A Generalized Spatial-Temporal Model of Rainfall Based on a Clustered Point Process

Author(s): Paul S. P. Cowpertwait

Source: *Proceedings: Mathematical and Physical Sciences*, Jul. 8, 1995, Vol. 450, No. 1938 (Jul. 8, 1995), pp. 163-175

Published by: Royal Society

Stable URL: https://www.jstor.org/stable/52664

A generalized spatial-temporal model of rainfall based on a clustered point process

By Paul S. P. Cowpertwait

Departments of Engineering Mathematics and Civil Engineering, University of Newcastle, Newcastle upon Tyne NE1 7RU, UK

The objective in this paper is to present and fit a relatively simple stochastic spatial-temporal model of rainfall in which the arrival times of rain cells occur in a clustered point process. In the x–y plane, rain cells are represented as discs; each disc having a random radius; the locations of the disc centres being given by a two-dimensional Poisson process. The intensity of each cell is a random variable that remains constant over the area of the disc and throughout the lifetime of the cell, the lifetime being an exponential random variable. The cells are randomly classified from 1 to n with different parameters for the different cell types, so that the random variables of an arbitrary cell, e.g. radius and intensity, are correlated. Multi-site second-order properties are derived and used to fit the model to hourly rainfall data taken from six sites in the Thames basin, UK.

1. Introduction

Stochastic spatial-temporal models of rainfall have been formulated based on physical processes observed in precipitation fields (Waymire *et al.* 1984; Mellor 1992). These 'field' models, which usually incorporate bands of rain, regions of high intensity rain, and rain cells, can be used to simulate fine resolution data over a large geographical region, and are potentially useful in real-time forecasting. However, fitting a field model to time-series data is not straightforward, because of the complexity of the model and underlying physical processes. A simple spatial-temporal model, that can be fitted using second-order properties taken from time-series records, is therefore likely to be of more practical value in hydrological studies that require long records of multi-site rainfall to be generated.

A relatively simple model, in which storm centres arrive in a Poisson process in two-dimensional space and time, was developed by Cox & Isham (1988). However, empirical analysis of data has shown that rainfall events tend to arrive in 'clusters'. Consequently, a model based on Poisson arrival times fails to fit hourly data over a range of aggregation levels (Rodriguez-Iturbe *et al.* 1987). For a single-site time series, this inadequacy has been overcome by using a clustered point process for the arrival times of rain cells, such as the Bartlett–Lewis rectangular pulses model or the Neyman–Scott rectangular pulses model (Rodriguez-Iturbe *et al.* 1987). The objective in this paper is to formulate a relatively simple spatial-temporal model that has rain cells arriving in a clustered point process and has parameters that can be related, at least in some broad sense, to physical features observed and measured in precipitation fields.

Proc. R. Soc. Lond. A (1995) **450**, 163–175

Printed in Great Britain

TEX Paper

A generalization of the Neyman–Scott model, in which each rain cell is randomly classed as 1 of n types, has been formulated by Cowpertwait (1994). In the generalized model, the parameters of a rain cell depend on the cell type, so that the intensity and duration of an arbitrary rain cell are correlated. In this paper, the temporal process of cell arrivals will follow that of the generalized model, so that more than one type of cell can be present in the same storm. This will allow areas of heavy convective rain to occur within areas of lighter rain, which agrees with observational studies on frontal weather systems (Austin & Houze 1972; Browning 1985). For mathematical convenience, some simplifying assumptions, similar to those made by Rodriguez-Iturbe *et al.* (1986), will be made about the behaviour of rain cells in two-dimensional space. For example, the cells are represented as discs that have a random radius; the location of the disc centres being given by a spatial Poisson process. Furthermore, the cells are taken to have zero velocity. While some of these assumptions may be unsatisfactory from the meteorological standpoint, they allow the formulation of a model that may be of practical value in simulation studies. Moreover, the detailed physical properties of a rain cell are virtually undetectable in aggregated time-series data. Consequently, the model is likely to be particularly useful in studies that use aggregated data, e.g. hydrological studies that use hourly data.

2. Generalized spatial-temporal model

(a) *Model formulation*

Consider a region (or catchment) of finite area A in the x–y plane. Let the arrival times of storm origins occur in a Poisson process of rate λ, the arrival times being the same at any point in the x–y plane. The arrival of a storm origin at a catchment implies that the physical conditions necessary for rainfall have been met, e.g. a frontal system has arrived at the catchment. However, rain only occurs at points covered by rain 'cells', which are defined as follows.

Each storm origin generates discs (rain cells) over the whole x–y plane, so that a random number of disc centres are generated over any catchment with finite area A in the x–y plane. The spatial positions of the disc centres are given by a Poisson process of rate ρ, so that the number of disc centres in a catchment of area A is a Poisson random variable with mean ρA. Each cell is randomly classed as 1 of n types, with probability α_i that the cell is of type i ($i = 1, \ldots, n$). The spatial process of disc centres is thus equivalent to the superposition of n independent Poisson processes with rates $\rho_i = \alpha_i \rho$. The radius R_i of a type i cell is an independent exponential random variable with parameter γ_i. The intensity X_i of a type i cell is an independent random variable and remains constant over the area of the disc and throughout the lifetime of the cell; the lifetime being an independent exponential random variable with parameter η_i. If t is the arrival time of the storm origin and s ($> t$) is the starting time for a rain cell generated by the origin, then $s - t$ is taken to be an independent exponential random variable with parameter β, so that arrival times of rain cells follow a Neyman–Scott process. The total rainfall intensity at an arbitrary time t and point m is the summation of the intensities of all cells 'alive' at time t that overlap point m. In three-dimensional space, the intensity X_i at a point m with coordinates (x, y, z) is scaled by a factor φ_m (which can be a function of the altitude z of the point, to make an allowance for the effects of orography). It is convenient to define the dimensionless parameter $\nu_i = 2\pi\rho_i/\gamma_i^2$.

