Hydrology and Earth System Sciences, 1(2), *367–378* (1997) © EGS

Hydrology & Earth System Sciences

A combined Pòlya process and mixture distribution approach to rainfall modelling

E. Todini[1] and M. di Bacco[2]

[1] Department of Earth and Geo-Environmental Sciences, University of Bologna—Bologna—Italy
[2] Department of Statistics, University of Bologna—Bologna—Italy

Abstract

A new probabilistic interpretation of at site rainfall sequences is introduced for the development of a stochastic model of rain. The model, is divided into two sub models; the first one describing the total number of rainfall spells within a window of time is described by a Pòlya process in order to reproduce better the variable probability of occurrence of rainfall during storm events (due to the presence of different numbers of rainfall cells); the second sub model, conditional on the first one, describes the total quantity of rainfall in the time window, given a number of rainfall spells.

The probabilistic rainfall model, which has shown interesting properties in reproducing the probability distribution of observed data at time scales ranging from one hour to twenty-four hours, may be the basis for a number of applications which include the development of a conditional stochastic generator of rain, within the frame of real-time flood forecasting, and the derivation of a probabilistic distribution of rainfall extremes at the various time scales.

Introduction

In order to interpret observed rainfall time series, several authors have used the assumptions of Poisson arrivals of rainfall spells associated with a probability distribution of rainfall quantities (Todorovic and Yevjevic, 1969; Gupta and Duckstein, 1975). Other models use two stage processes to represent the different transition probabilities between rainfall/rainfall, rainfall/no-rainfall, no-rainfall/rainfall and no-rainfall/no-rainfall conditions (Woolhiser and Pegram, 1979; Smith and Karr, 1983; Foufoula-Georgiou and Lettenmaier, 1987; Smith 1987). Others, again, are based upon the Neyman-Scott model (Kavvas and Delleur, 1981; Rodriguez-Iturbe *et al.*, 1984; 1987a,b,c; Cowpertwait, 1994) better to account for the physical features of rain fields. For a detailed description of the point models see also Cox and Isham (1980), Waymire and Gupta (1981a,b,c) or Rodriguez-Iturbe *et al.* (1987a).

Rainfall records are generally available as totals sampled at fixed time intervals (e.g. one hour, three hours, twenty-four hours); this, together with the need for relatively short horizon rainfall forecasts (from twelve to twenty-four hours) conditional upon the latest observations, motivated the development of the discrete-time model.

The advantage of the proposed formulation lies in the analytical derivation of the probability distributions of rainfall totals from which, when short forecasting horizons are needed as in the case of real time forecasting, the conditional distributions can be obtained directly using the conditional probability theorem.

The model is based upon the concept of a time window, that is a window of fixed time length within which a number of time intervals Δt can be rainy or non-rainy.

To develop the model, it is assumed that the knowledge of the number of rainy intervals is informative on the total quantity of rainfall while the total quantity of rainfall is not informative on the number of rainy intervals; the same amount could be originated by short duration severe storms or longer duration low intensity precipitation. Therefore, it is possible to treat separately (a) the probability of rainy intervals and (b) the probability of rain totals conditional upon the number of rainy intervals. Similarly to what was done by Thompson (1984) in the case of monthly rainfall totals, the model is thus derived in two stages: the first one is the derivation of the probability distribution of the number of rainy time intervals within a time window of specified length, while the second one is the derivation of the probability distribution function of the total quantity of rain, conditional upon the number of rainy intervals in the time window.

The derivation of the mixture distributions

Suppose that a sufficiently long record of rainfall is available and that it has been divided into windows of length

$n\Delta t$. For each window, the number of rainy time intervals d, with $0 \leq d \leq n$, is computed. If one assumes that P_o, the probability of no rainfall occurring in one single interval of time Δt in a window of length $n\Delta t$, is constant in time, the following Binomial distribution:

$$Pr\{d \mid P_0, n\} = \binom{n}{d} P_0^{n-d}(1-P_0)^d \qquad (1)$$

should reproduce the observed frequency distribution of the number d of rainfall spells in a window of length n. When there are no clouds in the sky the probability of rainfall occurring is far lower than when the sky is covered by thick clouds or when it is already raining, and so the time invariance of P_o, and the validity of Eqn. (1) is called into question. In Fig. 1, the frequency of d (varying from 0 to 6) obtained by using data for the raingauge station of Fornacina in Tuscany is shown for a window of $6\Delta t$, (Δt = 1 hour) together with the probabilities computed according to the Binomial model of Eqn. (1): the mismatch is evident and the same results have been found for a number of rainfall records from different parts of the world. An alternative and more plausible model to (1) is now derived.

DERIVATION OF A NEW MODEL FOR THE NUMBER OF RAINY SPELLS IN A WINDOW OF TIME

Given the results shown in Fig. 1, it seemed reasonable to define a new probabilistic model in which the Binomial model expressed by Eqn. (1) is assumed valid only when a value for P_o is known, i.e. it is taken as a model conditional upon P_o. Given the uncertainty on P_o, it is necessary to consider it as a random variable which is assumed to be distributed according to a Beta distribution between 0 and 1, that can be written, for a window of length n as:

$$f(P_0 \mid r_n, s_n) = \frac{\Gamma(r_n + s_n)}{\Gamma(r_n)\Gamma(s_n)} P_0^{r_n - 1}(1-P_0)^{s_n - 1} \qquad (2)$$

where parameters r_n and s_n are dependent on the window size n. The choice of the Beta distribution, a natural conjugate prior of the Binomial distribution, is not really binding because the Beta is formed by a very wide family of distributions and is quite adequate to reproduce the shape of the distribution of a quantity such as a probability.

The new model can then be derived as:

$$Pr\{d \mid n, r_n, s_n\} = \int_0^1 Pr\{d \mid P_0, n\} f(P_0 \mid r_n, s_n) dP_0$$
$$= \frac{n!}{d!\,(n-d)!} \frac{\Gamma(r_n + s_n)}{\Gamma(r_n)\Gamma(s_n)} \frac{\Gamma(n-d+r_n)\Gamma(d+s_n)}{\Gamma(n+r_n+s_n)} \qquad (3)$$

with two parameters r_n and s_n that have to be estimated from the observations.

This model, which corresponds to observing an average process combination of an infinite number of different Binomial processes owing to the assumed variability of P_o, is the classical Pòlya urn (Feller, 1970; Ord, 1972) process model, a non-markovian non-stationary birth process, extensively used for representing the spread of contagious diseases, where the probability of possible new ill people increases with the number of observed illness cases.

