



Imperfect scaling in distributions of radar-derived rainfall fields

M. J. van den Berg¹, L. Delobbe², and N. E. C. Verhoest¹

¹Laboratory of Hydrology and Water Management, Ghent University, Coupure links 653, 9000 Ghent, Belgium

²Royal Meteorological Institute, Avenue Circulaire 3, 1180 Uccle, Belgium

Correspondence to: M. J. van den Berg (martinus.vandenberg@ugent.be)

Received: 28 July 2013 – Published in Hydrol. Earth Syst. Sci. Discuss.: 5 September 2013

Revised: 4 October 2014 – Accepted: 2 November 2014 – Published: 19 December 2014

Abstract. Fine-scale rainfall observations for modelling exercises are often not available, but rather coarser data derived from a variety of sources are used. Effectively using these data sources in models often requires the probability distribution of the data at the applicable scale. Although numerous models for scaling distributions exist, these are often based on theoretical developments, rather than on data. In this study, we develop a model based on the α -stable distribution of rainfall fields, and tested on 5 min radar data from a Belgian weather radar. We use these data to estimate functions that describe parameters of the distribution over various scales. Moreover, we study how the mean of the distribution and the intermittency change with scale, and validate and design functions to describe the shape parameter of the distribution. This information was combined into an effective model of the distribution.

1 Introduction

Rainfall is one of the most important drivers of hydrological processes and is an important data source for hydrological modelling. These models typically operate on a spatial scale of less than 100 km (Ferraris et al., 2003), and a temporal scale of about an hour, requiring data at a similar spatio-temporal scale. The availability of suitable data, especially for prediction, is often not guaranteed as the output of circulation models and weather prediction models are typically of a much coarser resolution. Furthermore, the variability beneath the scale of simulation has been found to be important in hydrological modelling (Harris and Foufoula-Georgiou, 2001; Gires et al., 2012a).

Whenever suitable data are not available, the scaling behaviour in rainfall can be exploited to yield a statistical esti-

mate of the rainfall at a finer scale. At a very basic level, this behaviour leads to a cascade of scales,

$$\varepsilon_n = \varepsilon_0 \prod_{j=1}^n \mu\varepsilon_j, \quad (1)$$

where ε_0 is a coarse-scale field or the field average, and $\mu\varepsilon$ are multiplicative increments drawn from some distribution. To increase the number of pixels, each pixel at scale n is split into several pixels at scale $n + 1$, as illustrated in Fig. 1. A variety of such cascades have been proposed, starting from Kolmogorov (1941), who described homogeneous turbulence based on the Navier–Stokes equation, often referred to as simple scaling. These scaling laws have been modified in a variety of ways, leading to more complex scaling fields such as isotropic multifractal cascades (Parisi and Frisch, 1985) and their anisotropic counterpart, generalized-scale invariance (Lovejoy and Schertzer, 2013). Furthermore, the parametrization of these models has been eased by innovations such as the universal multifractal model (cascade) (Schertzer and Lovejoy, 1987) and the fractionally integrated flux (Schertzer and Lovejoy, 1987, 1997). Various other models and methods exist, based on different generators (roughly the increments $\mu\varepsilon$), such as log-Poisson generators (Deidda, 2000), log- β generators (Menabde and Sivapalan, 2000) and a other simulation methods including wavelets (Venugopal et al., 2006). More of these models can be found in Gupta and Waymire (1993), Menabde et al. (1997) and Koutsoyiannis et al. (2010), and a good general introduction to these methods can be found in Schertzer et al. (2002) and Tchiguirinskaia et al. (2000).

The above multifractal models can be employed to simulate (non-zero) rainfall according to a few parameters (i.e. they are universal; see Sect. 2). Generally, these models assume that $\mu\varepsilon$ is identically and independently distributed

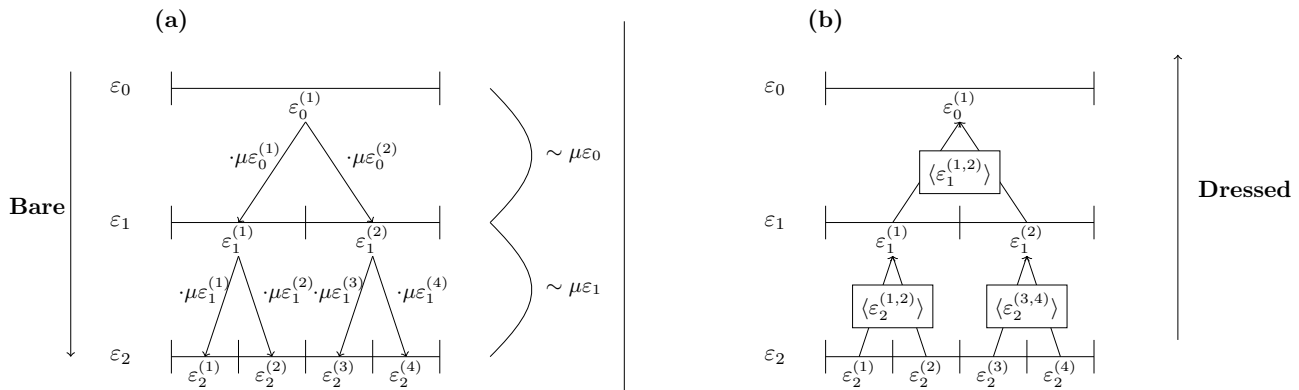


Figure 1. A basic rainfall model, graphically illustrated. The left-hand side of the image is the dressing procedure, whereas the right-hand side is the generation.

(i.i.d.). That is, the distribution of $\mu\epsilon$ is the same at every scale, and the draws are independent of other variables within the large-scale cascade. However, in recent years some criticism on these models has arisen, stating that real rainfall does not “perfectly” scale, but violates the underlying assumption of the i.i.d. $\mu\epsilon$. Empirical investigation of the scaling behaviour does indeed show that not all rainfall fields obey the basic assumption that the increments of ϵ between scales are i.i.d. Divergences from this behaviour were described by various authors who observed that the increments were dependent on factors such as large-scale rainfall intensity (Deidda, 2000; Over and Gupta, 1994) and pixel size (Menabde et al., 1997; Over and Gupta, 1994; Paulson and Baxter, 2007). Additionally, scaling behaviour was found to differ with the intensity of storms (Venugopal et al., 2006) and thus the non-raining intervals do not scale (Olsson, 1998). These deviations from perfect scaling are further examined in Veneziano et al. (2006), Serinaldi (2010), and Rupp et al. (2009), who showed that it is possible to model these imperfections in scaling through empirical functions of the parameters of various downscaling models.

