conditional on previously generated information, is determined or approximated via conditional moments. It is assumed that the sequence of X_i has certain properties allowing the calculation of conditional moments, e.g. it is an autoregressive sequence.

(d) The generation of X_i is performed by the so called **bisection procedure**, which can take several forms depending on the particular marginal distribution of the low-level variables.

The realization of the model includes two parts, the conditional moments determination and the bisection procedure, which can be studied separately. The former is influenced by the type of the stochastic structure of the successive low-level variables, while the latter is affected mainly by its marginal distribution type.

Though the model is general in its formulation, the configurations studied concern only single-site problems, described by Markov sequences, with Gaussian or gamma marginal distributions. Therefore it may be used in any single-site hydrological application with variables fulfilling or approximating these conditions; its applicability to short-scale rainfall disaggregation, characterized by intermittence and J-shaped distributions, is emphasized.

A brief presentation follows of the model equations, oriented towards model application. The relevant proofs may be found partly in Koutsoyiannis (1988).

Conditional moments determination

Let the low-level variables X_i, $i = 1, .., k$, add up to the high-level variable, Z, that is:

$$X_1 + X_2 + \ldots + X_k = Z \tag{1}$$

The low-level variables are considered as a sub-set of an infinite stochastic sequence, i.e. $(\ldots, X_{-1}, X_0, X_1, \ldots, X_k, X_{k+1}, \ldots)$; this particular sub-set characterizes the current **stage** of disaggregation. It is assumed that the disaggregation procedure has already been completed at the previous stages; thus all previous X_i's have known values: $(X_0 = x_0, X_{-1} = x_{-1}, \ldots)$.

Consider, as **initial parameters** of the model, at the current stage, the first and second moments of the low-level variables, forming the following groups:

(a) mean values of X_i, μ_i;
(b) variances of X_i, σ_i^2;
(c) covariances between X_i, X_j $(i, j > 0)$, σ_{ij}; and
(d) covariances between X_i $(i > 0)$ with variables X_j $(j \leqslant 0)$ of previous stages.

The number of independent parameters of groups (a), (b) and (c) is k, k, and $k(k - 1)/2$, respectively, and in total, $(k^2 + 3k)/2$. If k previous variables are considered as affecting the current stage, then the number of parameters of group (d) is k^2. Thus the total number of initial parameters is $3(k^2 + k)/2$.

Consider now the ith disaggregation step of the current stage, concerning the generation of the low-level variable X_i based on:

$$X_i + S_{i+1} = S_i \tag{2}$$

where the amount still to go:

$$S_i = X_i + X_{i+1} + \ldots + X_k = Z - X_1 - \ldots - X_{i-1} \tag{3}$$

has a known value, given that previous steps have been completed. Consider next, as **intermediate parameters of the *i*th step**, the first and second moments of the remaining low-level variables X_j, $i \leqslant j \leqslant k$, conditional on the previously generated information, i.e. $\Omega_{i-1} = (X_{i-1} = x_{i-1}, \ldots, X_1 = x_1, X_0 = x_0, X_{-1} = x_{-1}, \ldots)$. These parameters form groups similar to the groups (a), (b) and (c) of the initial parameters.

Finally, consider, as **final parameters of the *i*th step** at the current stage of the model, the first and second moments of the variables X_i, and S_i, conditional on the previously generated information. These final parameters are:

$$\mu_X = \mathrm{E}[X_i|\Omega_{i-1}],\ \sigma_X^2 = \mathrm{Var}[X_i|\Omega_{i-1}],\ \mu_S = \mathrm{E}[S_i|\Omega_{i-1}],$$

$$\sigma_S^2 = \mathrm{Var}[S_i|\Omega_{i-1}],\ \sigma_{XS} = \mathrm{Cov}[X_i, S_i|\Omega_{i-1}]$$

and they are fully determined by linear combinations of the intermediate parameters. These parameters are the link with the bisection procedure; at the ith step, their values (as well as the known value of S_i) are passed to the bisection procedure, which proceeds to the generation of the X_i value (as well as S_{i+1}).

Take now the case in which the sequence of low-level variables is "wide sense" Markov (first order autoregressive, Papoulis, 1965, p.420). The relevant analysis may be easily extended to higher order autoregressive sequences. The following relation is a consequence of the Markovian property:

$$\mathrm{Cov}[X_i, X_j]\,\mathrm{Cov}[X_j, X_l] = \mathrm{Cov}[X_i, X_l]\,\mathrm{Var}[X_j] \qquad i < j < l \tag{4}$$

This property reduces the number of the initial parameters. Thus the initial parameters of group (c) can be determined in terms of the parameters of group (b) and the $(k - 1)$ lag-one correlation coefficients:

$$\rho_i = \mathrm{Corr}[X_i, X_{i-1}] = \sigma_{i,i-1}/(\sigma_i\sigma_{i-1}) \tag{5}$$

with $i = 2, ..., k$. Similarly, the independent parameters of group (d), are reduced to one, since the covariances with the low-level variables of previous stages can be determined in terms of the parameters of groups (b) and (c) of the current and previous stages, and the lag-one correlation coefficient ρ_1, given by (5) for $i = 1$. Therefore the total number of parameters in this case is $3k$. Any covariance between low-level variables is given by:

$$\sigma_{ij} = \rho_j \ldots \rho_{i+1}\sigma_j\sigma_i \qquad i < j \leqslant k \tag{6}$$

which is a consequence of equations (4) and (5).

The intermediate parameters of the ith step are easily derived, considering that the wide sense Markov sequence X_i satisfies the difference equation (Papoulis, 1965, p.421):

$$X_i - a_iX_{i-1} = V_i \tag{7}$$

where V_i is a sequence of uncorrelated random variables and a_i is a sequence of constants. The resulting equations are:

$$\mathrm{E}[X_j|X_{i-1} = x_{i-1}] = \mu_j + \rho_j \ldots \rho_i\sigma_j\tilde{x}_{i-1} \tag{8}$$

$$\mathrm{Var}\ [X_j|X_{i-1} = x_{i-1}] = \sigma_j^2(1 - \rho_j^2 \ldots \rho_i^2) \tag{9}$$

$$\mathrm{Cov}[X_l,X_j|X_{i-1} = x_{i-1}] = \rho_l \ldots \rho_{j+1}(1 - \rho_j^2 \ldots \rho_i^2)\sigma_l\sigma_j \tag{10}$$

where $1 \leqslant i < j < l \leqslant k$; and $\tilde{x}_{i-1} = (x_{i-1} - \mu_{i-1})/\sigma_{i-1}$.