The consistency of the model hypotheses with the data

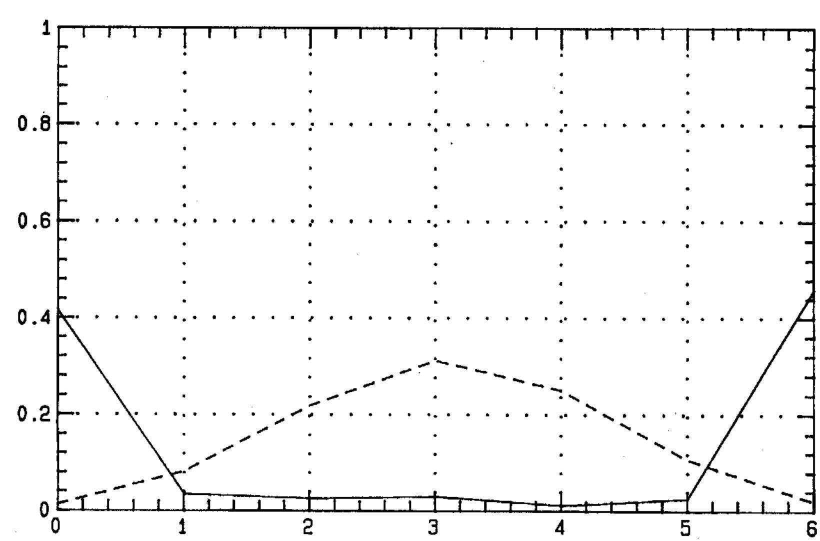

Fig. 1 *Comparison between frequencies of rainy spells in a window of 6 hours, observed at Fornacina gauge (solid line) and computed using the Binomial model (dashed line).*

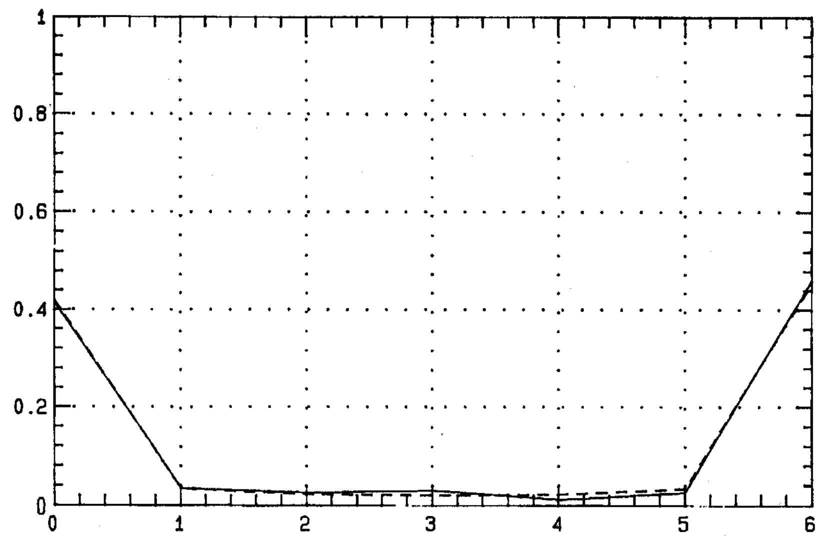

Fig. 2 *Comparison between frequencies of rainy spells in a window of 6 hours, observed at Fornacina gauge (solid line) and computed using the Polya process model (dashed line).*

can be seen in Fig. 2 where the results of the probabilistic model are plotted as in Fig. 1. These results are presented here to justify the approach taken but are dependent on the estimation of r_n and s_n.

DERIVATION OF THE MODEL OF RAINFALL TOTALS

It is common practice to assume that, when sampling on sufficiently small time intervals the rainfall quantity x is distributed according to a negative exponential distribution with parameter θ:

$$f(x \mid \theta, n = 1) = \theta e^{-\theta x} \qquad \forall x > 0 \tag{4}$$

A more complex problem is to derive the probability distribution of the rainfall totals for a number n of intervals.

Several probability distributions conditional upon the sampling interval have been proposed for the non-zero daily or hourly rainfall amounts. Todorovic and Woolhiser (1971) and Richardson (1981) used an exponential distribution, but Skees and Shenton (1974) and Mielke and Johnson (1974) suggested that the exponential distribution has a thinner tail than the one observed in the daily amounts. The mixed exponential distribution was explored by Smith and Schreiber (1973), Woolhiser and Pegram (1979), and Roldan and Woolhiser (1982), among others.

The probability of $x = 0$ in windows of length n, can easily be derived from Eqn. (3) for $d = 0$, given that when the total amount of rainfall x is equal to zero, also the number of rainy spells d must be null.

$$Pr\{x = 0 \mid n\} = Pr\{d = 0 \mid n, r_n, s_n\} = \frac{\Gamma(r_n + s_n)\Gamma(n + r_n)}{\Gamma(r_n)\Gamma(n + r_n + s_n)} \tag{5}$$

On the contrary, when $x > 0$, the probability of the rainfall totals can only be derived if a probability density function of x, conditional upon d and n, can be derived. To do this, the probability distribution of the sum of d independent, equally distributed negative exponential random variables is first expressed as a Gamma distribution:

$$f(x \mid \theta_n, d, n) = \frac{\theta_n^d x^{d-1} e^{-\theta_n x}}{\Gamma(d)} \tag{6}$$

where θ_n is a parameter which depends on the window length n. To obtain the probability density of the rainfall totals, according to rainfall observations, Eqn. (6) is assumed to be valid for all n, once the value for θ_n, which represents the inverse of the average rainfall total in the time window, is known. In other words, it is assumed here that, conditional upon the knowledge of the value for θ_n, the rainfall quantities falling in the different time intervals are independent and identically distributed: this hypothesis is known in statistics as 'exchangeability' (Berger, 1985). The fact that the rainfall quantities do not appear to be independent is expressed here by the assumption (similar to what was done for the probability of the number of rainy intervals) that what is really observed is the result of a mixture of different processes, all expressed by equation (6), but each with a different value of the parameter θ_n, which is now taken as a random variable distributed according to a natural conjugate prior of the Gamma distribution, that is a Gamma distribution on θ_n:

$$f(\theta_n \mid d, n, a_{d,n}, b_{d,n}) = \frac{a_{d,n}^{b_{d,n}} \theta_n^{b_{d,n}-1} e^{-a_{d,n}\theta_n}}{\Gamma(b_{d,n})} \tag{7}$$

Equation 7 expresses the fact that the growth rate of the inverse of the average value of θ_n is a function of two sets of parameters $a_{d,n}$ and $b_{d,n}$ which are a function of the number of rainy spells d and of the dimension of the time window n. The actual distribution for the rainfall totals is then derived as the expected value of the mixture of different processes. For $x > 0$ the following expression is therefore obtained:

$$f(x \mid d, n, a_{d,n}, b_{d,n}) = \int_0^\infty f(x \mid \theta_n, d, n) f(\theta_n \mid d, n, a_{d,n}, b_{d,n}) d\theta_n = g(x, d, a_{d,n}, b_{d,n}) \quad (8)$$

where:

$$g(x, d, a, b) = \frac{\Gamma(d+b)}{\Gamma(d)\Gamma(b)} \frac{a^b x^{d-1}}{(a+x)^{d+b}} \quad (9)$$

The resulting expression is in practice another special case of the Generalized Beta of the Second Kind (GB2), which was also found by Mielke and Johnson (1974) to be appropriate for rainfall modelling, the parameters of which are here expressed as a function of d and n.

