Hydrology and Earth System Sciences, 2(1), *19–30* (1998) © EGS

Hydrology & Earth System Sciences

Evaluation of a scaling cascade model for temporal rainfall disaggregation

Jonas Olsson

Department of Water Resources Engineering, Lund University, Box 118, SE–22100 Lund, Sweden.
Currently at: Department of Civil Engineering, Kyushu University, 6–10–1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan.

Abstract

The possibility of modelling the temporal structure of rainfall in southern Sweden by a simple cascade model is tested. The cascade model is based on exact conservation of rainfall volume and has a branching number of 2. The weights associated with one branching are 1 and 0 with probability $P(1/0)$, 0 and 1 with $P(0/1)$, and $W_{x/x}$ and $1 - W_{x/x}$, $0 < W_{x/x} < 1$, with $P(x/x)$, where $W_{x/x}$ is associated with a theoretical probability distribution. Furthermore, the probabilities P are assumed to depend on two characteristics of the rainy time period (wet box) to be branched: rainfall volume and position in the rainfall sequence. In the first step, analyses of 2 years of 8-min data indicates that the model is applicable between approximately 1 hour and 1 week with approximately uniformly distributed $W_{x/x}$ values. The probabilities P show a clear dependence on the box characteristics and a slight seasonal nonstationarity. In the second step, the model is used to disaggregate the time series from 17- to 1-hour resolution. The model-generated data reproduce well the ratio between rainy and nonrainy periods and the distribution of individual volumes. Event volumes, event durations, and dry period lengths are fairly well reproduced, but somewhat underestimated, as was the autocorrelation. From analyses of power spectrum and statistical moments the model preserves the scaling behaviour of the data. The results demonstrate the potential of scaling-based approaches in hydrological applications involving rainfall disaggregation.

Introduction

Characterising the temporal rainfall process by its scaling, i.e., scale-invariant, behaviour is an approach that receives a constantly increasing attention. Scaling refers to a symmetry across scales, i.e., absence of characteristic scales, and recently a large number of analyses have supported the presence of scaling properties of temporal rainfall (e.g. Hubert and Carbonnel, 1989; Olsson *et al.*, 1992, 1993; Hubert *et al.*, 1993; Tessier *et al.*, 1993, 1996; Olsson, 1995; Burlando and Rosso, 1996; Cârsteanu and Foufoula-Georgiou, 1996; Harris *et al.*, 1996; Onof *et al.*, 1996; Svensson *et al.*, 1996; Menabde *et al.*, 1997). Scaling implies that statistical properties of the process observed at different scales, i.e., resolutions, are governed by the same relationship, and an obvious application of rainfall scaling is for disaggregation or downscaling. In principle, parameters of the scaling relationship could be obtained from larger scales and then used to estimate the process properties at smaller scales.

In this study, the scaling behaviour of (temporal) rainfall is modelled by a cascade process, an approach argued for by, e.g. Schertzer and Lovejoy (1987), Gupta and Waymire (1990, 1993), and Lovejoy and Schertzer (1990). This conceptual model was originally used in statistical turbulence (e.g. Yaglom, 1966; Mandelbrot, 1974). In principle, a large-scale structure associated with an originally uniformly distributed *mass* or *volume* is decomposed into successively smaller structures to which the mass is transferred. As a consequence, the mass becomes concentrated in smaller and smaller units of the available space (e.g. Schertzer and Lovejoy, 1987). For rainfall, cascade models have mainly been used to describe the spatial structure (e.g. Lovejoy and Schertzer, 1990; Gupta and Waymire, 1993; Tessier *et al.*, 1993; Over and Gupta, 1994), but also the temporal structure has been explored (e.g. Hubert *et al.*, 1993; Olsson, 1995; Menabde *et al.*, 1997). To the author's knowledge, cascade models have up to now not been used for rainfall disaggregation.

This study is focused on temporal disaggregation of approximately daily values into finer time steps. Some work has already been devoted to this, notably by Hershenhorn and Woolhiser (1987). They developed a model for simulating the number, depths, durations, and starting times of events during a day based on the total rainfall during the day and the preceding and following days. The method was somewhat modified by Econopouly *et al.* (1990), who also investigated the spatial transferability of the model parameters. Bo *et al.* (1994) used the Bartlett-Lewis rectangular pulses, BLRP, model developed by Rodriguez-Iturbe *et al.* (1987, 1988) to, among other things, disaggregate daily rainfall into hourly values. Bo *et al.* (1994) argued that the successful result was due to a scaling (power-law) behaviour of the power spectrum. Another approach was proposed by Glasbey *et al.* (1995), who modified the BLRP model to simulate hourly data consistent with observed daily totals. For spatial rainfall, scaling-based disaggregation was performed by Perica and Foufoula-Georgiou (1996). They developed a disaggregation model based on the (empirically observed) scaling of probability distributions of rainfall fluctuations and correlation between the scaling parameters and the convective available potential energy (CAPE).

The aim of the present study was twofold. Firstly, to evaluate the applicable scale range of a cascade model designed to represent the temporal structure of rainfall in southern Sweden. Secondly, to test the possibility of using the model for temporal rainfall disaggregation within this range.

Rainfall data

During 1979–81, a detailed observation programme of short-term rainfall properties was performed in the city of Lund, southern Sweden. The rainfall intensity was measured with a time resolution of 1 min by small tipping-bucket gauges with an intensity resolution of 0.033 mm/min. The longest continuous measurement period was 2.5 years (January 1979 to July 1981), and the most complete time series was used in the present study. However, winter periods when the occurrence of snow may have introduced errors in the measurements were omitted from the series in the present analysis and, because of this, the length of the series analysed is 2 years. During this time there were no missing values. For further details about the database and observation area see Niemczynowicz (1986a, b). It should be remarked that the tipping-bucket gauge was used somewhat differently from common practice. Instead of recording the point of time of each 'tip', the number of 'tips' that occurred during each minute was recorded. By this measurement strategy, constant rainfall intensities lower than the intensity resolution of the gauge will in the series be represented by '1-tip' (0.033 mm/min) rainfall registrations separated by one or more spurious zero-registrations. To overcome or at least reduce the influence of this problem, the original 1-min registrations were aggregated into 8-min values. These aggregated values will, in general, comprise both the '1-tip' registrations and the erroneous zero-registrations, i.e., low-intensity rainfall will be evened out and be represented more accurately.

The final series thus consists of 2 years of 8-min rainfall observations, i.e. 131 072 or 2^{17} values. Due to the cascade structure employed in the present study, the time scales are expressed in 8-min periods multiplied by powers of 2. This is the reason for the values of the main time scale limits employed in the description of the results of this study below: 1 hour ($\approx 2^3$ 8-min periods), 17 hours ($\approx 2^7$ 8-min periods), and 5.7 days ($\approx 2^{10}$ 8-min periods). General scaling and multifractal properties of the series were investigated by Olsson (1995).