The final parameters of the ith step, calculated by using equations (3) and (8) to (10), are:

$$\mathrm{E}[X_i|X_{i-1} = x_{i-1}] = \mu_i + \rho_i\sigma_i\tilde{x}_{i-1} \tag{11}$$

$$\mathrm{Var}[X_i|X_{i-1} = x_{i-1}] = \sigma_i^2(1 - \rho_i^2) \tag{12}$$

$$\mathrm{E}[S_i|X_{i-1} = x_{i-1}] = \textstyle\sum_{j=i}^{k}\mu_j + \tilde{x}_{i-1}\sum_{j=i}^{k}\rho_j \dots \rho_i\sigma_j \tag{13}$$

$$\mathrm{Var}[S_i|X_{i-1} = x_{i-1}] = \textstyle\sum_{j=i}^{k}\sigma_j^2(1 - \rho_j^2 \dots \rho_i^2) + 2\sum_{j=i}^{k-1}\sum_{l=j+1}^{k}\rho_l \dots \rho_{j+1}(1 - \rho_j^2 \dots \rho_i^2)\sigma_l\sigma_j \tag{14}$$

$$\mathrm{Cov}[S_i, X_i|X_{i-1} = x_{i-1}] = \sigma_i(1 - \rho_i^2)\left[\sigma_i + \textstyle\sum_{j=i+1}^{k}\rho_j \dots \rho_{i+1}\sigma_j\right] \tag{15}$$

where $1 \leqslant i \leqslant k$. The above equations may also be used in the first step of the first stage, or any other similar case, where no condition is known, by setting $\rho_i = 0$. For application purposes, the following computational equations, equivalent to equations (13) to (15), are suggested, as reducing the required computational time, by avoiding repetitions.

$$\mathrm{E}[S_i|X_{i-1} = x_{i-1}] = E_i + \rho_i(\sigma_i + D_i)\tilde{x}_{i-1} \tag{16}$$

$$\mathrm{Var}[S_i|X_{i-1} = x_{i-1}] = A_i - B_i \tag{17}$$

$$\mathrm{Cov}[S_i, X_i|X_{i-1} = x_{i-1}] = (1 - \rho_i^2)\sigma_i(\sigma_i + D_i) \tag{18}$$

where:

$$E_i = \mu_i + E_{i+1} \tag{19}$$

$$D_i = \rho_{i+1}(\sigma_{i+1} + D_{i+1}) \tag{20}$$

$$A_i = \sigma_i^2 + 2\sigma_i D_i + A_{i+1} \tag{21}$$

$$B_i = \rho_i^2(\sigma_i^2 + 2\sigma_i D_i + B_{i+1}) \tag{22}$$

with $1 \leqslant i \leqslant k$, and $E_{k+1} = D_{k+1} = A_{k+1} = B_{k+1} = \rho_{k+1} = 0$. Equations (19) to (22) are applied successively for $i = k$ to $i = 1$, at the start of each stage, and the resulting values are stored. Then, at each step, equations (11), (12) and (16) to (18) are used to calculate the final parameters.

Bisection procedures

The bisection problem, i.e. the generation of variables X and Y, such that

$$X + Y = S \tag{23}$$

where S has a known value, s, can be studied independently of the other part of the model. What is required here is to determine the conditional distribution of $X|S$. The variable X can then be generated by this conditional distribution, and Y is obtained by equation (23). Due to the difficulties of the determination of conditional distributions, a simpler moments-based approach is preferable. This is done by assuming a proper auxiliary random variable W, and an explicit form $\mathrm{R}(S, W)$, such as:

$$X = \mathrm{R}(S, W) \tag{24}$$

with parameters being determined via the marginal and joint moments of (X, S).

The linear bisection scheme:

$$X = \mathrm{R}(S, W) = a\ S + W \tag{25}$$

where W is a random variable independent of S, is ideal for jointly normal variables. If W is assumed normal, then the bisection scheme preserves completely the distribution function of (X, Y, S). Moreover, when this scheme is combined with the other part of the model (conditional moments determination), the complete joint distribution function of low-level variables is preserved, if it is multidimensional normal. However, this bisection scheme is not proper for skewed distributions, since it cannot preserve non-zero skewness coefficients.

Another simple bisection scheme, the so-called **proportional** one, is defined by:

$$X = \mathrm{R}(S, W) = W \cdot S \tag{26}$$

where W is a random variable, generally dependent on S, referred to as proportional variable. The degree of correlation between W and S is much lower than that between X and S, and this simplifies the problem. The proportional scheme is ideal for gamma distributed variables, since it has been shown that when X and Y are independent gamma distributed, having equal scale parameters, and W is assumed independent of S and beta distributed, the complete distribution $\mathrm{F}_{XYS}()$ is preserved. This preservation expands to the whole sequence of low-level variables under the same assumptions. In the general case of dependent gamma marginal variables, with different scale parameters, the proportional scheme still gives satisfactory approximations of the gamma marginal distributions.

The parameters of the proportional scheme, i.e. the moments of W conditional on S, for the general gamma case are given by:

$$\mu_{W|S} = \theta - \eta \left[1 + \frac{\mu_S^2}{\sigma_S^2} \right] + \eta \frac{\mu_S^2}{\sigma_S^2} \frac{s}{\mu_S} \tag{27}$$

$$\sigma_{W|S}^2 = \frac{\sigma_X^2 + \theta^2 \sigma_S^2 - 2\theta \sigma_{XS}}{\sigma_S^2 + \mu_S^2} - \eta^2 \left[3 + \frac{\mu_S^2}{\sigma_S^2} \right] \tag{28}$$

where:

$$\theta = \mu_X / \mu_S \tag{29}$$

$$\eta = \frac{\sigma_{XS} - \theta \sigma_S^2}{\mu_S^2 + \sigma_S^2} \tag{30}$$

The above equations have been obtained under the assumption of a linear dependence between S and W. When X and Y are independent with common scale parameter, S and W should be assumed independent; thus $\eta = 0$ and equations (27) and (28) are apparently simplified. It is noted that equations (28) and (30) have different forms for non-gamma distributions (generalized equations are in Koutsoyiannis, 1988).