Given Eqns (3) and (9), the derivation of the probabilistic model of rainfall totals in a window of length $n\Delta t$ is now straightforward:

$$f(x \mid n, r_n, s_n, a_{d,n}, b_{d,n}) = \begin{cases} Pr\{d = 0 \mid n, r_n, s_n\} & x = 0 \\ \sum_{d=1}^{n} Pr\{d \mid n, r_n, s_n\} g(x, d, a_{d,n}, b_{d,n}) & x > 0 \end{cases} \quad (10)$$

For $x > 0$, the probability distribution function of the rainfall totals conditional upon the number of rainy time intervals can be derived by integrating Eqn. (8) to give:

$$F(x \mid d, n, r_n, s_n, a_{d,n}, b_{d,n}) = \int_0^x g(\chi, d, a_{d,n}, b_{d,n}) d\chi = G(x, d, a_{d,n}, b_{d,n}) \quad (11)$$

where:

$$G(x, d, a, b) = 1 - \frac{\Gamma(d+b)}{\Gamma(d)\Gamma(b)} \sum_{i=0}^{d-1} (-1)^i \binom{d-1}{i} \frac{1}{b+i} \left(\frac{a}{a+x}\right)^{b+i} \quad (12)$$

Using Eqns (3) and (12), the probability distribution function of the rainfall totals in a window of length $n\Delta t$ can also be derived as:

$$F(x \mid n, r_n, s_n, a_{d,n}, b_{d,n}) = \begin{cases} Pr\{d = 0 \mid n, r_n, s_n\} & x = 0 \\ Pr\{d = 0 \mid n, r_n, s_n\} + \sum_{d=1}^{n} Pr\{d \mid n, r_n, s_n\} G(x, d, a_{d,n}, b_{d,n}) & x > 0 \end{cases} \quad (13)$$

The dependence of the model parameters on the window size n gives rise to a potentially intractable parameter estimation problem. However, this can be resolved by taking advantage of the scaling properties of rainfall.

The parameter estimation problem

The model represented by Eqns (10) and (13) requires the estimation, for each window length n, of the two parameters r_n and s_n required for the definition of the probability distribution of d given n, and, in addition, $2n$ parameters $a_{d,n}$ and $b_{d,n}$ are required to define the probability distribution of x conditional upon d. The number of parameters required to define all the window lengths from *1* to n hours is $np = 2n + 2n(n + 1)/2 = n^2 + 3n$. For instance if $n = 24$ the number of parameters to be estimated becomes $n_p = 648$.

Fortunately, as reported in several works on the fractal, or multi-fractal, nature of rainfall (Lovejoy and Mandelbrot, 1985; Lovejoy and Schertzer, 1993; de Lima and Bogardi, 1995), the scaling properties of rainfall can be used to reduce the number of parameters. In addition, although Maximum Likelihood (ML) would have been the most appropriate estimation technique, the large number of available observations (in this case > 26,000) made it unnecessary and allowed the method of moments to be used, thus avoiding ML computational complexity. As a result, only three moments and three scale parameters are needed for all the possible time scales (or values of n) at which the validity of the model can be proved. The estimation of r_n and s_n at all time scales can thus be reduced to the estimation of one moment and one scale parameter, and the estimation of $a_{d,n}$ and $b_{d,n}$ can be reduced to the estimation of two moments and two scale parameters.

SCALE INVARIANCE PROPERTIES

Sampling the rainfall records on the basis of increasing size windows, a scale effect can be observed, both in terms of number of rainfall intervals and of rainfall totals. Figures 3 and 4, where the increasing order moments of variables d and x are presented as a function of the natural log of n, show the following properties:

- The growth of the first order moment (both of d and x) is linear as a function of n.
- The growth of all moments (both of d and x) is linear as a function of the natural log of the window dimension n, allowing for the variability of moments at high values of n.

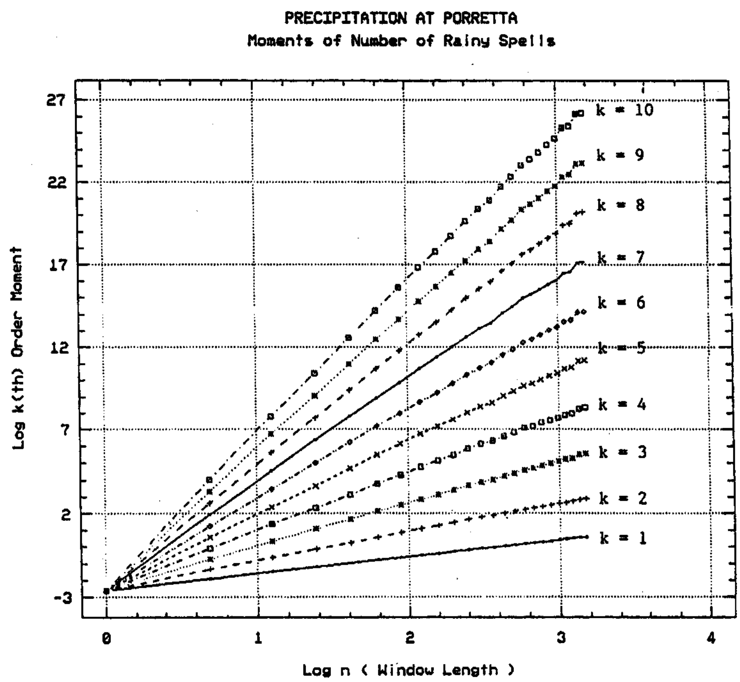

Fig. 3 *Increasing k-order moments of number d of rainy intervals versus ln n, with n the size of the time window.*

– All the moments of d are identical when $n = 1$, because in that case the series becomes a succession of zeroes and ones and therefore $0^k = 0$, $1^k = 1$.

