Hydrol. Earth Syst. Sci. Discuss., 7, 5267–5297, 2010
www.hydrol-earth-syst-sci-discuss.net/7/5267/2010/
doi:10.5194/hessd-7-5267-2010


Hydrology and Earth System Sciences Discussions



# Reconstruction of sub-daily rainfall sequences using multinomial multiplicative cascades

**L. Wang, C. Onof, and C. Maksimovic**

Department of Civil and Environmental Engineering, Imperial College London, SW7 2AZ, UK

Received: 10 July 2010 – Accepted: 13 July 2010 – Published: 3 August 2010

Correspondence to: L. Wang (li-pen.wang08@imperial.ac.uk)

Published by Copernicus Publications on behalf of the European Geosciences Union.

## Abstract

This work aims to develop a semi-deterministic multiplicative cascade method for producing reliable short-term (sub-daily) rainfall sequences. The scaling feature of sub-daily rainfall sequences is analysed over the timescales of interest (i.e., 5 min to hourly in this research) to help derive the crucial parameters, i.e., the fragmentation ratios, for the proposed method. These derived ratios are then further used to stochastically disaggregate hourly rainfall sequences to 5-min using the multiplicative cascade process. The log-Poisson distributed cascade method is involved in this work to validate the proposed methodology by comparing certain statistics of the generated rainfall sequences over a specific range of timescales. The results demonstrate that the proposed methodology in general has the ability to reproduce the patterns of sub-daily rainfall observations from Greenwich raingauge station in London.


## 1 Introduction

In the summer of 2007, over 55 000 homes and businesses in the UK were flooded due to the fact that urban drains, river channels, and flood defences were unable to cope with the extreme quantities of water resulting from both fluvial and pluvial causes (EA, 2007). Despite the construction of new flood infrastructure, flooding is still repeatedly causing damage, particularly induced by the surface (pluvial) floods over urban areas. The Pitt Review (Pitt, 2008), which calls for integrated solutions to this problem, addressing both technical and socio-economical issues in Integrated Urban Flood Management (IUFM), highlights the importance of appropriately tackling the surface floods mainly induced by pluvial causes. The existing short-term surface flood modelling requires at least a couple of hours lead time to provide reliable estimates of the distribution of floods over urban scales; however, the associated achievable lead time of high-resolution rainfall prediction is usually up to 45–60 min (Sokol, 2006). This predictability is limited mainly due to the limited capabilities of existing short-term rainfall



HESSD

7, 5267–5297, 2010

Reconstruct sub-daily rainfall using multinomial cascades

L. Wang et al.

forecast models, which generally rely on extrapolating the rainfall measurements conducted by the networks of raingauges and meteorological radars. Although these models are able to carry out spatially and temporally high-resolution rainfall forecasts (or nowcasts), the lead time is however too short compared to the reasonable response time to extreme floods.

The method of integrating rainfall models over multiple scale ranges has been widely developed as the solution to carrying out high-resolution rainfall prediction with longer lead time (Pierce, 2009; Bowler et al., 2006; Sokol, 2006; Golding, 1998). The notion is to blend the coarser information (e.g., for a 6-h time interval and 10 km grid) predicted by Numerical Weather Prediction (NWP) models with the techniques of radar-raingauge-based nowcasting (for an 1-h time interval and 1–3 km grid) and stochastically- (or statistically-) based downscaling methods (5-min and 100-m resolution). The improvement in stochastic downscaling techniques therefore plays a key role in resolving the difficulty in modelling the high non-linearity of precipitation at urban/street scales.

The aim of this work therefore is twofold. The first is to develop a stochastic disaggregation methodology to reproduce the structure of sub-daily rainfall time series. Based upon these stochastically-generated rainfall series, the second aim is to produce reliable "deterministic" rainfall information, which can be used as inputs for the corresponding hydrological modelling, paricularly for the short-term surface flooding modelling over urban areas.

## 2 Stochastic rainfall disaggergation

There are a multitude of rainfall disaggregation models in the literature which can be classified into three categories based upon the theories respectively used. The first category of disaggregation models are purely statistical models, among which for example Generalised Linear Models (GLMs) have been successfully applied to the studies of disaggregation processes of rainfall sequences, particularly at daily





HESSD
7, 5267–5297, 2010

timescales (Segond et al., 2006; Yang et al., 2005; Chandler and Wheater, 2002; Stern and Coe, 1984). GLMs are extended from the simple linear regression models, of which the idea is to predict a probability distribution for a rainfall sequence at a specific site of interest by linking the mean of that distribution to the values of a certain number of associated quantities known as "covariates" or "predictors." The advantage of GLMs is that they enable not only to profile the daily variability of rainfall sequences but also to reflect the effects of long-term climatic trends and orographic variability on rainfall patterns by considering the seasonal cycles and location-dependent coefficients in covariates.

Second, Poisson-cluster models were developed based upon a theory of stochastic point processes by considering physically-based parameters which constitute the hierarchical structure of the rainfall process. The Bartlett-Lewis rectangular pulse models are one specific class of the widely-used individual-site rainfall models based upon a Poisson-cluster process (Koutsoyiannis, 2003; Onof et al., 2000). Poisson-based models have been widely improved and applied to rainfall modelling. They are also useful to disaggregate daily rainfall sequences to hourly ones (Koutsoyiannis and Onof, 2001; Onof and Wheater, 1993, 1994; Onof and Arnbjerg-Nielsen, 2009). For example, Onof and Wheater (1993, 1994) improved the basic Bartlett-Lewis rectangular pulse model by randomising the duration of cells and applying a gamma instead of an exponential distribution to the cell intensity. These two modifications respectively address the general problems in Poisson-based models, i.e., the overestimation of the proportions of dry periods and relatively poor performance in reproducing the extreme values of the distribution of rainfall depth.

The third type of disaggregation models are multiplicative cascade methods, which are developed by investigating the scale-invariant behaviour of complex nonlinear processes. Random multiplicative cascade methods have been successfully applied to rainfall modelling, particularly focusing on high-resolution rainfall data in hourly and sub-hourly intervals (Molnar and Burlando, 2005; Over and Gupta, 1996; Pathirana et al., 2003; Onof and Arnbjerg-Nielsen, 2009). For example, aiming to integrate disaggregation models at various timescales, Onof and Arnbjerg-Nielsen (2009) adopted a canonical cascade generator, based upon a log-Poisson variable, to generate 5-min rainfall sequences from hourly ones. The common idea of this type of methods is to distribute rainfall volumes on successive regular subdivisions of an interval in a randomly multiplicative manner by constructing generators, a process of which the scaling feature has been theoretically explained by the theory of universal multifractals (Tessier et al., 1993). The results of these works showed the methods' ability to preserve specific statistical features, particularly for the complex sub-daily rainfall patterns.

The cascade-based method therefore has the potential to satisfactorily reproduce the non-linear structure of sub-daily rainfall patterns, and an enhanced cascade-based methodology is developed in this work based upon the notion of micmicing the construction of multinomial multiplicative cascade processes. The associated theory and derivation of the developed methodology are detailed in Sect. 4.

## 3 Rainfall data

This work uses rainfall data from the Greenwich raingauge station in London to validate the simulation results. The rainfall depth was measured with a time resolution of 5 min using tipping bucket raingauges with a depth resolution of 0.01 mm. The observation period is 14 years (from 21 February 1987 to 21 February 2001, i.e., approximately 1 500 000 data with the 5-min interval). Some statistics of 5-min rainfall intensity (mm/h) are summarised in Table 1 for each month. The mean intensity at the Greenwich station is not high for each month; however, due to the high proportion of dry periods, it is expected that rainfall concentrates within several short time intervals. This can be seen from the high "wet" mean, which is obtained by averaging merely the non-zero rainfall values. It is therefore very crucial to timely capture the rainfall patterns in these short-term "wet" periods.