Let m be a point in the x–y plane, which, without loss of generality, can be the origin O, and let $N^{(i)}(u, u+\delta u; t, t+\delta t)$ be the number of type i disc centres occurring in the region $\{(r,\theta) : u \leqslant r \leqslant u+\delta u; 0 \leqslant \theta < 2\pi\}$ with starting times in the interval $(t, t+\delta t)$. Then, in the limit as δt and $\delta u \to 0$, $\mathrm{d}N^{(i)}(u;t)$ is unity if and only if there is a type i cell with disc centre a distance u from m and starting time t; otherwise $\mathrm{d}N^{(i)}(u;t)$ is zero. If $Y_i^{(m)}(t)$ is the total rainfall intensity at time t and point m, due to type i cells only, and $X_{m,t}^{(i)}(u,s)$ is the rainfall intensity at time t and point m due to a type i cell with centre a distance u from m and starting time $t-s$, then

$$X_{m,t}^{(i)}(u,s) = \begin{cases} \varphi_m X_i & \text{with probability } \mathrm{e}^{-\gamma_i u}\mathrm{e}^{-\eta_i s}, \\ 0 & \text{otherwise,} \end{cases} \tag{2.1}$$

and

$$Y_i^{(m)}(t) = \int_{u=0}^{\infty}\int_{s=0}^{\infty} X_{m,t}^{(i)}(u,s)\,\mathrm{d}N^{(i)}(u;t-s)\,, \quad \text{for } i = 1,\ldots,n. \tag{2.2}$$

The total intensity at time t and point m, $Y^{(m)}(t)$, is the sum of the intensities of all cells alive at time t and overlapping m, i.e.

$$Y^{(m)}(t) = \sum_{i=1}^{n} Y_i^{(m)}(t). \tag{2.3}$$

If $S_k^{(m)}(h)$ is the aggregated rainfall depth in the kth time interval of duration h, then

$$S_k^{(m)}(h) = \int_{(k-1)h}^{kh} Y^{(m)}(t)\,\mathrm{d}t \tag{2.4}$$

and $\{S_k^{(m)}(h) : k = 1, 2, \ldots\}$ is a rainfall time series at aggregation level h.

Second-order properties of the time series are needed to fit the model to data. Under a suitable transformation, weak stationarity in space and time will be assumed. In particular, the expectation $E\{S_k^{(m)}(h)/\varphi_m\} = hE\{Y^{(m)}(t)\}/\varphi_m$ is the same for all m and k, where, from (2.2) and (2.3),

$$\begin{aligned} E\{Y_i^{(m)}(t)\} &= 2\pi\rho_i\lambda\int_{u=0}^{\infty}\int_{s=0}^{\infty} E\{X_{m,t}^{(i)}(u,s)\}u\,\mathrm{d}u\,\mathrm{d}s \\ &= 2\pi\lambda\varphi_m\rho_i E(X_i)/(\gamma_i^2\eta_i). \end{aligned} \tag{2.5}$$

(*b*) *Properties of a single storm*

In this section, the temporal process of cell arrival times is ignored and properties given for the total rainfall falling during a single storm at some point in a catchment. The spatial process is similar to the spatial models of total storm rainfall studied by Rodriguez-Iturbe *et al.* (1986).

Consider an arbitrary point m in the x–y plane. Without loss of generality, let m be the origin O. In polar coordinates (r,θ), consider the circular region $\{(r,\theta) : u \leqslant r \leqslant u+\delta u; 0 \leqslant \theta < 2\pi\}$. Let $N^{(i)}(u, u+\delta u)$ be the number of type i disc centres occurring in the region, which has a Poisson distribution with mean

$2\pi\rho_i u\delta u$, and let $X_m^{(i)}(u)$ and $D_m^{(i)}(u)$ be the intensity and duration (respectively) at m of a type i cell with cell centre a distance u from m. Then, the total rainfall depth Z_m at m, is given by

$$Z_m = \sum_{i=1}^{n} \int_{u=0}^{\infty} X_m^{(i)}(u) D_m^{(i)}(u)\, \mathrm{d}N^{(i)}(u). \tag{2.6}$$

The expectation and variance follow directly as:

$$E(Z_m) = \sum_{i=1}^{n} \int_{u=0}^{\infty} 2\pi\rho_i \eta_i^{-1} \varphi_m E(X_i) u \mathrm{e}^{-\gamma_i u}\, \mathrm{d}u = \varphi_m \sum_{i=1}^{n} \nu_i E(X_i)/\eta_i \tag{2.7}$$

and

$$\operatorname{Var}(Z_m) = 2\varphi_m^2 \sum_{i=1}^{n} \nu_i E(X_i^2)/\eta_i^2. \tag{2.8}$$

Suppose, the intensity X_i of a type i cell is an independent exponential random variable with parameter ξ_i and the intensity and radius of an arbitrary cell, selected at random, are denoted as X and R respectively, then the covariance is given by

$$\operatorname{Cov}(X, R) = -\left\{ \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\alpha_i \alpha_j}{\xi_i \gamma_j} - \sum_{i=1}^{n} \frac{\alpha_i}{\xi_i \gamma_i} \right\}. \tag{2.9}$$

Similar expressions can be given for the covariance of the intensity and duration and the covariance of the duration and radius of an arbitrary cell.

