Development and analysis of a simple model to represent the zero rainfall in a universal multifractal framework

Auguste Gires, Ioulia Tchiguirinskaia, D Schertzer, S. Lovejoy

To cite this version:

Auguste Gires, Ioulia Tchiguirinskaia, D Schertzer, S. Lovejoy. Development and analysis of a simple model to represent the zero rainfall in a universal multifractal framework. Nonlinear Processes in Geophysics, European Geosciences Union (EGU), 2013, 20, pp.343-356. <hal-00869291>
Development and analysis of a simple model to represent the zero rainfall in a universal multifractal framework

A. Gires1,*, I. Tchiguirinskaia1, D. Schertzer1, and S. Lovejoy2

1Université Paris Est, École des Ponts ParisTech, LEESU, Marne-la-Vallée, France
2McGill University, Physics department, Montreal, QC, Canada

*Invited contribution by A. Gires, recipient of the EGU Outstanding Student Poster (OSP) Awards 2012

Correspondence to: A. Gires (auguste.gires@leesu.enpc.fr)

Received: 29 November 2012 – Revised: 10 April 2013 – Accepted: 22 April 2013 – Published: 24 May 2013

Abstract. High-resolution rainfall fields contain numerous zeros (i.e. pixels or time steps with no rain) which are either real or artificial – that is to say associated with the limit of detection of the rainfall measurement device. In this paper we revisit the enduring discussion on the source of this intermittency, e.g. whether it requires specific modelling. We first review the framework of universal multifractals (UM), which are commonly used to analyse and simulate geophysical fields exhibiting extreme variability over a wide range of scales with the help of a reduced number of parameters. However, this framework does not enable properly taking into account these numerous zeros. For example, it has been shown that performing a standard UM analysis directly on the field can lead to low observed quality of scaling and severe bias in the estimates of UM parameters. In this paper we propose a new simple model to deal with this issue. It is a UM discrete cascade process, where at each step if the simulated intensity is below a given level (defined in a scale invariant manner), it only has a predetermined probability to survive and is otherwise set to zero. A threshold can then be implemented at the maximum resolution to mimic the limit of detection of the rainfall measurement device. While also imperfect, this simple model enables explanation of most of the observed behaviour, e.g. the presence of scaling breaks, or the difference between statistics computed for single events or longer periods.

1 Introduction

Obtaining a consistent, robust and reliable mathematical/physical representation of rainfall at every scale is a kind of Holy Grail for hydrologists. Indeed, they would need it to simulate realistic long time series that they commonly use to design structures such as levee, dams, or sewer networks, or to improve various rainfall-processing algorithms such as those merging between radar and rain gauge measurement devices, or nowcasting, that are more and more commonly used to improve real-time management of river or sewer systems. Such a representation of rainfall should be able to grasp all the features of the rainfall variability, i.e. the intrinsic variability exhibited over a wide range of spatio-temporal scales and also the succession of dry and rainy periods. This is an important point since there are many zeros (a pixel or time step where no rain has been recorded) in a high-resolution rainfall data set. For instance, in France typical long (many years) high-resolution (5 min) rain gauge time series contain roughly 96–98% zeros (Hoang et al., 2012) The nature of these zeros is especially tricky since some are real ones (i.e. it actually did not rain) whereas others are artificial (i.e. it actually did rain, but nothing was recorded because the intensity was below the limit of detection of the rainfall measurement device). To stress the importance of this threshold, let us simply mention that a typical rain gauge limit of detection is 0.2 mm h⁻¹ (one tipping of a bucket during an hour), and is roughly 2–3 times greater than the average rainfall in Paris (roughly 650 mm per year).

Mainly two types of stochastic rainfall models have been considered. The first one corresponds to point processes
where a rainfall time series is represented by a succession of events whose timing, duration and intensity are modelled by stochastic processes. The simplest model used for representing a succession of events is a Poisson process (see Waymire and Gupta, 1981, for a review). Further refinements relying on two-level Poisson processes were later introduced to better deal with rainfall extremes (Onof et al., 2000 for a review; Rodriguez-Iturbe et al., 1987). The main drawbacks of these approaches is that they are mainly valid at the scale for which they are defined and calibrated (although Olsson and Burlando, 2002, showed that such processes exhibit some scale-invariance properties on a limited range of scales) and they are intrinsically 1-D models with difficulties for generalisation in space or space/time (see however Wheater et al., 2005), and they require numerous parameters which have to be fitted.

The second type of stochastic models is based on scaling. At first, it was done in a phenomenological, fractal framework with the help of a “fractal sum of pulses” (Lovejoy and Mandelbrot, 1985; Lovejoy and Schertzer, 1985), which was a broad generalisation of the Bartlett-Lewis and Neyman-Scott point processes. However, scaling was not yet related to the physical process of interacting (multiplicative) cascades of the dynamics and the water content (Schertzer and Lovejoy, 1987). This yields a multifractal rainfall field, i.e. the rain accumulation is no longer a pointwise quantity but a (multi-)singular measure, and this explains the strong scale dependency of the rain rate, which is not pointwise defined either (Schertzer et al., 2010).