In previous investigations, imperfect scaling has been studied by fitting and refitting various cascade models and studying the dependence of the parameters on coarse-scale intensity and other variables (e.g. Serinaldi, 2010; Rupp et al., 2009; Veneziano et al., 2006). Some other studies have investigated the dependence of breakdown coefficients, i.e. under the assumption that $\mu\epsilon \in [0,1]$, splitting the mass at the coarse scale at each scale step (see e.g. Rupp et al., 2009). In this study we directly investigate the dependence of the empirically observed distributions of the increments on scale and coarse-scale intensity. To do this, we investigate the empirically found distributions of $\log \mu\epsilon$ and $\log \epsilon$ for variations between scales. Furthermore, we investigate how between scale correlations and scale variance behave, for a variety of storms, by characterizing them with a suitable set of equations.

We start by explaining the simulation of rainfall (Sect. 2), followed by a description of the data and some investigation into its basic scaling behaviour (Sect. 3). Then, the α -stable distribution is described in some detail in Sect. 4. This is followed by the methodology in Sect. 5 and the result (Sect. 6). Finally, we conclude with some discussion and conclusions.

2 Simulating and investigating multifractals

Scale-invariant processes and their generation are perhaps easiest understood in the context of discrete-in-scale (discrete) cascades (Parisi and Frisch, 1985; Schertzer and Lovejoy, 1987). In the discrete cascade, a multifractal process is constructed at n discrete scales by perturbing a coarse-scale field ϵ_0 with i.i.d. multiplicative increments $\mu\epsilon_j$; see Eq. (1). The increments $\mu\epsilon$ “inject” energy into the “flux” causing the field to become more volatile at finer scales. Due to its multiplicative nature, this field is highly singular, having many small values and only a few (very) large values. The moments of the cascade behave as (Schertzer and Lovejoy, 1987)

$$\langle \epsilon_\lambda^q \rangle = \lambda^{K(q)}, \quad (2)$$

where $K(q)$ is a moment scaling function and $\lambda = L_{\text{eff}}/l$ with $l = 2^n$. Here, L_{eff} is an outer scale at which the moments converge. Furthermore, $\langle \cdot \rangle$ denotes a field or ensemble average.

The field described above is multifractal and no longer has a single fractal dimension, but rather an infinity of fractal dimensions, each associated with a specific singularity. Evidently, this is problematic, requiring an infinity of parameters to describe the behaviour. In practice, these cascades converge to a universal multifractal if the increments $\mu\epsilon$ are from a log-stable distribution (Veneziano and Furcolo, 1999; Schertzer and Lovejoy, 1987), i.e. $\log \mu\epsilon \sim S_\alpha(\cdot)$ where $S_\alpha(\cdot)$ is the (Lévy) α -stable distribution (see Sect. 4).

If this is the case, $K(q)$ has the form

$$K(q) = \frac{C_1}{\alpha - 1} (q^\alpha - q), \quad (3)$$

where there are only two parameters, the co-dimension of the mean C_1 and a parameter α which controls the tail of the α -stable distribution, where $\alpha = 2$ leads to a normal distribution for $\log \mu \varepsilon$ and decreasing values for α lead to increasingly heavy tails.

Fields simulated with the above method are “conservative”: they are the direct outcome of multiplicative cascades and the realizations themselves are scale invariant. However, for most observed rainfall fields only the fluctuations of the field scale, i.e. $|\varepsilon(x) - \varepsilon(x + \Delta x)|$ (in one dimension) scale rather than the direct realizations ε themselves. Such fields are termed non-conservative and have an additional scaling component. This additional component can be modelled as

$$\varphi_\lambda = D^{-H} \varepsilon_\lambda; \quad (4)$$

i.e. the field is fractionally integrated to an order of H , the non-conservation parameter. The fluctuations of this field scale with (Davis et al., 1994)

$$\langle |\Delta \varphi_\lambda(\Delta x)|^q \rangle = \Delta x^{\zeta(q)}, \quad (5)$$

where the structure function $\zeta(q)$ has a direct relation to $K(q)$:

$$\zeta(q) = qH - K(q), \quad (6)$$

if the field is isotropic. Since $K(1) = 0$, the non-conservation can easily be estimated from the first-order structure function.

A further convenient way to diagnose whether measured fields are non-conservative is the relation to the slope of the Fourier power spectrum. The power spectrum of scaling fields behaves as

$$E(k) = |k|^{-\beta}, \quad (7)$$

where k is the wave number, and the exponent β relates to $K(q)$ and H as

$$\beta = 1 + 2H - K(2). \quad (8)$$

Since $K(2) \geq 0$, conservative fields will always have $\beta \leq 1$, and non-conservative fields generally have $1 \leq \beta \leq 3$ since H is generally between 0 and 1 for rainfall fields (Lovejoy and Schertzer, 2013; Davis et al., 1994).

Multifractal fields generated with the above model produce non-zero values everywhere, and thus they are only appropriate to simulate regions where it is raining everywhere. To overcome this, rainfall is often assumed to be the result of two separate processes, one to determine where it is raining, the support of the rainfall field, and another to determine the observed rain rates. Several different methods of introducing

zero values have been proposed in literature, generally there are those which simulate a separate (mono-)fractal rainfall support (e.g. Reborá et al., 2006) and those that set values below a certain threshold to zero (e.g. Ferraris et al., 2002). Both methods have their merit, however, practical analysis has proven difficult and it remains unclear whether the methods are correct or which is best. In this paper, we assume that the support of the rainfall is a monofractal field and analyse it as such.

Evidently, the above fields are simulated only to a finite scale. In contrast, if the observed fields were simulated according to such a model, they would be developed to an infinite scale and then integrated back up by the radar. This distinction is referred to as a dressed cascade, i.e. it has been developed to an infinite scale and then integrated back up. Fields simulated only to a finite scale, without integration, are referred to as bare cascades with fields in between being partially dressed. This difference is shown in Fig. 1. Although important for a variety of statistical measures, it is impossible to remove these effects (and thus get a direct view of the bare process) and we are left with having to estimate the bare process from the dressed cascade. Lovejoy and Schertzer (2013) showed that, for a variety of approaches, this is indeed valid.