Cascade model

Fig. 1 shows how a cascade scheme is used in this study to represent a time series of rainfall volumes at successively doubled temporal resolution, and will be used for defining some of the notions introduced below. *Cascade level*, l, refers to the time series at a certain resolution. The transition from one cascade level to the nearest higher level, corresponding to a doubling of the resolution, will be called a *modulation*. A time interval T at an arbitrary cascade level (i.e. time scale) will be termed a *box*, and its associated rainfall volume denoted V. If $V = 0$ the box is said to be *dry*, and *wet* if $V > 0$. A break-up of a wet box T into two equally sized sub-boxes T_1 and T_2, with volumes V_1 and V_2, is denoted a *branching* (i.e. the *branching number* is 2 in this study). In one branching, the total volume V is redistributed according to the two multiplicative weights W_1 ($0 \le W_1 \le 1$) and W_2 ($0 \le W_2 \le 1$), $W_1 + W_2 = 1$, i.e., $V_1 = W_1 \cdot V$ and $V_2 = W_2 \cdot V$. Concerning this redistribution, three principal possibilities exist on the condition that T_1 and T_2 are short enough to have a positive probability of zero rainfall. The first is that all of V occurred during T_1, i.e. $W_1 = 1$ and $W_2 = 0$. This possibility is in the following referred to as *1/0-division*. The

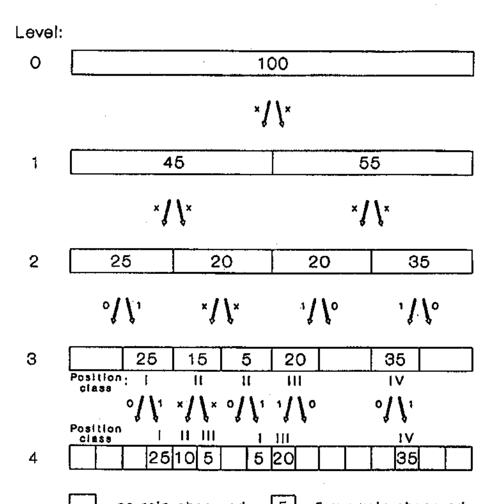

Fig. 1. *Schematic of a 1-dimensional cascade process, as defined in this study, representing a time series of rainfall volumes at different resolutions and visualizations of some of the notions introduced in the text.*

second is the reverse, i.e., $W_1 = 0$ and $W_2 = 1$ (*0/1-division*). The third possibility is that one non-zero part of V occurred during T_1 and the other non-zero part during T_2, i.e. $0 < W_1 < 1$ and $0 < W_2 < 1$, and this will be referred to as *x/x-division* (see Fig. 1).

In the modulation between two cascade levels, $P(1/0)$, $P(0/1)$, and $P(x/x)$ are defined, respectively, as the probability of having an 1/0-, 0/1-, and x/x-division, respectively, in the branching of a wet box. Naturally $P(1/0) + P(0/1) + P(x/x) = 1$, and when referring to these probabilities in general they will be jointly denoted P. In case of x/x-divisions, the corresponding weights $W_{x/x}$ are assumed to be associated with a theoretical probability distribution in the following denoted *$W_{x/x}$-distribution*. It is further assumed that the probabilities P and the $W_{x/x}$-distribution are scale-invariant, i.e. approximately constant over a range of cascade levels (i.e. time scales).

The model is thus a random cascade with the *generator* (e.g. Gupta and Waymire, 1993)

$$W_1, W_2 = \begin{cases} 1 \text{ and } 0 & \text{with probability } P(1/0) \\ 0 \text{ and } 1 & \text{with probability } P(0/1) \\ W_{x/x} \text{ and } 1-W_{x/x},\ 0 < W_{x/x} < 1 & \\ & \text{with probability } P(x/x) \end{cases} \quad (1)$$

where $W_{x/x}$ is associated with a theoretical probability distribution.

A main difference between the present model and most previously proposed cascade models is the exact conservation of rainfall volume between successive cascade levels. This has been termed a *microcanonical* property as opposed to *canonical* cascades where the volume is only on average conserved (e.g. Mandelbrot, 1974; Schertzer and Lovejoy, 1987). Another consequence of the present cascade structure is that the weights are not mutually independent, as is typically assumed, but W_1 and W_2 have a correlation of −1. However, the pairs of weights associated with different branchings are assumed to be mutually independent. It should be mentioned that a similar microcanonical approach was recently used by Cârsteanu and Foufoula-Georgiou (1996) to investigate the weights structure in temporal rainfall. They found a correlation of −0.2 between adjacent pairs of weights in individual rainfall events. The present model thus differs in some fundamental respects from the majority of random cascade models for rainfall developed and tested to date; the results of the present study cannot be compared directly with the results of previous analyses. However, this model has been preferred as it has the significant advantage of allowing an 'exact' reconstruction of the cascade generator from the time series (Cârsteanu and Foufoula-Georgiou, 1996). Because of this, the generator properties may be evaluated in a direct fashion rather than indirectly via derived relationships.

In this study, a further modification is introduced as compared to existing random cascade models. Although existing models preserve the empirically determined scaling behaviour of rainfall, they do not seem, entirely and convincingly, to reconcile on one hand the presence of zero-values and on the other hand the clustering of rainfall into continuous events or areas. A multifractal field produced by the log-Lévy model argued for by Lovejoy and Schertzer (1990) does not contain any zero-values unless some thresholding is employed. The model by Over and Gupta (1994) generates zero-values, but from a visual inspection of the simulated fields (page 1529) the rainfall areas appear significantly less connected than in the observed rainfall fields. In these models, it is assumed that the same generator parameters apply to all (temporal or spatial) units containing rain (wet boxes) in the modulation between two successive cascade levels.

One way to improve the model performance in the above respect could be to use a random cascade model with a general generator structure (Eqn. 1), but with generator parameters depending on characteristics of the wet box to be branched. In the case of temporal rainfall we propose (1) the rainfall volume of the box and (2) the position of the box in the rainfall sequence to be relevant characteristics. Concerning the rainfall volume, it may be assumed that, for a period receiving a large amount of rainfall, the probability of rain during both the first and the last half of the period, $P(x/x)$, is higher than for a period receiving a small amount of rainfall. This assumption requires that large rainfall volumes be generally produced by long durations rather than high intensities. This is likely to be the case in the present region where fairly evenly distributed rainfall associated with frontal passages constitutes most of the temporal rainfall occurrence (Ångström, 1974). Concerning the position of the rainy period in the rainfall sequence, it may be assumed that for a period surrounded by rainy periods, the probability of rain during both the first and the last half of the period, $P(x/x)$, is higher than for a period surrounded by dry periods. This assumption is motivated by the clustering of rainfall that has been observed at different scales (e.g. Austin and Houze, 1972). It should be mentioned that differences in cascade parameters might also arise due to discretization effects in the data. For example, since the duration of a rainfall event generally decreases with increasing resolution, as the starting and stopping time approach their true values, $P(0/1)$ ($P(1/0)$) for a period in the beginning (end) of a sequence may be expected to be significantly higher than $P(1/0)$ ($P(0/1)$).

Evaluation

The applicability of the model was evaluated by reconstructing the cascade generator from the time series. The range of cascade levels possible to study was 0–17, where $l = 0$ corresponds to 2 years (the length of the series) and $l = 17$ to 8 min (the temporal resolution). The reconstruction was performed by aggregating the time series values two by two starting from $l = 17$ to successively lower cascade levels. To test the hypothesis of a dependence of the generator parameters on box characteristics, the wet boxes at each cascade level were divided into classes defined both by the position in the rainfall sequence and the observed rainfall volume. For the position, four classes were defined (see Fig. 1): (I) period in the beginning of a sequence, i.e., preceded by a dry period and succeeded by a rainy period (this class will be referred to as *starting boxes*), (II) period within a sequence, i.e., preceded and succeeded by rainy periods (*enclosed boxes*), (III) period in the end of a sequence, i.e., preceded by a rainy period and succeeded by a dry period (*ending boxes*), and (IV) isolated period, i.e., preceded and succeeded by dry periods (*isolated boxes*). For the volume, each position class was further sub-divided into two volume classes: above and below the mean volume of the position class.