Since then stochastic multifractals have been extensively used to analyse, model and simulate rainfall fields (understood here as a time series in 1-D, a matrix of pixels in 2-D, or a succession of 2-D matrix in 3-D) and more generally geophysical fields extremely variable over a wide range of scales (de Lima and de Lima, 2009; de Lima and Grasman, 1999; Harris et al., 1997; Lovejoy and Schertzer, 2007; Marsan et al., 1996; Nykanen, 2008; Olsson and Niemczynowicz, 1996; Royer et al., 2008; Schertzer and Lovejoy, 1987, 1997). In the discrete case, which is used in this paper, a step of a cascade consists in dividing a “parent” structure into “daughter” structures and affecting to them the value of the “parent” structure multiplied by a random factor. Cascades are scale invariant in that the way structures are divided, and the probability distribution of the random multiplicative increment are the same at each step of the process. This process is physically based in the sense that it is in agreement with the scale invariance properties of the Navier–Stokes equations governing atmospheric behaviour, which is assumed to be transmitted to the unknown equations for rainfall processes (Hubert, 2001). In the specific framework of universal multifractals (Schertzer and Lovejoy, 1987, 1997), the underlying cascade process is fully characterised with the help of only three scale invariant parameters:

- $H$, the degree of non-conservation, which measures the scale dependency of the average field ($H = 0$ for a conservative field);
- $C_1$, the mean intermittency co-dimension, which measures the clustering of the (average) intensity at smaller and smaller scales. $C_1 = 0$ for an homogeneous field;
- the multifractality index $\alpha$ ($0 \leq \alpha \leq 2$), which measures the clustering variability with regards to the intensity level.

The aim of the paper is to improve the modelling of the numerous zeros of the rainfall fields in this multifractal framework. This topic has been discussed for more than 20 yr. For example, in the late 1980s there was a theoretical debate over whether zero values should be considered separately from the other values (Keddem and Chiu, 1987) or not (Lovejoy and Schertzer, 1989). Since then numerous authors have suggested methods which are often contradictory between themselves. Let us first mention the issue of how to model the zeros. Some authors suggest that the rainfall field results from the multiplication of a multifractal field by an independent binary field corresponding to the rainfall support (i.e. the portion of time where some rain is recorded). Different models have been published for the rainfall support: a so-called random cascade $\beta$ model in a discrete (Over and Gupta, 1996) or continuous (Schmitt et al., 1998) form, or a two-state renewal process to represent wet and dry periods (Schmitt et al., 1998). Olsson (1998) introduces the zeros within the cascade process by explicitly affecting a strictly positive probability for zero values at each step of the process. Contrary to the other previously mentioned models, this probability depends on the properties of the simulated rainfall field. More precisely, it is adjusted according to the intensity of the time step (it is a 1-D model for time series) and its position in the sequence of dry and wet periods. Other authors (de Montera et al., 2009; Lovejoy et al., 2008) disagree and state that a simple threshold implemented at the maximum resolution enables reproduction of the observed properties. Some authors report that the threshold introduces a break in the scaling behaviour (de Montera et al., 2009; Larnder, 1995), i.e. that a unique scaling law does not hold on the whole range of available scales, and that two different scaling laws are valid over two distinct range of scales. There is a consensus on the fact that the numerous (either real or artificial) zeros affect the estimates of UM parameters by introducing a bias that leads to an underestimation of $\alpha$ and an overestimation of $C_1$. Nevertheless, there is no agreement on how to retrieve the correct underlying UM parameters, and various techniques are found in the literature: taking into account only the rainfall support which is fractal (Schmitt et al., 1998), doing event-based analysis to limit the number of zeros (de Montera et al., 2009; Verrier et al., 2010), a kind of bootstrap method relying on a first-guess estimate and using multifractal simulations (Lovejoy et al., 2008), or
implementing an iterative process whose main step consists
in adding a random simulated multifractal field to simulated
the unobserved small values (Gires et al., 2012b). This ite-
|rative process enables improvement of the estimations of the
underlying UM parameters, but a serious limitation is that the
convergence is not ensured and therefore must be stopped af-
|ter a predetermined number of steps.

In Sect. 2 the theoretical framework of UM is recapped.
Discrepancies between this framework and observed rainfall
data are presented in Sect. 3. A new simple model to gen-
erate a rainfall field with its associated zeroes is presented
in Sect. 4. Results for a given parameter set are discussed in
Sect. 5. Finally the sensitivity of the new simple model to its
parameters is analysed in Sect. 6.

2 Recapitulation of universal multifractals

Before going on, we will briefly present the theoreti-
cal framework of UM. For more details see Lovejoy and
Schertzer (2007) and Schertzer and Lovejoy (2011). The ex-
planations are given in 1-D, but they are also valid in 2-D or
3-D. If the rainfall support is fractal (Lovejoy and Mundel-
brot, 1895; Olsson et al., 1993), the number \( N_\lambda \) of raining
time steps at resolution \( \lambda \) (= \( L/l \), where \( L \) is the observation scale
and \( L \) the outer scale of the phenomenon) scales as
\[
N_\lambda \approx \lambda^{D_F},
\]
where \( D_F \) defines the fractal dimension which is a scale
invariant notion that characterises how much space a geomet-
rical set occupies (\( D_F \) is smaller than the space in which
the studied geometrical set is embedded). But in a multi-
fractal framework, \( D_F \) becomes strongly dependent (Lovejoy et
al., 1987; Hubert et al., 1995) on the threshold defining the
occurrence or not of rainfall (\( D_F \) decreases with increasing
thresholds), pointing out that more than one fractal dimen-
sion is needed to fully characterise the rainfall field.

If \( \varepsilon_\lambda \) denotes a multifractal field at resolution \( \lambda \), then
the probability of exceeding a scale dependent threshold \( \lambda^{y} \)
defined with the help of a singularity \( y \) scales with the resolu-
tion as
\[
\text{Pr}(\varepsilon_\lambda \geq \lambda^{y}) \approx \lambda^{-c(y)},
\]
where \( c(y) \) is the codimension function (Schertzer and
Lovejoy, 1987), and it can be shown that this is equivalent to the
scaling of the statistical moment of order \( q \):
\[
\langle \varepsilon_\lambda \rangle \approx \lambda K(q),
\]
where \( K(q) \) is the moment scaling function. The functions
\( K(q) \) and \( c(y) \) fully characterise the variability across scales
of the field \( \varepsilon_\lambda \), and are linked by a Legendre transform (Parisi
and Frish, 1985).