3 Data

The data for this study were acquired by a C-band weather radar near Wideumont, Belgium, operated by the Belgian Royal Meteorological Institute (RMI). This installation covers a circular area with a radius of 240 km, producing a multilevel scan every 5 min. The region covered includes coastal landscapes to the west, and a low mountain range, the Ardennes, to the east with land cover mostly composed of forests, urban development and agriculture. The entire region has a temperate climate and receives about 800 mm of rain annually, almost uniformly distributed throughout the year (De Jongh et al., 2006) and a mean monthly temperature which varies between 18 °C in June and 3 °C in January.

The actual 5 min radar images are taken from large events during 2009, with 9 winter storms and 17 summer storms. These images were extracted from a 6-month time series during which larger storm episodes were selected to ensure sufficient data. These images correspond to the basic 5 min interval images; however, to reduce the data load, we opted to use only the first image of each hour. The images used were not aggregated in order to retain the basic spatial scaling behaviour as well as to avoid ripple effects (Delobbe et al., 2006) and possible temporal scaling.

The raw radar data are produced by a 5-elevation scan performed every 5 min. Measurements are collected up to 240 km with a resolution of 250 m in range and 1° in azimuth. A time-domain Doppler filtering is applied for ground clutter removal. An additional treatment, based on a static clutter

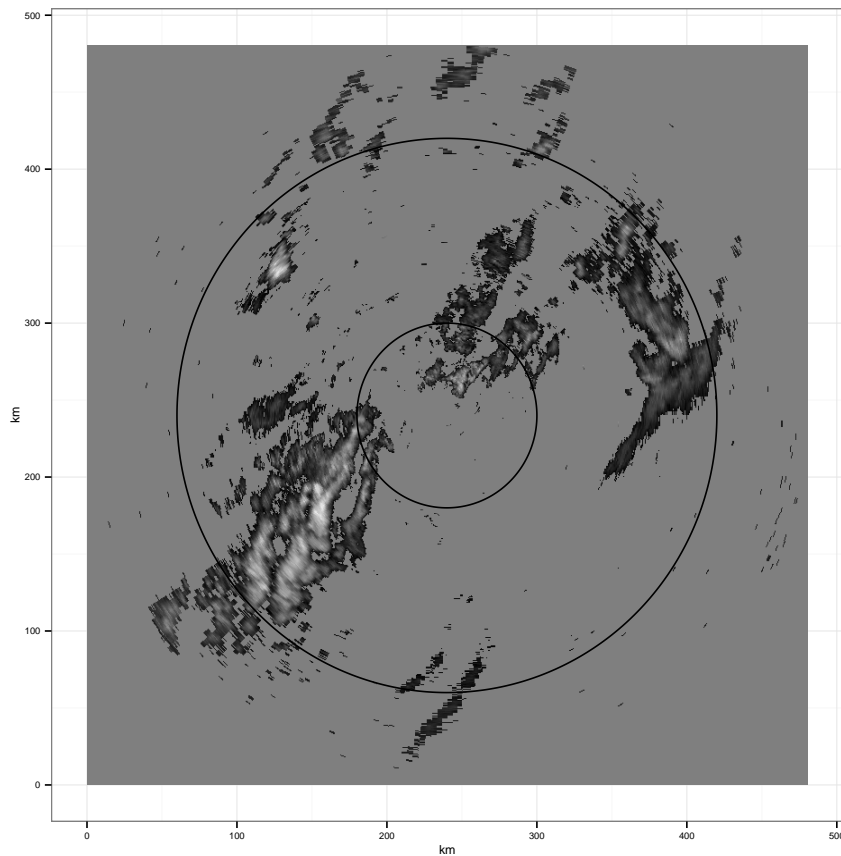


Figure 2. A log-transformed rainfall field, together with the radius of reliable observations (circles) at 60 and 180 km.

ter map, is applied to eliminate residual permanent ground clutter (e.g. buildings). The radar data are then stored as digital numbers representing the reflectivity values ranging from -31.5 to 95.5 dB in steps of 0.5 dB. A two-dimensional radar product is then extracted from the three-dimensional polar data on a Cartesian grid with a resolution of $0.6 \text{ km} \times 0.6 \text{ km}$ (Goudenhoofd and Delobbe, 2009). Reflectivity values are then converted into precipitation rates using the Marshall–Palmer relation

$$\varphi_{\text{obs}} = \sqrt[b]{\frac{10^{0.1 \cdot Z_{\text{dB}}}}{a}}, \quad (9)$$

where Z_{dB} is the reflectivity in decibels and a and b are dimensionless parameters, respectively equal to 200 and 0.6.

As with all weather radars, not all measurements are suitable for quantitative analysis. Firstly, the radar cannot accurately measure rain rates below $\sim 0.1 \text{ mm h}^{-1}$, dependent on distance from the radar. Moreover, the measurements within 60 km of the radar were found to be strongly corrupted with speckle as well as those further than 180 km from the radar. Because of this, the rain rates below 0.1 mm h^{-1} were set to zero, and values closer than 60 km or further than 180 km were discarded. An example of a rainfall field together with the radii between which points were kept is shown in Fig. 2.

3.1 Power density spectrum and multifractal analysis

We analysed the rainfall fields both individually and for each of the storms (by averaging the power spectra of each image in the storm), prior to any changes made to the image, i.e. the raw fields φ_{obs} . The spectra of all storms are shown in Fig. 3, together with a straight line fitted to the linear portion of the power spectrum. In each of these storms, linear behaviour is easily visible, with a break at about 15 km for summer storms and no clear observed scaling break for winter storms. This difference in the range over which the image scales is easily explained by the generally smaller scale of convective summer storms, in contrast to the large-scale stratiform systems typical of winter precipitation.

Furthermore, the summer storms tend to have a β in excess of 3, suggesting that the non-conservation coefficient H is larger than 1 (see Eq. 8). Although not often observed, this is possible in the sense that a fractional integration is not restricted to $H \leq 1$. However, the observed images contain a lot of noise, and generally only relatively few images were available in these series, suggesting that these results may be spurious.