It must be emphasized that in the above analysis and the relevant equations, all variables are in their initial form (no differences from means). Hence, if the variables are positive, W should be bounded in [0,1]. The two-parameter beta distribution is a proper representation for the distribution of $W|S$. Finally, if X and S have three parameter gamma distributions, they can be replaced in the above analysis with the respective differences from their lower bounds, and the same bisection procedure used.

RAINFALL MODEL

The rainfall model used in the present study, in combination with the dynamic disaggregation model, was based on the historical data of two rain recorder stations in the Aliakmon river basin, northern Greece. It represents the rainfall process in discrete time, from an hourly to a monthly time scale, using as a base the intermediate scale of a rainfall event. The main parts of the rainfall model are summarized as follows.

Rainfall event - rainfall occurrence

A rainfall event is considered as an individual entity which can be identified in a historical rainfall record as in the studies of Grace & Eagleson (1966) and Restepo-Posada & Eagleson (1982). Successive rainfall events were defined as statistically independent, with starting points forming a Poisson process. Values of the **separation time** (c), i.e. the minimum dry time interval for two successive rainfall pulses to be considered as independent events, were obtained by a developed criterion, based on the Kolmogorov-Smirnov test, and were found to lie in the range c = 5–7 h.

The complete description of the rainfall occurrence process requires that the joint distribution function $F_{VDB}(v, d, b)$ is known, where V is the rainfall inter-arrival time, D is the duration of the event and B is the time between events (dry interval). This was based on:

(a) the obvious relation:

$$D + B = V \tag{31}$$

(b) the consequence of the event definition (a property of the Poisson process) that the marginal distribution of $V - c$ is exponential, that is:

$$f_V(v) = \omega\, e^{-\omega(v-c)} \qquad v \geqslant c \tag{32}$$

(c) the assumption that the conditional distribution of D, given V, comprises two additive parts, an exponential part independent of V, and a triangular part dependent on V, that is:

$$f_{D|V}(d,v) = \begin{cases} \delta\, e^{-\delta d} + 2d\, e^{-\delta(v-c)}/(v-c)^2 & 0 \leqslant d \leqslant v-c \\ 0 & \text{elsewhere} \end{cases} \tag{33}$$

The marginal densities of the distributions of event duration, D, and time between storms, B, derived theoretically from these assumptions are:

$$f_D(d) = (\delta + 2\omega)\, e^{-(\delta+\omega)d} - 2\omega(\delta+\omega)d\; \epsilon[(\delta+\omega)\, d] \qquad d \geqslant 0 \tag{34}$$

$$f_B(b) = \frac{\omega\delta}{\delta+\omega}\, e^{-\omega(b-c)} - 2\omega e^{-(\delta+\omega)(b-c)} + 2\omega[1 + (\delta+\omega)(b-c)]\; \epsilon[(\delta+\omega)(b-c)] \qquad b \geqslant c \tag{35}$$

where:

$$\epsilon(x) = \int_x^{\infty} (e^{-\xi}/\xi)\, .\, d\xi \tag{36}$$

As illustrated in Figs 2 and 3, $f_D(d)$ is quite similar to the exponential distribution, but $f_B(b)$ deviates, mainly in its lower tail, from the exponential, Weibull and gamma distributions.

The distribution of the event rain depth, H, conditional on D, has been

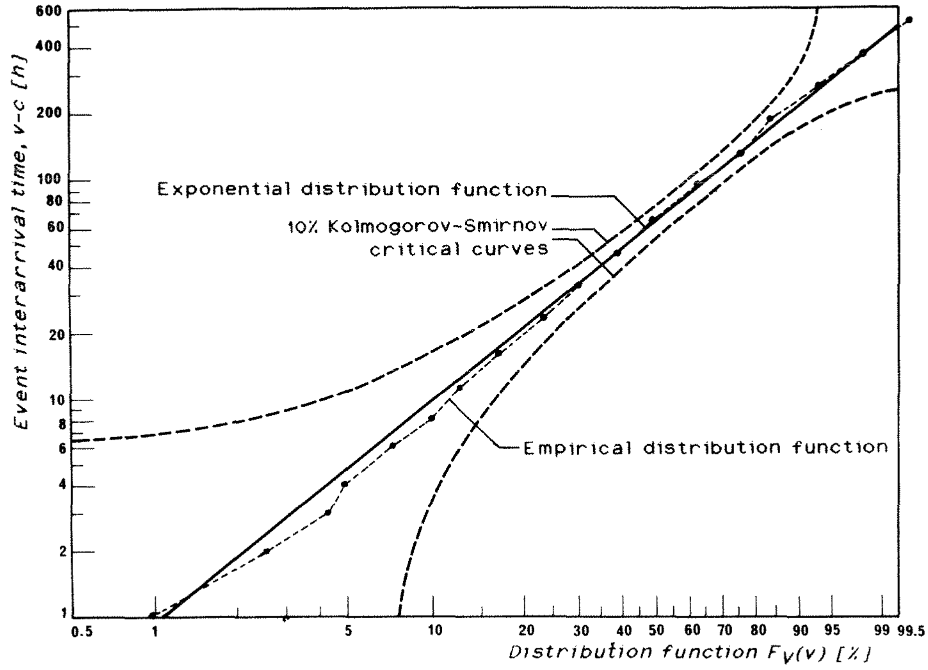

Fig. 1 Distribution function of rainfall event inter-arrival time, V.