In the specific framework of universal multifractals, to-
wards which most multiplicative processes converge (this a
broad generalisation of the central limit theorem, Schertzer
and Lovejoy, 1987, 1997), \( K(q) \) is given by
\[
K(q) = \frac{C_1}{\alpha - 1} (q^\alpha - q)
\]
for a conservative field :\( \langle \varepsilon_\lambda \rangle = 1 \), i.e. \( K(1)=0 \),
where \( C_1 \) and \( \alpha \) have been defined in the introduction.
Equation (4) yields a more mathematical definition for UM
parameters: \( K'(1) = C_1 \) and \( K''(1) = \alpha C_1 \).

UM parameters \( C_1 \) and \( \alpha \) are estimated with the help of the
doUBLE trace moment (DTM) technique (Lavallée et al.,
1993). This technique is based on the fact that if \( \varepsilon_\lambda \) is a mul-
tifractal field, then the field \( \varepsilon^{(n)}_\lambda \), obtained by upsampling
the \( n \)th power of the field at maximum resolution, scales like
\[
\left\langle \left( \varepsilon^{(n)}_\lambda \right)^q \right\rangle \approx \lambda^{K(q,n)}
\]
with
\[
K(q,n) = K(nq) - qK(n).
\]
For universal multifractals, this yields
\[
K(q,n) = n^\alpha K(q).
\]
Therefore the multifractality index \( \alpha \) corresponds to the slope
of the so-called DTM curve, which is the log–log plot of
\( K(q,n) \) vs \( n \) for fixed \( q \).

Finally, let us mention the standard framework to deal with
a non-conservative field, denoted \( \varphi_\lambda \) (i.e. we have \( \langle \varphi_\lambda \rangle \neq 1 \).
In that case it is usually assumed that it can be written as
(with an equality in probability distribution)
\[
\varphi_\lambda = \varepsilon_\lambda \lambda^{-H},
\]
where \( \varepsilon_\lambda \) is a conservative field (\( \langle \varepsilon_\lambda \rangle = 1 \)) of moment scaling
function \( K_c(q) \) (the sub-index “c” refers to the conservativity
of \( \varepsilon_\lambda \)), and \( H \) the non-conservation parameter. \( K_c(q) \) only
depends on UM parameters \( C_1 \) and \( \alpha \). \( H \) characterises the
scale dependence of the average field, i.e.
\[
\langle \varphi_\lambda \rangle \approx \lambda^{-H}.
\]
\( H \) is equal to zero for a conservative field. The moment scal-
ing function \( K(q) \) of \( \varphi_\lambda \) is given by
\[
K(q) = K_c(q) - Hq.
\]
The DTM technique should theoretically be implemented on
\( \varepsilon_\lambda \); however, if \( H < 0.5 \), it can be implemented directly on
\( \varphi_\lambda \) and will not generate biased estimates. In case of greater
\( H \), \( \varepsilon_\lambda \) should be used. Retrieving \( \varepsilon_\lambda \) from \( \varphi_\lambda \) theoretically
requires a fractional integration of order \( H \) (equivalent to a
multiplication by \( k^H \) in the Fourier space). A common
approximation, which provides reliable results (Lavallée et al.,
consists in taking $\varepsilon_\lambda$ as the absolute value of the fluctuations of $\varphi_i$ at the maximum resolution and renormalising it, i.e.:

$$\varepsilon_\lambda = \frac{|\varphi_\lambda(i + 1) - \varphi_\lambda(i)|}{|\varphi_\lambda(i + 1) - \varphi_\lambda(i)|} \times 1 - D. \quad \text{(11)}$$

Then $\varepsilon_\lambda$ is obtained by upscaling $\varepsilon_\lambda$.

$H$ can be estimated with the help of the following formula (Tessier et al., 1993), in which $\beta$ is the spectral slope that characterises the power spectrum of a scaling field, which follows a power law over a wide range of wave numbers:

$$E(k) \propto k^{-\beta}. \quad \text{(12)}$$

### 3 Discrepancies between observed behaviour and standard UM theoretical framework

The first discrepancy with this theoretical framework, which assumes scale invariance, is the occurrence of scaling breaks, enabling the distinguishing of two or more ranges of scales over which the theoretical framework is validated with different parameters. For instance, most authors report a break in Eq. (1) (de Lima and Grasman, 1999; Hubert et al., 1995; Olsson et al., 1993). It appears that for the large scales, the fractal dimension is equal to the dimension of the embedding space (1 in 1-D with time series, 2 in 2-D with maps), indicating that there is some rain everywhere. A break is also often observed in Eq. (3). For instance, Gires et al. (2011) reported one at 16 km on radar data of a heavy rainfall event (a Cevenol episode) that occurred in the south of France 5–9 September 2005. Such a break was also observed on 3 rainfall events in the Paris area (Tchiguirinskaia et al., 2012). In temporal analysis, Fraedrich and Larnder (1993) reported a break at 2–3 h, and one at 3 days on Postdam (Germany) rainfall time series. De Lima and Grassman (de Lima and Grassman, 1999) also reported one at roughly 0.1 h on a 15 min resolution long rainfall time series of Val Poroso, Portugal. Concerning the UM parameter estimates, it appears that $\alpha$ is greater and $C_1$ smaller for small scales than for large scales.

