Hydrol. Earth Syst. Sci. Discuss., 8, 7261–7291, 2011 www.hydrol-earth-syst-sci-discuss.net/8/7261/2011/ doi:10.5194/hessd-8-7261-2011 © Author(s) 2011. CC Attribution 3.0 License.



This discussion paper is/has been under review for the journal Hydrology and Earth System Sciences (HESS). Please refer to the corresponding final paper in HESS if available.

# Multiplicative cascade models for fine spatial downscaling of rainfall: parameterization with rain gauge data

D. E. Rupp<sup>1</sup>, P. Licznar<sup>2</sup>, W. Adamowski<sup>3</sup>, and M. Leśniewski<sup>3</sup>

<sup>1</sup>Cooperative Institute for Marine Resources Studies, Oregon State University, Hatfield Marine Science Center, Newport, Oregon, USA

<sup>2</sup>Institute of Environment Protection Engineering, Wrocław University of Technology, Wrocław, Poland

<sup>3</sup>Municipal Water Supply and Sewerage Company in Warsaw, PI. Starynkiewicza 5, 02-015 Warsaw, Poland

Received: 14 July 2011 - Accepted: 19 July 2011 - Published: 25 July 2011

Correspondence to: D. E. Rupp (david.rupp@oregonstate.edu)

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



# Abstract

Capturing the spatial distribution of high-intensity rainfall over short-time intervals is critical for accurately assessing the efficacy of urban stormwater drainage systems. In a stochastic simulation framework, one method of generating realistic rainfall fields is by <sup>5</sup> multiplicative random cascade (MRC) models. Estimation of MRC model parameters has typically relied on radar imagery or, less frequently, rainfall fields interpolated from dense rain gauge networks. However, such data are not always available. Furthermore, the literature is lacking estimation procedures for spatially incomplete datasets. Therefore, we proposed a simple method of calibrating an MRC model when only data from a moderately dense network of rain gauges are available, rather than from the

- from a moderately dense network of rain gauges are available, rather than from the full rainfall field. The number of gauges need only be sufficient to adequately estimate the variance in the ratio of the rain rate at the rain gauges to the areal average rain rate across the entire spatial domain. In our example for Warsaw, Poland, we used 25 gauges over an area of approximately 1600 km<sup>2</sup>. MRC models calibrated using the
- <sup>15</sup> proposed method were used to downscale 15-min rainfall rates from a 20 by 20 km area to the scale of the rain gauge capture area. Frequency distributions of observed and simulated 15-min rainfall at the gauge scale were very similar. Moreover, the spatial covariance structure of rainfall rates, as characterized by the semivariogram, was reproduced after allowing the probability density of the random cascade generator to vary with spatial scale.

## 1 Introduction

25

Urban catchments, due to their diminished damping properties relative to rural and natural catchments, are particularly responsive to bursts of local, high intensity rainfall. This makes characterization of the spatial distribution of rainfall at small time scales critical to evaluating the efficacy of urban stormwater drainage systems. Traditionally, design storms have been used to evaluate these systems in conjunction with



rainfall-runoff and hydrodynamic models, but in recent years there has been a push towards stochastically downscaling long (e.g., multi-decadal) time series of coarse (e.g., daily) rainfall to higher resolution (e.g., minutes) with which to force models of stormwater drainage systems (e.g., Hingray and Ben Haha, 2005; Molnar and Burlando, 2005;

Licznar et al., 2011a). Advantages of using long time series are that they allow for 5 a statistical analysis of system performance and they eliminate the problem of defining the appropriate initial catchment water storage for a design storm (Hingray and Ben Haha, 2005). Furthermore, long time series of daily rainfall are already abundant and readily available, and time series of high-resolution rainfall with which to develop downscaling models are becoming more prevalent. 10

While using long time series data provides advantages, their still remains the issue that the rainfall field is continuously evolving through time. While one might simplify the problem by using a predefined and static dimensionless rainfall field, this takes away a key strength of the stochastic simulation approach. An alternative is to stochastically downscale the rainfall field as well as the time series. For this purpose, a number 15 of models for stochastically downscaling rainfall fields have been developed. Following Ferraris et al. (2003), most can be grouped into three general types: autoregres-

sive models, point-process models that randomly position rainfall "cells", and fractal and multifractal cascade models. Additionally, there are hybrid models that combine features of these different approaches. For an overview of these various types, see 20 Ferraris et al. (2003) and references therein. We focus on multifractal cascade models because, as noted by Veneziano et al. (2006), multifractal models are simpler and have fewer parameters, and furthermore, though we do not consider these properties in this study, one can deduce the frequency distribution of rainfall intensities and rainfall extremes from their multi-fractal structure.

25

Parameter estimation for spatial downscaling models requires observations of the rainfall field. With multifractal cascade models, parameter estimation has mostly been done using radar-derived rainfall fields, though in a small number of cases rainfall fields were generated by interpolating rain gauge data (Svensson et al, 1996; Jothityangkoon



et al., 2000; Sharma et al., 2007). However, when the gauge density is coarse relative to the final spatial resolution of interest, the interpolation methods will fail because they smooth out the fine-scale variability.

In the absence of calibrated radar-derived rainfall scans or a very dense rain gauge network, the literature is lacking parameter estimation techniques for the applied hydrologist. Therefore, we propose a simple method of calibrating a multifractal cascade model for generating rainfall fields of short-duration rainfall (e.g., 15 min) when information across the full field is not available, or specifically, when only data from a network of rain gauges is available. The number of gauges need only be sufficient to adequately estimate the variance in the ratio of the rain rate at the rain gauges to the areal average rain rate across the entire spatial domain. The particular number will depend on the degree of spatially variability across the domain of interest.

In this study, we do not consider temporal evolution of the rainfall fields, which is required for a complete space-time downscaling model. Various cascade-based space-

time models have been proposed (e.g., Over and Gupta, 1996; Venugopal et al., 1999; Deidda, 2000; Jothityangkoon et al., 2000; Kang and Ramirez, 2010). Parameterization of a space-time model will be a topic of a subsequent paper.