After analyzing several rainfall records, it has been found that a very general expression can be used to reproduce all the k^{th} order sample moments, namely $m^{(k)}$, at the different time scales (values of n) as a function of the sample moments estimated for $n = 1$:

$$\begin{cases} m_n^{(k)} = m_{n=1}^{(k)}\ n & k = 1 \\ m_n^{(k)} = m_{n=1}^{(k)}\ n^{1+(k-1)[u+v\ln(k-1)]} & k > 1 \end{cases} \tag{14}$$

where u and v are parameters. Fig. 5 shows the relation-

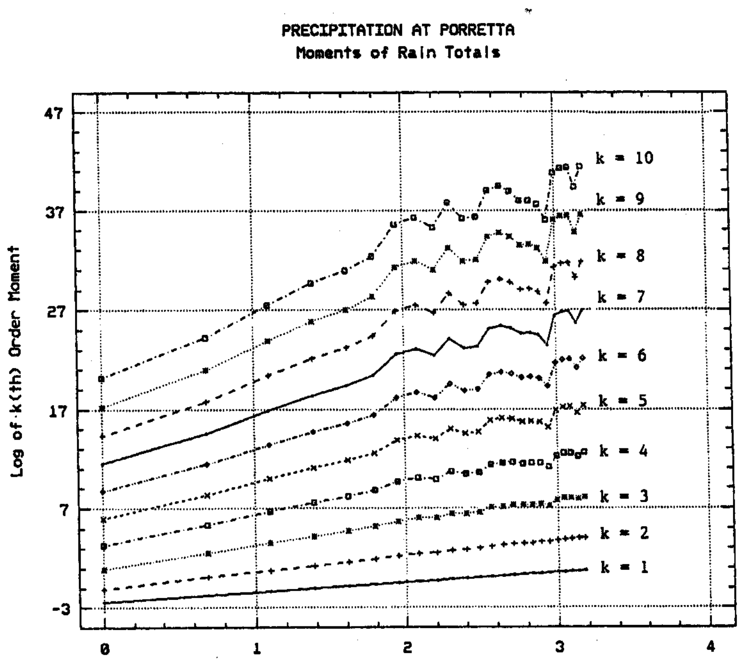

Fig. 4 *Increasing k-order moments of rainfall totals x versus ln n, with n the size of the time window.*

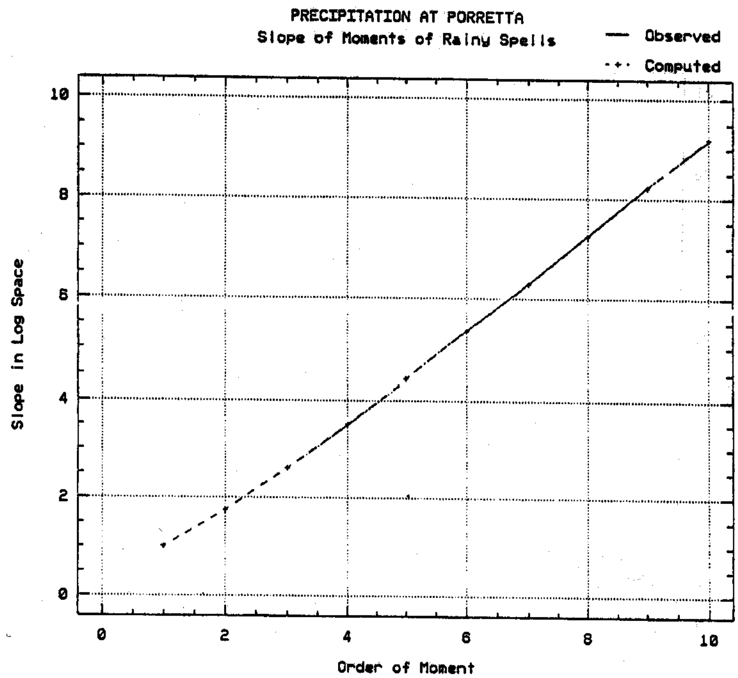

Fig. 5 *Comparison between observed (solid line) and modelled (dashed line) scaling function of equation (14) for increasing order moments of d, the number of rainy spells.*

ship, expressed by Eqn. (14), which was derived for the moments of d while Fig. 6 shows the one derived for the moments of x; it should be noted that, in both cases, the explained variance is over .9999 and that the relationship found for x, strictly linear, corresponds to a special case with $v = 0$.

For the reduction of the number of the rainfall model parameters the scaling properties of the first two moments

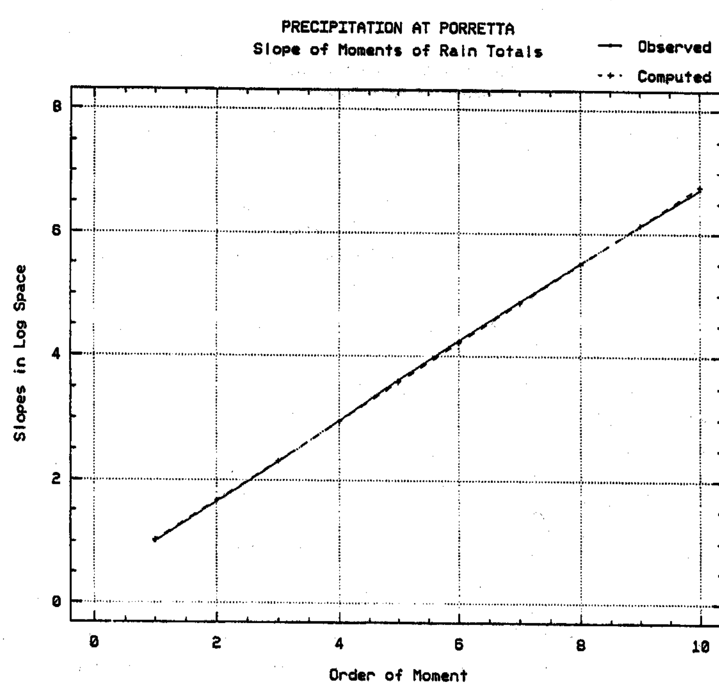

Fig. 6 *Comparison between observed (solid line) and modelled (dashed line) scaling function of equation (14) for increasing order moments of x, the rainfall totals.*

in terms of d and the first two in terms of x are necessary and can be expressed as follows:

$$\begin{cases} m^{(1)}_{d|n} = m^{(1)}_{d|n=1}\, n \\ m^{(2)}_{d|n} = m^{(2)}_{d|n=1}\, n^{1+\alpha} \end{cases} \tag{15}$$

$$\begin{cases} m^{(1)}_{x|n} = m^{(1)}_{x|n=1}\, n \\ m^{(2)}_{x|n} = m^{(2)}_{x|n=1}\, n^{1+\beta} \end{cases} \tag{16}$$

The two parameters α and β can easily be estimated by means of the following least squares fitting based upon the second order moment. In the case of α this gives :

$$\min_{\alpha} \sum_{n=1}^{N} \left[m^{(2)}_{d|n} - m^{(2)}_{d|n=1} n^{1+\alpha} \right]^2 \tag{17}$$

with N a sufficiently large integer (in this paper N was taken as 24), and where $m^{(2)}_{d|n}$ is defined as:

$$m^{(2)}_{d|n} = \frac{1}{n_w} \sum_{i=1}^{n_w} d_i^2 \tag{18}$$

n_w being the number of non-overlapping windows of length n in the record and d_i the number of rainy spells in the generic window.