Secondly, let us mention the issue of $H > 0$, indicating a non-conservative, smoother and more correlated field. Indeed, numerous authors report $H > 0$ for various rainfall events. For instance, de Montera et al. (2009) estimated $H$ roughly equal to 0.5 for high-resolution time series of different French rainfall events. Verrier et al. (2010) found $H$ roughly equal to 0.4 in a multifractal spatial analysis on African monsoon radar data. Nykanen (2008) and Nykanen and Harris (2003) analysed radar data of 5 heavy rainfall events in the Rocky Mountains. In these papers they computed the time evolution of $H$, which ranges from 0.31 to 0.61. Gires et al. (2012a, 2011) found $H$ roughly equal to 0.3–0.6 for a Cevenol episode and rainfall event over the London area. It appears that no physical explanation for this non-zero $H$ is provided by these authors. One of the goals of the simple model presented in Sect. 4 is to suggest one.

Thirdly, it appears that the UM parameters found in the spatial analysis of radar rainfall maps are quite different from the ones found in temporal analysis of rain gauge time series. Indeed, in spatial analysis, common values are $\alpha \approx 1.5–1.7$, $C_1 \approx 0.5–0.2$ and $H \approx 0.3–0.6$ (Gires et al., 2011, 2012a; Tessier et al., 1993; Verrier et al., 2010), whereas in temporal analysis they are $\alpha \approx 0.5–0.7$, $C_1 \approx 0.3–0.5$ and $H \approx 0–0.3$ (de Lima and de Lima, 2009; de Lima and Grassman, 1999; Fraedrich and Larnder, 1993; Ladoy et al., 1993; Olsson, 1995; Tessier et al., 1996). Since time series and maps are simply a temporal or spatial cut of the same underlying spatio-temporal process, the results of the analysis performed on both should be related. The simplest space–time-scaling model of rainfall, which corresponds to the linear general scale invariance (Schertzer and Lovejoy, 1985), relies on a unique anisotropy scaling exponent $H_1$ between space and time. In this framework (Deidda, 2000; Gires et al., 2011; Macor et al., 2007; Marsan et al., 1996; Radkevich et al., 2008), one should find

$$K_\text{space}(q) = K_\text{time}(q) \frac{1}{1 - H_1}, \quad \text{(13)}$$

and hence

$$\alpha_\text{space} = \alpha_\text{time} \quad \text{(14)}$$

and

$$\frac{C_1,\text{space}}{C_1,\text{time}} = \frac{H_\text{space}}{H_\text{time}} = \frac{1}{1 - H_1}. \quad \text{(15)}$$

With the help of Kolmogorov theory (Kolmogorov, 1941) and assuming that rain cells have the same lifetime as eddies, it can be shown that $H_1$ is expected to be equal to 1/3 (Marsan et al., 1996). The standard UM parameters found in spatial and temporal analysis are not in agreement with this simplest model. This might be due to a bias inherent to the fact that studied radar data have been selected to correspond to rainfall events, whereas rain gauge time series are usually long (a few months or years) and therefore include many dry periods, which are not appropriately represented in the standard UM framework or correspond to low values of the multifractality index $\alpha$. When the studied period corresponds to a rainfall event in a time series, temporal UM parameters similar to “usual” spatial ones are retrieved. For instance, Montera et al. (2009) analysed 30 s resolution rainfall time series of a few rainfall events and found $\alpha = 1.7$, $C_1 = 0.13$ and $H = 0.53$.

Lastly, it should be mentioned that implementing a threshold on a UM field has a significant impact on the estimates of the UM parameters one can retrieve from this field. To illustrate this, let us consider two simulated conservative UM field with different sets of UM parameters: $\alpha = 1.7$ and...
Two hundred independent samples of mean (on all the samples) equal to one are generated for both cases (a sample is the output field obtained by a run of the discrete cascade process; here, 12 cascade steps are implemented leading to samples of size 4096). A threshold of respectively 9 and 10 (which corresponds to a threshold singularity of respectively 0.26 and 0.27) is implemented in case 1 and 2, leading to 98.1% of zeros in both cases. This percentage is commonly observed on high-resolution (typically time steps lower than 5 min) long (few years) rain gauge time series. The DTM is then applied to these thresholded fields. The obtained estimates are $\alpha = 0.55$ and $C_1 = 0.38$ for case 1, and $\alpha = 0.45$ and $C_1 = 0.44$ for case 2. This is due to the fact that a threshold implemented on a multifractal field results in a multifractal phase transition (on this notion, see Schertzer and Lovejoy, 2011) that affects low moments and introduces a strong bias in the estimates (see Gires et al., 2012b, for a detailed analysis of this effect). It should be noted that the scaling observed (linearity of Eq. 3 in a log–log plot) for both cases is rather good, with coefficients of determination greater than 0.99. This illustrates that thresholding two UM processes with very different statistical properties yields rather similar parameter estimates. More precisely, thresholding can be seen as introducing a bias in the original UM parameter estimates that can be easily much more effective for $\alpha < 1$ than for $\alpha < 1$, which is not surprising because, as we have already mentioned, $\alpha < 1$ corresponds to a UM field with numerous zeros. At first thought the case 1 with $\alpha = 1.7$ and $C_1 = 0.2$ seems a better candidate to explain the parameter estimates for both the event analysis and the longer time series analysis. However, a simple threshold (associated either with a limit of detection or a physical limit) at the maximum resolution does not explain in a satisfactory way the presence of a scaling break, meaning that there is a need for developing a new model that explains both the scaling properties and the zeros.