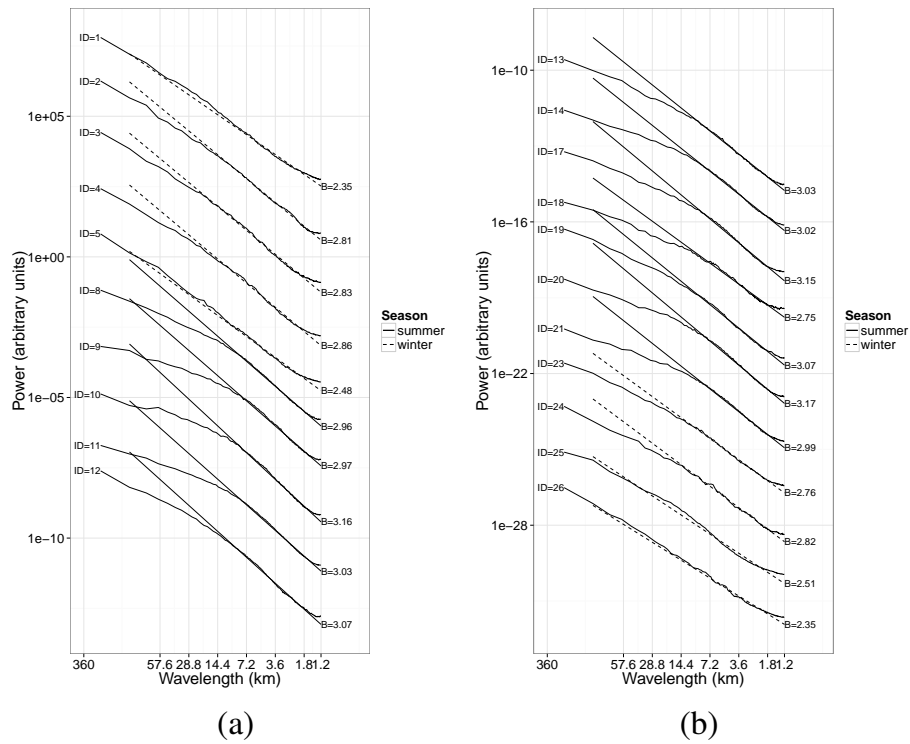


Figure 3. The power spectra of all rainstorms, up to a range of 180 km from the radar, for images which have more than 10 % active pixels and storms with at least 10 valid images. The storm spectra are found by averaging together the spectra of each of the images. The number at the end of each line is the slope β , whereas the number at the beginning is an ID number for each of the storms. Summer storms show a higher slope β , and a short range of scaling than do winter storms. This is possibly explained by the generally smaller scale of convective summer precipitation. The short scaling break at the smallest scales is likely a result of non-raining areas.

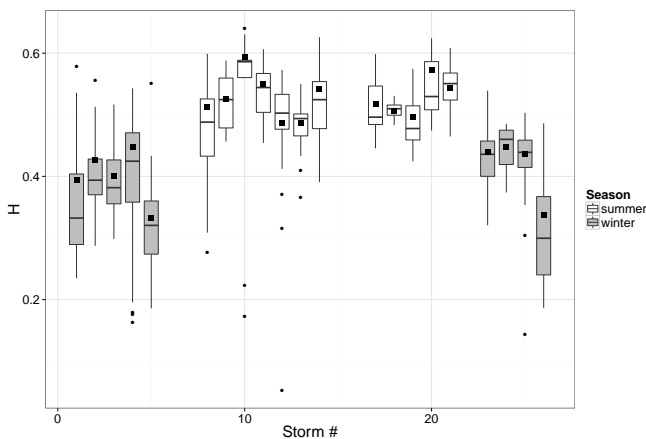


Figure 4. A boxplot of the non-conservation parameter H for all images. The overall parameter for each storm is marked as a black square. Summer storms show a higher non-conservation parameter than winter storms, suggesting a smoother behaviour within the storm (as structure functions are computed only over raining areas).

To find corroboration for the slopes with $\beta \geq 3$, the first-order structure functions were computed. The slope of these functions, up to a break, is equal to the non-conservation co-

efficient H . These slopes are shown in Fig. 4, for each storm as a whole and for all images individually. Note that even though the summer storms do have a higher H , they do not exceed one. This lends credence to the notion that the overly large β are due to speckle and other problems, and not due to characteristics of the rainfall.

3.2 Singular moment analysis

The moment scaling functions of the rainfall images and storms φ were analysed to determine the parameter α . The co-dimension of the mean C_1 is dependent on the outer scale L_{eff} . However, the truncation of the field due to the lower detection limit introduces a spurious outer scale (Lovejoy and Schertzer, 2013), and thus C_1 cannot be determined accurately. As it holds no importance for the remainder of the paper, it was not generally determined.

Instead of the regular double trace moment, the fields were analysed using the weighted multifractal analysis (WMA) (see Gires et al., 2012b; Verrier et al., 2011). This is the usual double trace moment analysis, with the following differences:

- The averages taken in upscaling are only over raining pixels.

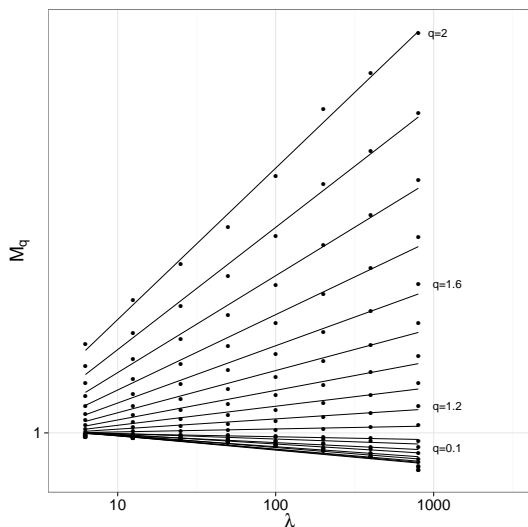


Figure 5. The empirical moments M_q as a function of scale for storm 1 in its entirety. The moments show good scaling characteristics as evidenced by the good fit of the linear function for each of the moments q .