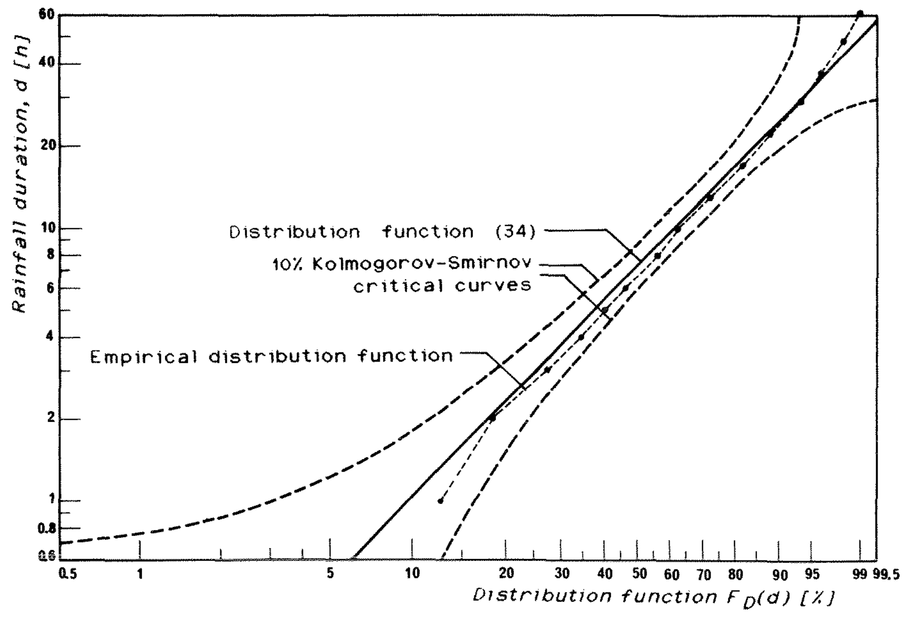

Fig. 2 Distribution function of rainfall event duration, D.

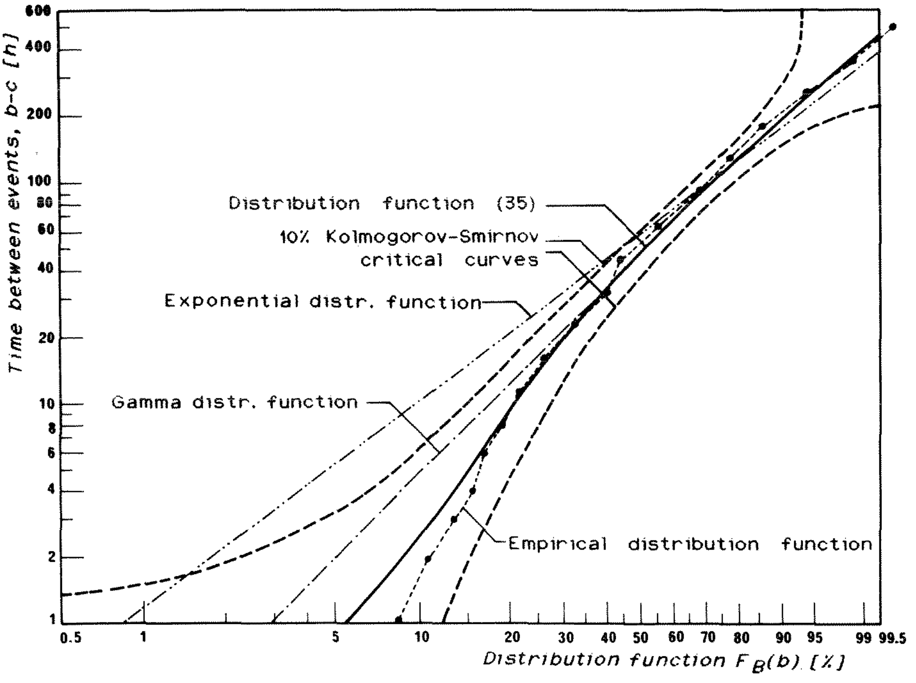

Fig. 3 Distribution function of time between events (dry interval), B.

assumed gamma, with mean and standard deviation linearly depending on the duration, i.e.:

$$E[H|D] = (d + a)\ \mu_\phi - b \tag{37}$$

$$\{\mathrm{Var}[H|D]\}^{1/2} = (d + a)\ \sigma_\phi \tag{38}$$

where a, b, μ_ϕ and σ_ϕ are constants. This is a modification of a scheme widely used in rainfall analysis and synthesis (e.g. Bras & Rodriguez-Iturbe, 1976; Eagleson, 1978), which assumes that conditional mean and standard deviation are proportional to the duration.

Internal rainfall event structure

Given a specific rainfall event, with duration D (considered as an integer multiple of Δ = 1 h), and total depth H, the sequence of hourly depths X_i, in the interior of the event, is related with H by:

$$X_1 + X_2 + \dots + X_{D/\Delta} = H \tag{39}$$

The following main assumptions concerning the structure of the hourly depths have been used:

(a) The sequence of hourly rain depths is non-stationary; the statistics of a specific X_i depend on the duration of the event, as well as on its time position in the event. It is accepted that these two influences are separable, and can be described by:

$$X_i = \mathrm{k}(D) \cdot \mathrm{g}(\theta_i) \cdot Z_i \tag{40}$$

where θ_i is the non-dimensionalized time position (t_i/D); Z_i is a sequence of dependent, identically distributed random variables, referred to as **homogenized hourly rain depths**, with mean μ_Z and standard deviation σ_Z; k(), and g() are properly defined functions.

(b) The covariance structure of Z_i, in a specific event is assumed to be stationary Markovian:

$$\mathrm{Cov}[Z_i, Z_{i+j}] = (\rho_1)^j \sigma_Z^2 \tag{41}$$

The lag-one correlation coefficient, ρ_1, generally depends on duration. The covariance between variables of different events is zero.

Secondary assumptions concern the form of $\mathrm{g}(\theta)$, which has been assumed linear, i.e.:

$$\mathrm{g}(\theta) = g_0 + g_1\theta \tag{42}$$

where g_0 and g_1 are constants, and the distribution of Z, which is J-shaped and has been considered as gamma or Weibull, depending on the fit to historical data.