4 Development of a simple model (UM + 0) to represent zero rainfall in a UM framework

In the previous section we highlighted the fact that the standard framework of UM implemented on rainfall data does not properly explain the presence of scaling breaks, the observation of non-conservative fields and the discrepancies between spatial (usually event based) or temporal analysis (usually on long time series). In this section we present a new simple model, denoted UM + 0, to generate rainfall and its zeros in a multifractal framework. Before introducing it, let us briefly remind how to build a discrete universal multifractal field. We will denote it as $W$ here (see below for an explanation, Fig. 1 for an illustration).

At each step of the cascade process, each time step is divided into $\lambda_1$ time steps. Although it is not mandatory $\lambda_1$ is usually equal to two. After $n$ steps the field is denoted

$$W_1 = \mu W_0,$$

$$W_2 = \mu W_1,$$

$$W_3 = \mu W_2,$$

$$\cdots$$

$$W_n = (\lambda_1^{-1})^{n-1} W_1.$$
$W_{n,i}$, where $i = 1, \ldots, \lambda_1^n$. The value affected to each sub-pixel ($W_{n,i}$) is the value of the parent pixel ($W_{n-1,j}$) multiplied by a random value $\mu w_{n,i}$, i.e.

$$W_{n,i} = W_{n-1,j} \mu w_{n,i}. \quad (16)$$

$i$ is an index that identifies a structure after $n$ steps of cascade process, and hence it varies from 1 to $\lambda_1^n$. $j$ is an index that identifies a structure after $n-1$ steps of cascade process, and hence it varies from 1 to $\lambda_1^{n-1}$. $i$ and $j$ are chosen to ensure the connection between parent and daughter structures, i.e. for a given $j$, $i$ takes $\lambda_1$ different values (for simplicity sake of notation, we do not make explicit the liaison between $i$ and $j$). The random variables corresponding to each multiplicative increment $\mu w_{n,i}$ are independent and identically distributed. To generate a UM field (Eqs. 3 and 4 valid) with parameters $C_1$ and $\alpha$, one must choose as random multiplicative increment $\mu w = \exp \left[ \frac{C_1 \ln(\lambda_1)}{\alpha} L(\alpha) \right] / \lambda_1^{1/\alpha}$ with $L(\alpha)$ being an extremal Lévy-stable random variable of index $\alpha$ (i.e. $\langle \exp(q L(\alpha)) \rangle = \exp(q^\alpha)$), that can be generated with the help of the procedure given by Chambers et al. (1975) or the generalised central limit theorem for Levy variables (Schertzer and Lovejoy, 1987). More details about generation of discrete and the more general continuous UM fields can be found in Pecknold et al. (1993) and Lovejoy and Schertzer (2010).

Let us now explain how the zeros of rainfall are introduced to simulate a rain rate field $R$ from the UM field $W$. As said in the introduction, mainly two techniques have been suggested, either to multiply by an independent random binary field that would correspond to the rainfall support or to threshold the simulated field at the maximum resolution. Both techniques have serious limitations: the independence assumption in the first one is theoretically inconsistent, and both treat in a similar way real and artificial zeros despite their different origin. In this paper we suggest to introduce the zeros within the cascade process (as done in Olsson, 1998) and not independently from the simulated values, and then to threshold the field at the maximum resolution. The two underlying assumptions are the following:

- The conserved quantity is not directly the rain rate but the total flux of water in all its phases (A. Tuck, private communication, 2011) in the atmosphere, and only a portion of it corresponds to rain. The UM + 0 cascade process provides a scale-invariant way of determining which portion and where. This is why the UM field is denoted $W$, and the one with the zeros $R$. They stand respectively for “water flux” and “rain rate”. This should not be taken accurately but more as a general idea.

- At each step of the cascade process, if the rainfall process is below a certain intensity (defined with the help of a singularity $\gamma_0$), then it is not certain to “survive” (defined with the help of a probability $p_0$). This would represent a kind of physical limit to rainfall processes.

More precisely $R$ is derived from $W$ as follows ($\mu_r$ and $\mu_w$ denote the random multiplicative increment for, respectively, $R$ and $W$):

- $R_0 = W_0$, $R_1 = W_1$.
- If $W_{n-1} \leq (\lambda_1^{-1})^{\gamma_0}$, then $\mu_r_{n,i} = \mu w_{n,i}$ with probability $p_0$ or $\mu_r_{n,i} = 0$ with probability $1 - p_0$.
- If $W_{n-1} > (\lambda_1^{-1})^{\gamma_0}$, then $\mu_r_{n,i} = \mu w_{n,i}$ anyway.

$\gamma_0$ and $p_0$ are parameters that characterise the process generating the actual zeros of the rainfall field. It should be noted that since it is a multiplicative process, once the rainfall process has been set to zero at a given resolution, it will remain equal to zero at the higher resolutions. This aims at representing long dry periods. Finally, a threshold (T) is implemented at the maximum resolution to mimic the limit of detection of any rainfall measurement device.