Similarly, β is estimated by minimizing

$$\min_{\beta} \sum_{n=1}^{N} [m^{(2)}_{x|n} - m^{(2)}_{x|n=1} n^{1+\beta}]^2 \tag{19}$$

with:

$$m^{(2)}_{z|n} = \frac{1}{n_w} \sum_{i=1}^{n_w} x_i^2 \tag{20}$$

x_i being the rainfall total in a generic non-overlapping window of length n.

ESTIMATION OF THE PÒLYA PROCESS PARAMETERS R_n AND S_n BY THE METHOD OF MOMENTS.

The Pòlya process parameters r_n and s_n can be performed by matching moment estimates given by Eqn. (15), to the expected values of the number d of rainy time intervals in a window of length n, and of its square:

$$\begin{cases} E\{d \mid n\} = \dfrac{ns_n}{(r_n + s_n)} \\ E\{d^2 \mid n\} = \dfrac{ns_n}{(r_n + s_n)} \left[\dfrac{r_n + n(s_n + 1)}{r_n + s_n + 1} \right] \end{cases} \tag{21}$$

to give:

$$\begin{cases} m^{(1)}_{d|n=1} n = \dfrac{ns_n}{(r_n + s_n)} \\ m^{(2)}_{d|n=1} n^{1+\alpha} = \dfrac{ns_n}{(r_n + s_n)} \left[\dfrac{r_n + n(s_n + 1)}{r_n + s_n + 1} \right] \end{cases} \tag{22}$$

The values of the two parameters for $n > 1$ are then obtained as:

$$\begin{cases} r_n = \dfrac{(1 - m^{(1)}_{d|n=1})(n^\alpha - n)}{m^{(1)}_{d|n=1}(n-1) - (n^\alpha - 1)} \\ s_n = \dfrac{m^{(1)}_{d|n=1})(n^\alpha - n)}{m^{(1)}_{d|n=1}(n-1) - (n^\alpha - 1)} \end{cases} \tag{23}$$

For the case of $n = 1$, Eqn. (23) are undetermined and so de l'Hôpital's rule is used to take the limit for $n \to 1$ to give:

$$\begin{cases} r_1 = \lim\limits_{n \to 1} r_n = \dfrac{(1 - m^{(1)}_{d|n=1})(\alpha - 1)}{m^{(1)}_{d|n=1} - \alpha} \\ s_1 = \lim\limits_{n \to 1} s_n = \dfrac{m^{(1)}_{d|n=1}(\alpha - 1)}{m^{(1)}_{d|n=1} - \alpha} \end{cases} \tag{24}$$

It is worthwhile noticing that the ratio between the two parameters:

$$\frac{r_n}{s_n} = \frac{1 - m^{(1)}_{d|n=1}}{m^{(1)}_{d|n=1}} \tag{25}$$

is constant and therefore independent of n.

ESTIMATION OF PARAMETERS $a_{d,n}$ AND $b_{d,n}$ BY THE METHOD OF MOMENTS

To estimate the parameters of the probability distribution of the rainfall totals conditional upon the number of rainy spells, it is necessary to reduce their number; therefore, two simplifying hypotheses have to be formulated. The first is that the parameter $b_{d,n}$ is constant for a given window of length n, so that only the variability of the expected value of θ_n is taken into account. The second is that there is a relationship between the expected values of the parameters θ_n and d. After several trials, a suitable form for this relationship has been found to be:

$$E\{\theta_n \mid d, n\} = \frac{1}{a_{d,n}} = \frac{1}{a_n d^\gamma} \tag{26}$$

This hypothesis corresponds to the following growth of the first order sample moment of x with d:

$$m^{(1)}_{x|d,n} = m^{(1)}_{x|d=1,n} d^{\gamma+1} \tag{27}$$

In practice it allows for a linear growth of the first order moment of the rainfall totals with $ln\ d$, similarly to what was observed as a function of n, but at larger rate, since only the rainy spells are accounted for in this case, after elimination of the zero events.

On the basis of this assumption, a value for γ can be derived by means of the following weighted least squares fitting, where the weights will account for the different