The distribution of X_i, marginal or conditional on D, may be approximated by the same type as the one of Z. The conditional mean and standard deviation are:

$$\mathrm{E}[X_i | D{=}d] = \mathrm{k}(d)\ \mathrm{g}(\theta_i)\ \mu_Z; \quad \text{and} \tag{43}$$

$$\{\mathrm{Var}[X_i | D{=}d]\}^{1/2} = \mathrm{k}(d)\ \mathrm{g}(\theta_i)\ \sigma_Z \tag{44}$$

$\mathrm{k}(d)$ is completely determined by using equations (37), (39), (42) and (43):

$$\mathrm{k}(d) = 1 + (a - b/\mu_\phi)/d \tag{45}$$

The following equation estimating the correlation coefficient, $\rho_1(d)$, can be derived from equations (38) to (41) and (44):

$$\sum_{i=1}^{d/\Delta-1} \xi_i(d) \cdot [\rho_1(d)]^i = \xi_0(d) \tag{46}$$

where:

$$\xi_0(d) = \frac{(d + a)^2 \sigma_\phi^2}{2\mathrm{k}^2(d)\, \sigma_Z^2} - (1/2) \sum_{i=1}^{d/\Delta} \mathrm{g}^2(\theta_i) \tag{47}$$

$$\xi_i(d) = \Sigma_{j=1}^{d/\Delta-1} g(\theta_j)\, g(\theta_{j+1}) \qquad i = 1, 2, ..., d/\Delta - 1 \tag{48}$$

Thus, $\rho_1(d)$ may be computed numerically from equation (46); it is an increasing function of d (see Fig. 6), and this has been confirmed by the historical data.

Monthly rainfall

The three variables describing the monthly rainfall are the number of rainfall events, N, the monthly rainfall depth, S, and the monthly rain duration, U.

The marginal distribution of N is a modified Poisson, the modification caused by the lower bound (c) of the inter-arrival time; given the month duration, τ, the probability function is accurately approximated by:

$$P_n = \Pr(N = n|\tau) = (1 + \kappa)\, \frac{(\lambda - \kappa n)^n}{n!}\, e^{-(\lambda - \kappa n)} \qquad n = 0, 1, ..., m \tag{49}$$

where:

$$\lambda = \omega\tau = \tau/(\mu_V - c) \tag{50}$$

$$\kappa = \omega c = c/(\mu_V - c) \tag{51}$$

$$m = [\tau/c] = [\lambda/\kappa] \tag{52}$$

As a satisfactory approximation for simulation purposes, confirmed by the data, the gamma distribution was used for both S and U; their moments are completely determined by the corresponding moments of V, D and H.

Model parameters

All the rainfall model parameters can be expressed in terms of four main independent parameters, namely the separation time (c), and the mean values of rainfall inter-arrival time (μ_V), event duration (μ_D) and event depth (μ_H), and five secondary independent parameters, namely a, b, σ_ϕ, σ_Z and g_1. Three more secondary parameters may be introduced concerning the mean and standard deviation of events with duration equal to 1 h, (μ_{H1}, σ_{H1}) because it was found that equations (37) and (38) may not apply to these events, and the probability that hourly depth equals zero (p_0), a possibility permitted by the event definition. This probability may be represented by the value of the continuous distribution function of X or Z (gamma or Weibull) at the point x = 0.05 mm, since, in fact, values less than 0.05 mm are interpreted as zero (see Fig. 5). However, if this representation is not satisfactory, then p_0 should be used as an independent parameter. Conclusively, the maximum number of independent parameters is 12, and this

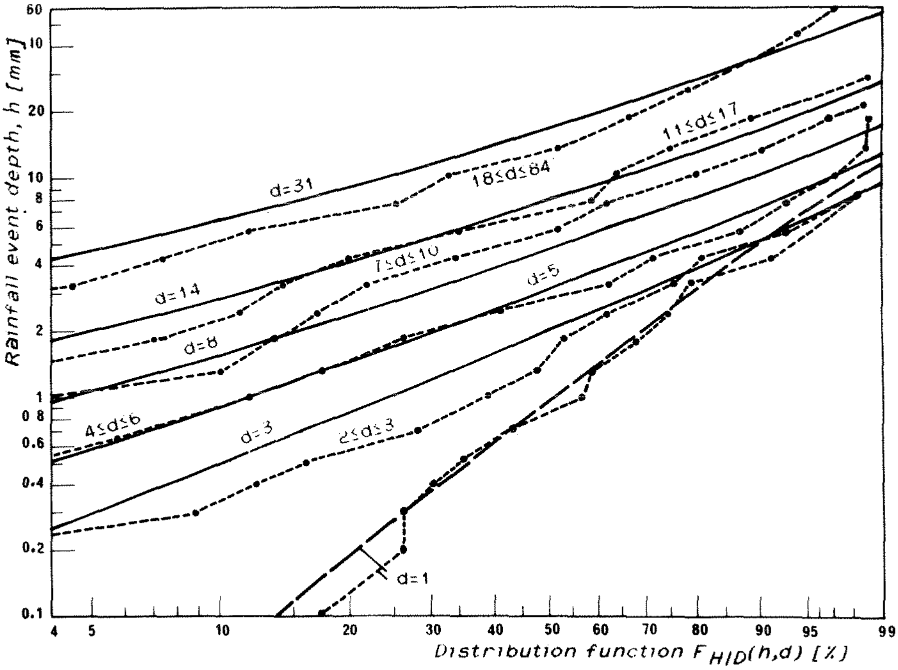

Fig. 4 *Distribution function of rainfall event depth, H, conditional on duration, D. Continuous lines represent the gamma distribution function, and dashed lines the empirical distribution functions of the synthetic sample.*

number may be reduced to 4, by omitting secondary parameters. All parameters are season-dependent.

RAINFALL DISAGGREGATION

Consider now the problem of disaggregation of monthly rainfall into hourly depths. Because of the intermittent nature of the rainfall process, a two-phase disaggregation procedure has been adopted. The first phase, **external disaggregation,** is to generate the rainfall events, while the second, **internal disaggregation,** generates hourly depths within each event.