5 Results of UM + 0 for a set of parameters

The aim of this section is to discuss the results of a multifractal analysis performed on a field generated with the UM + 0 cascade process. The UM parameters used are $C_1 = 0.1$ and $\alpha = 1.9$, the resolution of the cascade is $\Lambda = 4096$. The choice of such a high value of $\alpha$ is theoretically motivated to better assess the estimate bias resulting from the thresholding and corresponds to an upper bound of empirical estimates on portions of time series or maps where there are (almost) no zeros. For example, de Montera et al. (2010) reported $C_1 = 0.13$ and $\alpha = 1.7$ for a 30 s time series, Mandapaka et al. (2009) reported $C_1 = 0.18$ and $\alpha = 1.9$ for spatial and temporal analysis performed on high-resolution (2.5 m, 1 s) lidar data, and Verrier et al. (2010) reported $C_1 = 0.12$
and $\alpha = 1.78$ for spatial analysis of radar data (1 km resolution). Concerning the additional parameters introduced in the new model, $\gamma_0 = 0.1$ and $p_0 = 0.5$ are tested in this section. Tests are performed with or without a threshold of 10 implemented on the field $\text{UM} + 0$ previously normalised to one on average and which corresponds to a singularity $\gamma_T = 0.27$. The thresholded field is denoted $\text{UM} + 0 + T$ in the following. The influence of these parameters will be discussed in Sect. 5. The analysis is performed on an ensemble of 1000 independent 1-D samples (i.e. the output field of a realisation of the UM + 0 cascade process) of length 4096. Unless otherwise mentioned, the analysis is performed on the dressed field, i.e. the one obtained by re-aggregating the field from its maximum resolution. Only few parameters are computed on the bare fields, i.e. the ones obtained at each resolution through the cascade process which is not accessible with actual rainfall data.

The percentage of zero is equal to 95% and 97% for the $\text{UM} + 0$ and $\text{UM} + 0 + T$ field, respectively. This would mean that with such a limit of detection, the rainfall measurement devices miss slightly more than a third of the rainy time steps due to the fact that $\gamma_T > \gamma_0$. Fractal analysis (Eq. 1 in a log–log plot) is performed and displayed in Fig. 2 for the UM + 0 field. By taking into account the whole range of scales, $D_F$ is estimated to 0.67 (with a coefficient of determination $R^2$ greater than 0.99), which is in agreement with standard values. However, when considering a break at $\lambda = 128$, one finds $D_F = 0.78$ for small scales and $D_F = 0.59$ for large scales. This does not reproduce the behaviour of actual rainfall fields, which usually exhibit greater fractal dimensions for large scales than for small ones. This is a serious limitation of this model. As expected the estimated fractal dimensions are slightly smaller with a threshold (0.72 and 0.57 for respectively small and large scales), but they exhibit similar overall behaviour.

TM (Eq. 3 in a log–log plot) and DTM analysis are performed on the field dressed from the maximum resolution of 4096, as it is done with actual rainfall data. Figure 3a displays TM analysis for the UM field. As expected the scaling is excellent ($R^2 < 0.997$ for all the statistical moment orders), and the DTM analysis (Fig. 3d) yields $C_1 = 0.094$ and $\alpha = 1.91$, which is very close to the parameters inputted in the simulations. With the $\text{UM} + 0$ field the scaling is worsened, as can be seen on Fig. 3b, where a curvature is visible for all the statistical moment orders. For example, the $R^2$ for $\eta = 2$ without considering any break is equal to 0.95. In the following this parameter will be taken as an indicator of the quality of the scaling. When taking into account a break at $\lambda = 128$, this same $R^2$ is equal to 0.98 and 0.97 for small and large scales, respectively. $K(q)$ (see Fig. 3c) of the $\text{UM} + 0$ field for small scales remains rather similar to the one of UM field except for small moments where the multifractal phase transition associated with numerous zeros introduces a bias. For large scales, the discrepancies are much greater.

Fig. 3. (a) Scaling of $W_{\lambda}$ (Eq. 3 in log–log plot) (b) scaling of $R_{\lambda}$, (c) $K(q)$ for $R_{\lambda}$, and (d) determination curve of $C_1$ and $\alpha$, for simulation parameters $C_1 = 0.1$, $\alpha = 1.9$, $\gamma_0 = 0.1$ and $p_0 = 0.5$. 

even for great moments, and the curvature of $K(q)$, which reflects the multifractality of the associated field, is almost lost. The determination curve of UM parameters in the DTM analysis (Eq. 7 in a log–log plot) are displayed Fig. 3d, where the linear portion used to estimate $C_1$ and $\alpha$ is highlighted. This leads to $C_1 = 0.32$ and $\alpha = 0.69$ for large scales and $C_1 = 0.11$ and $\alpha = 1.33$ for small scales. When a threshold is implemented, the quality of the scaling remains similar and the UM parameters are slightly more affected. Indeed, one finds $C_1 = 0.33$ and $\alpha = 0.65$ for large scales and $C_1 = 0.13$ and $\alpha = 1.22$ for small scales. This simple model reproduces the scaling break and the main variations of UM parameters between small and large scales. Finally, it should be noted the behaviour observed on small scales enables much more accurate retrieval of the estimates of the underlying UM field.

One of the main features of the UM + 0 cascade process is that it is not conservative, i.e. since some time steps (with small singularities) are set to zero at each step, the average mean decreases with the number of cascade steps. To analyse more precisely this effect, Fig. 4a displays the mean of the process computed on the bare field, for the UM and UM + 0 process, vs. the resolution in a log–log

Fig. 4. Estimation of the non-conservation parameter Bare $H$ (a) and Dressed $H$ through the spectral behaviour (b) for simulation parameters $C_1 = 0.1$, $\alpha = 1.9$, $\gamma_0 = 0.1$ and $p_0 = 0.5$.