Consider the aggregation from level j to level $i = j - 1$. For each wet box at level i, the class was determined and the weights W_1 and W_2 were extracted by dividing the total volume by both its corresponding volumes at level j. For a certain class, denote the total number of wet boxes at level i by N_i, and the number of wet boxes associated with 1/0-, 0/1- and x/x-divisions by $N_i(1/0)$, $N_i(0/1)$, and $N_i(x/x)$. The empirical estimates of $P(1/0)$, $P(0/1)$ and $P(x/x)$ associated with the modulation from level i to level j for the class may then be obtained as $f(1/0) = N_i(1/0)/N_i$, $f(0/1) = N_i(0/1)/N_i$, and $f(x/x) = N_i(x/x)/N_i$. The corresponding $W_{x/x}$-distribution may be estimated from a histogram of the weights $W_{x/x}$. When all cascade levels have been investigated, i.e. the 8-min values have been aggregated to 2 years, the $W_{x/x}$-histograms at different levels may be compared. Also, f may be plotted as a function of l. This way the scale invariance of the cascade generator properties may be directly evaluated, and the range where the properties are approximately constant will be termed the *scaling regime*. The mean value of f over a range of cascade levels i-k will be denoted $f_{i\text{-}k}$.

Figure 2 shows $W_{x/x}$-histograms associated with modulations between the time scales 8 min and 22.8 days (i.e., cascade levels 17 and 5) for enclosed boxes below mean volume, obtained from the time series. Histograms corresponding to lower cascade levels are not shown since the number of $W_{x/x}$ values, n, was too small to obtain reliable results. In Fig. 2, n ranges from 1142 for 8–16 min to 32 for 11.4–22.8 days, and the number of histogram intervals, k, was determined by $k = 1 + 3.3\log_{10}n$ (Haan, 1977). The histograms are naturally symmetrical since every value of $W_{x/x}$ between 0 and 0.5 corresponds to a value $1\text{-}W_{x/x}$ between 0.5 and 1. For the three modulations at the smallest time scales, 8 min—1 hour (Figs. 2a–c), the histograms exhibit a pronounced peak around 0.5 and the shape suggests an approximately normal distribution. For the rest of the histograms, however, the more or less flat shape suggests a uniform distribution (Figs. 2d–l). Histograms corresponding to enclosed boxes above mean volume exhibit a similar appearance.

The time scale where the distribution changes from approximately uniform to approximately normal corresponds to the duration of rainfall events in the series. Olsson *et al.* (1992) found that, for the present data, the mean duration is around 40 min. When studying the successive modulations between levels corresponding to time scales smaller than 40 min, the results to an increasing degree reflect the internal temporal structure of rainfall events. In the present region, the rainfall during an event is generally rather evenly distributed. Thus, for time periods shorter than the mean duration, it may be expected that the rainfalls during the first and the last half of the period often are approximately the same, i.e., both $W_{x/x}$ values are close to 0.5. For longer time periods, however, the first and the last half may well contain different events associated with different total volumes, e.g., due to different durations. Thus, it is reasonable that the $W_{x/x}$ values at larger time scales do not cluster around 0.5 but exhibit a larger spread.

For position classes other than enclosed boxes, the histogram shapes were similar although a uniform distribution was often suggested also for the smallest scales.

Figure 3 shows the empirical probabilities $f(1/0)$, $f(0/1)$, and $f(x/x)$ as a function of modulation time scales between 8 min and 11.4 days for all four position classes below mean volume, obtained from the time series. At the largest scale, only the class of enclosed boxes has f values (Fig. 3b). This is because in the present data zero-values do not exist at time scales larger than 2.8 days. Thus, for the present data, the modulation between 5.7 and 2.8 days constitutes an upper time scale limit for the proposed modelling approach. Between 1–2.1 hours and 2.8–5.7 days, for all classes, the f values remain fairly constant. However, at modulations below 1–2.1 hours the f values deviate, particularly for enclosed and isolated boxes (Figs. 3b and d). In general, $f(x/x)$ increases whereas $f(1/0)$ and $f(0/1)$ decrease. Similarly as for the histograms, this behaviour may also be explained by considering the rainfall event duration. At time scales smaller than the mean duration, the wet boxes are to an increasing degree clustered into continuous sequences. Thus it is reasonable to expect the observed increase of $f(x/x)$.

The empirical probabilities $f(1/0)$, $f(0/1)$, and $f(x/x)$ generally differ distinctly within each class, save perhaps for enclosed boxes (Fig. 3). This justifies a generator with different probabilities for 1/0-, 0/1-, and x/x-divisions, respectively. It is also evident that the probabilities differ between the position classes. This is also true for the volume classes. For boxes above mean volume the principal variation of $f(1/0)$, $f(0/1)$, and $f(x/x)$ with scale is similar to Fig. 3. However, for each position class the values of the probabilities corresponding to volumes below and above the mean, respectively, differ distinctly (see Table 1 below for the probability values used in the disaggregation). This verifies the hypothesis of a dependence of the cascade generator parameters on box characteristics.

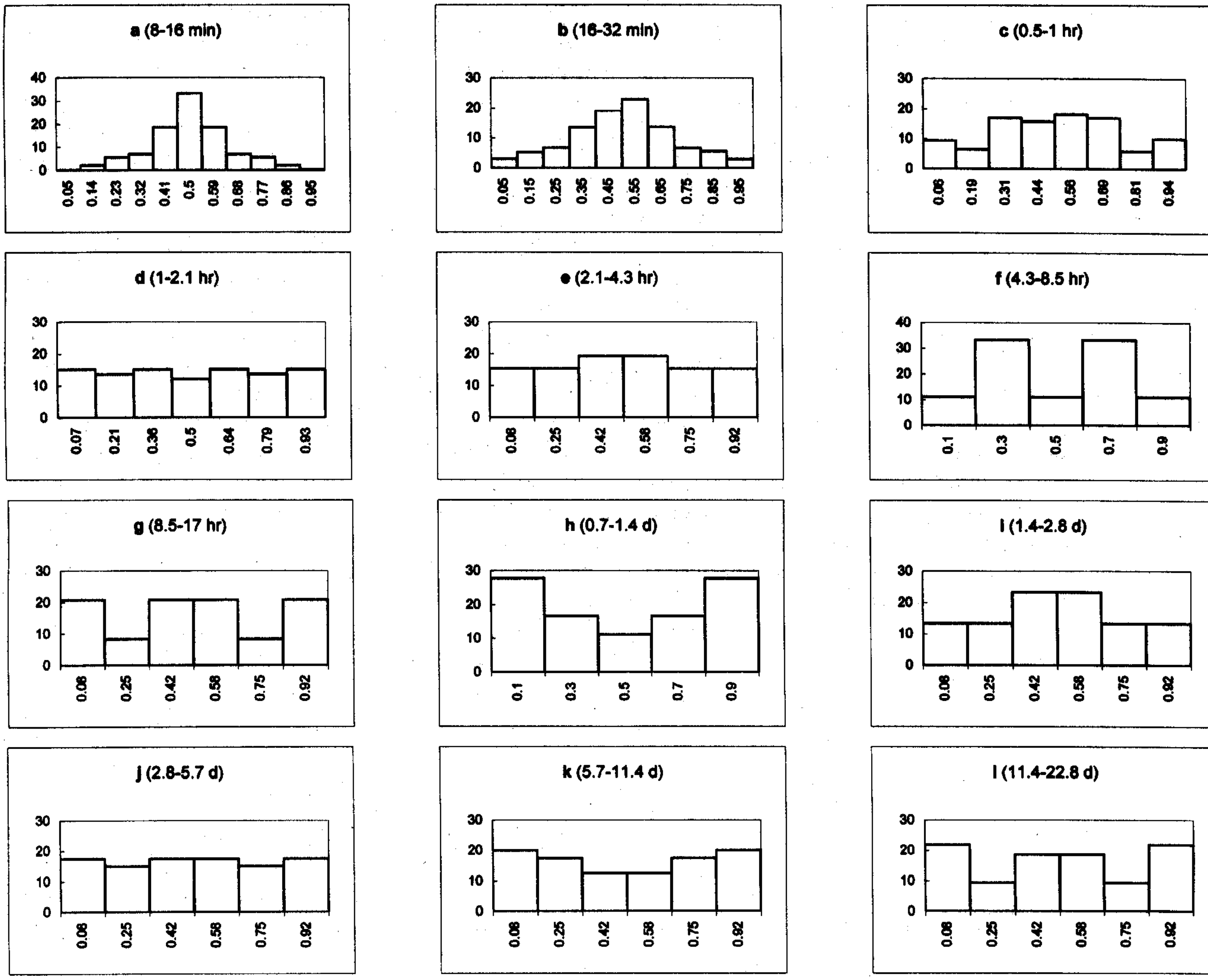

Fig. 2. *Histograms (in percentages) of the $W_{x/x}$ values for modulations ranging from time scales 8–16 min (a) to 11.4–22.8 days (l), for enclosed boxes below mean volume.*