Now let p be the probability that a type i cell with disc centre in the region $\{(r,\theta) : 0 \leqslant r \leqslant u;\ 0 \leqslant \theta < 2\pi\}$ overlaps m, so

$$p = 2(1 - u\gamma_i \mathrm{e}^{-u\gamma_i} - \mathrm{e}^{-u\gamma_i})/(u^2\gamma_i^2),$$

after integration. Then, the probability that j type i cells, out of all cells generated over the whole x–y plane, overlap m is

$$\lim_{u\to\infty} \sum_{k=j}^{\infty} \binom{k}{j} p^j (1-p)^{k-j} (\pi u^2 \rho_i)^k \frac{\mathrm{e}^{-\pi u^2 \rho_i}}{k!} = \nu_i^j \frac{\mathrm{e}^{-\nu_i}}{j!} \tag{2.10}$$

so that the number of type i rain cells overlapping an arbitrary point m in the x–y plane, $C_i^{(m)}$, is a Poisson random variable with mean ν_i, and the expected dry area for a catchment of area A is $A\exp(-\sum_{i=1}^{n} \nu_i)$.

Consequently, by taking the number of cells associated with a storm origin as a Poisson random variable with mean $\mu_C = \sum_{i=1}^{n} \nu_i$, the second-order properties of $S_k^{(m)}(h)$ can be found directly from Cowpertwait (1994, eqns (2.12)–(2.14)). These properties, which are used to fit the model in § 4, are given below together with a modified expression for the probability of a dry interval.

(*c*) *Single-site properties*

The expectation of the aggregated process (2.4) is given by

$$E\{S_k^{(m)}(h)\} = h\lambda\varphi_m \sum_{i=1}^{n} \frac{\nu_i E(X_i)}{\eta_i}. \tag{2.11}$$

In this section and § 2 *d* it will be useful to define, for all $i, j = 1, \ldots, n$:

$$C_{ij}(\tau) = \tfrac{1}{2}\lambda\beta\nu_i\nu_j E(X_i)E(X_j) \times \left[\frac{2\beta \mathrm{e}^{-\eta_j\tau}}{(\beta^2-\eta_j^2)(\eta_i+\eta_j)} - \frac{\mathrm{e}^{-\beta\tau}}{(\beta-\eta_j)(\beta+\eta_i)}\right] = A_{ij}\mathrm{e}^{-\eta_j\tau} + B_{ij}\mathrm{e}^{-\beta\tau}, \text{ say.} \tag{2.12}$$

Then, from Cowpertwait (1994, eqn (3.9)),

$$\mathrm{Cov}\,\{Y_i^{(m)}(t), Y_j^{(m)}(t+\tau)\} = \begin{cases} \varphi_m^2 C_{ij}(\tau), & i \neq j, \\ \lambda\nu_i\varphi_m^2 E(X_i^2)\mathrm{e}^{-\eta_i\tau}/\eta_i + \varphi_m^2 C_{ii}(\tau), & i = j. \end{cases} \tag{2.13}$$

For $l \geqslant 0$, after routine integration of terms in (2.13), the variance ($l = 0$) and autocovariance ($l > 0$) are given by

$$\mathrm{Cov}\,\{S_k^{(m)}(h), S_{k+l}^{(m)}(h)\} = \varphi_m^2 \sum_{i=1}^{n}\sum_{j=1}^{n}\{A_j(h,l)A_{ij} + B(h,l)B_{ij}\} + \varphi_m^2\sum_{i=1}^{n}\{\lambda\nu_i E(X_i^2)A_i(h,l)/\eta_i\}, \tag{2.14}$$

where

$$A_j(h,l) = \begin{cases} 2(h\eta_j + \mathrm{e}^{-\eta_j h} - 1)/\eta_j^2, & l = 0, \\ (1-\mathrm{e}^{-\eta_j h})^2\mathrm{e}^{-\eta_j h(l-1)}/\eta_j^2, & l = 1, 2, \ldots \end{cases} \tag{2.15}$$

and

$$B(h,l) = \begin{cases} 2(h\beta + \mathrm{e}^{-\beta h} - 1)/\beta^2, & l = 0, \\ (1-\mathrm{e}^{-\beta h})^2\mathrm{e}^{-\beta h(l-1)}/\beta^2, & l = 1, 2, \ldots \end{cases} \tag{2.16}$$

To obtain an expression for $\mathrm{pr}\{S_k^{(m)}(h) = 0\}$, let ω_i be the probability that a type i cell overlapping point m with arrival time in $(0, t)$ terminates before t, and $p_h(t)$ be the probability of no rain in $(t, t+h)$ due to a storm origin at time zero. Then,

$$\omega_i = 1 - \beta(\mathrm{e}^{-\beta t} - \mathrm{e}^{-\eta_i t})/\{(\eta_i - \beta)(1-\mathrm{e}^{-\beta t})\}. \tag{2.17}$$

If ω is the probability that an arbitrary cell with arrival time in $(0,t)$ terminates before time t, then $\omega = \sum_{i=1}^{n}\alpha_i\omega_i$. In previous work, the total number C of cells associated with a storm origin has been a geometric random variable. However, in § 2 *b* C is a Poisson random variable with mean $\mu_C = \sum_{i=1}^{n}\nu_i$. Therefore,

$$p_h(t) = \sum_{j=0}^{\infty}\sum_{k=0}^{j}\binom{j}{k}(1-\mathrm{e}^{-\beta t})^k\omega^k\mathrm{e}^{-\beta(t+h)(j-k)}{\mu_C}^j\frac{\mathrm{e}^{-\mu_C}}{j!} = \exp\{-\mu_C + \mu_C\mathrm{e}^{-\beta(t+h)} + \omega\mu_C(1-\mathrm{e}^{-\beta t})\}. \tag{2.18}$$

The probability that an arbitrary time interval of duration h (hours) is dry is then given by (Cowpertwait 1994, eqn (2.16))

$$\mathrm{pr}\,\{S_k^{(m)}(h) = 0\} = \exp\left[-\lambda\int_0^{\infty}\{1-p_h(t)\}\,\mathrm{d}t - \lambda\int_0^{h}\{1-p_t(0)\}\,\mathrm{d}t\right]. \tag{2.19}$$

(d) Multi-site properties

Consider two points (sites) P_1 $(m=1)$ and P_2 $(m=2)$ in the x–y plane separated by a spatial distance d with polar coordinates $(0,0)$ and $(d,0)$ respectively. Divide the plane into four quadrants bounded by the perpendicular bisector of the line joining P_1 and P_2 and the x-axis, i.e. bounded by the lines with equations: $2r\cos\theta = d$ and $\theta = 0$.