The disaggregation in both phases is a combination of the dynamic disaggregation model and the rainfall model. The latter calculates the initial parameters for the former which performs the generation. The Markovian configuration of the model studied, with the proportional bisection scheme, is satisfactory for both phases; for cases in which low-level variables are independent, the correlation coefficients are set to zero. Particular additional procedures, causing proper side effects on the generated variables, have been designed and used along with the disaggregation model. The step-by-step course of the model permits the use of such side procedures.

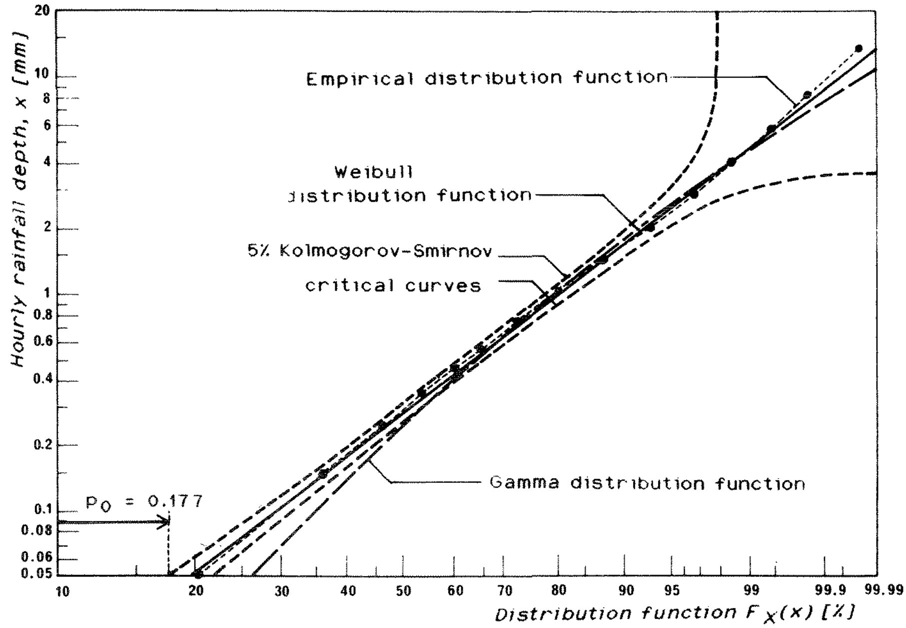

Fig. 5 Distribution function of hourly rainfall depth, X.

It is supposed that, at the start of use of the model, monthly rainfall variables, i.e. the number of rainfall events, N, the monthly depth, S, and the monthly rain duration, U, have known values. Nevertheless, the implementation was designed to include a separate part, generating, if needed, one or more of these values, in order to be a complete rainfall generator, from monthly through hourly time scales.

Below is a coded description of each disaggregation phase.

1(a) External disaggregation - section a

Input: Number of events, $N = n$

Output: Inter-arrival times V_i (= low-level variables) determining the starting points of events.

Basic relation:

$$\Sigma_{i=1}^{n} (V_i - c) = T^* \tag{53}$$

where $T^* = \tau - nc - A + B$, is the high-level variable, A is the time distance of the starting point of the first event of the current month, and B is the same quantity for the next month.

Remarks: The distribution of $(V_i - c)$ is exponential, a particular case of the gamma. Successive low-level variables are independent.

Side procedures: A and B are generated separately, using the same exponential distribution of V_i.

1(b) External disaggregation - section b

Input: Number of events, $N = n$, monthly duration, $U = u$ (= high-level

variable), event inter-arrival times $V_i = v_i$.
Output: Event durations D_i (= low-level variables).
Basic relation:

$$\sum_{i=1}^{n} D_i = U \tag{54}$$

Remarks: The disaggregation model uses only the exponential part of the conditional distribution of equation (33), while the triangular part is left to a side procedure. Successive low-level variables are independent.
Side procedures: If the value, d_i, generated by the disaggregation model is greater than $v_i - c$, it is rejected and a new one is generated in the range $(0, v_i - c)$, using the triangular distribution.

1(c) **External disaggregation - section c**
Input: Number of events, $N = n$, monthly rain depth, $S = s$ (= high-level variable), event durations, $D_i = d_i$.
Output: Event rain depths H_i (= low-level variables).
Basic relation:

$$\sum_{i=1}^{n} H_i = S \tag{55}$$

Remarks: Distribution of H_i, conditional on D_i, is gamma, with moments depending on d_i. Successive low-level variables are independent.
Side procedures: None.

2 **Internal disaggregation**
Input: Event rain depth, $H_i = h_i$ (= high-level variable), event duration, $D_i = d_i$.
Output: Hourly rain depths X_{ij} (= low-level variables).

$$\sum_{j=1}^{d_i/\Delta} X_{ij} = H_i \tag{56}$$

Remarks: Distribution of X_{ij}, conditional on D_i, is gamma or Weibull, with moments depending on d_i and j. Both distribution types are treated with the proportional bisection scheme, and the adjustment of the Weibull distribution is left to a side procedure. The covariance structure of low-level variables is Markovian. The correlation coefficients depend on d_i.
Side procedures: An empirical procedure, based on the generation of uniform random numbers, handles the probability of zero depth, p_0; also it adjusts the short interval tail of $F_X(x_{ij})$, when it is Weibull. Moreover, the side procedure handles the number of successive zero rain depths, disallowing exceedance of the value (c/Δ).

The rainfall generating model has been coded in the Pascal programming language. Conventional microcomputers were used to develop and run the model.

RESULTS AND TESTING OF THE MODEL

To illustrate the model and test its reliability, a complete application with observed rain data will be presented. The illustration is for testing the disaggregation model as a rainfall generator, and not of the rainfall model (for testing of the latter see Koutsoyiannis, 1988). Thus, all tests are formulated as comparisons of the generator results with the quantities theoretically anticipated by the rainfall model.