- Each pixel has a weight associated with the fraction of rainy pixels within the disjoint boxes at the finest-scale level.

This analysis is similar to the scaling described by Eq. (13), but over disjoint boxes rather than a moving average. This results in an overweighting of the pixels with more rainy values providing more accurate values for α and C_1 .

The results of this analysis are shown in Fig. 5 for the entire first storm. First, the resulting moments have been taken over a range over scales with $\eta = 1$ to determine the scaling of the moments. It is easily observed that most of these moments do indeed show a straight line for a large portion of their entire scaling range. The fit to the empirical moment scaling function $K(q)$ is shown in Fig. 6. The moment scaling function appears to be well captured by the theoretical form of Eq. (3), suggesting that the field is indeed multifractal. The α parameters for all storms (not shown) are all close to 2. Hence, it appears that the cascade is log-normal and multifractal in nature for all storms.

4 α -stable distributions

As mentioned in the introduction and Sect. 2, the logarithm of the rainfall fields ε_λ and their increments $\mu\varepsilon_\lambda$ are assumed to be distributed according to the α -stable distribution. The α -stable family of distributions allows for a large variety in behaviour, including right- and left-skewed behaviour, as well as symmetric behaviour. Furthermore, the distribution allows either a heavy tail or a light, vanishing, tail on either side, or on both sides, of the mode. Due to this highly flexible behaviour, it includes several well-known distributions

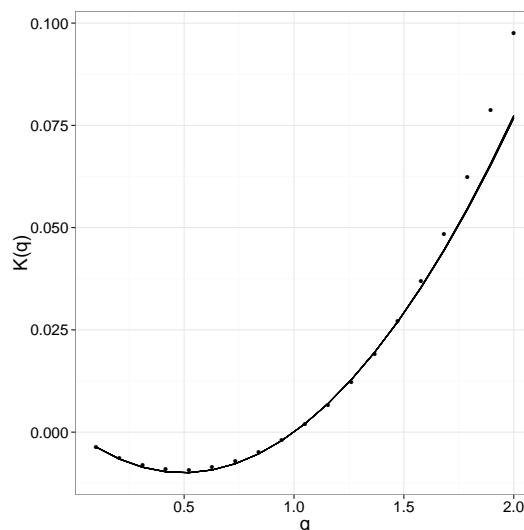


Figure 6. The empirical function $K(q)$, together with the fitted function using the WMA method. The empirical function crosses zero for $K(0)$ because of the use of the WMA method. Above $q \sim 1.5$ the fit starts to diverge from the empirical function, suggesting that above $q = 1.5$ the scaling breaks and the empirical $K(q)$ takes on a linear character.

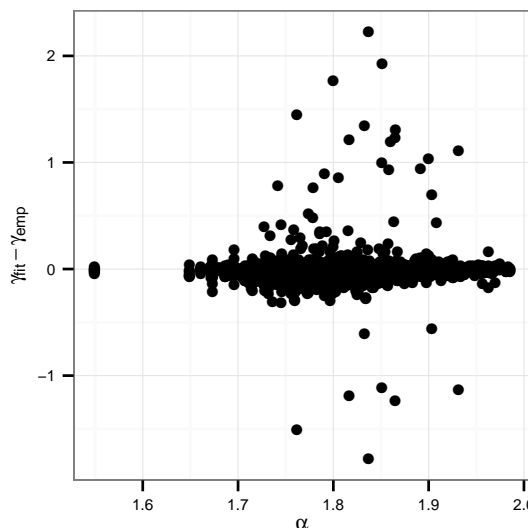


Figure 7. The difference between the parameter γ as fitted, and as predicted with correlations versus the tail parameter α . The few very large values are due to the largest scales, but generally the approximations appears to behave without bias or response to deviations from normal when $\alpha \ll 2$.

such as the normal distribution and the Cauchy distribution. The α -stable distribution does not have a closed form, but rather expresses its density as an integral of the characteristic (moment-generating) function over all moments ranging from $-\infty$ to $+\infty$. This would result in an indefinite integral that only has a closed form in a few special cases. Hence,

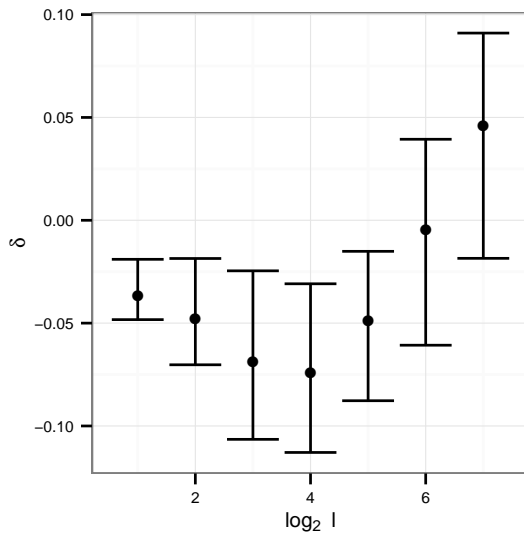


Figure 8. The empirical means of the increments, averaged over all images, and its fit. The error bars denote the 25th and 75th percentiles. There does appear to be some steady behaviour, but it appears highly complex, with relatively small values suggesting that the mean might be sufficient to model its behaviour.

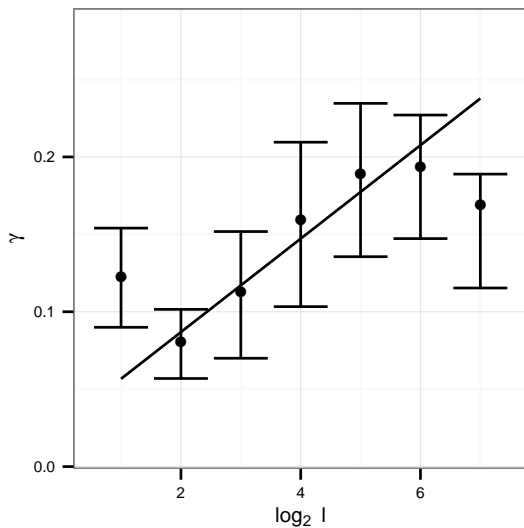


Figure 9. The empirically fitted γ of the increments, averaged over all images, and its fit. The error bars denote the 25th and 75th percentiles. The behaviour of the function is linear in the middle of the scaling range, but breaks above $\log_2 l \approx 6$ or 30 km, and for the first scale, which is roughly in agreement with the scaling range found in the power spectrum.

an approximation is required. Although different approximations exist, they are all roughly equivalent and here we used that of Nolan (1997), as implemented in the R package *stabledist* (Wuertz et al., 2012).