HESSD
7, 5267–5297, 2010

## 4 A semi-deterministic cascade method

### 4.1 The multinomial multiplicative cascade process

The multiplicative cascade is a single process to generate fine-scale data by subdividing a unit set into smaller and smaller subsets according to a fixed set of fragmentation (contracting) ratios and at the same time subdividing the associated unit measure by another set of fragmentation ratios. Random process can then be implemented based upon a self-similar process with deterministic sets of fragmentation ratios (Mandelbrot, 1989), such that the mass of measures is exactly conserved. Figure 1 is the schematic of a multinomial multiplicative cascade process, of which each component (or box) is subdivided into $b$ (termed the branching number) smaller sub-components between two successive levels. The subdivision is carried out using a specific set of fragmentation ratios (denoted $p_j^L$, where $j$ is the relative position of each box in sub-components and $L$ is the level index), which is shown in Fig. 2, where the constraint, i.e., $\sum_{j=1}^{b} p_j^L = 1$, is applied to each branching in order to conserve the rainfall volumes.

This process has been widely used to construct/analyse rainfall sequences over a certain range of timescales. The components at level $g$ in Fig. 1, for example, can be assumed to be an hourly rainfall sequence, and each component of it thus represents an 1-h rainfall volume. Subdivision is then continuously carried out for each component at level $g$ to obtain $b$ sub-components at level $g+1$, which could be regarded as the corresponding sub-hourly rainfall sequence. Repeat the same fragmentation several times till level $h$; the associated 5-min rainfall sequence then is obtained. It is therefore crucial to estimate these fragmentation ratios ($p_i^L$) when cascade-based methods are applied to rainfall modelling.

Many efforts have been made to characterise these ratios and to apply them to spatially- or temporally-distributed rainfall modelling in the literature (Deidda, 2000; Deidda et al., 1999; Gupta and Waymire, 1993; Molnar and Burlando, 2005; Onof and Arnbjerg-Nielsen, 2009; Onof et al., 2005; Over and Gupta, 1994, 1996; Pathirana and Herath, 2002; Pathirana et al., 2003; Shrestha et al., 2004), among which the binomial (i.e., $b$=2) random cascade methods are the mainstream and have been widely developed. The general idea is to construct rainfall generators (generally denoted $W_i^L$) based upon analysing statistical features (or probability distributions) of $p_i^L$ values. The associated parameters of generators can be empirically estimated from historical rainfall observations. Particularly, to simplify, $W_i^L$ are in general assumed to be independently and identically distributed (i.i.d.), and the expected value of $W_i^L$ is assigned to be unity to macrocanonically preserve rainfall volume (this means that the volume is merely conserved on average).

Onof et al. (2005), for example, referring to the derivation in Deidda et al. (1999), used a log-Poisson cascade to disaggregate hourly rainfall sequences to 5-min. The associated generator is formed as,

$$W = A\beta^N, \tag{1}$$

where $N$ is a log-Poisson distributed random variable, and $A$ and $\beta$ are two parameters that can be identified by fitting the structure function, which is empirically plotted based upon scale-invariant features of historical raingauge measurements. The results showed that certain statistics were satisfactorily reproduced; however, due to the fact that these variables are strictly positive, some transformation of the simulation results is necessary to produce zero-values (e.g., thresholding). Similarly, the log-Lévy distributed cascade methods have difficulty in producing zero-values (Lovejoy and Schertzer, 1990). Over and Gupta (1994) introduced a zero "atom" to the generator to produce zero-values; however, the dichotomy structure of the generator, analogous to the well known $\beta$-model (Over and Gupta, 1996), could not satisfactorily reproduce the generally-observed clustering nature of real rainfields. These works entail that the simulation results of random cascade based methods and their forms and parameters rely significantly on the type of probability distributions used. A generator, therefore, may radically alter when a different rainfield or rainfall sequence starts.



HESSD
7, 5267–5297, 2010
Reconstruct sub-daily rainfall using multinomial cascades
L. Wang et al.

Olsson (1998) developed a microcanonical generator with a relatively general structure based upon binomial cascades. The fragmentation ratios for each pair of successive timescale levels were firstly analysed to determine a proper type of probability distribution, of which the parameters were characterised using the volume and the whereabouts of the divided component. This shows that, unlike conventional cascade methods, Olsson (1998) does not apply the same generator parameters for overall cascade levels of interest. Moreover, the probabilities of $1/0$ and $0/1$ divisions were studied and applied to the generator to produce zero-value rainfall. They represent the condition that the volume of a certain component at level $g$, for example, is totally redistributed to one of the associated subcomponents at level $g+1$. Olsson (1998) used the rainfall data in Lund (southern Sweden) to demonstrate that the probability distribution of fragmentation ratios altered over different pairs of successive timescale levels, and the modelling results showed that certain statistical and scaling features of observed 1-h rainfall data are satisfactorily reproduced.

Whereas the random cascade models have been widely applied, deterministic cascade models are seldom discussed in the literatures. Sivakumar (2001) developed a simple chaotic model to disaggregate rainfall. The results suggest the presence of a deterministic chaotic process in rainfall, but this is controversial. Wang et al. (2009b), instead of calculating the probability distribution of $p_i^L$ values, developed a deterministic cascade method to directly estimate the values of $p_i^L$, which were further used to generate the rainfall sequences over the time scales of interest, based upon the theory of left-sided multifractals (Mandelbrot et al., 1990; Mandelbrot, 1990). This method is implemented by numerically solving the generating equation (Hentschel and Procaccia, 1983), which is theoretically derived by the self-similar feature of a cascade process,

$$\sum_{i=1}^{b} s_i^{-\tau(q)} p_i^q = 1, \tag{2}$$

where $s_i$ is the contracting ratio to subdivide timescales and usually is assumed to be $1/b$ for the convenience of practical use, and $\tau(q)$ is a scaling feature curve derived





HESSD
7, 5267–5297, 2010

from its associated multifractal spectrum (Evertsz and Mandelbrot, 1992). Although the results showed that this method was unable consistently to reproduce specific statistics of real rainfall observations due to its oversimplified procedures of generating rainfall sequences in a deterministic scheme and its controversial assumption that scale-invariance exists over whole timescales of interest (5 min to 1 day), some breakthrough was made. For example, Wang et al. (2009b) demonstrated that a quad-nomial ($b$=4) cascade method had the potential to generate complex rainfall patterns more satisfactorily at the high resolution, including the production of more diverse rainfall patterns and zero-value precipitation, than the conventional binomial ($b$=2) cascade-based methods.

### 4.2 Methodology

Based upon Olsson (1998) and Wang et al. (2009b), the proposed methodology is developed by re-investigating the construction process of multinomial multiplicative cascades. A quad-nomial cascade is thus used, aiming to enhance the complexity of rainfall patterns conventionally generated by binomial cascade-based methods; and a general structure of cascade is used, where the weight distribution of fragmentation ratios for each pair of successive timescale levels is empirically estimated. However, instead of assuming these empirical weights to fit the associated theoretical probability distribution as Olsson (1998), this work directly estimates the "optimal" set of fragmentation ratios for each pair of branching timescale levels, and applies these optimal ratios to the reproduction of rainfall sequences via the random disaggregation process over the timescale of interest.

The proposed methodology is implemented in two steps. In the first step, the fragmentation ratios are deduced by analysing their weight distribution over the timescales of interest. In this work, the targeted scales range from hourly to sub-hourly. At the second step, the derived optimal fragmentation ratios are used to carry out stochastic rainfall disaggregation and an ensemble scheme is further applied to generating deterministic sub-hourly (e.g., 5-min) rainfall sequences.





HESSD
7, 5267–5297, 2010

### 4.2.1 Derivation of fragmentation ratios

Figure 2 is the schematic of a specific pair of successive cascade levels (level $g-1$ and $g$). An arbitrary component at level $g-1$ (the parent level) can be formed, for example, as the product:

$$V_i^{g-1} = \underbrace{p_i^1 p_i^2 \cdots p_i^{g-1}}_{\text{total of } g-1 \text{ ratios}}, \tag{3}$$

and an associated sub-component can be formed as:

$$V_j^g = \underbrace{p_i^1 p_i^2 \cdots p_i^{g-1} p_j^g}_{\text{total of } g \text{ ratios}}. \tag{4}$$

The associated fragmentation ratio at level $g$ thus is calculated by

$$p_j^g = \frac{V_j^g}{V_i^{g-1}} \tag{5}$$

when $V_i^{g-1}$ is non-zero.