Table 1. Values of the empirical probabilities f_{7-14} and the probabilities P used in the disaggregation, for the three types of division and eight classes.

Position	Volume	0/1-division f_{7-14}	0/1-division P	1/0-division f_{7-14}	1/0-division P	x/x-division f_{7-14}	x/x-division P
Starting	Below mean	0.61	0.61	0.23	0.21	0.17	0.18
	Above mean	0.31	0.29	0.13	0.10	0.56	0.61
Enclosed	Below mean	0.34	0.34	0.34	0.34	0.32	0.32
	Above mean	0.13	0.13	0.13	0.13	0.74	0.74
Ending	Below mean	0.19	0.21	0.61	0.61	0.20	0.18
	Above mean	0.07	0.10	0.28	0.29	0.65	0.61
Isolated	Below mean	0.44	0.44	0.45	0.44	0.11	0.12
	Above mean	0.25	0.24	0.23	0.24	0.52	0.52

In terms of constant $W_{x/x}$ distributions and f values, a reasonable scaling regime thus exists between 5.7 days (l = 7) and 1 hour (l = 14). These limits agree with limits found by analysing similar data using monofractal techniques (Olsson *et al.*, 1992, 1993), whereas analyses of power spectrum and statistical moments of the present series have indicated multifractal scaling from approximately the same upper limit down to the 8-min resolution (Olsson, 1995). As previously mentioned, however, the results of the present study are not directly comparable with most previous analyses due to the microcanonical framework employed.

To verify the applicability of the model with uniformly distributed $W_{x/x}$ values in the scaling regime, some statistical tests and calculations were carried out. To test the hypothesis of a uniform distribution, the probability-plot correlation-coefficient, PPCC, test proposed by Vogel and Kroll (1989) was employed. For the seven modulations between 5.7 days and 1 hour for all eight classes (56 histograms), the uniform PPCC was 0.987±0.009 (mean±std). At significance level 0.01, the hypothesis of a uniform distribution could not be rejected for any histogram. At significance level 0.1, only one histogram was rejected. Concerning the probabilities f, for the model to be applicable, it is required that these exhibit a limited variation and no significant trend. To quantify the variation, the standard deviation of f in the scaling regime was calculated. For the three probability types and eight classes (24 probabilities), the standard deviation was 0.082±0.035, which indicates a limited variation. The hypothesis that the slope of a straight regression line of f versus modulation scale equals zero in the scaling regime, i.e., absence of trend, was tested using the t-ratio (estimated slope divided by its standard error) as test statistic. At significance level 0.01, the hypothesis of zero slope of a straight regression line could not be rejected for any probability. At significance level 0.1, eight probabilities were rejected.

Table 1 shows f_{7-14}, i.e., the mean of f in the scaling regime, for all probabilities; for all position classes, $f_{7-14}(x/x)$ for volumes above mean is substantially higher than for volumes below mean. This confirms the assumption about the influence of the volume on the probabilities made in the previous section. Also the lower quartile, the median, and the upper quartile were tested as limit between the volume classes. However, this did not change the result significantly, and in the disaggregation below the mean will be used for simplicity. Concerning the influence of the position, $f_{7-14}(x/x)$ is higher for enclosed than for isolated boxes, and $f_{7-14}(0/1)$ ($f_{7-14}(1/0)$) for starting (ending) boxes is substantially higher than $f_{7-14}(1/0)$ ($f_{7-14}(0/1)$). Thus, these assumptions were also confirmed.

Table 1 shows that some simplifications of the model are possible. Firstly, starting and ending boxes have reversed $f_{7-14}(0/1)$ and $f_{7-14}(1/0)$ values (i.e., $f_{7-14}(0/1)$ for starting boxes nearly coincided with $f_{7-14}(1/0)$ for ending boxes and vice versa) and similar $f_{7-14}(x/x)$ values (see also Figs 3a and c). This is valid for both volume classes. Thus, starting and ending boxes may be viewed as being of the same principal 'edge-of-rainfall-sequence' class having the same but reversed $f_{7-14}(0/1)$ and $f_{7-14}(1/0)$ values, and identical $f_{7-14}(x/x)$ values. Secondly, for enclosed and isolated boxes there are almost no difference between $f_{7-14}(0/1)$ and $f_{7-14}(1/0)$ (see also Figs 3b and d). This means that if all rainfall originated from one half of the period, there is an equal chance of this being the first or the last half, which appears reasonable. This behaviour is also valid for both volume classes.

In order to assess the applicability of the model for disaggregation, a number of complementary analyses were performed. An important issue is how accurately the P values can be estimated using only larger scales, since this is the course of action in a real-world application where only larger-scale data are available. This issue was investigated by comparing f_{7-14} to the average of f in the interval 5.7 days to 17 hours (l = 10), f_{7-10}. The absolute difference between f_{7-14} and f_{7-10} for all 24 probabilities was 0.051±0.040, which supports the potential for disaggregation. Other crucial issues are the temporal stationarity and robustness of the probabilities. Although the seasonal variability of rainfall in the region is limited, its effect on the probabilities may be significant and must therefore be evaluated. For this purpose, f_{7-14} for spring, summer, and autumn data, respectively, were calculated separately, and then compared with f_{7-14} obtained from the entire series and with each other. Table 2 shows that the absolute difference between f_{7-14} estimated for separate seasons and the entire series, respectively, is about 0.05±0.05, whereas the difference between different seasons is about 0.07±0.08. This indicates a slight temporal nonstationarity of the generator parameters. It should be stressed that the probabilities estimated from single seasons are associated with larger uncertainties due to the smaller amount of data. To investigate the temporal robustness, f_{7-14} was estimated using only the first and the second years of the time series and compared with f_{7-14} obtained from the entire series. Table 2 shows that the absolute difference is only about 0.02 ± 0.02.

Disaggregation

Since the proposed model proved to be applicable between 5.7 days and 1 hour, tests of rainfall disaggregation within this range were performed. In hydrological applications, the available rainfall data are often 1-day or 12-hour volumes, whereas a higher resolution, e.g. 1 hour, is required for model input. Therefore, the original 8-min values were firstly aggregated into 17-hour (1024-min) values, which were then disaggregated to 1-hour (64-min) values by using the cascade model generator (Eqn. 1). In the disaggregation, the values of f_{7-14} obtained in the evaluation were used as estimates of the probabilities P. The simplifications described in the previous section were imple-

Table 2. Absolute differences between the 24 empirical probabilities estimated from different time periods. The subscripts correspond to different time periods and are self-explanatory.