# 2 Data and methodology

# 2.1 Data

Rainfall data were collected from a network of 25 rain gauges distributed throughout Warsaw, Poland. The gauges were installed by Warsaw Waterworks in the fall of 2008 to better characterize storm systems with the specific objective of modeling combined sewer- stormwater systems. Individual gauges were located to obtain best representative meteorological observations in urban settings (Oke, 2006) and to have approximately constant gauge density over the entire city (Fig. 1). The gauges were connected to a single data acquisition system by means of general packet radio service



(GPRS) modems. The data used in this study were recorded with a temporal resolution of 1 min and covers the period from the 38th week of year 2008 up to the 49th week of year 2010.

All gauges were weighing-type instruments suitable for both liquid and solid precipitation (MPS systém Ltd., model TRwS 200E). The manufacturer's claimed accuracy was 0.1 % and the resolution was 0.001 mm. Field tests of the installed gauges were conducted prior to operational use. Good agreement between total depth of known and recorded precipitation was observed. However, at a 1-min resolution, the output signal was detectably damper and broader than the input signal. As a consequence, rain was at times still being recorded for up to a few minutes after water was no longer being added to the gauge funnel. To reduce the relative error caused by this modulation of the signal, we aggregated the data to 15-min intervals.

#### 2.2 Spatial downscaling model

Our downscaling model is based on a discrete multiplicative random cascade (MRC). In the discrete MRC model of rainfall fields, the small-scale rainfall per unit area in a square cell  $\Delta$  at the *n*th cascade level is given by

$$R_n(\Delta_{n,k}) = R_0 \prod_{j=1}^n W_{j,k}$$
(1)

where the area of  $\Delta_n$  is given by  $R_0 L_0^2 b^{-n}$ . Here the large-scale rainfall  $R_0$  is the rainfall amount per unit area over the host cell with area  $L_0^2$ . The constant *b* is the branching number, or number of sub-cells (in our case, 4) into which rainfall from a cell is partitioned at the next level in the cascade (Fig. 1). For each level, the index pair (*j*, *k*) represents the cell along the path to the *n*-th level cell. The cells at the *n*-th cascade level are indexed by  $\Delta_{n,k}$ ,  $k = 1, 2, ..., 4^n$  (see Over and Gupta, 1996). The cascade weight *W* is a random variable with a prescribed distribution function, of which various



have been proposed in the context of rainfall (e.g., Schertzer and Lovejoy, 1987; Gupta and Waymire, 1993; Over and Gupta, 1996; Deidda et al., 1999; Ahrens, 2003).

The weight W is generated as a random quantity with the following probability density:

5 P(W = 0) = 1 - p

 $P(W = p^{-1}W^+) = p$ 

where *P* denotes probability, *p* is a parameter and  $W^+$  are the non-zero (positive) weights (Over and Gupta, 1994). Eqs. (2a) and (2b) comprise the cascade generator: Eq. (2a) generates the intermittency in the rainfall field (subareas of zero rainfall), while Eq. (2b) generates the rainfall volumes greater than zero.

The non-zero weights  $W^+$  have a log-stable density, which is to say that  $X = \ln(W^+)$  has a stable distribution with four parameters: the stability index  $0 < \alpha = 2$ , the skewness parameter  $-1 \le \beta = 1$ , the scale parameter  $\sigma > 0$ , and the shift parameter  $-\infty < \mu < \infty$ . We denote the stable distribution by  $S(\alpha, \beta, \sigma, \mu)$ . While the shift parameter can be defined in several ways, we follow the definition as given in Samorodnitsky and Taqqu (1994). Properties of stable distributions in the context of multifractal rainfall fields have been discussed by Schertzer and Lovejoy (1987), Lovejoy and Schertzer (1990), and Gupta and Waymire (1990, 1993), for example.

To ensure that the moments of  $W^+$  are finite, we set  $\beta = -1$  (Samorodnitsky and Taqqu, 1994). Furthermore, to conserve mass, on average, throughout the entire cascade process, we impose the condition that  $E[W^+] = 1$ . This means that

 $\mu = \sigma^{\alpha} \sec(\pi \alpha/2)$ 

15

(McCulloch, 1996), which leaves two free parameters  $\alpha$  and  $\sigma$  to describe the distribution. When  $\alpha = 2$ , distribution becomes normal with mean  $\mu$  and variance  $\sigma_N^2$ , where  $\sigma_N^2 = \sigma_N / \sqrt{2}$ . Discussion Paper **HESSD** 8, 7261-7291, 2011 Parameterization with rain gauge data D. E. Rupp et al. Discussion Paper **Title Page** Abstract Introduction References Conclusions **Discussion** Paper **Figures** Tables 14 Back Close Full Screen / Esc **Discussion Paper Printer-friendly Version** Interactive Discussion

(2a)

(2b)

(3)

## 2.3 Parameter estimation

Typically, estimation of spatial cascade model parameters relies on an analysis of the spatial scaling of the statistical moments of the observed rainfall quantities (e.g., Over and Gupta, 1996; Deidda, 2000; Jothityangkoon et al., 2000; Pathirana and Herath, 2002; Sharma et al., 2007; Kang and Ramirez, 2010). The *q*-th moment *M* at each spatial scale  $\lambda$  is calculated as

$$M(\lambda_n, q) = \sum_{k} \left[ R_n \left( \Delta_{n,k} \right) \right]^q \tag{4}$$

where the spatial scale  $\lambda_n$  is given by  $L_n/L_0$ . Rainfall quantities  $R_n(\Delta_{n,k})$  at a particular scale  $\lambda_n$  are determined by aggregating observed rainfall into grids with cells of area  $\Delta_n$ . The relationship between the moments and scale is made through log-log plots of  $M(\lambda_n, q)$  versus  $\lambda_n$  for various q. Linearity of the individual moments versus scale in log-log space implies either mono- or multifractality. The moment-scaling behavior of a fractal field has the form

 $M(\lambda_n, q) = (\lambda)^{\tau(q)}$ 

- where  $\tau(q)$  versus q is either a line (monofractal) or a curve (multifractal). Finally, the parameters of the cascade generator are estimated by fitting a distribution-dependent theoretical function to the empirical relationship  $\tau(q)$ . For examples on how the moment-scaling estimation method would be applied to the MRC model such as the one described in Sect. 2.2, see Pathirana and Herath (2002) and Serinaldi (2010).
- <sup>20</sup> The above estimation method requires observed quantities of  $R_n$  across a range of scales  $\lambda_n$ . Unfortunately, we are hindered by a low gauge density (~0.25 gauges km<sup>-2</sup>) relative to a desirable grid cell density (on the order of 10<sup>3</sup> cells km<sup>-2</sup>, or a resolution of 30 by 30 m).