Consider the quadrant $Q = \{(r,\theta) : r \geqslant d/(2\cos\theta); 0 \leqslant \theta \leqslant \frac{1}{2}\pi\}$. A cell with disc centre in Q overlaps P_1 and P_2 if and only if the cell overlaps point P_1. Therefore, the expected number of type i cells, generated from a single storm origin, simultaneously overlapping both P_1 and P_2, denoted as ϖ_i, is given by

$$\begin{aligned}\varpi_i &= 4\int_0^{\pi/2}\int_{d/2\cos\theta}^{\infty} \rho_i r \mathrm{e}^{-\gamma_i r}\,\mathrm{d}r\,\mathrm{d}\theta \\ &= \nu_i + 2d\rho_i\gamma_i^{-1}K_0(\gamma_i d/2) - 2\pi^{-1}\nu_i\int_0^{\gamma_i d/2} K_0(y)\,\mathrm{d}y, \end{aligned} \tag{2.20}$$

where $K_u(\cdot)$ is a Bessel function of imaginary argument. Hence, the conditional probability that a type i cell overlaps P_1 given that the cell overlaps P_2, denoted as κ_i, is given by

$$\kappa_i = 1 + d\gamma_i\pi^{-1}K_0(\gamma_i d/2) - 2\pi^{-1}\int_0^{\gamma_i d/2} K_0(y)\,\mathrm{d}y. \tag{2.21}$$

Asymptotically, for large d,

$$\kappa_i \approx (d\gamma_i\pi^{-1}\mathrm{e}^{-\gamma_i d})^{1/2}. \tag{2.22}$$

Cells that overlap P_1 can be partitioned into those that overlap P_1 only (with expected number $(1-\kappa_i)\nu_i$ for type i cells), and those that overlap both P_1 and P_2 (with expected number $\kappa_i\nu_i$ for type i cells). Cells that overlap P_2 can be partitioned in a similar way, with the same expectations, i.e. the expected number of type i cells overlapping P_2 only is $(1-\kappa_i)\nu_i$, and the expected number of type i cells overlapping both P_2 and P_1 is $\kappa_i\nu_i$. Then, denoting the total intensities at P_1 and P_2 as $Y^{(1)}(t)$ and $Y^{(2)}(t)$ respectively (defined by (2.3)) and following Cowpertwait (1994, § 3), the lag τ cross-covariance is given by

$$\begin{aligned}\operatorname{Cov}\{Y^{(1)}(t), Y^{(2)}(t+\tau)\} &= \varphi_1\varphi_2\sum_{i=1}^{n}\sum_{j=1}^{n} C_{ij}(\tau) \\ &\quad + \varphi_1\varphi_2\sum_{i=1}^{n}\left\{\lambda\kappa_i\nu_i E(X_i^2)\frac{\mathrm{e}^{-\eta_i\tau}}{\eta_i}\right\}. \end{aligned} \tag{2.23}$$

Let $S_k^{(1)}(h)$ and $S_{k+l}^{(2)}(h)$ be the aggregated rainfall totals in the kth and $(k+l)$th intervals of duration h for points P_1 and P_2 respectively (defined by (2.4)). Then, the aggregated cross-covariance function follows directly as, for $l \geqslant 0$:

$$\begin{aligned}\operatorname{Cov}\{S_k^{(1)}(h), S_{k+l}^{(2)}(h)\} \;=\; &\varphi_1\varphi_2\sum_{i=1}^{n}\sum_{j=1}^{n}\int_{h(k-l)}^{hk}\int_{h(k+l-1)}^{h(k+l)} C_{ij}(|t-s|)\,\mathrm{d}s\,\mathrm{d}t \\ &+ \varphi_1\varphi_2\sum_{i=1}^{n}\int_{h(k-l)}^{hk}\int_{h(k+l-1)}^{h(k+l)}\left\{\lambda\kappa_i\nu_i E(X_i^2)\frac{\mathrm{e}^{-\eta_i|t-s|}}{\eta_i}\right\}\mathrm{d}s\,\mathrm{d}t\end{aligned}$$

$$= \varphi_1\varphi_2 \sum_{i=1}^{n}\sum_{j=1}^{n}\{A_j(h,l)A_{ij} + B(h,l)B_{ij}\}$$
$$+ \varphi_1\varphi_2 \sum_{i=1}^{n}\left\{\lambda\kappa_i\nu_i E(X_i^2)\frac{A_i(h,l)}{\eta_i}\right\}, \qquad (2.24)$$

where A_{ij}, B_{ij}, $A_j(h,l)$ and $B(h,l)$ are given by (2.12), (2.15) and (2.16). This can be expressed in terms of the single-site autocovariance (2.14) as, for $l \geqslant 0$:

$$\operatorname{Cov}\{S_k^{(1)}(h), S_{k+l}^{(2)}(h)\} = \varphi_1\varphi_2 \operatorname{Cov}\{S_k^{(m)}(h), S_{k+l}^{(m)}(h)\}/\varphi_m^2$$
$$- \varphi_1\varphi_2 \sum_{i=1}^{n}\left\{\lambda(1-\kappa_i)\nu_i E(X_i^2)\frac{A_i(h,l)}{\eta_i}\right\}. \qquad (2.25)$$