Fig. 5. Considering only the heaviest rainfall: (a) scaling of $R_\lambda$, (b) $K(q)$ and (c) determination curve of $C_1$ and $\alpha$, for simulation parameters $C_1 = 0.1$, $\alpha = 1.9$, $\gamma_0 = 0.1$ and $p_0 = 0.5$. 

plot. Linear regressions are furthermore performed to evaluate non-conservation parameters $H$ as defined in Eq. (8) (it will be called bare $H$ in the following). As expected for the UM field which is conservative, bare $H$ is roughly equal to 0. For the UM + 0 field the mean naturally decreases with the resolution. However, the curve does not exhibit a linear behaviour in this log–log plot, indicating that the standard framework of Eq. (8) with a single $H$ for the whole range of scales is not appropriate for this model (and may be not also for rainfall). Estimating the bare $H$ on two ranges of scales (which is not perfectly accurate since the curve exhibits more a curvature than two linear portions), one finds $H = 0.09$ for small scales and $H = 0.25$ for large ones. This is not what is observed in the literature where the estimates of $H$ are usually smaller for large scales. It should be mentioned that smaller values of $H$ are found when estimated on the dressed field with the help of the spectral slope (Eq. 12) (it will be called dressed $H$ in the following). Figure 4b displays the power spectrum for the UM + 0 field. A break was considered (for $k = 128$, which corresponds to $\lambda = 128$) to be consistent with the TM analysis, but similar values of $\beta$ are found for both ranges of scales (roughly $\beta = 1$), which indicates that there is basically no need for a break. This corresponds to the theoretical fact that we did not introduce a characteristic scale in the $R$ process. However, we numerically find $H = 0.09$ for small scales and $H = 0.17$ for large ones. These differences between both ranges of scales are not due to a change in $\beta$ but to a change of the multifractal correction $K_{\alpha}(2)$ in Eq. (12). Furthermore, the estimates of $\beta$ are basically equal to those on the UM field, indicating that the simple model does not modify the spectral behaviour. Similar results are found on the UM + 0 + T fields. This simple model does not reproduce the spectral behaviour and the values of $H$ for the dressed field according to the range of scales reported in the literature. Nevertheless, this model provides a framework to explain the fact that $H \neq 0$ on many observed rainfall fields, which is not the case for the models consisting in a simple threshold or a multiplication by an independent support. A possible improvement of this simple model could be to introduce a vertical velocity component, as suggested in Lovejoy and Schertzer (2008), where the authors developed a rain rate model, consisting in a product of a liquid water content and a vertical velocity, which reproduced the spectral properties of actual rainfall fields.

In order to mimic what is done when a rainfall-event-based analysis is performed (i.e. only considering a rainy portion of a time series corresponding to a rainfall event), a multifractal analysis is performed only on the portion of the field with the heaviest simulated rainfall. To achieve this, the 128 consecutive time steps with the maximum cumulated rainfall depth are selected for each of the 1000 independent samples of length 4096. Then only the 50 with the greatest depth are kept for the analysis. Figure 5a displays the TM analysis, and shows that no scaling break should be considered ($R^2 > 0.99$ for $q > 0.5$). Concerning $K(q)$, the curves obtained with only the heaviest rainfalls of the UM + 0 and UM + 0 + T fields are rather similar to the one of the UM field (see Fig. 5b). Concerning the UM parameters estimates (Fig. 5c), one finds $C_1 = 0.094$ and $\alpha = 1.60$ for the heaviest rainfall of the UM + 0 field. The estimates are slightly more biased for the heaviest rainfall of the UM + 0 + T field ($C_1 = 1.01$ and $\alpha = 1.49$), but remain much closer to the estimates of the underlying UM field than when the analysis is performed on the whole field. It is interesting to note that this simple model reproduces UM parameters estimates found in the literature for both long time series analysis (with its large-scale results when the whole field is considered) and event-based analysis (by selecting only the heaviest rainfalls). Finally, it should be mentioned that such analysis on the heaviest rainfall was performed on simulated UM fields with $C_1 = 0.5$ and $\alpha = 0.5$ as parameter values of the simulation and lead to $C_1 = 0.31$ and $\alpha = 0.57$ as estimates. These values are quite different from the standard one for rainfall events, which confirms that the standard UM parameters retrieved from long time series cannot be used to explain event-based ones.

6 Sensitivity of UM + 0 to its various parameters

In this section the sensitivity of the model to the parameters $\gamma_0$ and $p_0$ is analysed. All the possible combinations with $\gamma_0 = -0.1; 0; 0.1; 0.2$ and $p_0 = 0.1; 0.3; 0.5; 0.7; 0.9$ are tested, always keeping $C_1 = 0.1$ and $\alpha = 1.9$. Concerning the threshold, we proceed as follows: (i) normalisation of the generated UM + 0 field on average at the maximum resolution (as it is commonly done before performing a multifractal analysis) and (ii) implementation of a threshold $T = 10$ (the same for all the parameter sets), which corresponds to a singularity $\gamma_T = .27$. Figure 6a displays the percentage of zeros at the maximum resolution ($\Lambda = 4096$). As expected without any threshold, the percentage of zeros increases with $\gamma_0$ and decreases with $p_0$. For small $p_0$ (roughly < 0.5) the influence is not really significant. When a threshold is implemented, the differences basically disappear between the different fields, with percentage of zeros always greater than 95 %. The fractal dimensions for large and small scales are displayed in Fig. 6b and c, respectively. For small scales, without any threshold, as expected $D_F$ decreases with $\gamma_0$ and the influence of $p_0$ is not really significant. An inverse behaviour is observed with a threshold. The differences between fields thresholded and those not are much greater for small $\gamma_0$ and $p_0$, i.e. when the proportion of “real” zeros is greater. For example, for $p_0 = 0.1$, the fields with $\gamma_0 = 0.2$ (i.e. closer to $\gamma_T$) exhibit the same small-scale fractal dimension, whereas for $\gamma_0 = -0.1$, $D_F$ is equal to 0.93 for the unthresholded field and to 0.50 for the thresholded one. Basically, similar behaviour is observed for large scales, with values slightly smaller than for small scales, which is (as
mentioned in the previous section) contrary to the observations on actual rainfall.