If we imposed fine resolution grids over our gauge network, very few cells would contain enough gauges to adequately estimate the areal-average rain rate for those cells.

(5)

Though we could use deterministic spatial interpolation techniques (e.g., Thiessen polygons) to estimate the rainfall everywhere at every cell in the grid, this would likely result in a much too smooth rainfall surface.

Because of our inability to carry out a reliable analysis of moment scaling in space,
we assumed a priori that there is power law scaling of the statistical moments. We based this assumption on previous observations of multifractality in rainfall fields for spatial scales under 30 km (e.g., Kumar and Foufoula-Georgiou, 1993a, b; Perica and Foufoula-Georgiou, 1996; Pathirana and Herath, 2002; Kang and Ramirez, 2010). Multifractality implies that realistic rainfall fields could be reasonably reproduced, in a statistical sense, by a family of parsimonious multiplicative random cascade models.

As an alternative to moment-scaling analysis, we parameterized our model using only the final product of the weights  $W_1$  through  $W_n$ , which we express by the variable Y as

$$Y_n(\Delta_{n,k}) = \prod_{j=1}^n W_{j,k}$$

<sup>15</sup> We consider the cases for Y = 0 and Y > 0 separately.

From Eq. (2b),  $p_j$  is the probability that  $W_j > 0$  along a path in any j of n cascade levels. The probability that Y = 0 (which is to say that at least one  $W_j$  equals zero along the path down all n levels) can be calculated as 1 minus the probability that  $W_j$  is greater than zero in all n levels. From the binomial distribution function (Ross, 1998) we obtain the solution for the probability that Y = 0:

$$P(Y=0) = 1 - \prod_{j=1}^{n} p_j$$
(7)

To help us determine the distribution of Y when Y > 0, we defined the variable  $Y^+$ :

$$Y_n^+(\Delta_{n,k}) = \prod_{j=1}^n W_{j,k}^+$$

20



(6)

(8)

where  $Y^+ > 0$ . Noting from Eq. (2b) that  $W = W^+/p$ , Y can similarly be defined in terms of  $W^+$  as

$$Y_n(\Delta_{n,k}) = \prod_{j=1}^n \left( W_{j,k}^+ / p_j \right)$$

Combining Eqs. (8) and (9) and using the substitution P(Y > 0) = 1 - P(Y = 0) yields the definition of  $Y^+$  in terms of Y:

$$Y^+ = YP(Y > 0)$$

For constant  $\alpha$  the log-stable distribution parameters for  $Y^+$  can be easily determined from the log-stable parameters of  $W^+$  because the product of log-stable variables is also log-stable. Let  $Y^+ = W_1^+ W_2^+ \dots W_n^+$  for  $j = 1, \dots, n$ , where  $W_j^+$  are independent random variables given by  $W_j^+ = \exp(X_j)$ , with  $X_j \sim S_X(\alpha, -1, \sigma_j, \mu_j)$ . If  $Z = \ln(Y^+) =$  $X_1 + X_2 + \dots + X_n$ , then Z is distributed as  $Z \sim S_Z(\alpha, -1, \sigma_Z, \mu_Z)$ , where

$$\sigma_Z^{\alpha} = \sum_{j=1}^n \sigma_j^{\alpha} \tag{11}$$

ิล	n	d
u		u

10

20

$$\mu_Z = \sum_{j=1}^n \mu_j$$

for  $\alpha \neq 1$  (Samorodnitsky and Taqqu, 1994).

We estimated parameters for 8 variations, or versions, of the MRC model described in Sect. 2.2. The versions varied in complexity. The simplest version used the log-normal distribution for  $W^+$  with parameters that were scale-invariant and rainfallindependent, whereas the most complex used the log-stable distribution with parameters that depended on both scale and rainfall. Each version is summarized below:



(9)

(10)

(12)

- 1. SI/RI $\sigma$ /LN: the parameters of the cascade generator were scale invariant (SI) and the random variable  $W^+$  was log-normally (LN) distributed and independent of rainfall intensity (RI $\sigma$ ).
- 2. SI/RI $\sigma$ /LS: the parameters of the cascade generator were scale invariant and the random variable  $W^+$  had a log-stable (LS) distribution and was independent of rainfall intensity.

5

15

20

- 3. SI/RD $\sigma$ /LN: the parameters of the cascade generator were scale invariant and the random variable  $W^+$  was log-normally distributed and dependent on rainfall intensity (RD $\sigma$ ).
- 4. SI/RD $\sigma$ /LS: the parameters of the cascade generator were scale invariant and the random variable  $W^+$  had a log-stable distribution that was dependent on rainfall intensity.
  - 5. SD $\sigma$ /RI $\sigma$ /LN: the scale parameter  $\sigma$  of the cascade generator was scale dependent (SD) and the random variable  $W^+$  was log-normally distributed and independent of rainfall intensity.
  - 6. SD $\sigma/RI\sigma/LS$ : the scale parameter  $\sigma$  of the cascade generator was scale dependent and the random variable  $W^+$  had a log-stable distribution and was independent of rainfall intensity.
  - 7. SD $\sigma$ /RD $\sigma$ /LN: the scale parameter  $\sigma$  of the cascade generator was scale dependent and the random variable  $W^+$  was log-normally distributed and dependent on rainfall intensity.
  - 8. SD $\sigma$ /RD $\sigma$ /LS: the scale parameter  $\sigma$  of the cascade generator was scale dependent and the random variable  $W^+$  had a log-stable distribution was dependent on rainfall intensity.