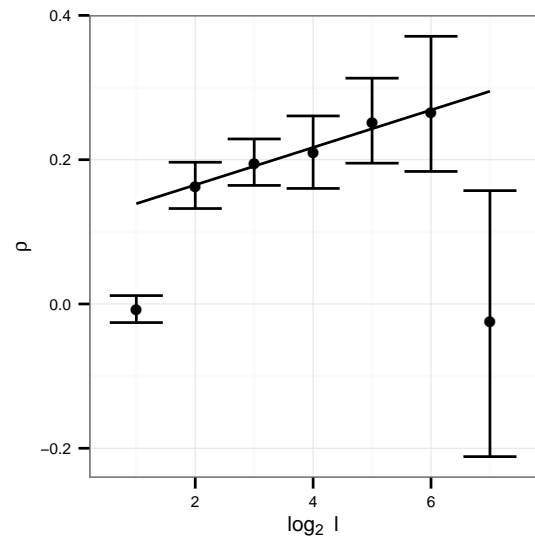


Figure 10. The empirical ρ of the increments, averaged over all images, and its fit. The error bars denote the 25th and 75th percentiles. Almost the exact same pattern is observed as for γ .

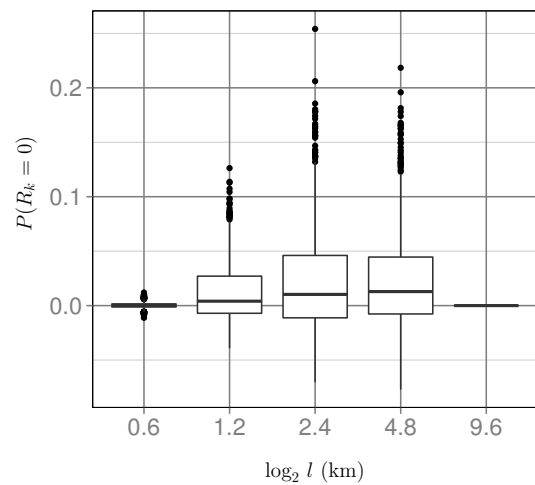


Figure 11. The difference between the probability of dry pixels as predicted, and as observed.

There are a number of different parametrizations available for the α -stable distribution, all suitable for different purposes; we opted for the S1 parametrization of Nolan (2012). In this parametrization, α determines the heaviness of the tail, a parameter β (note that this is not the same β as in Eq. 8) determines the skewness, and two parameters γ and δ determine shape and location. If two distributions X and Y have the same α and both have $\beta = -1$, their sum is also an α -stable distribution, Z , with shape and location parameters

$$\gamma_Z^\alpha = \gamma_X^\alpha + \gamma_Y^\alpha, \tag{10}$$

$$\delta_Z = \delta_X + \delta_Y. \tag{11}$$

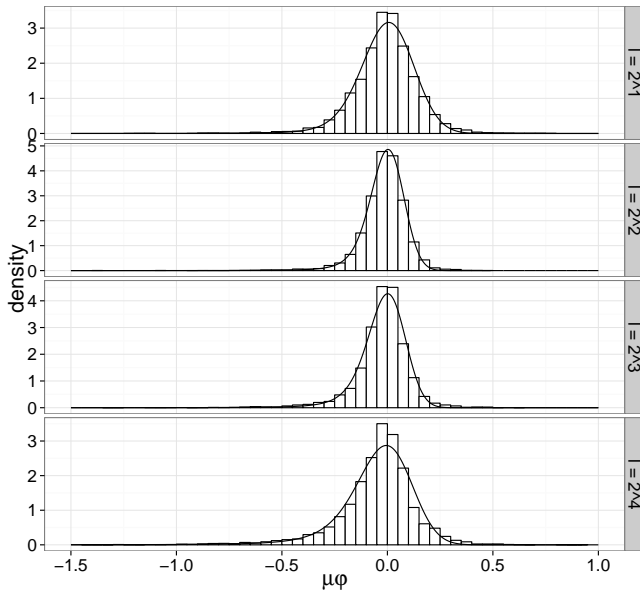


Figure 12. The weighted histograms of rainfall field increments at a range of scales. The lines are the fitted distributions without any preset parameters.

When $\alpha = 2$, the α -stable distribution becomes the normal distribution. As a result, the effect of the parameter β diminishes as $\alpha \rightarrow 2$ and has no effect when $\alpha = 2$ as the normal distribution is necessarily symmetric. Additionally, when $\alpha = 2$, the shift parameter δ is equal to the mean, and the shape parameter relates to the variance as $\sigma^2 = 2\gamma^2$. Additionally, the distribution only has moments that are smaller than α , hence, if $\alpha \geq 1$ the shift parameter is equal to the mean. Moreover, when $\beta = -1$, the distribution is entirely skewed to the left, meaning that it only has a “fat” tail towards the left (negative numbers). Consequentially, the positive moments converge, whereas the negative moments do not. Fortunately, as we generally only deal with positive moments in rainfall analysis, this property allows for an easy analysis.