Cross-over transition probabilities can be derived as follows. If $S_k^{(2)}(h)$ is zero then $S_k^{(1)}(h)$ is also zero if and only if (i) all cells overlapping P_1 only (i.e. P_1 but not P_2) terminate before the time interval $[(k-1)h, kh]$ and (ii) no cells overlapping P_1 only have arrival times in the interval. Therefore, from (2.19),

$$\phi_{1,2}(h) = \operatorname{pr}\{S_k^{(1)}(h) = 0 \,|\, S_k^{(2)}(h) = 0\}$$
$$= \exp\left[-\lambda\int_0^\infty \{1 - p_h(t)\}\,\mathrm{d}t - \lambda\int_0^h \{1 - p_t(0)\}\,\mathrm{d}t\right], \qquad (2.26)$$

where $p_h(t)$ is given by equation (2.18) but with $\mu_C = \sum_{i=1}^{n}(1-\kappa_i)\nu_i$ and $\omega = \sum_{i=1}^{n}\alpha_i\omega_i$.

The probability $\operatorname{pr}\{S_k^{(1)}(h) > 0 \,|\, S_k^{(2)}(h) > 0\}$ is given by

$$\operatorname{pr}\{S_k^{(1)}(h) > 0 \,|\, S_k^{(2)}(h) > 0\} = \{1 - 2\phi(h) + \phi_{1,2}(h)\phi(h)\}/\{1-\phi(h)\}, \qquad (2.27)$$

where $\phi(h) = \operatorname{pr}\{S_k^{(m)}(h) = 0\}$ for all k and m.

To derive the cross-over dry transition probability

$$\operatorname{pr}\{S_{k+1}^{(1)}(h) = 0 \,|\, S_k^{(2)}(h) = 0\},$$

consider storm origins with arrival times in the intervals: $(-\infty, (k-1)h]$, $((k-1)h, kh]$ and $(kh, (k+1)h]$. Interval $(kh, (k+1)h]$ is dry at P_1 due to storm origins with arrival times in $(kh, (k+1)h]$ with probability e_1 given by

$$e_1 = \exp\left[-\lambda\int_0^h \{1 - p(t)\}\,\mathrm{d}t\right], \text{ where } p(t) = \exp\left\{-(1-\mathrm{e}^{-\beta t})\sum_{i=1}^{n}\nu_i\right\}. \qquad (2.28)$$

Interval $(kh, (k+1)h]$ is dry at P_1 due to cells that overlap P_1, but not P_2, and have associated storm origins with arrival times in $(-\infty, kh]$, with probability e_2 given by

$$e_2 = \exp\left[-\lambda\int_0^\infty \{1 - p_h(t)\}\,\mathrm{d}t\right], \qquad (2.29)$$

where $p_h(t)$ is given by equation (2.18) with $\mu_C = \sum_{i=1}^{n}(1-\kappa_i)\nu_i$. Interval $(kh, (k+1)h]$ is dry at P_1 due to cells that overlap both P_1 and P_2, and have associated storm origins with arrival times in $((k-1)h, kh]$, with probability e_3 given by

$$e_3 = \exp\left[-\lambda h + \lambda h \exp\left\{-(1-\mathrm{e}^{-\beta h})\sum_{i=1}^{n}\kappa_i\nu_i\right\}\right]. \qquad (2.30)$$

Finally, cells that overlap both P_1 and P_2 and have arrival times in $(-\infty, (k-1)h]$ have all terminated before the interval $((k-1)h, kh]$. Hence, interval $(kh, (k+1)h]$ is dry at P_1 due to cells that overlap both P_1 and P_2, and have associated storm origins with arrival times in $(-\infty, (k-1)h]$, with probability e_4 given by

$$e_4 = \exp\left[-\lambda \int_0^\infty \{1 - p_h(t)\}\,\mathrm{d}t\right], \tag{2.31}$$

where $$p_h(t) = \exp\left\{-\mathrm{e}^{-\beta t}(1 - \mathrm{e}^{-2\beta h}) \sum_{i=1}^{n} \kappa_i \nu_i\right\}.$$

Therefore,

$$\mathrm{pr}\,\{S_{k+1}^{(1)}(h) = 0\,|\,S_k^{(2)}(h) = 0\} = e_1 e_2 e_3 e_4 \tag{2.32}$$

and

$$\mathrm{pr}\,\{S_{k+1}^{(1)}(h) > 0\,|\,S_k^{(2)}(h) > 0\} = \frac{\{1 - 2\phi(h) + e_1 e_2 e_3 e_4 \phi(h)\}}{\{1 - \phi(h)\}}. \tag{2.33}$$

3. Further generalizations

(a) *Classifying storm origins*

In the model formulation in § 2, the parameters of a cell depend on the type of cell, but not on the storm origin. A further extension of the model is to classify the storm origins, each origin being randomly classed from 1 to n types, and have different parameters associated with different types of storm. This is equivalent to the superposition of n independent Poisson processes with rates λ_i, so the second-order properties can be easily extended. For example, each type i storm origin could generate a random number C_i of cell origins, where each cell is randomly classified from 1 to n_i with probability α_{ij} that the cell is of type j, and the arrival times of the cells after the origin are independent exponential random variables with parameter β_i ($j = 1, \ldots, n_i; i = 1, \ldots, n$). The cell duration, intensity, and arrival times could then have different distributions for different storm types, which may be of some practical value because different rainfall systems can exhibit vastly different properties.