Discussion Paper	HES 8, 7261–7 Parameteri	HESSD 8, 7261–7291, 2011 Parameterization with		
_	rain gau	ige data		
Discu	D. E. Rupp et al.			
ssion P	Title Page			
aper	Abstract	Introduction		
_	Conclusions	References		
Discu	Tables	Figures		
JSSIO	14	►I.		
Pa		•		
oer	Back	Close		
—	Full Scre	Full Screen / Esc		
Discussion F	Printer-friendly Version			
aper		<b>O</b> BY		

For those cases when  $\sigma$  was scale invariant,  $\sigma$  was solved for uniquely in terms of  $\sigma_Z$  by inverting Eq. (11):

$$\sigma = n^{-1/\alpha} \sigma_z$$

When  $\sigma$  was scale dependent,  $\sigma$  varied as the following function of the length scale  $\lambda_{5} = L/L_{0}$ :

 $\sigma^{\alpha}(\lambda) = \sigma_{1}^{\alpha} \lambda^{\gamma}$ 

where  $\gamma$  is a constant and  $\sigma_1$  is the value of  $\sigma$  at  $\lambda = 1$ . For  $\gamma > 0$  as the scale  $\lambda$  decreases the variance of  $W^+$  decreases, which places it in the family of "bounded" cascade models (Marshak et al., 1994). Combining Eqs. (11) and (14) and using the substitution  $\lambda = (1/2)^{j-1}$ ,  $\sigma_1^{\alpha}$  can be solved for in terms of  $\sigma_Z$ :

$$\sigma_{1}^{\alpha} = \frac{1 - 2^{-\gamma}}{1 - 2^{-n\gamma}} \sigma_{Z}^{\alpha}$$
(15)

for  $\gamma > 0$ 

10

In all 8 model versions, the intermittency parameter was determined from P(Y > 0):

 $p = [P(Y > 0)]^{1/n}$ 

<sup>15</sup> Because it has been observed that spatial cascade parameters, and the intermittency parameter in particular, depend on large-scale rainfall (Over and Gupta; 1994; 1996; Deidda, 2000; Jothityangkoon et al., 2000; Pathirana and Herath, 2002; Deidda et al., 2004; 2006; Sharma et al., 2007), we allowed some of the parameters of the cascade generator to vary with the large-scale rainfall depth  $R_0$ . While it has been argued that for both space (Veneziano et al., 2006) and time (Veneziano et al., 2006; Rupp et al., 2009; Serinaldi, 2010) the parameters should vary with rainfall intensity at each scale (not just the largest scale), our dataset did not permit us to adequately examine rainfall dependency across scales, therefore we were restricted to making the parameters dependent on the large-scale rainfall only.

(13)

(14)

(16)

In all 8 versions, we allowed the intermittency parameter to depend on large-scale rainfall by varying P(Y > 0) in Eq. (16) with  $R_0$  as

$$P(Y > 0) = \frac{1}{2} \left\{ 1 + \operatorname{erf}\left[\frac{\ln(R_0) - m}{\sqrt{2s^2}}\right] \right\}$$

where erf is the error function with parameters *m* and *s* (Rupp et al., 2009). In 4 of the models, the scale parameter $\sigma$  was varied with rainfall by relating  $\sigma_z$  in Eqs. (13) and (15) to  $R_0$  as

$$\sigma_Z = c + f(R_0) \tag{18}$$

where f() is an arbitrary function and c is a constant. We used cubic splines to determine f().

Observations of rainfall at gauges were used to estimate values of Y,  $Y^+$ , and P(Y > 0). Combining Eqs. (1) and (6), we see that an estimate of Y from observations of R at a given rain gauge can be calculated as

$$\widehat{Y}_{i,k} = \frac{R_{n,i}}{R_{0,i}}$$
(19)

where  $\hat{Y}$  is the estimate of Y,  $i = 1, 2, ..., N_{obs}$  indexes the *i*-th observation in time, and  $k = 1, 2, ..., N_{gauges}$  indexes the rain gauge. The areal average rainfall  $R_{0,i}$  at the reference length  $L_0$  was approximated by taking the mean of the rainfall measured over all  $N_{gauges}$  at time *i*. To estimate P(Y = 0), we used

$$\hat{P}_i(Y > 0) = (\text{number of gauges with non-zero rain})_i / N_{\text{gauges}}$$
 (20)

Finally,  $Y^+$  was estimated with

20 
$$\hat{Y}_{i,k}^{+} = \hat{Y}_{i,k} \hat{P}_{i}(Y > 0)$$

10

Discussion Paper **HESSD** 8, 7261-7291, 2011 Parameterization with rain gauge data **Discussion** Paper D. E. Rupp et al. **Title Page** Abstract Introduction References Conclusions **Figures Discussion** Paper **Tables** 14 Back Close Full Screen / Esc **Discussion** Paper **Printer-friendly Version** Interactive Discussion

(17)

(21)

Because  $Y^+$  is bounded by zero and positive infinity, whereas the upper limit to  $\widehat{Y}^+$  is

 $N_{\text{gauges}}$ , the distribution of  $\widehat{Y}^{\dagger}$  is only approximately equal to the distribution of  $Y^{\dagger}$ . This limitation, plus instrument error at very low and very high rainfall intensities, introduces a bias into the estimation of  $\sigma_z$ . For now, we simply accept this bias as a shortcoming of the estimation procedure, though we discuss it further in Sect. 3.

Free software packages for estimating the parameters of the stable distribution are rare, and we found none that suited our particular needs. For this reason, we used

a simple procedure to estimate  $\alpha_z$  and  $\sigma_z$  from the "observed" values  $\ln Y$ . An optimization algorithm minimized the sum of squared differences between the following observed and theoretical quantiles: 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, and 0.95. For the normal distribution, we used the maximum likelihood method.