The 12 parameters of the rainfall model, shown in Table 1, were derived from a 13 year historical rainfall record. Adopting these parameters, 50 years of monthly, synthetic events and hourly rainfall depths were generated. Characteristic sizes of this sample are shown in Table 2. Various statistics and empirical distribution functions, obtained from the sample, were then compared to the theoretically expected ones. In particular, the test results are the following:

(a) Preservation of mean values and variances (standard deviations) Table 3 shows good agreement between theoretical and empirical values.

Table 1 *Parameters used for the application of the model (values computed from a historical sample: rain gauge station: Chalara, Aliakmon River basin; month: April)*

Parameter		*Value*	
Main parameters:			
Separation time	c	*7*	*h*
Mean inter-arrival time	μ_V	*95.51*	*h*
Mean event duration	μ_D	*10.84*	*h*
Mean event depth	μ_H	*7.02*	*mm*
Secondary parameters:			
Parameters of the expression	a	*3.00*	*h*
of H, conditional on D	b	*1.02*	*mm*
	σ_ϕ	*0.326*	$mm\ h^{-1}$
Standard deviation of Z	σ_Z	*0.872*	*mm*
Parameter of the linear expression of non-stationarity	g_1	*-0.417*	
Additional parameters:			
Moments of 1h rainfall events	μ_{H1}	*1.830*	*mm*
	σ_{H1}	*2.330*	*mm*
Probability that hourly depth equals zero	P_0	*0.177*	

Table 2 *Characteristic sizes of the synthetic sample obtained as an output of the model*

Years of simulation	*50*
Number of rainfall events	*371*
Hourly rain depths	
(a) total	*4079*
(b) with duration > 1 h	*4033*

Statistical tests (where possible) at a 5% significance level showed that the equality hypotheses could not be rejected. Even the seemingly large difference 17.5% in σ_U (Table 3) is not significant at the 5% level; note that it refers to the non-disaggregation part of the model, and the sample size is relatively small (n = 50).

Table 3 *Comparison of empirical values of various statistics with their corresponding theoretical values*

Variable		Mean: Statistic	Theoretical value	Sample value	Deviation	Standard deviation: Statistic	Theoretical value	Sample value	Deviation
Monthly rainfall									
Number of events	N	μ_N	7.53	7.42	-1.5%	σ_N	2.54	2.54	0.0%
Rain depth	S	μ_S	52.92	53.31	0.7%	σ_S	29.57	30.70	3.8%
Rain duration	U	μ_U	81.71	81.58	-0.2%	σ_U	41.03	48.21	17.5%
Event rainfall									
Inter-arrival time	V	μ_V	95.51	100.64	5.4%	σ_V	88.51	91.57	3.5%
Rain duration	D	μ_D	10.84	10.99	1.4%	σ_D	11.07	11.98	8.2%
Time between events	B	μ_B	83.67	89.65	7.1%	σ_B	87.39	90.56	3.6%
Rain depth	H	μ_H	7.02	7.18	2.3%	σ_H	8.58	8.79	2.4%
Hourly rainfall (events with d > 1 h)									
Rain depth	X	μ_X	0.635	0.641	0.9%	σ_X	0.975	0.970	-0.5%
Homogenized rain depth	Z	μ_Z	0.575	0.582	1.2%	σ_Z	0.872	0.881	1.0%

(b) Preservation of distribution functions Figs 1, 2, and 3 demonstrate the preservation of the marginal distributions of the time intervals V, D, and B, respectively. The curves corresponding to the limits of the Kolmogorov-Smirnov test at a 10% significance level are drawn; all points of the empirical distributions lie inside the relevant bands. Similarly the empirical distributions of hourly depth, X, lie inside the 5% Kolmogorov-Smirnov bands in Fig. 5. Note the shifting (by a side procedure) of the empirical distribution function towards the Weibull. Figure 4 illustrates the preservation of the distribution of the event depth, H, conditional on duration, D. Empirical distributions were computed from separate classes of the synthetic sample, defined by certain bounds of the duration; the plotted theoretical distributions refer to the mean duration of each class.

(c) Preservation of correlation coefficients Significant correlation coefficients appear only between hourly depths in the interior of an event. Exact statistical tests are not possible due to the variation of correlation coefficients with the event duration. Nevertheless, Fig. 6 is an indication of the preservation of the lag-one correlation coefficient of the homogenized hourly depth. The same sample classes mentioned above were used to compute empirical values. Figure 7 is similar, depicting the preservation of higher lag coefficients (here only two samples classes were used).

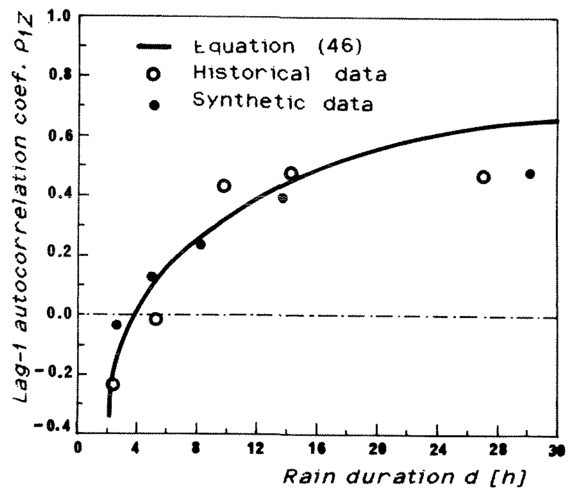

Fig. 6 Functional dependence of lag-one autocorrelation coefficient of the homogenized hourly rainfall depth, ρ_{1Z}, on rain duration, d. Empirical values were calculated from separate classes of the sample.