By repeating this procedure, ratios for all components can be obtained and their weight (or termed empirical probability to be distinguished from the theoretical probability) distribution is illustrated as Fig. 3, of which the upper is estimated between the timescale levels of 80 and 20 min and the lower is for 20 and 5 min. It is found that the empirical probability distribution of fragmentation ratios for each position is evenly distributed. An assumption is therefore made herein that during the cascade process the order in which derived ratios are applied to generate subcomponents is fully random. Moreover, slight changes of weight distribution for two consecutive scale ranges are observed. This depicts that scaling features alter between these two timescale ranges and that the assumption of using the same generator parameters for whole





HESSD
7, 5267–5297, 2010

timescale range of interest, made in previous literature like Wang et al. (2009a), may be insufficient to charaterise the scaling behaviour of sub-daily rainfall sequences.

A set of optimal fragmentation ratios for each pair of successive levels is derived based upon the analysis of their empirical probability distribution shown in Fig. 3. The set of ratios, having the maximal probability to occur, is selected, which can be formed as:

$$\begin{cases} \text{obj. max} \sum_{i=1}^{b} \Pr(p_i^L) \\ \text{s.t.} \sum_{i=1}^{b} p_i^L = 1 \end{cases}, \tag{6}$$

where $\Pr(p_i^L)$ represents the associated empirical probability of a specific value of fragmentation ratios equal to $p_i^L$. In order to remain mass conserved, a constraint of $\sum_{i=1}^{b} p_i^L = 1$ is applied. The problem is therefore transformed to an optimisation question. In this work, the optimal sets of fragmentation ratios are deduced based upon the Genetic Algorithm (Goldberg, 1989), and a C++ library, GAlib, developed by Matthew Hall is used.

### 4.2.2 Rainfall disaggregation

The derived $p_i^L$ values for each timescale level are then used to carry out disaggregation. The optimal set of fragmentation ratios over levels $g-1$ and $g$, for example, is obtained and termed as $\mathbf{P}^g = (p_1^g, p_2^g, \ldots, p_b^g)$. Among $b$ sub-components at level $g$ of a specific parent component (at level $g-1$), each rainfall volume can be estimated by multiplying this parent component's rainfall volume by a value randomly and unrepeatedly selected from $\mathbf{P}^g$. Although the fragmentation ratios are deterministic for each pair of successive levels, the process to generate rainfall sequences is random (stochastic). This means that the generated rainfall sequences would be different each time the





HESSD
7, 5267–5297, 2010

disaggregation is carried out. However, it is inappropriate to provide stochastic rainfall information to the corresponding short-term flooding modelling due to the limited response time. An ensemble of rainfall disaggregations is thus produced to average a specific number of randomly generated rainfall sequences, based upon the proposed methodology, to synthesise high-resolution rainfall sequences.

The proposed methodology is validated in the following section using observed rainfall data for the Greenwich raingauge station in London.

## 5 Validation

The procedure of validation is to aggregate the observed 5-min rainfall data to the hourly scale; then, these aggregated data are disaggregated back to the 5-min sequence based upon the proposed methodology. The synthetic 5-min sequence generated by the log-Poisson distributed random cascade method is used herein as benchmark. Certain statistics over 5-, 10-, 20-, and 30-min timescales are then analysed (shown in Figs. 4–9) and compared to demonstrate that the proposed methodology, based upon non-random fragmentation ratios, has the ability to reproduce certain statistics as the log-Poisson distributed cascade method has. Due to the limitation of using a quad-nomial ($b$=4) cascade process, in the proposed methodology the observed rainfall data are aggregated to the timescale of 80-min ($\approx$1.3 h) and then disaggregated to the timescale of 5-min; while in the log-Poisson cascade method, due to the limitation of the sofeware used, the rainfall data are aggregated to the scale of 60-min (=1 h), then disaggregated to the 3.75-min, and then linearly interpolated to the scale of 5-min.

### 5.1 Log-poisson cascade methods

In this work, the *Cascade* computer programme for disaggregation (Onof, 2009) is used, in which the generator for log-Poisson distributed cascades (formed as Eq. 1) is



HESSD
7, 5267–5297, 2010

implemented. The scaling features of observed rainfall data is firstly investigated and the associated sturcture function ($K(q)$) is empirically estimated. By introducing the relation of $K(q)$ and the generator $W$, expressed as:

$$K(q) = -C\frac{q(1-\beta)+\beta^q-1}{\ln 2}. \tag{7}$$

The associated parameters ($\beta$ and $C$) for the generator then are derived, where the parameter $A$ in Eq. (1) can be obtained using $A = \exp^{C(1-\beta)}$. The optimal parameters based upon the monthly analyse are thus shown in Table 2.

### 5.2 Reproducibility of statistics

In Figs. 4–6, the dark solid lines are drawn based upon the statistics of observed rainfall data, the grey lines are based upon the proposed method (the dashed-dotted lines for the ensemble of 10 stochastic simulations, the solid for 30 simulations, and the dashed for 100 simulations) and the dark dashed lines are based upon the log-Poisson distributed cascade method. To be mentioned here, all rainfall depths (before they are estimated to rainfall intensity) less than 0.01 mm are regarded as 0.00 mm when the statistics are estimated.

In Fig. 4a–d, overestimation of standard deviation (Std-dev) from the simulation results based upon the log-Poisson cascade method is in general observed particularly in August and September, which are the very two months having largest Std-dev values of observed rainfall data; while underestimation is found based upon the proposed methodology. Among the latter, the performance of the estimates based upon the ensemble of 10 simulations is better than the other two ensemble scenarios, particularly when the 5-min rainfall intensity is studied, and the performance of ensembles gradually converges when the intensity with a longer time interval is investigated.

In Fig. 5a–d, both the log-Poisson cascade method and the proposed methodology with 10-simulation ensembles are observed having the ability to satisfactorily reproduce the lag-1 autocorrelation variables, particularly for 10- and 20-min rainfall intensity.





HESSD
7, 5267–5297, 2010

However, for the cases of 30- and 100-simulation ensembles, lag-1 autocorrelation variables are in general overestimated, and furthermore the monthly variation is not obvious. This may be because the variability of time series is smoothened while the ensemble number increases.

In Fig. 6a–d, the proportion of dry periods is investigated. Both the log-Poisson cascade method and the proposed methodology exhibit the promising ability to reproduce this variable, although it is somewhat underestimated. Moreover, the monthly variation of this variable is found to be shown, but similarly, for the cases of 30- and 100-simulation ensembles, the number of dry periods is less due to the smoothened time series.

In Fig. 7a–d, the MSE (Mean Squared Error) is used to quantify the difference between the observed rainfall data and the simulated rainfall sequences. The results show that the proposed methodology has the ability to generate the rainfall sequences of which the patterns are similar to the observed rainfall data. This is very crucial and useful for the corresponding short-term flood modelling.