Note that except for at the largest spatial scale, we did not areally average the precipitation data. We also did not consider the particular location in space of the observed rainfall. Both of these characteristics distinguish our study from others. However, we did use the number of cascade levels n that brought us to the scale of the rain gauge

<sup>15</sup> did use the number of cascade levels *n* that brought us to the scale of the rain gauge itself. Given that the rain gauges have a diameter of approximately 0.15 m, and that the spatial extent of our rain gauge network corresponds to approximately  $L_0 = 20\,000$  m, approximately n = 17 cascade levels are needed.

## 2.4 Model evaluation

5

- <sup>20</sup> Direct comparison of the stochastically downscaled data to the observed rainfall requires disaggregation down to the capture area of the rain gauge through n = 17 cascades levels, which would result in a grid with over  $17 \times 10^9$  cells. This would be impractical, especially given that very many such spatial fields would be generated for time increments of as little as 15 min. Instead of generating complete rainfall fields at the spatial resolution of a rain gauge, we followed the cascade process down a subset
- of the total number of possible paths down the cascade. Along a path at each subdivision, we randomly chose 1 of the 4 cells (with equal weight given to each cell), and



tracked the position (x, y) of the cell for each of the n = 17 cascade levels. In total, we followed 240 paths per 15-min increment. We used Eq. (1) to calculate the 15-min rainfall amounts at each of these 240 locations. The large-scale areal-averaged rainfall depth  $R_0$  at each time increment was approximated by the mean of all the available observed rainfall depths.

To further reduce the computational burden, downscaling was done on a sub-sample of all available 18723 15-min time steps where  $R_0 > 0$ . We selected 2000 time steps such that the cumulative distribution function (CDF) of the sub-sampled  $R_n$  was similar to the CDF of the full record.

The spatial structures of the observed and simulated 15-min rainfall were compared using semivariograms of log  $(R_n)$  for  $R_n \ge 0.001$  mm (the minimum rainfall amount recorded). The spatial structure of the intermittency was examined with semivariograms of presence (1) and absence (0) of rain.

In addition to the spatial structure, we assessed the ability of the models to reproduce the cumulative distribution frequency (CDF) of 15-min rainfall for  $R_n > 0$ .

## 3 Results and discussion

The proportion of gauges with zero rain in a 15-min period, P(Y = 0), was found to be strongly dependent on the large-scale rainfall rate  $R_0$  (Fig. 3). Consistent with many other studies (e.g., Over and Gupta, 1994, 1996; Jothityangkoon et al., 2000; Pathirana and Herath, 2002; Sharma et al., 2007), the sparseness of the rain field was much greater when  $R_0$  was low, while at high rainfall rates the tendency was for it to be raining everywhere. The sigmoidal shape of Eq. (10) appears suitable for simulating the rainfall intermittency, as it allows for P(Y = 0) to go to 1 as  $R_0$  goes to 0, and to go to 0 as  $R_0$  goes to  $+\infty$ .

The empirical histograms of  $\ln \gamma^+$  were rightward skewed, thus more similar to a logstable density with  $\beta = -1$  than to a log-normal distribution (Fig. 4). At progressively lower values of  $R_0$  (e.g., <~0.01 mm), the empirical histograms were progressively



more dominated by  $\ln Y^+ = 0$ , such that neither the log-stable nor the log-normal densities matched the observations. It is clear that by fitting theoretical distributions to  $\ln Y^+$  at low values of  $R_0$ , we are merely fitting to a data artifact and not to true rainfall behavior.

<sup>5</sup> The value of the stable parameter  $\alpha_Z$  showed a general increasing trend with increasing  $R_0$  (Fig. 5). However, when  $\alpha_Z$  was fixed at a constant value of 1.47, the fits of the log-stable distributions were only marginally degraded (Fig. 4). This was fortunate because it allowed us to keep  $\alpha_Z$  as a constant parameter and only have to vary the scale parameter  $\sigma_Z$  with large-scale rainfall. The particular value of  $\alpha_Z = 1.47$  is the average  $\alpha_Z$  using all values of  $\ln \hat{Y}^+$  for  $R_0 \ge 0.004$  mm. The threshold of 0.004 mm was selected because below this value the empirical distributions were strongly influenced by data precision (Fig. 4).

The dependency of  $\sigma_z$  on  $R_0$  was complex (Fig. 5), though cubic splines with no more than 6 knots reproduced the empirical relationship of  $\sigma_z$  with  $R_0$  well. The relationship was similar in form for both the log-stable ( $\alpha_z = 1.47$ ) and log-normal ( $\alpha_z = 2$ ) distributions (Fig. 5). Variability of  $\sigma_z$  with  $R_0$  was high, but much of the variability was due to low values of  $\sigma_z$  at low values of  $R_0$ , where we have already determined the empirical distributions to be unreliable.

If we ignore  $\sigma_Z$  for  $R_0 < 0.01$ , the pattern in Fig. 5 (lower panel) implies a smoother field at intermediate rainfall rates (~ 0.1 mm 15 min<sup>-1</sup>) and a more variable field at lower and higher rates. This trend toward higher variability at the highest rainfall rates could be the result of localized, high-intensity rainfall generated from strong convective storm cells. This trend is not evident in the studies of Over and Gupta (1996), Jothityangkoon et al. (2000), Pathirana and Herath (2002), or Sharma et al. (2007), who only observed the scale parameter to increase and then decrease with increasing  $R_0$ . The difference

between our and previous results may be due to differences in scales between studies: Jothityangkoon et al. (2000) and Sharma et al. (2007) analyzed daily rainfall, while Over and Gupta (1996) and Pathirana and Herath (2002) analyzed radar scans with resolutions of 4 to 5 km.



The semivariogram of the observed 15-min rainfall shows covariability in rainfall intensity increasing strongly with increasing separation distance (Fig. 6). In contrast, all the scale-invariant models produced rainfall fields that showed little change in correlation with separation, though proximal rain was slightly more similar than distant rain (Fig. 6). When the scale parameter  $\sigma$  as allowed to decrease with decreasing scale 5 via Eq. (14), however, the general variogram pattern of the observed rainfall could be reproduced for separation distances of less than about 10 km by using a value of  $\gamma$  $\sim 0.8$  (Fig. 6). At separation distances above 10 km, the semivariances of the simulated rainfall become nearly constant, irrespective of the model or the value of  $\gamma$ . We believe this is an artifact of the discrete nature of the cascade procedure that was ap-10 plied. Note that at the first cascade level, the rainfall is first separated into four 10 km by 10 km cells. The rainfall simulated at a point in one of these first four cells will be equally correlated with the rainfall simulated at a point *anywhere* in one of the other three cells. One possible way to prevent this is to incorporate non-stationarity into the

<sup>15</sup> cascade process, which we return to briefly at the end of this section.