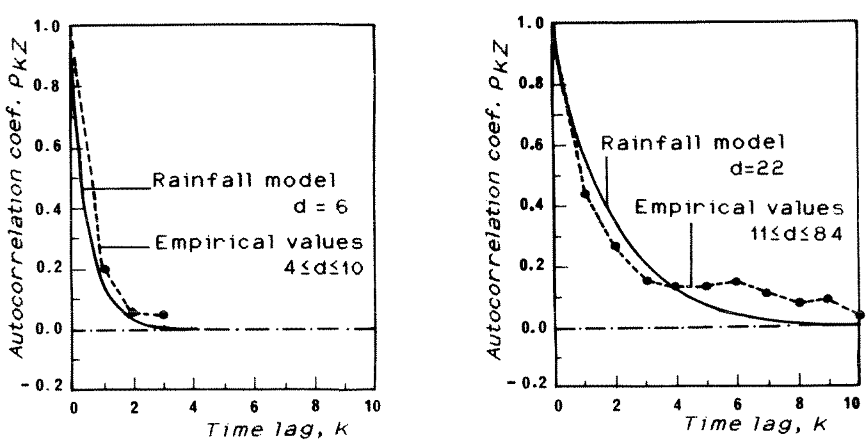

Fig. 7 Autocorrelation function of the homogenized hourly rainfall. Empirical values were calculated from two separate classes of the synthetic sample.

(d) Approximation of probabilities of zero hourly depths Without the use of the side procedure affecting the number of zero rainfall depths (in theory values less than 0.05 mm) lying inside rainfall events, their frequency would be equal to $p_0' = 0.267$, as resulted from the gamma distribution. The final frequency, obtained with the side procedure, was $p_0^* = 0.201$, which approximates the desired frequency $p_0 = 0.177$. While a statistical test gave the result that p_0 and p_0^* differed significantly, the approximation achieved is considerable.

CONCLUSIONS

The single-site dynamic disaggregation model developed is a generalized step-by-step approach to stochastic disaggregation problems. Forms studied concern low-level variables with Markovian structure and normal or gamma marginal distributions. Important features of the model are: (a) the modular structure (composed of two parts studied separately) allowing various configurations of the model; (b) the flexible step-by-step approach allowing the use of side procedures, adjusting properly the generated values in each step without loss of the additive property; and (c) the simple analytical equations allowing a varying number of low-level variables and varying time scales.

A combination of the dynamic disaggregation model with a developed rainfall model gave a point rainfall generator, performing with monthly through hourly time scales. The rainfall model can incorporate a varying number of parameters (4 to 12), depending on the desired accuracy.

An advantage of the combined model is that the same disaggregation procedure is used for four different purposes: the determination of the starting points of rainfall events, the generation of rain durations and event rain depths, and the disaggregation of the latter into hourly depths.

The application of the combined model using all 12 parameters derived from historical records gave satisfactory results. Various statistics, as well as marginal distribution functions of the associated variables computed from synthetic samples derived by the model, exhibited agreement with the theoretical expectations; the relevant statistical tests at 5% or 10% significance levels were positive.

The rainfall generator models the total rainfall regardless of intensity. The internal disaggregation part of the model (phase 2) may be applied independently to severe storms in order to simulate their time profiles. Thus the model may be useful for the simulation of severe flood-producing storms and the estimation of design floods.

Finally, the dynamic disaggregation model may be combined with other rainfall models. Moreover it can perform with other time scales, larger than monthly (e.g. from annual into seasonal or monthly depths) or shorter than hourly.

Acknowledgements The authors wish to thank the reviewers and Professor S. Burgess for their constructive comments.

REFERENCES

Bras, R. L. & Rodriguez-Iturbe, I. (1976) Rainfall generation: A nonstationary time-varying multidimensional model. *Wat. Resour. Res.* **12** (3), 450-456.

Eagleson, P. S. (1978) Climate, soil, and vegetation: 2, The distribution of annual precipitation derived from observed storm sequences. *Wat. Resour. Res.* **14** (5), 713-721.

Grace, R. A. & Eagleson, P. S. (1966) The synthesis of short-time increment rainfall sequences. Report no. 91, Hydrodynamics Laboratory, Massachusetts Institute of Technology, Cambridge, Massachussetts, USA.

Hoshi, K. & Burges, S. J. (1979) Disaggregation of streamflow volumes. *J. Hydraul. Div. ASCE*

105 (HY1), 27-41.
Koutsoyiannis, D. (1988) A point rainfall disaggregation model. PhD thesis, National Technical University of Athens, Greece (in Greek).
Marien, J. L. & Vandewiele, G. L. (1986) A point rainfall generator with internal storm structure. *Wat. Resour. Res.* **22** (4), 475-482.
Mejia, J. M. & Rousselle, J. (1976) Disaggregation models in hydrology revisited. *Wat. Resour. Res.* **12** (2), 185-186.
Papoulis, A. (1965) *Probability, random variables and stochastic processes.* McGraw-Hill.
Restepo-Posada, P. J. & Eagleson, P. S. (1982) Identification of independent rainstorms. *J. Hydrol.* **65**, 303-319.
Schaake, J. C., Ganslaw, M. J., Fothergill, J. W. & Harbaugh, T. E. (1972) Multivariate rainfall generator for annual, seasonal, monthly and daily events. *Proc. Symp. on Mathematical Modeling Techniques in Water Resources Systems,* vol. 2, 437-460, Ottawa, Canada.
Stedinger, J. R. & Vogel, R. M. (1984) Disaggregation procedures for generating serially correlated flow vectors. *Wat. Resour. Res.* **20** (1) 47-56.
Tao, P. C. & Delleur, J. W. (1976) Multistation, multiyear synthesis of hydrologic time series by disaggregation. *Wat. Resour. Res.* **12** (6), 1303-1312.
Todini, E. (1980) The preservation of skewness in linear disaggregation schemes. *J. Hydrol.* **47**, 199-214.
Valencia, D. & Schaake, J. C. (1972) A disaggregation model for time series analysis and synthesis. Report no. 149, Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics, Massachusetts Institute of Technology, Cambridge, Massachussetts, USA.
Valencia, D. & Schaake, J. C. (1973) Disaggregation processes in stochastic hydrology. *Wat. Resour. Res.* **9** (3), 211-219.
Woolhiser, D. A. & Osborn, H. B. (1985) A stochastic model of dimensionless thunderstorm rainfall. *Wat. Resour. Res.* **21** (4) 511-522.

Received 2 February 1989; accepted 26 September 1989