(b) *Modelling rainfall in mountainous regions*

In the model formulation in § 2, a scaling factor φ_m was introduced to allow for the effects of orography. However, in mountainous regions, a cell may deposit rain at a high altitude but not at a low altitude, because the rain due to the cell completely evaporates before reaching the ground level at low altitude. Consequently, it seems reasonable to introduce a probability $\psi_{m,i}$ that a type i cell overlapping point m survives to ground level, with high probabilities being associated with points at high altitudes. Then, in the notation of § 2 *a*

$$X_{m,t}^{(i)}(u, s) = \begin{cases} \varphi_m X_i & \text{with probability } \psi_{m,i}\mathrm{e}^{-\gamma_i u}\mathrm{e}^{-\eta_i s}, \\ 0 & \text{otherwise,} \end{cases} \tag{3.1}$$

and the second-order properties are easily modified by replacing $E(X_i)$ and $E(X_i^2)$ in equations (2.11)–(2.14) and (2.24) with $\psi_{m,i}E(X_i)$ and $\psi_{m,i}E(X_i^2)$ respectively.

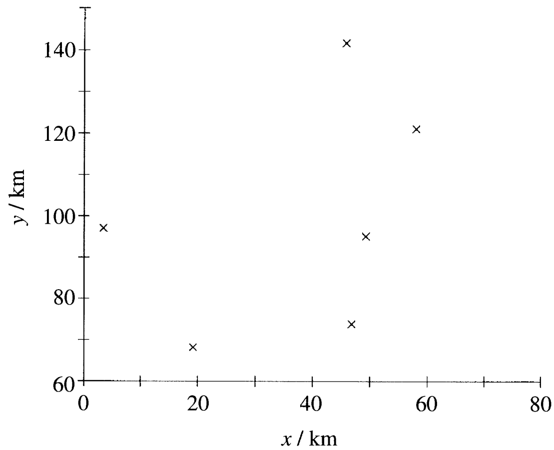

Figure 1. The relative positions of sites used in the empirical example.

To find $\mathrm{pr}\,\{S_k^{(m)}(h) = 0\}$, let ω_i be the probability that a type i cell surviving at point m with arrival time in $(0, t)$ terminates before t (given by (2.17)), ω be the probability that an arbitrary cell with arrival time in $(0, t)$ terminates before time t, and $p_h(t)$ be the probability of no rain in $(t, t+h)$ due to a storm origin at time zero. Then

$$\omega = \sum_{i=1}^{n} \alpha_i (1 - \psi_{m,i} + \omega_i \psi_{m,i}), \tag{3.2}$$

and, after some algebra,

$$p_h(t) = \exp\biggl\{-\mu_C + \mu_C \mathrm{e}^{-\beta(t+h)} + \omega \mu_C (1 - \mathrm{e}^{-\beta t}) + \mu_C (\mathrm{e}^{-\beta t} - \mathrm{e}^{-\beta(t+h)}) \sum_{i=1}^{n} \alpha_i (1 - \psi_{m,i}) \biggr\}, \tag{3.3}$$

where $\mu_C = \sum_{i=1}^{n} \nu_i$. The probability that an arbitrary interval is dry at point m, $\mathrm{pr}\,\{S_k^{(m)}(h) = 0\}$, is then given by substituting $p_h(t)$, from (3.3), into (2.19).

4. Empirical example

As an example, the model of § 2 was fitted to rainfall data taken from six sites in the Thames basin, UK (figure 1). For each site, the data for October were pooled to produce six records of hourly data.

There is a choice of many properties, each with a range of aggregation levels, that could be used to fit the model. As the focus in this paper is on model formulation, rather than model validation, a small subset of the many available properties was selected for estimation purposes, the sample values being given in table 1.

The mean 1 h rainfall was used to estimate the scaling factor φ_m for each site m $(m = 1, 2, \ldots, 6)$. This is equivalent to dividing each hourly series by its mean to produce six transformed series, each having the same mean and approximately the same variance (i.e. approximate weak stationarity in space), and then fitting to properties in the transformed series.

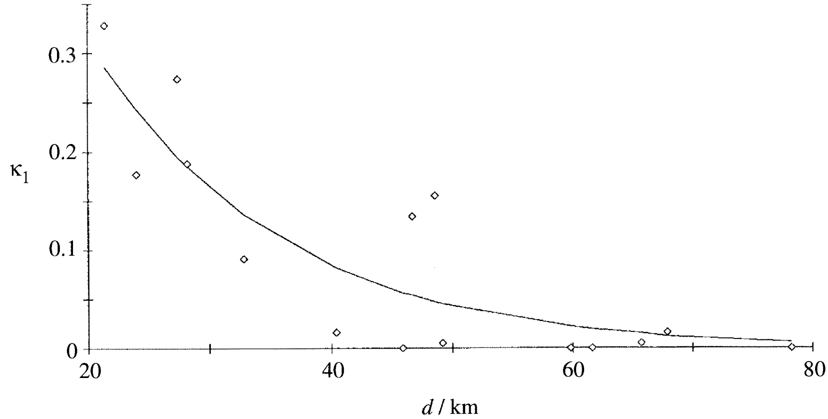

Figure 2. Proportion of type 1 rain cells (convective), $\hat{\kappa}_1$, occurring simultaneously at a pair of sites ($\diamond$), estimated from hourly lag 0 cross-correlation function (using equation (2.24)), and predicted values (represented by the line) using least squares estimate of γ_1 in equation (2.21) plotted against distance d between pairs of sites in the x–y plane. Standard deviation of residuals is 0.055.

Two types of rain cell were fitted: heavy convective cells and light stratiform cells, with the cell intensities X_1 and X_2 distributed exponentially with parameters ξ_1 and ξ_2 respectively. The durations of the heavy and light cell types were taken to be exponential random variables with parameter values 3.0 and 0.53 h^{-1} respectively. These were estimated in a previous paper using a long record of rainfall data taken from a relatively close site (see Cowpertwait 1994, § 4).