Compared probabilities	Absolute difference (mean±std)
$f_{spring} - f_{total}$	0.058±0.064
$f_{summer} - f_{total}$	0.044±0.043
$f_{autumn} - f_{total}$	0.059±0.058
$f_{spring} - f_{summer}$	0.073±0.074
$f_{summer} - f_{autumn}$	0.073±0.075
$f_{spring} - f_{autumn}$	0.071±0.096
$f_{year1} - f_{total}$	0.022±0.021
$f_{year2} - f_{total}$	0.016±0.010
$f_{year1} - f_{year2}$	0.025±0.017

mented and, because of these, the number of P values to be specified was reduced from 16 to 8. The P values employed in the disaggregation are shown in Table 1.

When disaggregating a certain wet box, its volume and position were used firstly to determine the class, i.e., the set of P values to be used. A random number r, uniformly distributed between 0 and 1, was then drawn to determine the type of division. If $r < P(1/0)$, a 1/0-division was made ($W_1 = 1$ and $W_2 = 0$). If $P(1/0) < r < (P(1/0) + P(0/1))$, a 0/1-division was made ($W_1 = 0$ and $W_2 = 1$). If $(P(0/1) + P(1/0)) < r$, a x/x-division was made ($W_1 = W_{x/x}$ and $W_2 = 1 - W_{x/x}$). In this third case, since the $W_{x/x}$-distribution was suggested to be uniform in the evaluation, $W_{x/x}$ was drawn from a uniform distribution between 0 and 1.

When all the 17-hour values had been disaggregated, a series of 8.5-hour values had thus been produced. This series was then disaggregated by the same procedure to 4.3-hour values, the 4.3-hour values to 2.1-hour values, and finally the 2.1-hour values to 1-hour values. The disaggregation procedure will thus preserve nonstationarities in the data at scales larger than the scale from which the disaggregation starts, whereas at smaller scales the generated data will be stationary. This is in agreement with the present data, which do not contain any pronounced diurnal nonstationarities, but may constitute a limitation of the model when applied to data from other rainfall regimes.

This disaggregation of the 17-hour values was repeated ten times, i.e. ten different realisations of the 1-hour series were produced. The performance was evaluated by comparing the disaggregated with the observed 1-hour values (aggregated 8-min values). This comparison was made for all time scales to which the 17-hour values were disaggregated, i.e. 8.5 hours, 4.3 hours, 2.1 hours, and 1 hour. The following five variables were considered in the first step: (1) percentage of zero-values, (2) rainfall volume of individual values, (3) rainfall volume of events, (4) duration of events, and (5) length of dry periods between events. At all scales, an event was defined as a sequence of consecutive non-zero values.

In Table 3, the results, in terms of mean and standard deviation of the above five variables, from the ten disaggregations of the 17-hour values performed by the model (M_{7-14}) are compared to the observed data (Obs). As expected, the agreement between observed and generated data generally decreases with scale. At the largest scale of 8.5 hours the agreement is nearly perfect for all variables, but also at the smallest scale of 1 hour the agreement appears satisfactory. In terms of percentage of zero-values and mean individual volume, the generated data almost perfectly match the observed at all scales. Thus, the division into rainy and non-rainy periods is accurately performed by the model. The standard deviation of individual volumes is, to an increasing degree, overestimated as the scale decreases. This indicates the presence of too high values in the generated data (confirmed by inspection), a fact that to some extent contradicts the common opinion that microcanonical cascades are too 'calm' to be relevant for rain (e.g. Lovejoy and Schertzer, 1995; Menabde *et al.*, 1997). The mean value and standard deviation of event volumes and durations are well reproduced at larger scales, but underestimated at smaller scales. Consequently, the dry period lengths will also be underestimated. Table 3 shows that this is the case but, generally, the dry periods are rather well represented, which is expected since long dry periods at the 17-hour scale are essentially preserved by the model.

Since the main aim of the study was disaggregation down to 1 hour, the 1-hour series produced by the model was studied in greater detail. This was done firstly by evaluating the agreement of the entire cumulative distribution function of the four last variables (of the five variables described above) using quantile-quantile plots, *q-q*-plots. The *q-q*-plots were constructed by interpolating the larger data set to the same size as the smaller set by using the Weibull plotting position (e.g. Helsel and Hirsch, 1992), and some typical examples are shown in Fig. 4. For non-zero 1-hour volumes (Fig. 4a), the overall agreement is good. There is, however, a systematic overestimation of the disaggregated values around 5 mm. Also, the highest values are somewhat overestimated by the model, as suspected from the overestimated standard deviation (Table 3). For event volumes and durations (Figs. 4b and c), respectively, the agreement between observed and disaggregated data is similar (note that in Fig. 4c, due to the small number of event durations present, a point value generally represents many identical pairs of values). For small values, the agreement is good. For medium-sized values, the agreement is satisfactory although the variables are somewhat underestimated by the model. For the largest values, the variables are significantly underestimated by the model. For dry period lengths (Fig. 4d), the overall agreement is good although a slight systematic underestimation by the model is apparent.

Figure 5 shows typical examples of the agreement

Table 3. Comparison between observed data (Obs), data generated by the model using f_{7-14} to approximate P (M_{7-14}), and data generated by the model using f_{7-10} to approximate P (M_{7-10}), in terms of five variables at all scales to which disaggregation was performed.

Scale	Data	Zero values %	Individual volume mm(mean±std)	Event volume mm(mean±std)	Event duration hrs(mean±std)	Dry period hrs(mean±std)
8.5 hrs	Obs	82.5	2.8 ± 3.8	5.0 ± 7.3	15.2 ± 9.65	72.0 ± 84.0
	M_{7-14}	82	2.7 ± 3.9	5.0 ± 7.1	15.6 ± 10.6	71.0 ± 84.0
	M_{7-10}	81	2.6 ± 3.8	5.0 ± 7.1	16.2 ± 10.9	71.0 ± 84.0
4.3 hrs	Obs	89	2.2 ± 2.8	3.9 ± 5.9	7.4 ± 4.6	60.0 ± 79.5
	M_{7-14}	88	2.1 ± 3.1	3.7 ± 5.4	7.5 ± 5.0	57.0 ± 78.5
	M_{7-10}	87.5	2.0 ± 2.9	3.7 ± 5.35	7.9 ± 5.45	55.5 ± 78.0
2.1 hrs	Obs	93	1.7 ± 2.0	3.2 ± 5.1	4.1 ± 2.7	52.0 ± 76.0
	M_{7-14}	92.5	1.6 ± 2.4	2.8 ± 4.2	3.7 ± 2.4	45.5 ± 73.0
	M_{7-10}	92	1.5 ± 2.3	2.65 ± 3.9	3.8 ± 2.5	42.0 ± 71.0
1 hrs	Obs	95	1.2 ± 1.45	2.7 ± 4.45	2.35 ± 1.9	44.0 ± 72.0
	M_{7-14}	95	1.25 ± 1.9	2.2 ± 3.4	1.9 ± 1.3	36.0 ± 67.0
	M_{7-10}	95	1.1 ± 1.8	2.0 ± 3.0	1.8 ± 1.2	32.0 ± 64.0