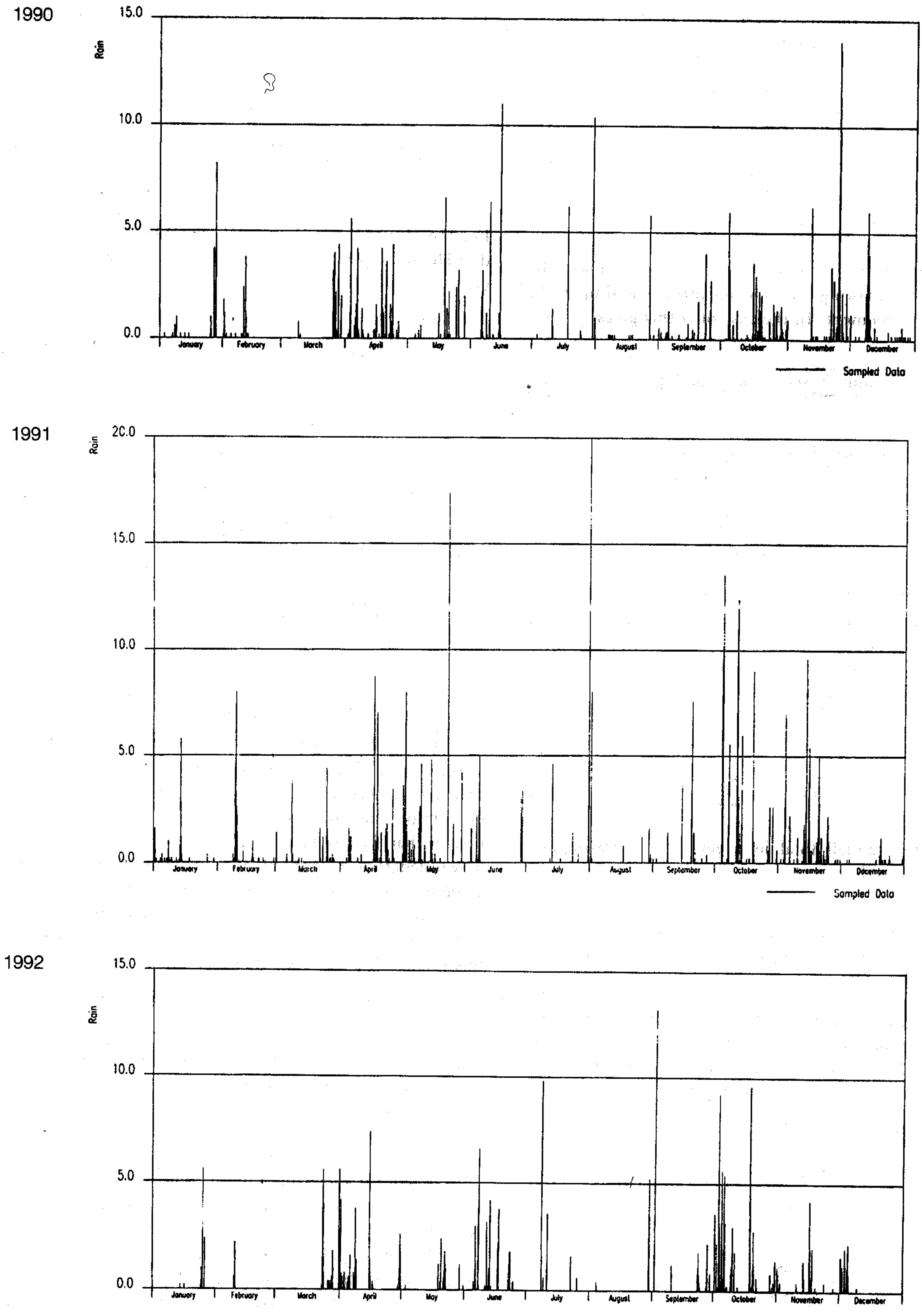

Fig. 7 *Hourly precipitation recorded at Porretta (1990–1991–1992).*

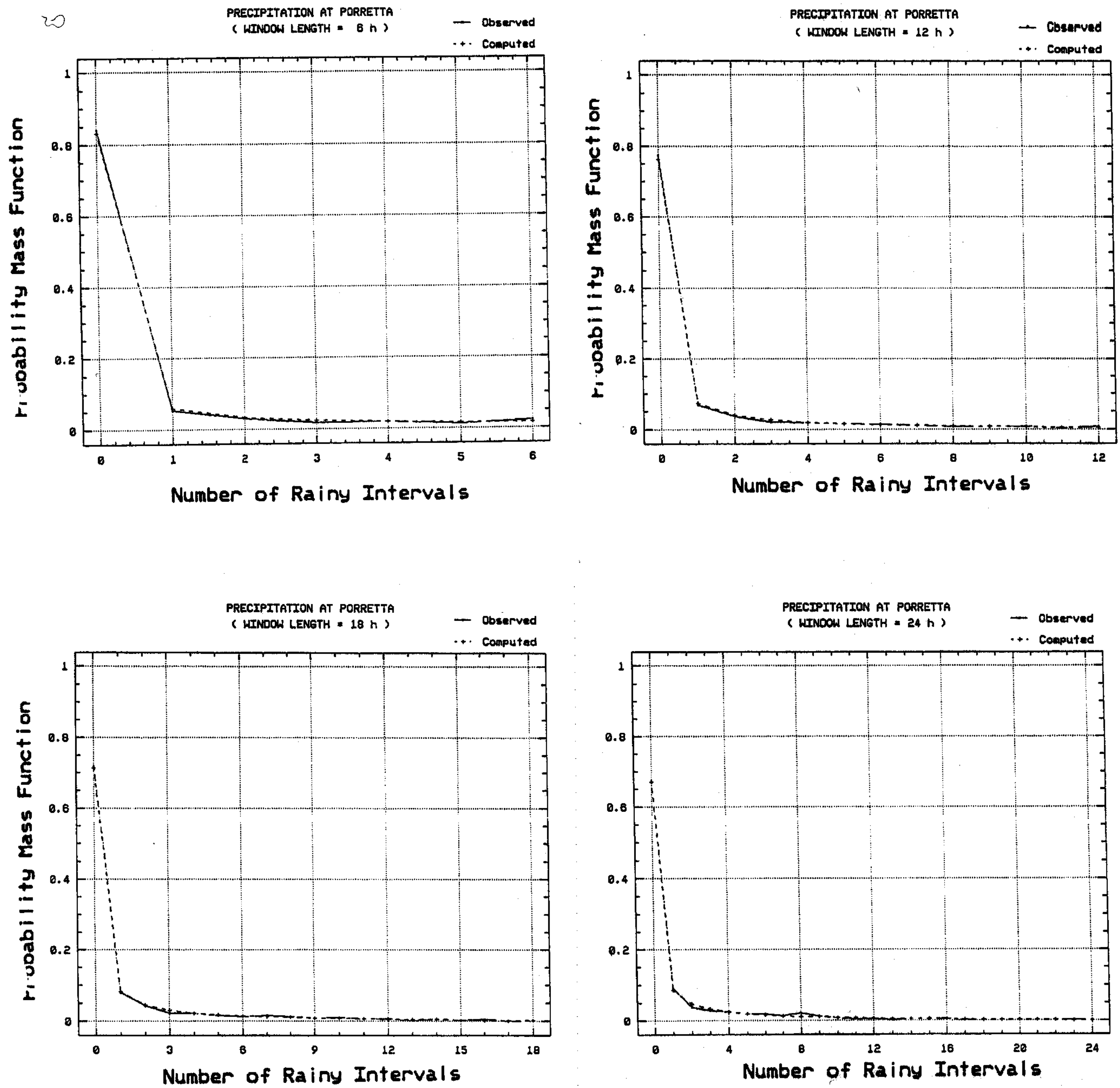

Fig. 8 *Comparison between the observed frequencies (solid line) and the probability of rainfall occurrences estimated using equation (3) (dashed line), for n = 6, 12, 18, 24.*

levels of probability $Pr\{d \mid n, r_n, s_n\}$ as well as for the different numbers of samples ($1/n$ is proportional to the number of non-overlapping windows in the record):

$$\min_{\gamma} \sum_{n=1}^{N} \frac{1}{n} \sum_{d=1}^{n} Pr\{d \mid n, r_n, s_n\}(m^{(1)}_{x|d,n} - m^{(1)}_{x|d=1,n})^2 \tag{28}$$

Given γ, the estimation of the parameters a_n and b_n can be performed by matching the sample moment estimates given by Eqn. (16), to the expected value of the total quantity of rainfall x in a window of length n, and of its square:

$$\begin{cases} E\{x \mid n\} = \dfrac{a_n E\{d^{\gamma+1} \mid n\}}{b_n - 1} \\ E\{x^2 \mid n\} = \dfrac{a_n^2 [E\{d^{2\gamma+2} \mid n\} + E\{d^{2\gamma+1} \mid n\}]}{(b_n - 1)(b_n - 2)} \end{cases} \tag{29}$$

which gives:

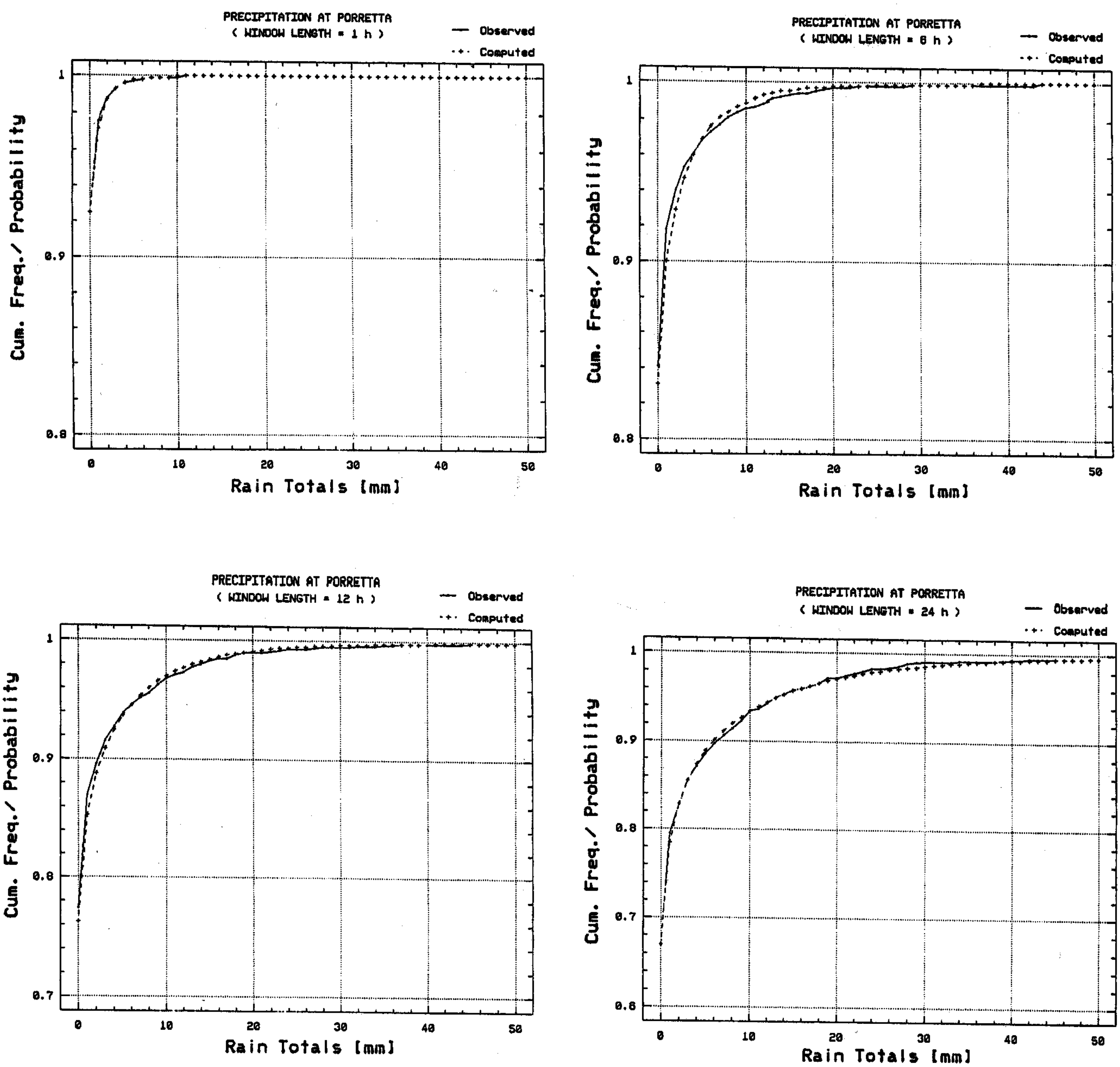

Fig. 9 *Comparison between the observed cumulated frequencies of the rainfall totals (solid line) and the probability distribution given by equation (13), for n = 1, 6, 12, 24.*

$$\begin{cases} a_n = \dfrac{m^{(1)}_{x|n}}{c_n E\{d^{\gamma+1} \mid n\}} \\ b_n = 1 + \dfrac{1}{c_n} \end{cases} \tag{30}$$

with c_n defined as:

$$c_n = 1 - \frac{(m^{(1)}_{x|n})^2}{m^{(2)}_{x|n}} \frac{E\{d^{2\gamma+1} \mid n\} + E\{d^{2\gamma+2} \mid n\}}{(E\{d^{\gamma+1} \mid n\})^2} \tag{31}$$

Finally, substituting for $m^{(1)}_{x|n}$ and for $m^{(2)}_{x|n}$ from eqn. 16, and bearing in mind that $E\{d^k|n\}$ can be estimated as

$$E\{d^k \mid n\} = \sum_{d=1}^{n} \Pr\{d \mid n, r_n, s_n\} d^k ,$$

where $Pr\{d|n,r_n,s_n\}d^k$ is given by Eqn. (3) as a function of r_n and s_n, one obtains:

$$\begin{cases} a_n = \dfrac{n m^{(1)}_{x|n=1}}{c_n \sum_{d=1}^{n} \Pr\{d \mid n, r_n, s_n\} d^{\gamma+1}} \\ b_n = 1 + \dfrac{1}{c_n} \end{cases} \tag{32}$$

where c_n is estimated as follows:

$$c_n = 1 - \frac{(m^{(1)}_{x|n=1})^2 \sum_{d=1}^{n} Pr\{d \mid n, r_n, s_n\} d^{2\gamma+1}(d+1)}{m^{(2)}_{x|n=1}\left(\sum_{d=1}^{n} Pr\{d \mid n, r_n, s_n\} d^{\gamma+1}\right)^2} n^{1-\beta} \quad (33)$$

Finally, under the stated assumptions, $a_{d,n}$ and $b_{d,n}$ can then be derived as:

$$\begin{cases} a_{d,n} = a_n d^{\gamma} \\ b_{d,n} = b_n \end{cases} \quad (34)$$

Summary of parameter estimation

The results can be now summarized. On the basis of the model hypotheses and of the scaling properties, all the parameters relevant to the different time windows are in fact be expressed in terms of the five parameters r_n, s_n, a_n, b_n and g which can be explicitly derived from three moment estimates: $m^{(1)}_{d|n=1}$, $m^{(1)}_{x|n=1}$, $m^{(2)}_{x|n=1}$, relevant to windows of duration $1\Delta t$, plus the three parameters: α, β, γ, which reflect the scaling properties of the rainfall records.

Application of the proposed model

The model assumptions have been tested on several rainfall records, but, for the sake of clarity, only data from a single recording raingauge, the station of Porretta, located in Central Italy on the Appennines, are used for illustrating the properties of the proposed model. A three year (1/1/1990–31/12/1992) long record of hourly sampled rainfall (see Figs 7, 8, 9) was used to estimate: $m^{(1)}_{d|n=1}$, $m^{(1)}_{x|n=1}$, $m^{(2)}_{x|n=1}$, α, β, γ. Using these parameters, the values for r_n, s_n, a_n and b_n were computed and used to evaluate the probability of the rainfall totals in increasing size windows.