The remaining parameters λ, β, ν_1, ν_2, ξ_1, and ξ_2 were estimated by minimizing the following sum of squares:

$$\sum_{m=1}^{6} \sum_{f_m \in F} w(\hat{f}_m)\{1 - f_m/\hat{f}_m\}^2 \tag{4.1}$$

subject to: λ, β, ν_1, ν_2, ξ_1, $\xi_2 > 0$, where $\hat{f}_m$ is a sample estimate taken from the hourly data for the mth site, and F denotes a set of aggregated single-site properties. In this example, F consisted of the 1 h mean, variance, and lag 1 autocorrelation, the 6 and 24 h variances, and the proportion of dry days (table 1).

The $w(\hat{f}_m)$ are weights that can be applied if some of the properties are to be fitted to a greater accuracy. For example, $w(\hat{f}_m)$ could be inversely proportional to $\mathrm{Var}(\hat{f}_m)$, or different weights could be subjectively chosen to ensure that the model fits accurately properties judged important in a particular application. In this example, all the weights were 1 except that for the term containing the mean which was given an arbitrary high weight of 100.

The proportion of type 1 and 2 cells simultaneously overlapping a pair of sites, κ_1 and κ_2 were estimated from the lag 0 cross-correlation function (using equation (2.24)) and are plotted against the distance between the sites in figures 2 and 3. Least squares estimates of the parameters for the radius of a heavy and light cell were obtained using equation (2.21) with series expansions substituted for the Bessel function and the integral of the Bessel function (figures 2 and 3).

The estimates for the mean radius of a heavy and light cell were 6.6 km and 22.6 km, so that the mean areas occupied by heavy and light cells are approx-

Table 1. *Rainfall for Thames catchment*

(For 1 October – 31 October 1989–93: sample estimates and fitted values (in parentheses) given by the spatial-temporal model.)

site location x, y (km)	alt (m)	mean 1 h (mm)	1 h variance (mm^2)	6 h variance (mm^2)	24 h variance (mm^2)	24 h proportion dry	1 h lag 1 correlation
46.8, 73.9	104	0.0962	0.243	4.16	25.3	0.561	0.554
		(0.0962)	(0.258)	(4.08)	(24.5)	(0.594)	(0.523)
58.1, 121.	66	0.0809	0.178	2.77	15.5	0.561	0.595
		(0.0809)	(0.182)	(2.88)	(17.3)	(0.594)	(0.523)
3.4, 97.1	94	0.0945	0.259	3.87	23.6	0.606	0.518
		(0.0945)	(0.249)	(3.94)	(23.6)	(0.594)	(0.523)
49.3,95.2	50	0.0820	0.196	2.90	17.8	0.606	0.461
		(0.0820)	(0.188)	(2.97)	(17.8)	(0.594)	(0.523)
45.8, 142.	91	0.0798	0.167	2.84	17.8	0.606	0.533
		(0.0798)	(0.177)	(2.81)	(16.8)	(0.594)	(0.523)
19.2, 68.3	160	0.1212	0.451	6.84	41.4	0.626	0.519
		(0.1212)	(0.410)	(6.48)	(38.9)	(0.594)	(0.523)

Parameter estimates:
$\hat{\lambda} = 0.0156\ \mathrm{h}^{-1}$, $\hat{\beta} = 0.186\ \mathrm{h}^{-1}$, $\hat{\xi}_1 = 0.0211$ h, $\hat{\xi}_2 = 0.114$ h,
$\hat{\nu}_1 = 1.22$, $\hat{\nu}_2 = 2.71$, $\hat{\eta}_1 = 3.0\ \mathrm{h}^{-1}$, $\hat{\eta}_2 = 0.53\ \mathrm{h}^{-1}$,
$\hat{\rho}_1 = 4.43 \times 10^{-3}\ \mathrm{km}^{-2}$, $\hat{\rho}_2 = 8.46 \times 10^{-4}\ \mathrm{km}^{-2}$, $\hat{\gamma}_1^{-1} = 6.62$ km, $\hat{\gamma}_2^{-1} = 22.6$ km,
$\hat{\varphi}_j$ = sample mean 1 h rainfall at jth site.

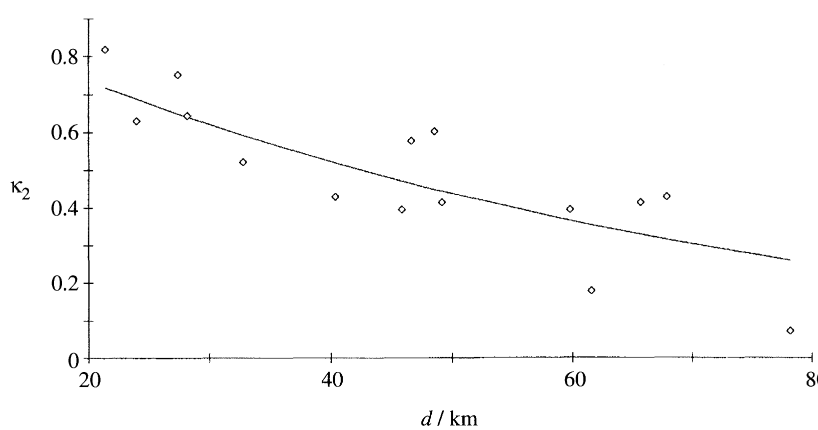

Figure 3. Proportion of type 2 rain cells (stratiform), $\hat{\kappa}_2$, occurring simultaneously at a pair of sites ($\diamond$), estimated from hourly lag 0 cross-correlation function (using equation (2.24)), and predicted values (represented by the line) using least squares estimate of γ_2 in equation (2.21) plotted against distance d between pairs of sites in the x–y plane. Standard deviation of residuals is 0.11.