The α -stable distribution can be fitted in a variety of ways, including the well-known maximum likelihood method. Nevertheless, fitting α -stable distributions is still a difficult exercise, partly due to the lack of a closed form. Despite these difficulties, numerous different approaches are available and a summary of these approaches can be found in Nolan (2001). For this study, the method of McCulloch (1986) is used together with general maximum likelihood fitting. Although faster methods do exist (e.g. Koutrouvelis, 1980), maximum likelihood fitting affords more flexibility such as taking into account the truncation in the rainfall fields. The method of McCulloch (1986) was used to generate an initial starting point for a boxed Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm used in the optimization. Although an in depth explanation is not within the scope

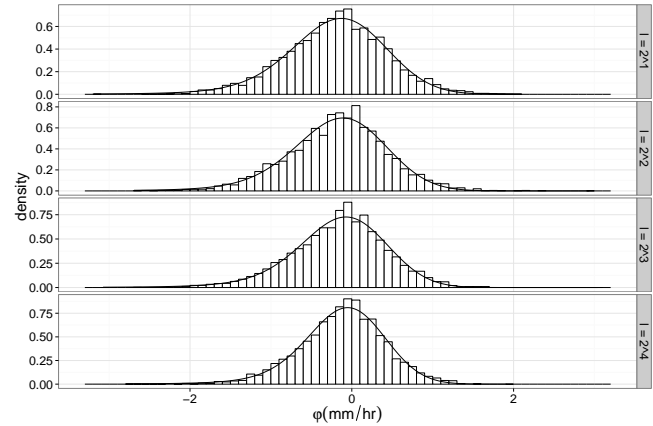


Figure 13. The weighted histograms of a rainfall field at a range of scales. The lines are the fitted distributions without any preset parameters. Note that the observed tail is a result of the local normalization, not a natural feature of the rainfall field.

of this paper, the method of McCulloch (1986) relies on a lookup table of quantiles and associated parameter values, which are interpolated to obtain a crude first-guess estimate of the parameters for the maximum likelihood fitting. For more details on the maximum Likelihood fitting of the α -stable distribution, please refer to Nolan (2001).

5 Methodology

The starting point for any analysis is the rainfall intensity field φ_{obs} . Prior to scaling the field, any local trends were first removed by normalizing the field over disjoint boxes

$$\varphi_{\text{obs}}^{\text{norm}}(B_{100}) = \frac{\varphi_{\text{obs}}(B_{100})}{\langle \varphi_{\text{obs}}(B_{100}) \rangle}, \quad (12)$$

where B_{100} is used to denote the disjoint box of 100 by 100 pixels, or 60 by 60 km. This approach to scaling is similar to that used in detrended fluctuation analysis (Kantelhardt et al., 2002). The size of the boxes was chosen such that the distributions scaled properly without bimodality, but that the size of the boxes was as large as possible to avoid effects on the scaling behaviour.

Subsequently, the data were coarse-grained for analysis. This scaling was done using a moving average (low-pass) filter using a box with sides of length l rather than the disjoint boxes common to multifractal analysis to allow for more points at higher scales and avoid spurious correlations due to a lack of points. This scaling is performed as

$$\varphi_l(i, j) = \frac{1}{\#\text{Rain}(i, j)} \sum_{x=i-l/2}^{i+l/2} \sum_{y=j-l/2}^{j+l/2} \varphi_{\text{obs}}(x, y), \quad (13)$$

where $\#\text{Rain}$ denotes the number of active pixels in φ_{obs} within the region over which the sum is performed. Each

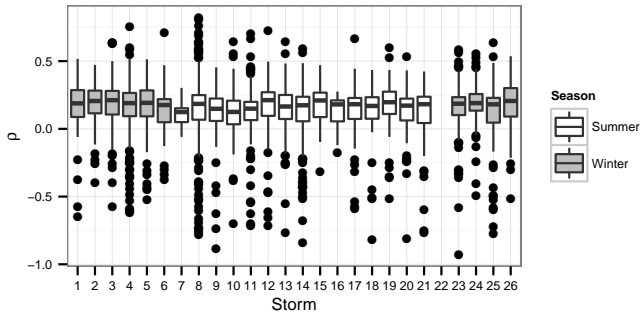


Figure 14. The correlations of all scales for each of the storms. Note that almost all rainfall fields exhibit correlations, and that almost all of them are positive.

point has a weight associated with it, equal to the fraction of rainy pixels within the area of averaging, i.e.

$$\omega_l(i, j) = \frac{l^2}{\#\text{Rain}(i, j)}. \quad (14)$$

To determine the distributions and correlation, only points with more than 90 % rainy pixels were selected for the analysis to further reduce the effects of non-rainy areas on the analysis.

The resulting set of rainfall images φ_l with $l \in 2^{(0..K)}$ with increasing (coarsening) scale were then used to extract the increments. The log-increments $\mu\varphi$ are extracted as

$$\log(\mu\varphi_l) = \log(\varphi_{l+1}) - \log(\varphi_l), \quad (15)$$

where $l+1$ is used to indicate the coarser scale. The resulting cascades can then be analysed by fitting an α -stable distribution to each of the fields, $\log(\mu\varphi_l)$ and $\log(\varphi_l)$, for each scale using the fitting method described in Sect. 4.

Moreover, as mentioned earlier, the parameter α should be the same for all these distributions. Therefore, the fit is done in two steps, first a preliminary step where all distributions are fitted separately resulting in a set $\alpha_{l=2^{0..K}}$, which contains all α values for both the increments and the fields. Then, the distributions are fitted a second time, forcing $\alpha = \langle \alpha_{l=2^{0..K}} \rangle$. Although no formal relationship exists between distributions with different α , it was found that the mean of a set was in good agreement with optimized values of α . Hence, this analysis results in a set of parameters $(\alpha_{\mu\varphi_l}, -1, \gamma_{\mu\varphi_l}, \delta_{\mu\varphi_l})$ for each scale level l , where it should be noted that $\delta_{\mu\varphi_l}$ is forced to be equal to $\langle \log(\mu\varphi_l) \rangle$. A similar set is found for each rainfall field $\log(\varphi_l)$, denoted by subscript φ_l .