Figures 10 to 13 compare the probability of rainfall occurrences estimated using Eqn. (3) and the observed frequencies for n = 6, 12, 18, 24 hours, while Fig. 9 shows a comparison between the observed cumulated frequencies of the rainfall totals and the probability distribution of eqn. 13 for n = 1, 3, 6, 12, 18, 24 hours.

The results show, at all time scales, an extremely high degree of agreement between model and observation.

Conclusions

The analytical model presented in this paper, which takes advantage of the time scaling properties of rainfall to reduce the number of model parameters to six, has reproduced the observed rainfall frequencies at all time scales ranging from one hour to twenty-four hours. Its probabilistic formulation allows for the development of a conditional probability model at the different time scales; this can be used either for the calculation of the expected conditional probabilities and conditional moments or for the generation of traces of rain, conditional on the latest observations, within the frame of real-time forecasting applications. At present research is being carried out in order to develop the at-site conditional generator as well as its multi-site version.

Finally, given that the proposed model performs particularly well in the tail of the distribution, research has being conducted in order derive the probability distribution of yearly maxima of precipitation from one hour to twenty-four hours, starting from the assumption that the probability distribution of rainfall at the different time scales is known and given by the model: interesting early results were found and will be reported after additional extended testing.

References

Berger J.O.,1985. *Statistical Decision Theory*, Springer-Verlag, New York, (Second Edition).

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., 1980. *Point Processes*, Chapman and Hall, London.

Daboni L., Wedlin A., 1982. *Statistica*, UTET (in Italian).

de Lima M.I.P. and Bogardi J.J., 1995. Multifractals in hydrological studies: the analysis of rainfall time series. Proc. UNESCO Conference on '*Statistical and Bayesian methods in hydrological sciences*'. In press.

Feller W., 1970. *An Introduction to Probability Theory and Its Applications*, John Wiley & Sons. New York.

Foufoula-Georgiou E. and Georgakakos K.P., 1991, Hydrologic Advances in Space-Time Precipitation Modeling and Forecasting. In D.S. Bowles and P.E.O'Connell (eds.), *Recent Advances in the Modeling of Hydrologic Systems*, 47–65.

Foufoula-Georgiou E. and Lettenmayer D.P., 1987, A Markov Renewal Model for Rainfall Occurrences. *Wat. Resour. Res.* 23(5), 875–884.

Gupta V.K. and Duckstein L., 1975. A Stochastic Analysis of Extreme Droughts, *Wat. Resour. Res.*, **12**(2), 221–228.

Hey J.D., 1991. *Data in Doubt*, Blackwell Basil.

Kavvas M.L. and Delleur J.W., 1991. A Stochastic Cluster Model for Daily Rainfall Sequences. *Wat. Resour. Res.* 17(4), 1151–1160.

Lindley D.V., 1985. *Making Decisions*, Wiley and Sons Inc., New York.

Lovejoy S. and Mandelbrot B.B., 1985. Fractal properties of rain, and a fractal model. *Tellus*, Vol.37A, 209–232.

Lovejoy S. and Schertzer D., 1993. Multifractals and rain. In Kunzewicz K.W. (ed.) *New Uncertainty Concepts in Hydrology and Water Resources*. Unesco series in Water Sciences. Cambridge University Press.

Mielke P.W.Jr. and Johnson E.S., 1974. Some Generalized Beta Distributions of the Second Kind Having Desirable Application Features in Hydrology and Meteorology, *Wat. Resour. Res.* **10**(2), 223–236.

Ord J.K., 1972. *Families of Frequency Distributions*, Griffin, London.

Richardson C.W., 1981. Stochastic Models of Daily Precipitation, Temperature and Solar Radiation, *Wat. Resour. Res.* 17(1), 182–190.

Rodriguez-Iturbe I., Gupta V.K., Waymire E., 1984. Scale Considerations in the Modeling of Temporal Rainfall, *Wat. Resour. Res.* **20**(11), 1611–1619.

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

Rodriguez-Iturbe I., Febres de Power B., Valdes J.B., 1987b. Rectangular Pulses Point Process models for Rainfall: Analysis of Empirical Data, *J. Geophys. Res.* **92**(D8), 9645–9656.

Rodriguez-Iturbe I., Cox D.R., Isham V., 1987c. A Point Process Model for Rainfall: further developments, *Proc. R. Soc. Lond.*, A.

Roldan J. and Woolhiser D.A., 1982. Stochastic Daily Precipitation Models, 2, A Comparison of Distribution Amounts. *Wat. Resour. Res.* **25**(7), 1667–1675.

Skees P.M. and Shenton L.R., 1974. Comments on the Statistical Distribution of Rainfall per Period Under Various Transformations, Proc. of the Symp. on Statistical Hydrology, Tucson, AZ, U.S. Dept. of Agr., Misc. Publ. N. 1275, 172–196

Smith J.A. and Karr A.F., 1983. A point Process Model of Summer Season Rainfall Occurrence. *Wat. Resour. Res.* **19**(1), 95–103.

Smith J.A., 1987. Statistical Modelling of Daily Rainfall Occurrences, *Wat. Resour. Res.* **23**(5), 855–893.

Thompson C.S., 1984. Homogeneity Analysis of Rainfall Series: an Application of the Use of a Realistic Rainfall Model. *J. Climatology*, **4**, 609–619.

Todorovic P. and Yevjevic V., 1969. *Stochastic Processes of Precipitation*, Colorado State University, Hydro Paper 36, 1–61.

Todorovic P. and Woolhiser D.A., 1971. *Stochastic Model of Daily Rainfall*, Proc. of the USDA-IASPS Symposium on Statist. Hydrology, Tucson, AZ..

Waymire E. and Gupta V.K., 1981a. The Mathematical Structure of Rainfall Representations, 1, A Review of Stochastic Rainfall Models, *Wat. Resour. Res.* **17**(5), 1261–1272.

Waymire E. and Gupta V.K., 1981b. The Mathematical Structure of Rainfall Representations, 2, A Review of the Theory of Point Processes, *Wat. Resour. Res.* **17**(5), 1273–1285.

Waymire E. and Gupta V.K., 1981c. The Mathematical Structure of Rainfall Representations, 3, Some Applications of the Point Process Theory to Rainfall Processes, *Wat. Resour. Res.* 17(5), 1287–1294

Woolhiser D.A. and Pegram G.G.S., 1979. Maximum Likelihood Estimation of Fourier Coefficients to Describe Seasonal Variations of Parameters in Stochastic Daily Precipitation Models, *J. Appl. Meteorol.* 8(1), 34–42.