20

The semivariogram of the presence/absence observations shows that if it is raining (or not raining) at one location, it is more likely to be raining (or not raining) nearby than it is further away (Fig. 7). The simulated rainfall fields have this property as well (at least below the 10 km separation distance), but not to the degree of the observed rainfall field. Figure 7 gives semivariograms of the simulated presence/absence data using Models SI/RI $\sigma$ /LN and SD $\sigma$ /RI $\sigma$ /LN only: semivariograms from all 8 models were similar because the models are identical in how they simulate intermittency.

Over most of the range of R, the log-normal models reproduced well the observed CDF of 15-min rainfall totals (Fig. 8). However, the simulated CDFs using the lognor-

<sup>25</sup> mal models diverged from the observed CDF below 0.02 mm 15 min<sup>-1</sup>. The rainfall-dependent (RD) models performed slightly better than the rainfall-independent (RI) models up to the very highest rainfall intensities. Above about 15 mm 15 min<sup>-1</sup>, the rainfall-dependent models overestimated the rainfall intensity at a given probability of occurrence by roughly a factor of two.



The log-stable models did better than the log-normal models at reproducing the overall shape of the observed CDF (Fig. 9), including matching the curvature for  $R < 0.02 \text{ mm } 15 \text{ min}^{-1}$ . The log-stable models also mimicked the upward curvature of the observed CDF for the highest values of R., though they did under-predict the probabilities of these extreme events. Making the scale parameter rainfall-dependent resulted in an improved CDF, though at the highest intensities these models still underestimated the rainfall intensity at a given probability of occurrence by as much as a factor of two.

5

As mentioned previously, the estimation method is biased due to artifacts of the data. For one, the measurement instrument has a detection threshold, which results in sparser measured than true rainfall, and hence an underestimation of p, particularly

- <sup>10</sup> In sparser measured than true rainfall, and hence an underestimation of *p*, particularly when the larger-scale rainfall  $R_0$  is low (Veneziano, et al., 2006). A second factor is the recording precision of the rainfall observations, which particularly affects parameter estimation at low rainfall intensities. The effect of precision can be seen in the preponderance of values of  $Y^+ = 1$  at low values of  $R_0$  (Fig. 4) which, in turn, results in an underestimation of  $\sigma$ . Similar observations were made by Rupp et al. (2009) and Licznar et al. (2011a, b) regarding the empirical weights  $W^+$  when analyzing rainfall time series. A third factor is the underreporting of high rainfall intensities due to instru-
- ment error, which reduces the variance of  $Y^+$ . A fourth factor, as already discussed in Sect. 2.3, arises from the finite number of rainfall gauges used in the estimation <sup>20</sup> method.

A variety of procedures can be used to partly account for the bias. One is to iteratively adjust the model parameters until the estimated parameters from the simulated dataset are nearly the same as those from the observed dataset (Veneziano et al., 2006). Another procedure is to exclude some data while estimating parameters. For example, in our study we left out data where  $R_0 < 0.004$  mm when estimating  $\sigma_{Z_i}$  for the case where  $\sigma_Z$  was assumed to be independent of  $R_0$ . However, this excluded only a relatively small amount of data and thus did not greatly affect the value of  $\sigma_Z$ independent of  $R_0$ . As another example, Licznar et al. (2011a), simply eliminated what would be analogous in our study to all values of  $Y^+ = 1$  from the empirical frequency



distribution, under the assumption that most of these values were artifactual. A third procedure to deal specifically with recording precision is to add random noise to the rainfall observations, with the intent of replacing the information lost by round-off error and thus removing the discretization that leads to an excess of certain values of  $W^+$  (or  $Y^+$ ) (Licznar et al., 2011b).

Bias-correcting procedures such as those above should be explored, and we expect that they would improve the fits of frequency distributions. We know, for example, that both data precision and the finite number of gauges serve to decrease the estimated value of the scale parameter  $\sigma_Z$ , in the former case by generating an overabundance of  $\hat{Y}^+ = 1$  and in the latter case by imposing a maximum value to  $\hat{Y}^+$  of  $N_{gauges}$ . A bias-correcting procedure that led to an increase in the value of the log-stable parameter  $\sigma_Z$  would produce more extreme events, resulting in a CDF more like the observed one in Fig. 9. It would also increase the semivariance overall, which was generally underpredicted by the log-stable models (Fig. 6, lower panels).

- Lastly, we have assumed stationarity in the rainfall field, though there may be longterm spatial patterns across the Warsaw metropolitan area. With our short record length (less than 3 yr) it would be difficult detect any but very clear and strong largescale patterns, which we did not see. Should continuing observations reveal deterministic patterns in the spatial distribution of rainfall, we could account for these within the MRC framework. Examples of how this might be done using a deterministic field
- of weights that are applied to the cascade generator are given by Jothityangkoon et al. (2000) and Pathirana and Herath (2002).

# 4 Conclusions

5

We have presented and evaluated a method for estimating the parameters of a multiplicative random cascade model for downscaling rainfall fields when observations of the full fields are not available either from radar imagery or from interpolation of very dense rain gauge network data. The estimation procedure still relies on rain gauge



data, but the density of the network need only be such that (1) the rainfall rate over a given time interval averaged over the entire spatial domain can be reasonably approximated by averaging the rainfall rate from all the gauges, (2) the number and the spatial coverage of the gauges are adequate for generating a semivariogram of rainfall 5 intensity.