Figure 7a displays the coefficient of determination $R^2$ of the linear regression which gives $K(q)$ (Eq. in a log–log plot) for $q = 2$ when no break is taken into account. It is an indicator of the quality of the scaling and the necessity to consider a break. It decreases with greater $\gamma_0$ and smaller $p_0$. There is a slight but not significant improvement when a threshold is implemented. In any case the values remain always rather low, and a break at $\lambda = 128$ is therefore taken into account in the following. For small scales (Fig. 7c and e), when there is no threshold the estimate of $\alpha$ decreases and the one of $C_1$ increases with greater $\gamma_0$ and smaller $p_0$. The differences are smaller for $C_1$ than for $\alpha$. Similar behaviour is retrieved for large scales with stronger bias (Fig. 7b and d). When a threshold is implemented, inverse behaviour is observed, indicating that estimates are more biased when the proportion of zeros due to the threshold (i.e. false ones) is greater. It should be noted that with a threshold the disparities due to $\gamma_0$ and $p_0$ are damped. Finally, it should be mentioned that the curves with $\gamma_0 = 0.2$ exhibit quite different patterns than the others, especially when a threshold is implemented. This is likely to be due to the fact that in that case $\gamma_0 > C_1$, meaning that a too significant proportion of the field was removed by the UM + 0 process and that therefore the threshold has a smaller influence.

The non-conservation bare and dressed $H$ exponents of the UM + 0 fields are displayed Fig. 8. For both small and large scales they naturally increase with greater $\gamma_0$ and smaller $p_0$ (i.e. with more real zeros generated). However, the dependency in $p_0$ for small scales is not significant. As discussed in the previous section, $H$ is greater for large scales than for small scales, which is contrary to what is observed on actual rainfall fields. The values found on the dressed fields are smaller than the bare ones and do not reach the ones found in the literature (0.3–0.6). Considering the bare ones, it seems that one should have $p_0 > 0.5$ and $\gamma_0 > 0.1$ for actual rainfall fields.

Figure 9 displays $C_1$ and $\alpha$ estimated only on the heaviest rainfalls. The same comment concerning the dependencies in $\gamma_0$ and $p_0$ as for the estimates of $C_1$ and $\alpha$ computed on the whole field (i.e. not only the heaviest event) could be made. The main difference is that the estimates are much less biased.

Finally, it should be said that the same analysis was performed with UM parameters $C_1 = 0.2$ and $\alpha = 1.7$ instead of $C_1 = 0.1$ and $\alpha = 1.9$, and that similar results were found. The main differences consist in a slightly better scaling (that nevertheless requires the introduction of a scaling break), smaller estimates of $H$ (which suggests the UM parameter set presented more in details in this paper is more relevant even though obtained values of $H$ are smaller needed), and of course slightly smaller values of $\alpha$ and greater ones for $C_1$. 

Fig. 6. Analysis of the sensitivity to $\gamma_0$ and $p_0$: percentage of zero (a), $DF$ for large scales (b) and $DF$ for small scales (c).
Fig. 7. Analysis of the sensitivity to $\gamma_0$ and $p_0$: coefficient of determination $R^2$ of the linear regression in the TM analysis for $q = 2$ (a), $\alpha$ and $C_1$ for large and small scales (b, c, d, e).

Fig. 8. Analysis of the sensitivity to $\gamma_0$ and $p_0$: Bare $H$ and Dressed $H$. 
7 Conclusions

In this paper a new simple model to simulate the numerous zeros of rainfall fields in the framework of discrete universal multifractal cascades is presented. The real zeros are explicitly distinguished from the artificial ones. These zeros are generated within the cascade process, in a scale invariant way, and not independently from the rainfall values. It basically consists in setting to zero with a probability 1\(-p_0\) the rainfall field if it is smaller than a given singularity \(\gamma_0\), at the previous step of the cascade process. This process is called UM + 0. The underlying assumptions are that the conserved quantity is the total flux of water in all its phases in the atmosphere and not the rain rate, and that if there is not enough water, then the rainfall process is not certain to occur. In that way the model can be physically justified. Finally, the field is thresholded at its maximum resolution to mimic the limit of detection of any rainfall measurement device.

This rather simple scale-invariant model enables us to retrieve many of the properties observed on actual rainfall fields: a scaling break, non-conservation (only a portion of it is explained), differences between UM parameters for small and large scales and discrepancies between event-based analysis and long time series analysis. Nevertheless, some of the properties of the generated fields are not in agreement with those of actual rainfall fields: the spectral behaviour is not modified, the exponent of non-conservation is greater for large scales than for small scales and its dressed values are too small, and the fractal dimension is greater for small scales than for large ones. This means that this UM + 0 model should be considered as nothing more than an interesting and encouraging simple model.

More generally the results found with this simple model suggest that the underlying UM parameters for rainfall are rather constant (\(\alpha = 1.7 - 1.9\) and \(C_1 = 0.1 - 0.2\)), and what would change from one place to another is the way the zeros are generated. This process would explain most of the discrepancies between the theoretical multifractal framework and observed rainfall behaviour.

To confirm these results further investigation are needed: (i) empirical multifractal analysis of the total water flux in all its phases in the atmosphere, and not only rain rate should be performed; (ii) the model generating zeros should be refined to retrieve all the observed properties; (iii) a way to accurately estimate the parameters of the model should be developed. A goal of this paper is to foster research on these issues.

Acknowledgements. The authors greatly acknowledge partial financial support from the Chair “Hydrology for Resilient Cities” (sponsored by Veolia) of Ecole des Ponts ParisTech, and EU NWE Interreg IV RainGain project (www.raingain.eu).

Edited by: N. S. Erokhin
Reviewed by: two anonymous referees

References

A. Gires et al.: Zero rainfall in a UM framework