### 5.3 Reproducibility of patterns

Since the aim of this work is to generate the high-resolution rainfall data that can be used as inputs for short-term surface flood modelling and warning, a further assessment of the ability to reproduce patterns of the observed rainfall data is carried out herein, particularly for extreme rainfall values. The method used in this work is to quantify the probability of that the simulated rainfall intensity (within a certain time interval) is larger than a specific threshold given that the associated observed rainfall intensity is larger than this threshold. In other words, the ability of the proposed methodology to timely capture short-term rainfall patterns is evaluated. The Hit Rate ($H$), which has been widely used in forecast verification, is thus employed. The definition of $H$ can be explained using Table 4, and formed as (Mason, 2002),

$$H = \frac{a}{a+c}. \tag{8}$$





HESSD
7, 5267–5297, 2010

In Fig. 8a–d, a lower threshold of 0.1 Wet Mean (≈0.2–0.4 mm/h) is firstly used, and it is investigated that the Hit Rates of rainfall sequences generated based upon the proposed methodology in general satisfy the benchmark obtained based upon log-Poisson cascade method, except the rainfall sequences generated by the proposed method with the ensemble of 10 simulations. In Fig. 9a–d, a higher threshold of 0.5 Wet Mean (≈1–2 mm/h, which is a somewhat heavy rainfall intensity range) is used, and the Hit Rates obtained based upon the proposed methodology also exhibit the ability to satisfy the benchmark. These results entail that in general the proposed methodology has acceptable abilities to capture the rainfall quantity in the right timing.

## 6 Conclusions

In this work, a semi-deterministic cascade methodology is developed for synthesising sub-daily rainfall sequences, which may be further used as inputs for the corresponding hydrological modelling. The construction of a multiplicative cascade process and its applications to rainfall modelling are detailed in this work to help identify the improvements that may be carried out in the state-of-the-art cascade-based methods. Based upon these understandings, the methodology is derived, and its further validation is carried out by comparing certain statistics of the generated rainfall sequences based upon the proposed method with those generated based upon the log-Poisson distributed random cascade method using the 5-min rainfall data from Greenwich raingauge station in London. The results entail that the proposed method has the ability to reconstruct the sub-daily rainfall sequences not only with feasible statistical features but also with similar geometric patterns. The latter is a seldom discussed issue in the literatures; however, it is crucial for the further use as outputs to carryout more reliable short-term flooding modelling and warning.

*Acknowledgements.* The research was conducted as part of the Flood Risk Management Research Consortium (FRMRC2, SWP3). The first author, Li-Pen Wang, would like to acknowledge the fully financial support from the Ministry of Education Taiwan for his postgraduate research programme.

## References


Bowler, N. E., Pierce, C. E., and Seed, A. W.: STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP, Q. J. Roy. Meteor. Soc., 132, 2127–2155, 2006. 5269

Chandler, R. E. and Wheater, H. S.: Analysis of rainfall variability using generalised linear models: a case study from the west of Ireland, Water Resour. Res., 38, 1192–1201, 2002. 5270

Deidda, R.: Rainfall downscaling in a space-time multifractal framework, Water Resour. Res., 36, 1779–1794, 2000. 5272

Deidda, R., Benzi, R., and Siccardi, F.: Multifractal modeling of anomalous scaling laws in rainfall, Water Resour. Res., 35, 1853–1867, 1999. 5272, 5273

EA: Environment Agency: Review of 2007 Summer Floods, 2007. 5268

Evertsz, C. J. G. and Mandelbrot, B. B.: Multifractal measures, in: Chaos and Fractals, edited by: Peitgen, H.-O., Jurgens, H., and Saupe, D., 922–953, Springer, New York, 1992. 5275

Goldberg, D. E.: Genetic Algorithms in Search, Optimization, and Machine Learning, Kluwer Academic Publishers, Boston, MA, 1989. 5277

Golding, B. W.: Nimrod: a system for generating automated very short range forecasts, Meteorol. Appl., 5, 1–16, 1998. 5269

Gupta, V. K. and Waymire, E. C.: A Statistical analysis of mesoscale rainfall as a random cascade, J. Appl. Meteorol., 32, 251–267, 1993. 5272

Hentschel, H. G. E. and Procaccia, I.: The infinite number of generalized dimensions of fractals and strange attractors, Physica D, 8, 435–444, 1983. 5274

Koutsoyiannis, D.: Rainfall disaggregation methods: theory and applications, in: Workshop on Statistical and Mathematical Methods for Hydrological Analysis, Rome, 2003. 5270

Koutsoyiannis, D. and Onof, C.: Rainfall disaggregation using adjusting procedures on a poisson cluster model, J. Hydrol., 246, 109–122, 2001. 5270

Lovejoy, S. and Schertzer, D.: Multifractals, universality classes and satellite and radar measurements of cloud and rain fields, J. Geophys. Res., 95, 2021–2034, doi:10.1029/JD095iD03p02021, 1990. 5273

Mandelbrot, B. B.: Multifractal measures, especially for the geophysicist, Pure Appl. Geophys., 131, 5–42, 1989. 5272

Mandelbrot, B. B.: New "anomalous" multiplicative multifractals: left sided f(alpha) and the

modelling of DLA, Physica A, 168, 95–111, 1990. 5274
Mandelbrot, B. B., Evertsz, C. J. G., and Hayakawa, Y.: Exactly self-similar left-sided multifractal measures, Phys. Rev. A, 42, 4528, 1990. 5274
Mason, I. B.: Binary events, in: Forecast Verification: a Practitioner's Guide in Atmospheric Science, edited by: Jolliffe, I. T. and Stephenson, D. B., Wiley, West Sussex, 37–76, 2002. 5280
Molnar, P. and Burlando, P.: Preservation of rainfall properties in stochastic disaggregation by a simple random cascade model, Atmos. Res., 77, 137–151, 2005. 5270, 5272
Olsson, J.: Evaluation of a scaling cascade model for temporal rain-fall disaggregation, Hydrol. Earth Syst. Sci., 2, 19–30, doi:10.5194/hess-2-19-1998, 1998. 5274, 5275
Onof, C.: Cascade computer programme for disaggregation, Imperial College London, 2009. 5278
Onof, C. and Arnbjerg-Nielsen, K.: Quantification of anticipated future changes in high resolution design rainfall for urban areas, Atmos. Res., 92, 350–363, 2009. 5270, 5271, 5272
Onof, C. and Wheater, H. S.: Modelling of British rainfall using a random parameter Bartlett-Lewis rectangular pulse model, J. Hydrol., 149, 67–95, 1993. 5270
Onof, C. and Wheater, H. S.: Improvements to the modelling of British rainfall using a modified random parameter Bartlett-Lewis Rectangular Pulse Model, J. Hydrol., 157, 177–195, 1994. 5270
Onof, C., Chandler, R. E., Kakou, A., Northrop, P., Wheater, H. S., and Isham, V.: Rainfall modelling using Poisson-cluster processes: a review of developments, Stoch. Env. Res. Risk A., 14, 384–411, 2000. 5270
Onof, C., Townend, J., and Kee, R.: Comparison of two hourly to 5-min rainfall disaggregators, Atmos. Res., 77, 176–187, 2005. 5272, 5273
Over, T. M. and Gupta, V. K.: Statistical analysis of Mesoscale rainfall: dependence of a random cascade generator on large-scale forcing, J. Appl. Meteorol., 33, 1526–1542, 1994. 5272, 5273
Over, T. M. and Gupta, V. K.: A space-time theory of mesoscale rainfall using random cascades, J. Geophys. Res., 101, 26319–26331, 1996. 5270, 5272, 5273
Pathirana, A. and Herath, S.: Multifractal modelling and simulation of rain fields exhibiting spatial heterogeneity, Hydrol. Earth Syst. Sci., 6, 695–708, doi:10.5194/hess-6-695-2002, 2002. 5272





HESSD
7, 5267–5297, 2010

Reconstruct sub-daily rainfall using multinomial cascades

L. Wang et al.

Pathirana, A., Herath, S., and Yamada, T.: Estimating rainfall distributions at high temporal resolutions using a multifractal model, Hydrol. Earth Syst. Sci., 7, 668–679, doi:10.5194/hess-7-668-2003, 2003. 5270, 5272

Pierce, C. E.: The short term ensemble prediction system: a guide to model design and performance for the Wet Bench, Tech. rep., Met Office Joint Center for Hydro-Meteorological Research, 2009. 5269

Pitt, M.: Lessons from the 2007 Floods: what people need, Tech. rep., 2008. 5268

Segond, M.-L., Onof, C., and Wheater, H. S.: Spatial-temporal disaggregation of daily rainfall from a generalized linear model, J. Hydrol., 331, 647–689, 2006. 5270