When the cascade generator is independent and identically distributed (iid) throughout the cascade, the parameters can be estimated solely from the frequency distribution of the ratios of the rain rate at each gauge to the large-scale average rain rate. We found, however, that an iid cascade generator failed to reproduce the spatial covariance structure of the rainfall over Warsaw, Poland: proximal rainfall was too dissimilar using an iid parameterization, and the simulated rainfall only showed a weak relationship between distance and covariability (or semivariance), whereas this relationship was strong in the observed data.

10

To better reproduce the spatial structure of actual rainfall fields (as summarized by the semivariogram), we added scale dependence to the cascade generator. The scaledependent generator introduced an additional model parameter ( $\gamma$ ) that could not be estimated directly from the rain rate ratios. We therefore treated  $\gamma$  as a tuning parameter that was estimated by matching the observed and simulated semivariograms. To keep the model simple (i.e., to one tuning parameter) for this study, we considered only

scale dependence in the generation of positive rainfall amounts, not in the generation of rainfall intermittency. A similar strategy, however, could be used for the intermittency parameter along with the semivariogram of rainfall presence/absence, though it would require the introduction of at least one additional parameter.

Overall, the scale-dependent MRC models generated the correct frequency distribution of short-duration rainfall intensities. We recommend, however, further research into bias in parameter estimation; we expect that through bias-correction procedures, improvements could be made at both the extreme lower and upper ends of the distribution.



We evaluated the model by using statistical properties of the rain gauge data as performance targets. This meant it was necessary to downscale to the approximate capture area of the rain gauge ( $15 \times 15$  cm). For most stormwater drainage system studies, generating fields at such a fine resolution would be impractical. However, an expedient property of the MRC model is that it lends itself nicely to downscaling to any spatial scale  $\lambda_n$ , which can conveniently be used to generate gridded rainfall fields for use as input to hydrologic/hydrodynamic models.

Acknowledgements. We thank Ryan Stewart for his comments and suggestions on an earlier version of this manuscript.

#### 10 **References**

5

- Ahrens, B.: Rainfall downscaling in an alpine watershed applying a multiresolution approach, J. Geophys. Res., 108, 8388, doi:10.1029/2001JD001485, 2003.
- Deidda, R.: Rainfall downscaling in a space-time multifractal framework, Water Resour. Res., 36, 1779–1794, doi:10.1029/2000WR900038, 2000.
- <sup>15</sup> Deidda, R., Benzi, R., and Siccardi, F.: Multifractal modeling of anomalous scaling laws in rainfall, Water Resour. Res., 35, 1853–1867, doi:10.1029/1999WR900036, 1999.
  - Deidda, R., Badas, M. G, and Piga. E.: Space-time scaling in high intensity Tropical Ocean Global Atmosphere Coupled Ocean-Atmosphere Response Experiment (TOGA-COARE) storms, Water Resour. Res., 40, W02506, doi:10.1029/2003WR002574, 2004.
- Deidda, R., Badas, M. G., and Piga, E.: Space-time multifractality of remotely sensed rainfall fields, J. Hydrol., 322, 2–13, doi:10.1016/j.jhydrol.2005.02.036, 2006
  - Ferraris, L., Gabellani, S., Rebora, N., and Provenzale, A.: A comparison of stochastic models for spatial rainfall downscaling, Water Resour. Res., 39, 1368, doi:10.1029/2003WR002504, 2003.
- <sup>25</sup> Gupta, V. and Waymire, E.: Multiscaling properties of spatial rainfall and river flow distributions, J. Geophys. Res., 95, 1999–2009, 1990.

Gupta, V. and Waymire, E.: A statistical analysis of mesoscale rainfall as a random cascade, J. Appl. Meteorol., 32, 251–267, 2003.



- Hingray B. and Ben Haha, M.: Statistical performances of various deterministic and stochastic models for rainfall series disaggregation, Atmos. Res., 77, 152–175, 2005.
- Jothityangkoon, C., Sivapalan, M., and Viney, N.: Tests of a space-time model of daily rainfall in southwestern Australia based on nonhomogeneous random cascades, Water Resour. Res., 36, 267–284, 2000.
- Kang, B. and Ramírez, J. A.: A coupled stochastic space-time intermittent random cascade model for rainfall downscaling, Water Resour. Res., 46, W10534, doi:10.1029/2008WR007692, 2010.

Kumar, P. and Foufoula-Georgiou, E.: A multicomponent decomposition of spatial rainfall fields:

10 **1.** Segregation of large- and small-scale features using wavelet transforms, Water Resour. Res., 29, 2515–2532, 1993a.

Kumar, P. and Foufoula-Georgiou, E.: A multicomponent decomposition of spatial rainfall fields: 2. Self-similarity in fluctuations, Water Resour. Res., 29, 2533–2544, 1993b.

- Licznar, P., Łomotowski, J., and Rupp, D. E.: Random cascade driven rainfall disaggregation for urban hydrology: An evaluation of six models and a new generator, Atmos. Res. 99(3–4),
- 563–578, doi:10.1016/j.atmosres.2010.12.014, 2011a.

5

15

- Licznar, P., Schmitt, T. G., and Rupp, D. E.: Distributions of microcanonical cascade weights of rainfall at small timescales, Acta Geophysica, early on-line publication, doi:10.2478/s11600-011-0014-4, 2011b.
- Lovejoy, S. and Schertzer, D.: Multifractals, universality classes and satellite and radar measurements of cloud and rain fields, J. Geophys. Res., 95, 2021–2034, 1990.

Marshak, A., Davis, A., Cahalan, R., and Wiscombe, W.: Bounded cascade models as nonstationary multifractals, Phys. Rev. E, 49, 55–69, 1994.

McCulloch, J. H.: Financial applications of stable distributions, in: Handbook of Statistics 14:

- Statistical Methods in: Elsevier Science B. V., edited by: Finance, Maddala, G. S. and Rao, C. R., Amsterdam, 394–425, 1996.
  - Molnar, P. and Burlando, P.: Preservation of rainfall properties in stochastic disaggregation by a simple random cascade model, Atmos. Res. 77(1–4), 137–151, doi:10.1016/j.atmosres.2004.10.024, 2005.
- <sup>30</sup> Oke, T.: Initial guidance to obtain representative meteorological observations at urban sites, Instruments and Observing Methods, Report no. 81, World Meteorological Organization, WMO/TD-No. 1250, 2006.