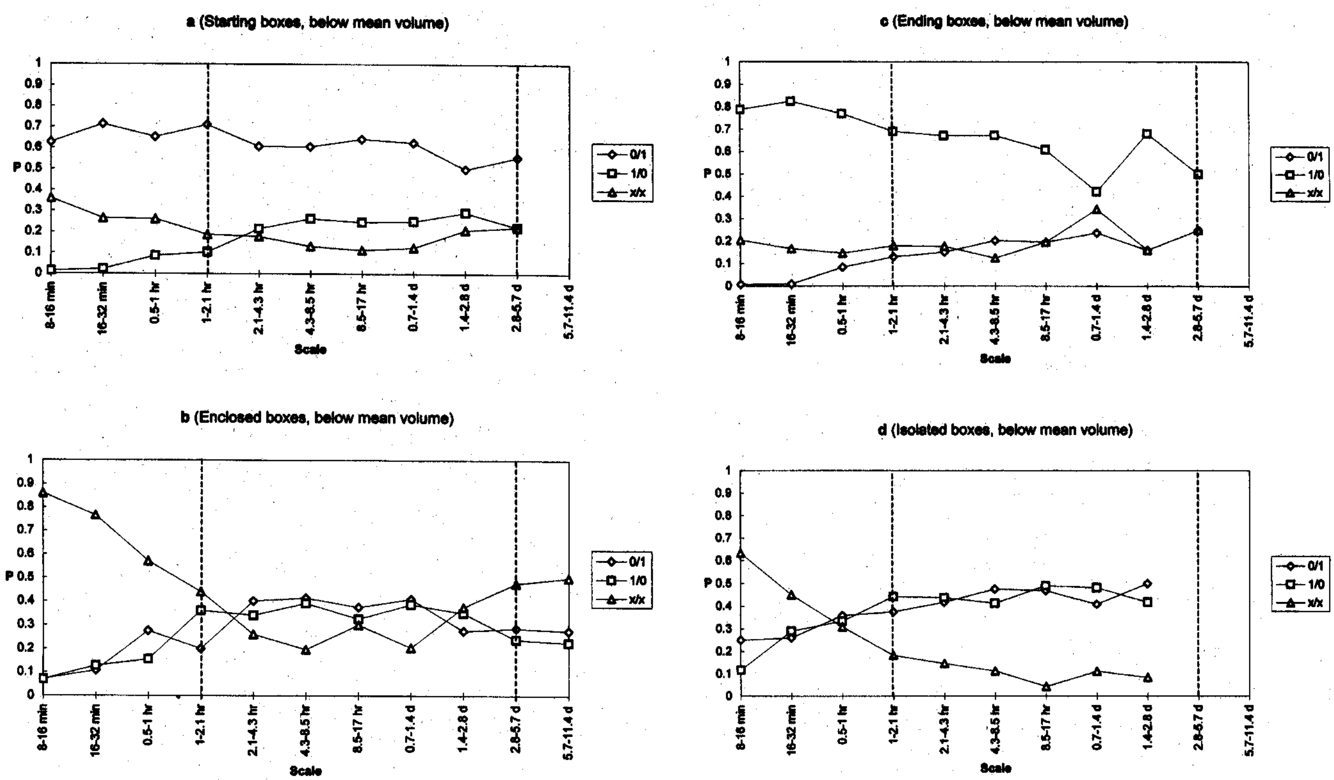

Fig. 3. *Empirical probabilities f(1/0), f(0/1), and f(x/x) as a function of modulation time scale for starting (a), enclosed (b), ending (c), and isolated (d) boxes below mean volume. The vertical dashed lines mark the time scale limits of applicability of the model as interpreted in the present study.*

between observed and generated 1-hour series in terms of power spectrum and autocorrelation. The spectrum from the generated series matches the approximately power-law shaped observed spectrum (Fig. 5a) well for time scales below 17 hours (for larger time scales the spectra naturally coincide). For short lags, the observed autocorrelation is underestimated by the model (Fig. 5b). The overall shape is, however, reasonably well reproduced considering that the disaggregation was performed without taking any correlation structure into account.

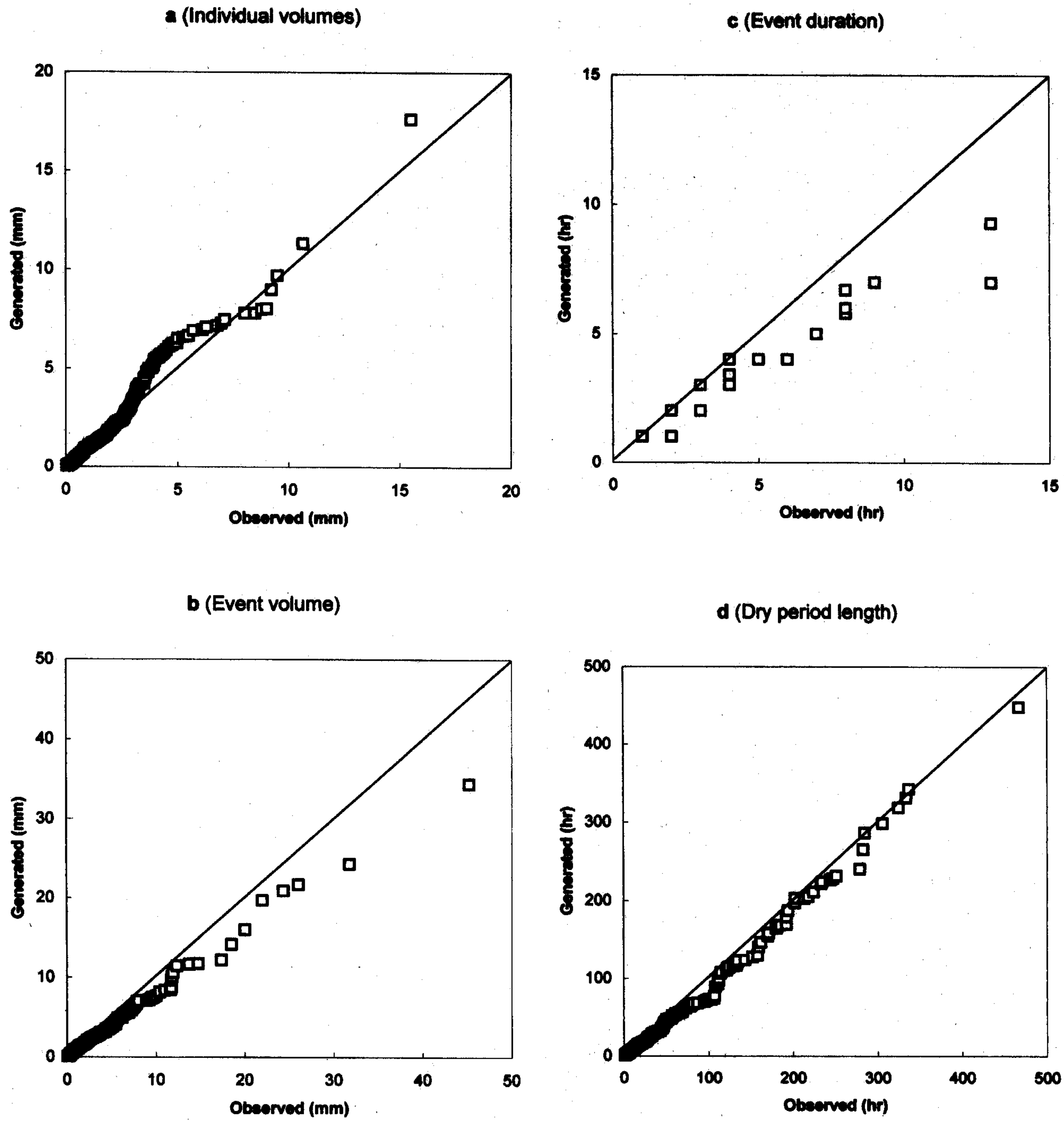

Fig. 4. *Quantile-quantile-plots comparing the cumulative distribution functions of four variables in observed 1-hour data and 1-hour data disaggregated from 17-hour data by the model M_{7-14}: individual 1-hour volume (a), event volume (b), event duration (c), and dry period length (d). The solid line represents x=y, i.e. perfect match.*