Shrestha, R., Tachikawa, Y., and Takara, K.: Downscaling spatial rainfall field from global scale to local scale using improved multiplicative random cascade method, Disaster Prevention Research Institute Annuals, Kyoto University, 47B, 235–251, 2004. 5272

Sivakumar, B.: Is a chaotic multi-fractal approach for rainfall possible?, Hydrol. Process., 15, 943–955, 2001. 5274

Sokol, Z.: Nowcasting of 1-h precipitation using radar and NWP data, J. Hydrol., 328, 200–211, 2006. 5268, 5269

Stern, R. D. and Coe, R.: A model fitting analysis of daily rainfall data, J. R. Stat. Soc., 147, 1–34, 1984. 5270

Tessier, Y., Lovejoy, S., and Schertzer, D.: Universal multifractals: theory and observations for rain and clouds, J. Appl. Meteorol., 32, 223–250, 1993. 5271

Wang, L. P., Maksimovic, C., and Onof, C.: An improved multifractal model for short-term rainfall forecast, in: Computing and Control in the Water Industry 2009, University of Sheffield, UK, 2009a. 5277

Wang, L. P., Maksimovic, C., and Onof, C.: A deterministic multiplicative cascade method for sub-daily rainfall forecast, in: Proceedings of the 8th International Workshop on Precipitation in Urban Areas, edited by: Molnar, P., Burlando, P., and Einfalt, T., St. Moritz, Switzerland, 2009b. 5274, 5275

Yang, C., Chandler, R. E., Isham, V. S., and Wheater, H. S.: Spatial-temporal rainfall simulation using generalised linear models, Water Resour. Res., 41, 1–13, 2005. 5270





HESSD

7, 5267–5297, 2010

**Reconstruct sub-daily rainfall using multinomial cascades**

L. Wang et al.

**Table 1.** Summary monthly statistics of 5-min rainfall intensity (mm/h) for Greenwich raingauge station.

| Month | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mean | 0.066 | 0.055 | 0.047 | 0.071 | 0.051 | 0.064 | 0.051 | 0.061 | 0.078 | 0.096 | 0.066 | 0.068 |
| Std-dev | 0.474 | 0.434 | 0.402 | 0.515 | 0.628 | 0.764 | 0.629 | 0.901 | 0.869 | 0.680 | 0.528 | 0.525 |
| Dry Prop. | 97.6 | 98.0 | 98.3 | 97.5 | 98.5 | 98.3 | 98.6 | 98.6 | 97.9 | 97.0 | 97.8 | 97.7 |
| $r_1$ | 0.455 | 0.438 | 0.401 | 0.466 | 0.505 | 0.580 | 0.539 | 0.698 | 0.676 | 0.556 | 0.536 | 0.473 |
| $r_2$ | 0.447 | 0.410 | 0.383 | 0.423 | 0.314 | 0.323 | 0.354 | 0.459 | 0.544 | 0.452 | 0.454 | 0.454 |
| Wet Mean | 2.794 | 2.754 | 2.724 | 2.852 | 3.305 | 3.818 | 3.762 | 4.361 | 3.694 | 3.235 | 3.001 | 2.900 |

Std-dev for standard deviation.

$r_1$ and $r_2$ respectively for lag-1 and 2 autocorrelation.

Dry Prop. for the proportion of dry volumes.





HESSD

7, 5267–5297, 2010

Reconstruct sub-daily rainfall using multinomial cascades

L. Wang et al.

HESSD
7, 5267–5297, 2010

Reconstruct sub-daily rainfall using multinomial cascades

L. Wang et al.





**Table 2.** Optimal parameters for Greenwich, where $\beta$ and $C$ are respectively the constant and scale parameters for the log-Poisson based random cascade generator.

| Month | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\beta$ | 0.174 | 0.167 | 0.276 | 0.274 | 0.325 | 0.248 | 0.192 | 0.181 | 0.301 | 0.276 | 0.237 | 0.257 |
| $C$ | 0.424 | 0.473 | 0.597 | 0.565 | 0.826 | 0.715 | 0.624 | 0.645 | 0.826 | 0.532 | 0.576 | 0.508 |

**Table 3.** Optimal fragmentation ratios (**P**) for Greenwich. A quad-nomial cascade on the 2-level hierarchy of time scales (80–20 min and 20–5 min) is used in this work.

| Level | Time Scale | Fragmentation Ratios (**P**) |
|---|---|---|
| 1 | 80–20 min | 0.0000, 0.0078, 0.0157, 0.9765 |
| 2 | 20–5 min | 0.0039, 0.0118, 0.0274, 0.9569 |





HESSD

7, 5267–5297, 2010

**Table 4.** Contingency table for categorical forecasts of a binery event. The numbers of observations in each category are represented by $a$, $b$, $c$ and $d$, and $n$ is the total.

| Forecast | Observed | | |
|---|---|---|---|
| | Yes | No | Total |
| Yes | $a$ | $b$ | $a+b$ |
| No | $c$ | $d$ | $c+d$ |
| Total | $a+c$ | $b+d$ | $a+b+c+d=n$ |





HESSD
7, 5267–5297, 2010

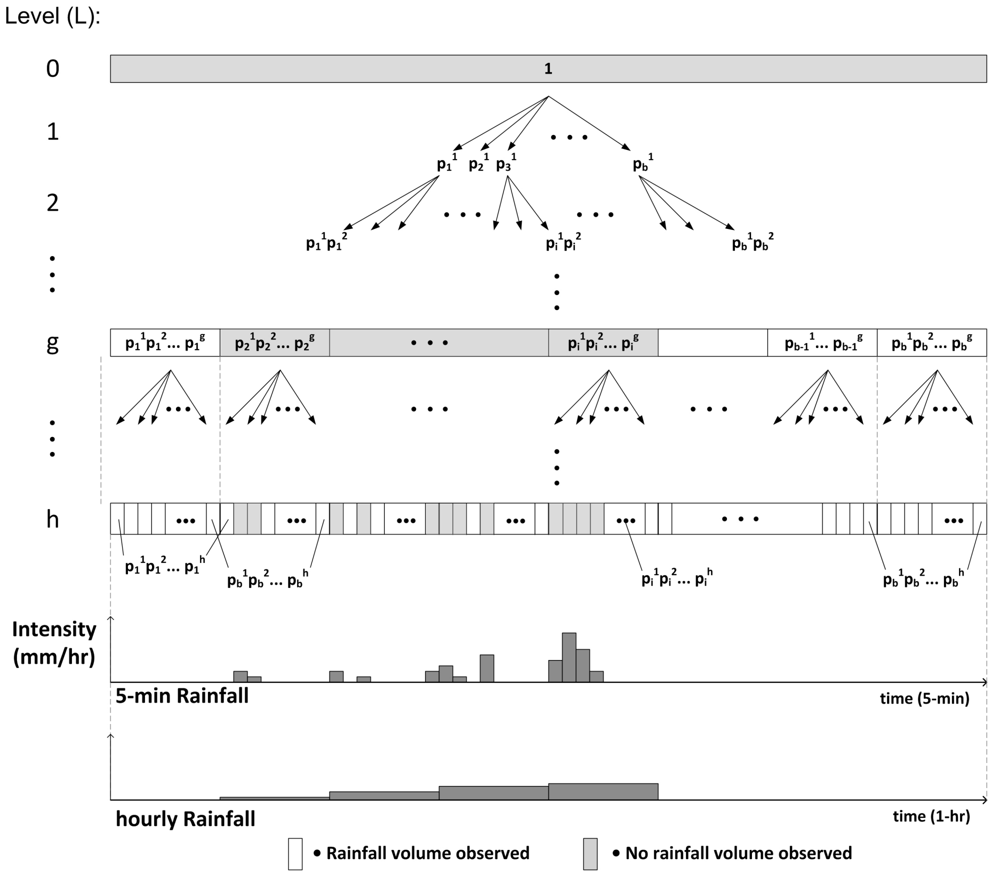


**Fig. 1.** Schematic of a multinomial ($b \geq 2$) cascade and the associated mappings to rainfall sequences with different resolution (e.g., the cascade at level $g$ for the 1-day rainfall sequence; level $h$ for the 5-min ones).