Over, T. M. and Gupta, V. J.: Statistical analysis of mesoscale rainfall: Dependence of a random



cascade generator on large-scale forcing, J. Appl. Meteorol., 33, 1526–1542, 1994.

Over, T. M. and Gupta, V. J.: A space-time theory of mesoscale rainfall using random cascades, J. Geophys. Res., 101, 26319-26331, 1996.

Pathirana, A. and Herath, S.: Multifractal modelling and simulation of rain fields exhibiting spa-

tial heterogeneity, Hydrol. Earth Syst. Sci., 6, 695–708, doi:10.5194/hess-6-695-2002, 2002. 5 Perica, S. and Foufoula-Georgiou, E.: Linkage of scaling and thermodynamic parameters of rainfall: Results from midlatitude mesoscale convective systems, J. Geophys. Res., 101, 7431-7448, 1996.

Ross, S.: A First Course in Probability, Prentice-Hall, Inc., Upper Saddle River, N.J., USA, 514 p., 1998.

Rupp, D. E., Keim, R. F., Ossiander, M., Brugnach, M., and Selker, J. S.: Time scale and intensity dependency in multiplicative cascades for temporal rainfall disaggregation, Water Resour. Res., 45, W07409, doi:10.1029/2008WR007321, 2009.

Samorodnitsky, G. and Taggu, M. S.: Stable Non-Gaussian Random Processes, Chapman & Hall, New York, 1994.

15

10

30

Schertzer, D. and Lovejoy, S.: Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes, J. Geophys. Res., 92, 9692–9714, 1987.

Serinaldi, F.: Multifractality, imperfect scaling and hydrological properties of rainfall time series simulated by continuous universal multifractal and discrete random cascade models, Nonlin.

Processes Geophys., 17, 697–714, 2010, 20

http://www.nonlin-processes-geophys.net/17/697/2010/.

- Sharma, D., Das Gupta, A., and Babel, M. S.: Spatial disaggregation of bias-corrected GCM precipitation for improved hydrologic simulation: Ping River Basin, Thailand, Hydrol. Earth Syst. Sci., 11, 1373–1390, doi:10.5194/hess-11-1373-2007, 2007.
- <sup>25</sup> Svensson, C., Olsson, J., and Berndtsson, R.: Multifractal properties of daily rainfall in two different climates, Water Resour. Res., 32, 2463-2472, doi:10.1029/96WR01099, 1996. Veneziano, D., Furcolo, P., and Iacobellis, V.: Imperfect scaling of time and space-time rainfall,
  - J. Hydrol. 322, 105–119, 2006.

Venugopal, V., Foufoula-Georgiou, E., and Sapozhnikov, V.: A space-time downscaling model for rainfall, J. Geophys. Res., 104, 19705-19721, 1999.

<b>Discussion</b> Pape	HE 8, 7261–7	<b>HESSD</b> 8, 7261–7291, 2011		
	Parameter rain ga	ization with uge data		
Discu	D. E. R.	D. E. Rupp et al.		
ussion P	Title Page			
aper	Abstract	Introduction		
	Conclusions	References		
Disc	Tables	Figures		
ussio	14	►I		
n Pap	•	•		
oer	Back	Close		
—	Full Scr	Full Screen / Esc		
Discussion	Printer-frie Interactive	Printer-friendly Version		
Paper	e	<b>O</b> BY		



Fig. 1. Location of rain gauges used in study in Warsaw, Poland.





**Fig. 2.** Schematic of two-dimensional multiplicative cascade with branching number b = 4.

Discussion Paper	HE 8, 7261–7 Parameter	HESSD 8, 7261–7291, 2011		
_	rain gau	uge data		
Discu	D. E. Rupp et al.			
ssion P	Title Page			
aper	Abstract	Introduction		
—	Conclusions	References		
Discu	Tables	Figures		
Issio	I	►I.		
n Pap		•		
Der	Back	Close		
—	Full Scr	Full Screen / Esc		
Discussion Pa	Printer-frie Interactive	Printer-friendly Version Interactive Discussion		
oer	$\odot$	BY		









**Fig. 4.** Empirical density (bars) of  $IN Y^+$  and densities of the fitted stable distribution with  $\alpha$  as a fitting parameter (thin green line) and with  $\alpha$  fixed at 1.47 (heavy blue line) and 2 (thin red line). Distributions are shown for various large-scale areal-averaged 15-minute rain depths  $R_0$  for  $R_0 > 0$ . The rainfall depth shown in each plot is the midpoint of the range of log-transformed rainfall depths used to bin that data in a given depth class.





**Fig. 5.** Estimated stable distribution parameters  $\alpha_z$  and  $\sigma_z$  for  $Z = \ln Y^+$  against large-scale areal-averaged 15-minute rain  $R_0$  for  $R_0 > 0$ . The open symbols indicate where estimation was clearly affected by artifacts arising from data precision. In the upper panel, the dashed line shows the estimate of  $\alpha_z$  using all data for  $R_0 \ge 0.004$  mm and the dotted line is  $\sigma_z = 2$  (the normal distribution). In the lower panel, values of  $\sigma_z$  assume constant  $\alpha_z$  of 1.47 and 2 for the stable and normal distributions, respectively. The solid curves are fitted cubic spline functions. The dashed and dotted horizontal lines indicate the values of  $\sigma_z$  averaged over all  $R_0 \ge 0.004$  mm for the stable and normal distribution, respectively.

















**Fig. 8.** Cumulative distribution function (CDF) of non-zero 15-min rainfall totals from all observations, of the subset of 2000 15-min intervals used to initialize the large-scale rainfall *R*0 for the simulations, and of the simulations. Simulated rainfall was generated using log-normal Models SI/RI $\sigma$ /LN, SD $\sigma$ /RI $\sigma$ /LN with  $\gamma$  = 0.8, SI/RD $\sigma$ /LN, and SD $\sigma$ /RD  $\sigma$ /LN with  $\gamma$  = 0.8.