To investigate to which degree the scaling properties of the disaggregated data obtained by the final model agree with the observed data, analyses of statistical moments were performed. This is a standard method in investigations of scaling and is often used in rainfall analyses (e.g. Over and Gupta, 1994; Svensson *et al.*, 1996). The analysis is generally done by averaging the values over different non-overlapping time intervals of length λ, where λ thus may be viewed as a scale parameter. To obtain the h:th order moment $M(\lambda,h)$, all averaged values are raised to the power h, and then summed. If the data exhibit scaling, $M(\lambda,h)$ is related to λ by

$$M(\lambda,h) \approx \lambda^{K(h)} \qquad (2)$$

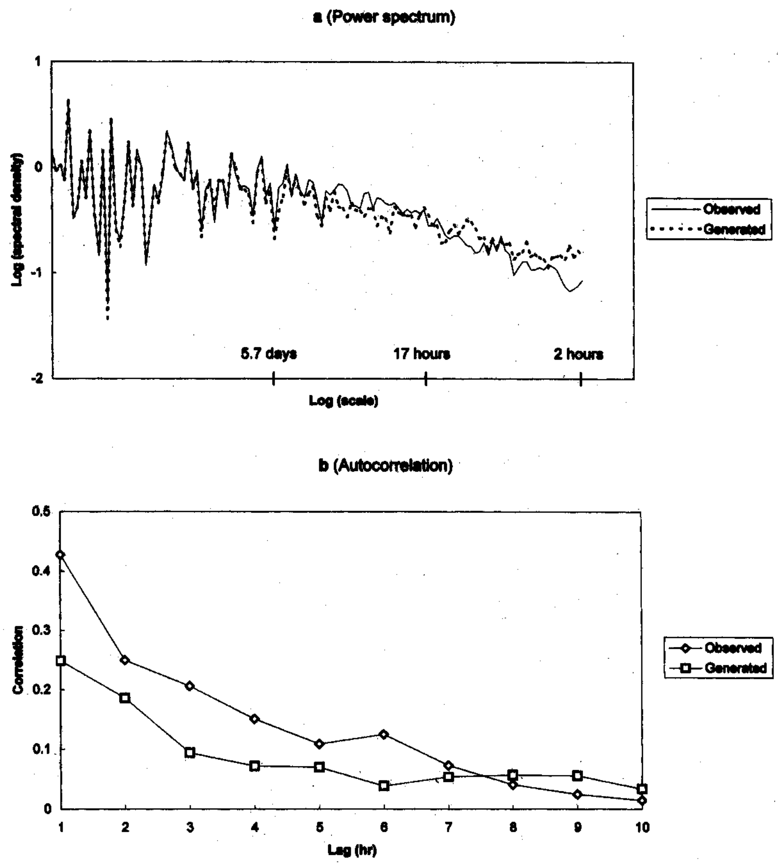

Fig. 5. *Power spectrum (a) and autocorrelation (b) for observed 1-hour data and 1-hour data disaggregated from 17-hour data by the model* M_{7-14}.

where $K(h)$ is a characteristic function of the scaling behaviour. The validity of Eqn. 2 is usually investigated by plotting $M(\lambda,h)$ as a function of λ in a log-log-plot. If scaling holds, the points fall on an approximately straight line. This is a theoretical property of cascade models (e.g. Schertzer and Lovejoy, 1987; Gupta and Waymire, 1990), such as the present. Fig. 6 shows a typical plot of the moments $M(\lambda,h)$ versus scale λ for observed and generated 1-hour data. As for the power spectrum, at scales larger than 17 hours the curves naturally coincide. For scales smaller than 17 hours, the curves are straight lines of the same slope as in the range 17 hours to 5.7 days. As is evident, the curves of the observed and generated data agree very well for all values of h. This result, and to some extent the result from the power spectrum analysis, confirm that the model is able to reproduce the scaling behaviour of the rainfall time series.

Finally, to investigate the sensitivity of the generated data to differences in P values, ten disaggregations were performed in which the P values were approximated by the f values from the scale interval 5.7 days to 17 hours, i.e. f_{7-10}, obtained in the evaluation. In Table 3, the results, in terms of mean and standard deviation of the above five variables, from the ten disaggregations of the 17-hour values performed by the model (M_{7-10}) are compared to the observed data (Obs). The agreement with the observed series is similar but in general slightly worse than for the generated data obtained using P values estimated from the entire range 5.7 days to 1 hour (M_{7-14}). However, the difference is small and even at the 1-hour scale the accuracy of the data generated by M_{7-10} decreased by only approximately 3% as compared to the data obtained by M_{7-14}.

The results of the disaggregation may be compared with the results obtained by Bo *et al.* (1994), who performed disaggregation of continuous rainfall time series from central Italy and Kentucky, USA, in the same scale interval using the modified Bartlett-Lewis rectangular pulses model. They fitted the model on a monthly basis to daily data and showed that it could be used also for reproducing smaller-scale statistics. The accuracy of the generated data in the present study appears comparable with the results obtained by Bo *et al.* (1994) despite the fact that seasonal nonstationarities were not taken into account here. Bo *et al.* (1994) found an upper time scale limit at 2 days, which were claimed to be associated with a break of the power spectrum. This limit is thus in some disagreement with the limits obtained using the present approach,

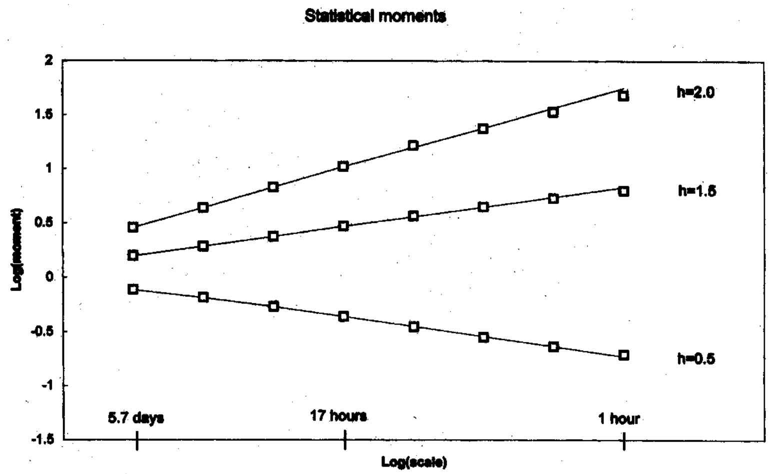

Fig. 6. *Statistical moment $M(\lambda,h)$ as a function of scale λ between 5.7 days and 1 hour for h=0.5, 1.5, and 2.0. The point values correspond to observed 1-hour data, and the solid lines to 1-hour data disaggregated by the model M_{7-14}.*

a fact probably related to geographical differences in statistical rainfall properties.

Summary and conclusions

A cascade scheme was employed to model the temporal small-scale structure of rainfall. The model was based on exact conservation of rainfall volume between cascade levels. The model parameters were assumed to depend on characteristics of the rainy time period (wet box) to be branched. For a 2-year rainfall time series observed in southern Sweden, the model was reasonably applicable between approximately 1 week and 1 hour. The parameters showed a pronounced dependence on the box characteristics, and were characterised by a slight seasonal nonstationarity. To investigate the potential for disaggregation, the series was disaggregated from 17-hour resolution to 1-hour resolution by the model. The observed division into rainy and non-rainy periods as well as the scaling structure were nearly perfectly reproduced by the model, whereas events and dry periods were somewhat underestimated.