HESSD
7, 5267–5297, 2010

Reconstruct sub-daily rainfall using multinomial cascades

L. Wang et al.

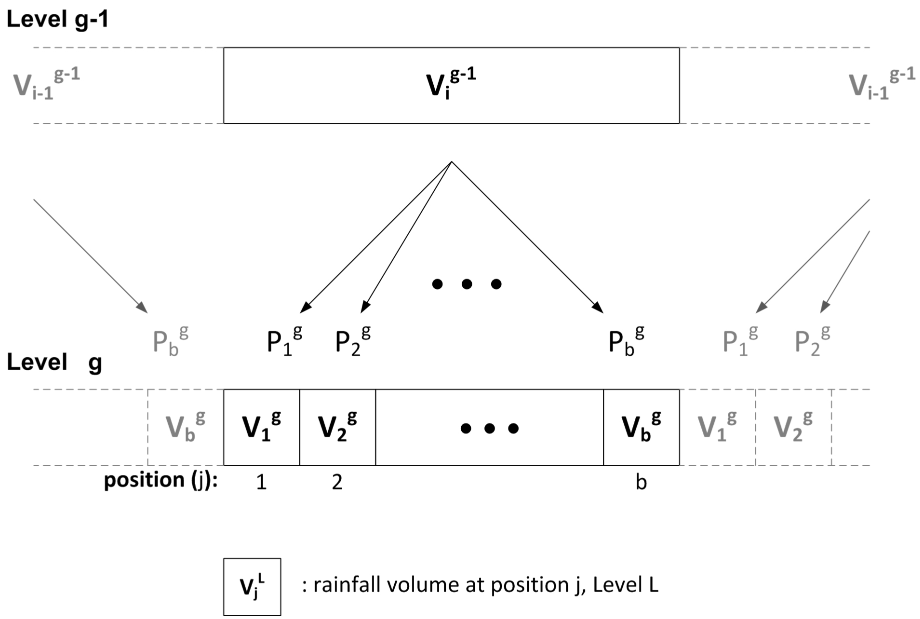


**Fig. 2.** A close look of a multinomial cascade over levels $g-1$ and $g$.



HESSD
7, 5267–5297, 2010

Reconstruct sub-daily rainfall using multinomial cascades

L. Wang et al.

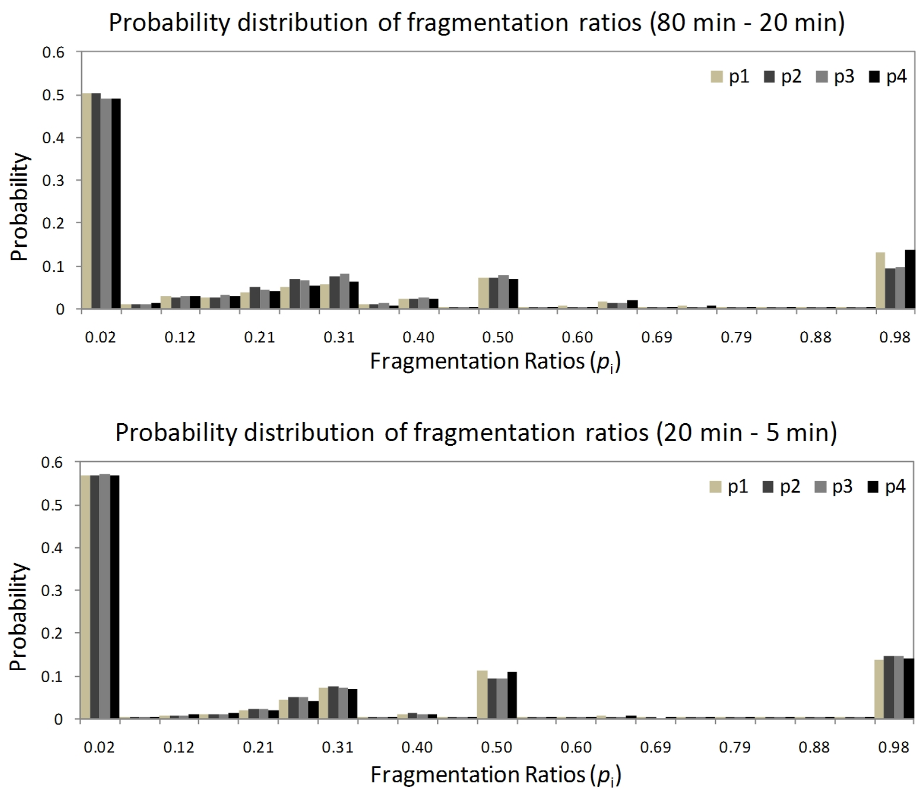


**Fig. 3.** Probability (weight) distribution of fragmentation ratios between successive timescale levels of interest for Greenwich (the upper: 80–20 min; and the lower: 20–5 min).

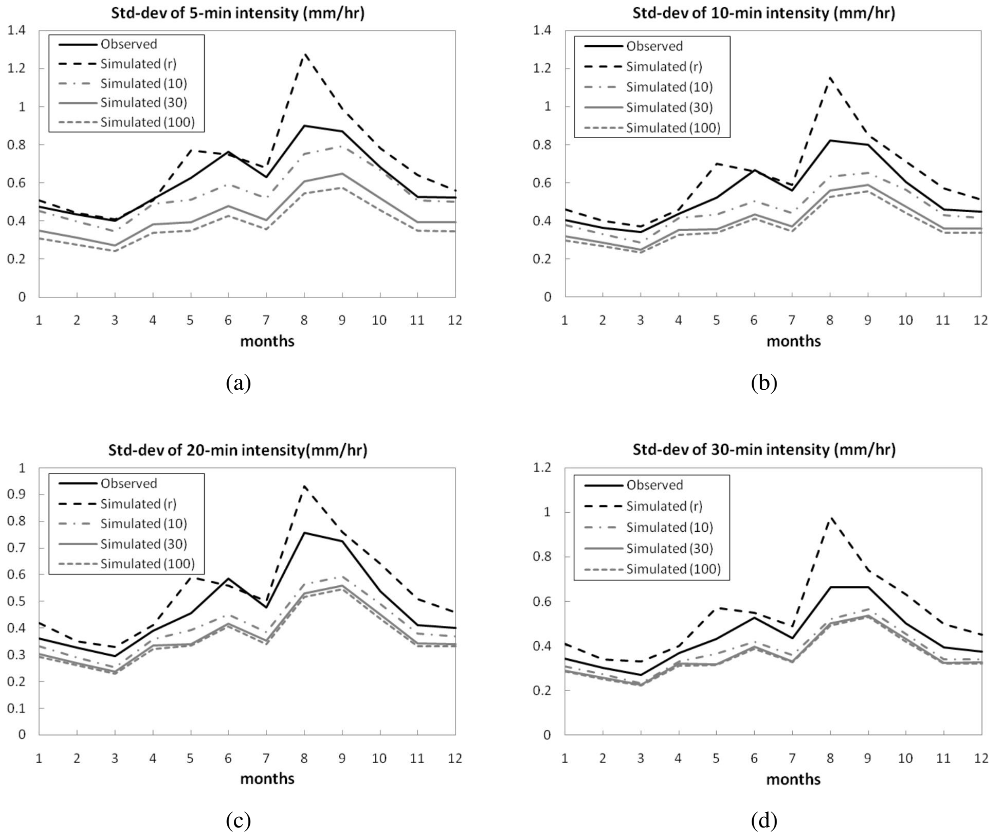


**Fig. 4.** Standard deviation of observed (the dark solid lines) and simulated 5 **(a)**, 10 **(b)**, 20 **(c)** and 30 **(d)** min rainfall intensities (mm/h), respectively based upon the proposed methodology (the grey lines) and log-Poisson based random cascade method (the dark dashed lines).