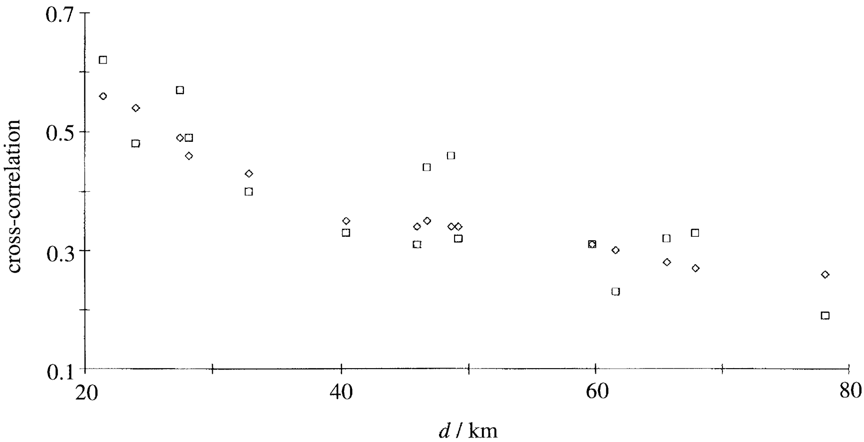

Figure 4. Lag 0 hourly cross-correlations for pairs of sites plotted against distance d between sites. Sample estimates (□) and fitted values (◇). Standard deviation of residuals is 0.061.

imately 140 km^2 and 1600 km^2 respectively. These are in broad agreement with the precipitation areas classified as small mesoscale areas and large mesoscale areas by Austin & Houze (1972) and Waymire *et al.* (1984), where the small areas correspond to regions of heavy convective cells, and the larger areas to lighter stratiform rain.

A comparison was made between the sample lag 0 hourly cross-correlations and the fitted values. They are plotted against spatial distance in figure 4, from which no obvious bias in the fitted values is evident. Furthermore, a satisfactory fit is suggested by a 75% explained variance (R^2). Although not shown here, a satisfactory fit was also obtained to sample cross-correlations at higher lags. For each pair of sites, the maximum sample cross-correlation occurred at lag zero.

5. Discussion

In this paper, the spatial-temporal model has been fitted, using second-order properties, to a relatively short record of multi-site data. However, there are many other properties of interest in applications, e.g. multivariate extreme values. An assessment of fit to these properties could be achieved using a longer record of multi-site data. The model has a flexible structure, via the generalization, so that a reasonable fit to multi-site extreme values could probably be achieved.

The range of applications to which the model could be applied may be limited because rain cells are taken to have zero velocity. For example, the model is not intended for real-time forecasting. However, the model could be used in longer-term simulation studies for which rainfall data are needed above some aggregation level. Current work suggests that the present form of the model may be representative of hourly rainfall data which are used in many hydrological applications.

The model could be used to simulate rainfall at all points in a catchment. This would require an estimate of the scaling factor, or mean rainfall, at any point. A simple regression model could be used to predict the mean rainfall from the altitude (e.g. see table 1, where the mean increases approximately linearly with altitude). Furthermore, if some of the parameters can be fixed regionally or nationally, then only a limited amount of data would be needed to fit the model. For example, the mean radii for heavy and light rain cells might be treated as approximately constant over a large geographical region. The model could then be fitted using data taken from a single gauge only.

An exciting new area of research is the use of point process models in climate change studies. Burlando & Rosso (1991) proposed a method of linking the parameters of the Neyman–Scott model with those of an atmospheric circulation model. The model presented in this paper could also be used in simulating future rainfall scenarios. One approach might be to find a relationship between model parameter estimates, e.g. proportions of cell (or storm) types, and the frequency of different weather types (Jones *et al.* 1993) in the present climate. This relationship could then be used to estimate the parameters of the rainfall model from weather types generated by a circulation model of a future atmosphere containing enhanced concentrations of greenhouse gases. The spatial-temporal model could then be used to assess the impact of future rainfall time series on river basins or urban drainage systems.

I thank Professor Enda O'Connell for stimulating my interest in multi-site analysis. I am also grateful to Thames Water for providing the data. This research was in part supported by the EC Environment Research Programme (Contract: EV5V-CT94-0510, Climatology and Natural Hazards).

References

Austin, P. M. & Houze, R. A. 1972 Analysis of the structure of precipitation patterns in New England. *J. appl. Meteorology* **11**, 926–935.

Browning, K. A. 1985 Conceptual models for precipitation systems. *Meteor. Mag.* **114**, 293–319.

Burlando, P. & Rosso, R. 1991 Extreme storm rainfall and climate change. *Atmos. Res.* **27**, 169–189.

Cowpertwait, P. S. P. 1994 A generalized point process model for rainfall. *Proc. R. Soc. Lond.* A **447**, 23–37.

Cox, D. R. & Isham, V. 1988 A simple spatial-temporal model of rainfall. *Proc. R. Soc. Lond.* A **415**, 317–328.

Jones, P. D., Hulme, M. & Briffa, K. R. 1993 A comparison of Lamb circulation types with an objective classification scheme. *Int. J. Climatology* **13**, 655–663.

Mellor, D. 1992 The modified turning bands model for space-time precipitation. Ph.D. thesis, Department of Civil Engineering, University of Newcastle upon Tyne.

Rodriguez-Iturbe, I., Cox, D. R. & Eagleson, P. S. 1986 Spatial modelling of total storm rainfall. *Proc. R. Soc. Lond.* A **403**, 27–50.

Rodriguez-Iturbe, I., Cox, D. R. & Isham, V. 1987 Some models for rainfall based on stochastic point processes. *Proc. R. Soc. Lond.* A **410**, 269–288.

Waymire, E., Gupta, V. K. & Rodriguez-Iturbe, I. 1984 A spectral theory of rainfall intensity at the meso-β scale. *Water Resour. Res.* **20**, 1453–1465.

Received 9 September 1994; revised 25 January 1995; accepted 28 February 1995