Although the model-generated series overall compared well with the observed series, it is clear that the approach used here can be improved in various respects. Obviously the disaggregation can be performed seasonally to take the slight nonstationarities of the cascade parameters into account, something that was not done here due to insufficient seasonal data. The systematic underestimation of events and dry periods may be due to an excessive tendency of the model to break up long rainfall events into two parts separated by a zero-value, at some modulation. This problem, in turn, may be related to correlations between the P values and rainfall generating mechanisms. For example, $P(x/x)$ is likely to be higher for rain produced by frontal passages (long continuous events) than by convective activity (showers separated by dry periods). Such correlations could be included in the model, for example by defining the position of a wet box not only by the characteristics of the preceding and succeeding box, respectively, but by using a number of preceding and succeeding boxes. Another way could be to include a correlation between the pairs of weights at each cascade level. Such a correlation is usually not considered in cascade models, but has been found in empirical analyses of rainfall time series (Cârsteanu and Foufoula-Georgiou, 1996).

A main concern of the present approach is naturally the connection between the cascade structure of the model on the one hand, and the underlying physics of the rainfall process on the other. Establishing this connection remains a crucial task of research in the present field. The fact that the conceptually simple model employed in the present study was able to capture fundamental features of the rainfall process accurately over a range of scales does, however, constitute yet another indication that a cascade type of behaviour is inherent in the rainfall producing mechanisms. It also shows that scaling-based rainfall models are potentially important tools in hydrological applications where some form of disaggregation or downscaling of rainfall is required.

Acknowledgements

The author wishes to acknowledge Ronny Berndtsson for fruitful discussions, the reviewers for providing helpful criticism, and the Swedish Natural Science Research Council for supporting this research.

References

Ångström, A., (1974). *Sveriges klimat*, Generalstabens Litografiska Anstalts Förlag, Stockholm.

Austin, P.M. and Houze, R.A., (1972). Analysis of structure of precipitation patterns in New England, *J. Appl. Meteorol.*, **11**, 926–935.

Bo, Z., Islam, S. and Eltahir, E.A.B., (1994). Aggregation-disaggregation properties of a stochastic rainfall model, *Wat. Resour. Res.*, **30**, 3423–3435.

Burlando, P. and Rosso, R., (1996). Scaling and multiscaling models of depth-duration-frequency curves for storm precipitation, *J. Hydrol.*, **187**, 45–64.

Cârsteanu, A. and Foufoula-Georgiou, E., (1996). Assessing dependence of weights in a multiplicative cascade model of temporal rainfall, *J. Geophys. Res.*, **101**, 26363–26370.

Econopouly, T.W., Davis, D.R. and Woolhiser, D.A., (1990). Parameter transferability for a daily rainfall disaggregation model, *J. Hydrol.*, **118**, 209–228, 1990.

Glasbey, C.A., Cooper, G. and McGechan, M.B., (1995). Disaggregation of daily rainfall by conditional simulation from a point-process model, *J. Hydrol.*, **165**, 1–9.

Gupta, V.K. and Waymire, E., (1990). Multiscaling properties of spatial rainfall and river flow distributions, *J. Geophys. Res.*, **95**, 1999–2009.

Gupta, V.K. and Waymire, E., (1993). A statistical analysis of mesoscale rainfall as a random cascade, *J. Appl. Meteorol.*, **32**, 251–267.

Haan, C.T., (1977). *Statistical Methods in Hydrology*, Iowa State University Press, Ames, Iowa.

Harris, D., Menabde, M., Seed, A. and Austin, G., (1996). Multifractal characterization of rain fields with a strong orographic influence, *J. Geophys. Res.*, **101**, 26405–26414.

Helsel, D.R. and Hirsch, R.M., (1992). *Statistical Methods in Water Resources*, Elsevier, Amsterdam.

Hershenhorn, J. and Woolhiser, D.A., (1987). Disaggregation of daily rainfall, *J. Hydrol.*, **95**, 299–322.

Hubert, P. and Carbonnel, J.P., (1989). Dimensions fractales de l'occurrence de pluie en climat soudano-sahélien, *Hydrol. Continent.*, **4**, 3–10.

Hubert, P., Tessier, Y., Lovejoy, S., Schertzer, D., Schmitt, F., Ladoy, P., Carbonnel, J.P., Violette, S. and Desurosne, I., (1993). Multifractals and extreme rainfall events, *Geophys. Res. Lett.*, **20**, 931–934.

Lovejoy, S. and Schertzer, D., (1990). Multifractals, universality classes and satellite and radar measurements of clouds and rain fields, *J. Geophys. Res.*, **95**, 2021–2034.

Lovejoy, S. and Schertzer, D., (1995). Multifractals and rain, in *New Uncertainty Concepts in Hydrology and Hydrological Modeling*, edited by A. W. Kundzewicz, pp. 61–103, Cambridge Univ. Press, New York.

Mandelbrot, B.B., (1974). Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier, *J. Fluid. Mech.*, **62**, 331–358.

Menabde, M., Harris, D., Seed, A., Austin, G. and Stow, D., (1997). Multiscaling properties of rainfall and bounded random cascades, *Wat. Resour. Res.*, **33**, 2823–2830.

Meneveau, C. and Sreenivasan, K.R., (1987). Simple multifractal cascade model for fully developed turbulence, *Phys. Rev. Lett.*, **59**, 1424–1427.

Niemczynowicz, J., (1986a). The dynamic calibration of tipping-bucket raingauges, *Nord. Hydrol.*, **17**, 203–214.

Niemczynowicz, J., (1986b). Storm tracking using raingauge data, *J. Hydrol.*, **93**, 135–152.

Olsson, J., (1995). Limits and characteristics of the multifractal behavior of a high-resolution rainfall time series, *Nonlin. Proc. Geophys.*, **2**, 23–29.

Olsson, J., Niemczynowicz, J., Berndtsson, R. and Larson, M., (1992). An analysis of the rainfall time structure by box-counting—some practical implications, *J. Hydrol.*, **137**, 261–277.

Olsson, J., Niemczynowicz, J. and Berndtsson, R., (1993). Fractal analysis of high-resolution rainfall time series, *J. Geophys. Res.*, **98**, 23265–23274.

Onof, C., Northrop, P., Wheater, H.S. and Isham, V., (1996). Spatiotemporal storm structure and scaling property analysis for modeling, *J. Geophys. Res.*, **101**, 26415–26425.

Over, T.M. and Gupta, V.K., (1994). Statistical analysis of mesoscale rainfall: dependence of a random cascade generator on large-scale forcing, *J. Appl. Meteorol.*, **33**, 1526–1542.

Perica, S. and Foufoula-Georgiou, E., (1996). Model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions, *J. Geophys. Res.*, **101**, 26347–26361.

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

Rodriguez-Iturbe, I., Cox, D.R., and Isham, V., (1987). A point process model for rainfall: further developments, *Proc. R. Soc. Lond. A*, **417**, 283–298.

Schertzer, D. and Lovejoy, S., (1987). Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes, *J. Geophys. Res.*, **92**, 9693–9714.

Svensson, C., Olsson, J. and Berndtsson, R., (1996). Multifractal properties of daily rainfall in two different climates, *Wat. Resour. Res.*, **32**, 2463–2472.

Tessier, Y., Lovejoy, S. and Schertzer, D., (1993). Universal multifractals: theory and observations for rain and clouds, *J. Appl. Meteorol.*, **32**, 223–250.

Tessier, Y., Lovejoy, S., Hubert, P., Schertzer, D. and Pecknold, S., (1996). Multifractal analysis and modeling of rainfall and river flows and scaling, causal transfer functions, *J. Geophys. Res.*, **101**, 26427–26440.

Vogel, R.M. and Kroll, C.N., (1989). Low-flow frequency analysis using probability-plot correlation coefficients, *J. Wat. Resour. Plann. Manage.*, **115**, 338–357.

Yaglom, Y., (1966). The influence of fluctuations in energy dissipation on the shape of turbulence characteristics in the inertial interval, *Sov. Phys. Dokl.*, **11**, 26–29.