HESSD
7, 5267–5297, 2010

Reconstruct sub-daily rainfall using multinomial cascades

L. Wang et al.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

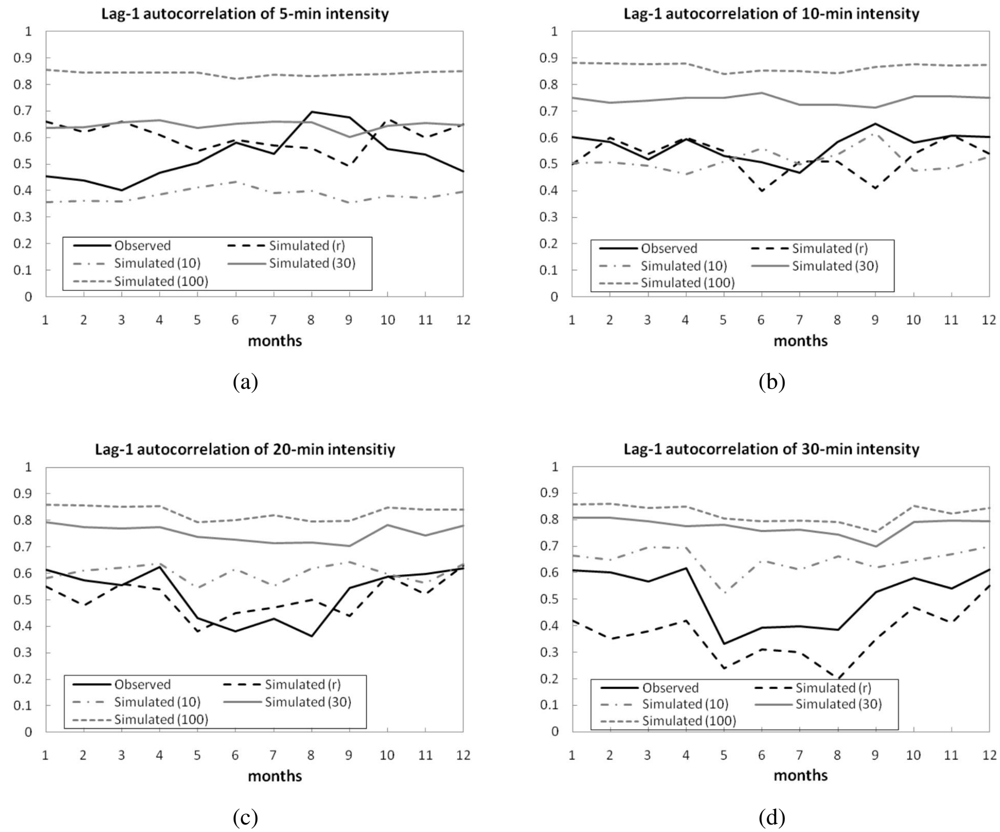


**Fig. 5.** Lag-1 autocorrelation of observed (the dark solid lines) and simulated 5 **(a)**, 10 **(b)**, 20 **(c)** and 30 **(d)** min rainfall intensities (mm/h), respectively based upon the proposed methodology (the grey lines) and log-Poisson based random cascade method (the dark dashed lines).





HESSD
7, 5267–5297, 2010

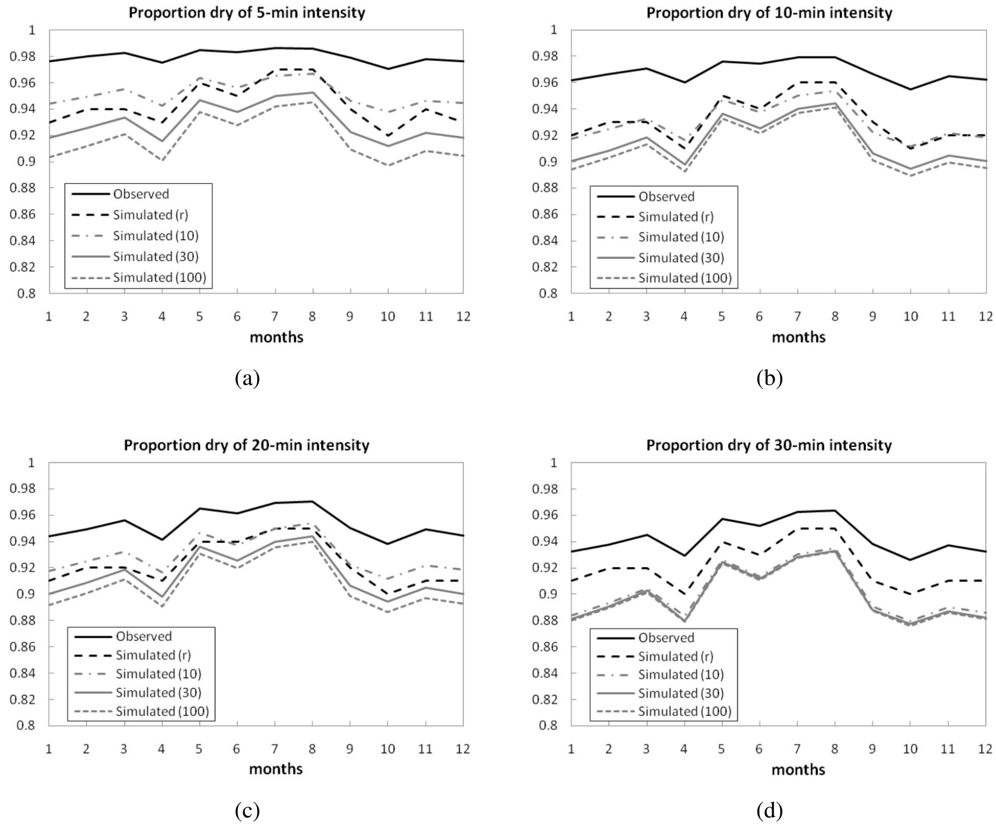


**Fig. 6.** Proportion of dry periods of observed (the dark solid lines) and simulated 5 **(a)**, 10 **(b)**, 20 **(c)** and 30 **(d)** min rainfall intensities (mm/h), respectively based upon the proposed methodology (the grey lines) and log-Poisson based random cascade method (the dark dashed lines).





HESSD
7, 5267–5297, 2010

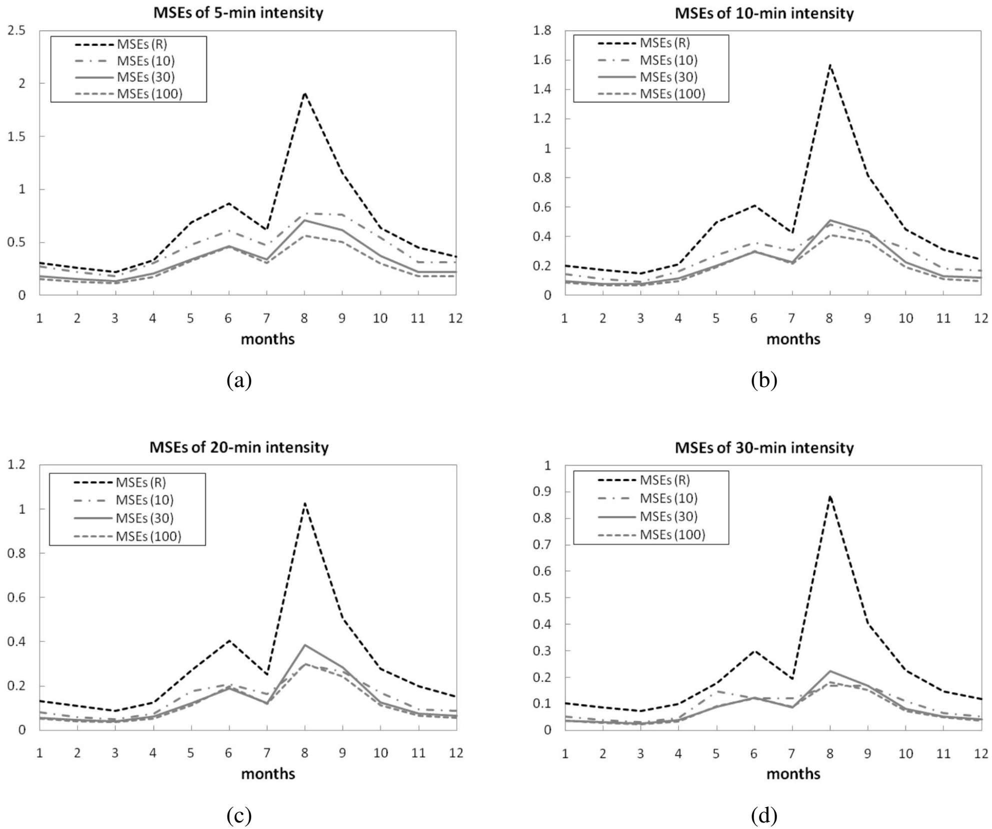


**Fig. 7.** MSEs (Mean Squared Errors) of observed (the dark solid lines) and simulated 5 **(a)**, 10 **(b)**, 20 **(c)** and 30 **(d)** min rainfall intensities (mm/h), respectively based upon the proposed methodology (the grey lines) and log-Poisson based random cascade method (the dark dashed lines).





HESSD
7, 5267–5297, 2010

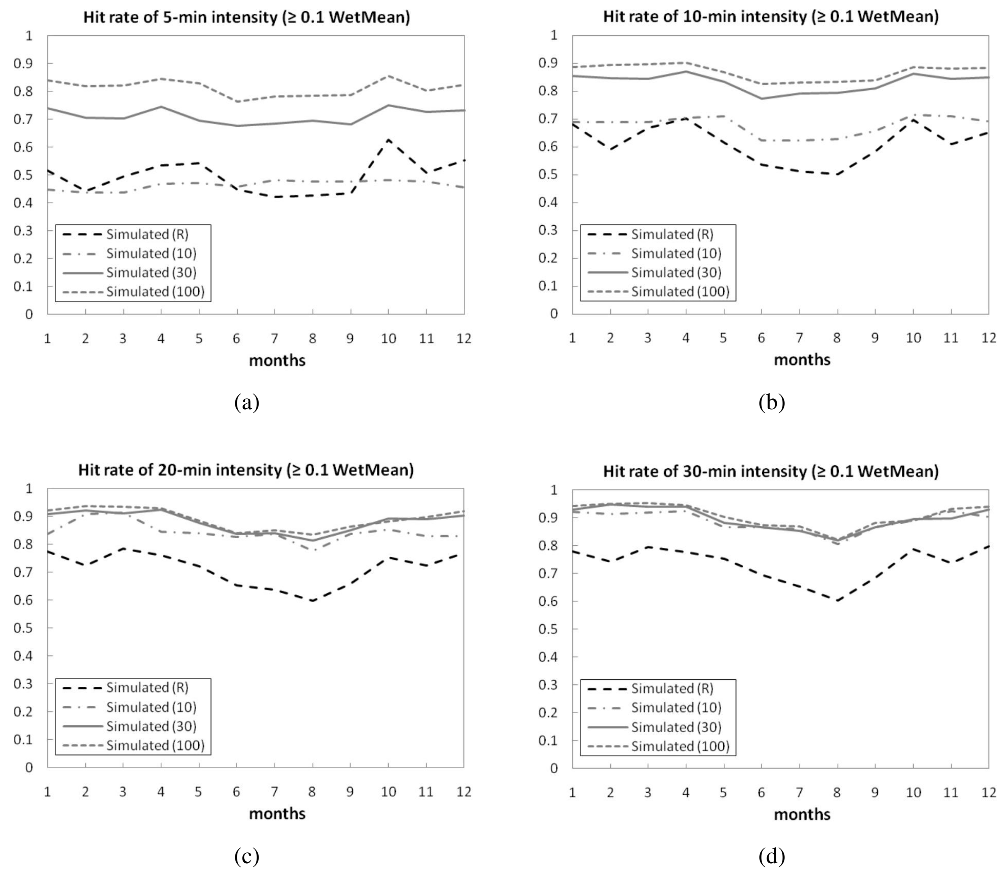


**Fig. 8.** Hit rates of simulated 5 **(a)**, 10 **(b)**, 20 **(c)** and 30 **(d)** min rainfall intensity (mm/h), respectively based upon the proposed methodology (the grey lines) and log-Poisson based random cascade method (the dark dashed lines), with the threshold of 0.1 Wet Mean.

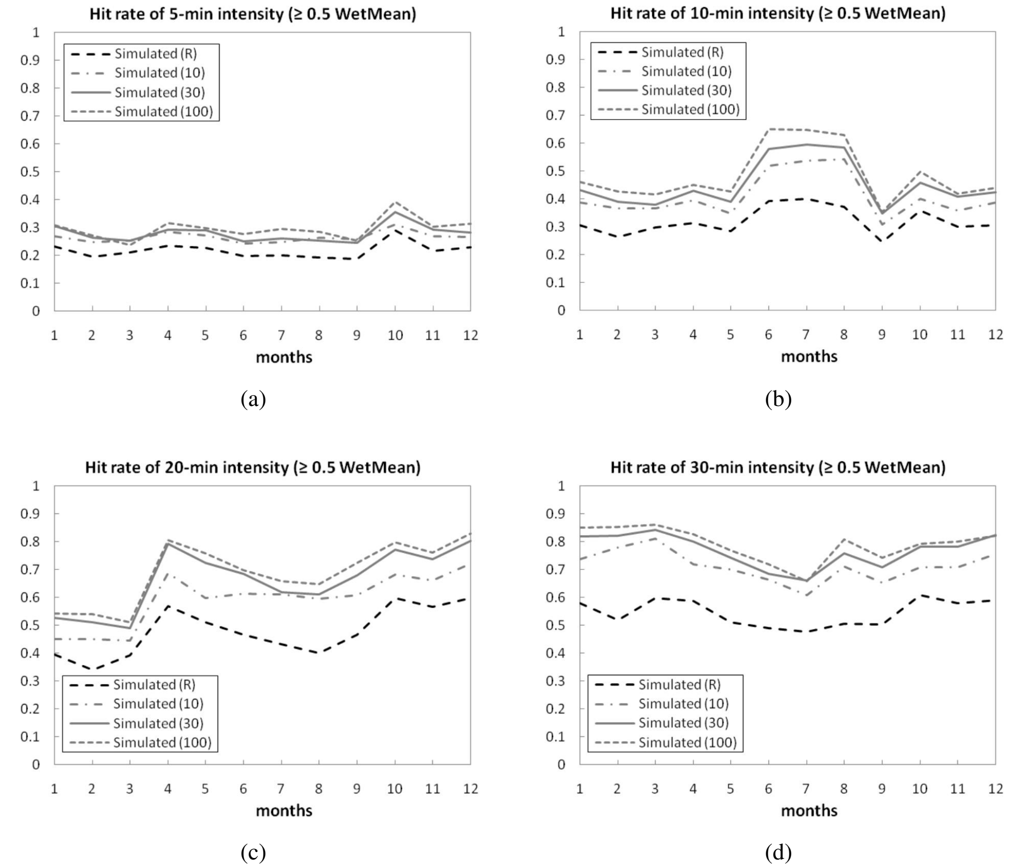


**Fig. 9.** Hit rates of simulated 5 **(a)**, 10 **(b)**, 20 **(c)** and 30 **(d)** min rainfall intensity (mm/h), respectively based upon the proposed methodology (the grey lines) and log-Poisson based random cascade method (the dark dashed lines), with the threshold of 0.5 Wet Mean.





HESSD
7, 5267–5297, 2010