Besides the basic parameters of the distribution, we are also interested in establishing whether or not the fields and their increments are actually i.i.d.. A simple test would be to use the correlation to assess whether or not these distributions are uncorrelated. However, the α -stable distribution with stability parameter α does not admit moments $q > \alpha$; hence, if $\alpha < 2$ the (Pearson) correlation does not exist. As

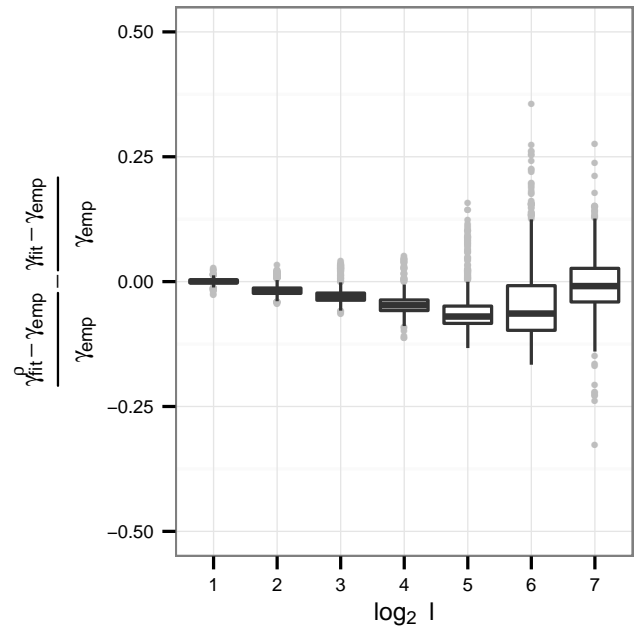


Figure 15. The difference between the relative error of the distribution without correlation and that with correlation propagated from the largest scale. The inclusion of correlations leads to small, but consistent, improvements (i.e. negative values). It is immediately clear that the coarsest scale is not well captured by the functions, evidenced by the large relative error.

a result, using raw correlations is not feasible, and a difficult problem in α -stable analysis arises. Many different measures have been suggested, but to the authors' knowledge all of these pertain to symmetric distributions, i.e. those with $\beta = 0$. Nonetheless, we adopt the correlation value of Garel and Kodia (2009) as it offers important benefits and presents a conceptually simple framework.

The basis of the correlation value of Garel and Kodia (2009) relies on the notion that, for properly scaled variables with finite second-order moments, the slope of the regression $E[\varphi|\mu\varphi] = E[\mu\varphi|\varphi]$ (note that the logarithm and the scale indicators have been dropped for notational convenience) is equal to the Pearson correlation ρ . However, the regression line and its slope always exist, in contrast to the Pearson correlation coefficient, even though we cannot generally say that it is finite or exchangeable (i.e. it could be that $E[\varphi|\mu\varphi] \neq E[\mu\varphi|\varphi]$). Hence, an appealing correlation measure is

$$\rho(\varphi, \mu\varphi) = \text{sign}(\theta_{\varphi|\mu\varphi}) \sqrt{\theta_{\varphi|\mu\varphi} \theta_{\mu\varphi|\varphi}}, \quad (16)$$

where $\theta_{\varphi|\mu\varphi}$ is the slope of the regression line $E[\varphi|\mu\varphi]$, and similarly for $\theta_{\mu\varphi|\varphi}$. Use of the square root is to ensure that if the second-order moment exists, the metric coincides with the Pearson correlation. Finally, the sign function is used to ensure that negative and positive correlations are differentiated. A proof for this metric is beyond scope of the paper, rather, we will investigate its practical skill. Furthermore,

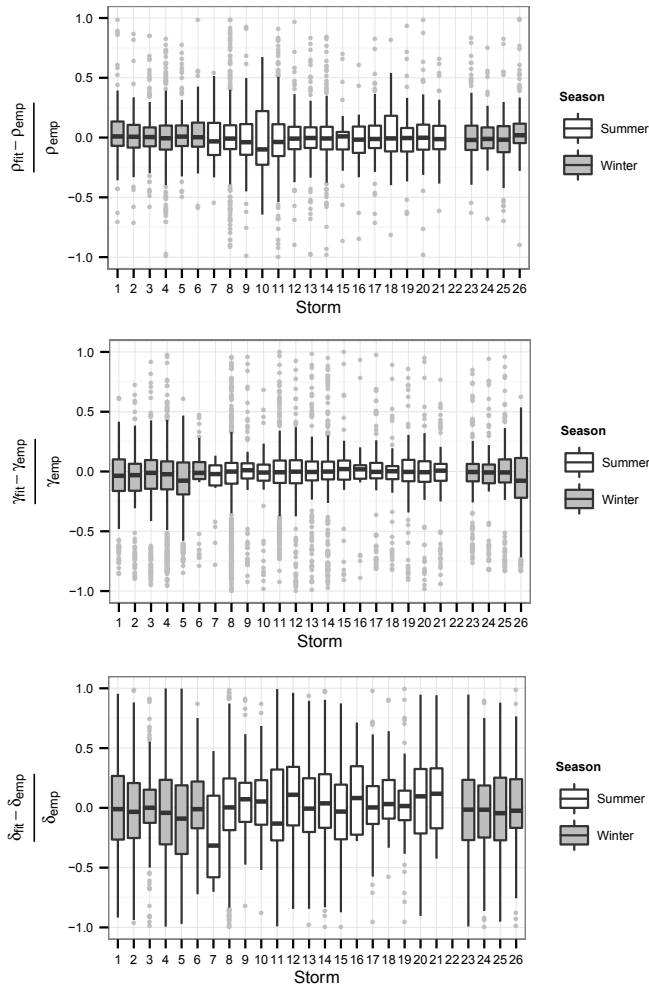


Figure 16. The relative errors of the mean, shape and correlation functions. All functions appear to behave relatively stable throughout winter and summer, with the exception of γ where the summer storms have a smaller error, but more outliers.

weighted and partial correlations are easily implemented by using either weighted regression, or by determining the correlation on residuals.

The relationship between the shape parameters of the rainfall field and its increments, $\gamma_{\mu\varphi}$ and γ_{φ} , with $\rho(\varphi, \mu\varphi) \neq 0$ is dependent on the entire bivariate distribution (Nolan, 2012). However, modelling such a distribution is highly cumbersome and not at all evident as multivariate stable distributions are an area of ongoing research; therefore, a simplification is needed. We observe that if $\alpha = 2$ the relationship between $\gamma_{\mu\varphi}$ and γ_{φ} is

$$\gamma_{\varphi+\mu\varphi}^{\alpha} = \gamma_{\varphi}^{\alpha} + \gamma_{\mu\varphi}^{\alpha} + \rho(2\sigma_{\varphi}^{\alpha})^{(1/\alpha)}(2\sigma_{\mu\varphi}^{\alpha})^{(1/\alpha)}, \quad (17)$$

where ρ denotes the Pearson correlation coefficient. The above is dependent on the notion that if $\alpha = 2$, the α -stable distribution becomes a normal distribution, with variance $2\gamma^{\alpha}$. Therefore, to simulate the effects of the summation of

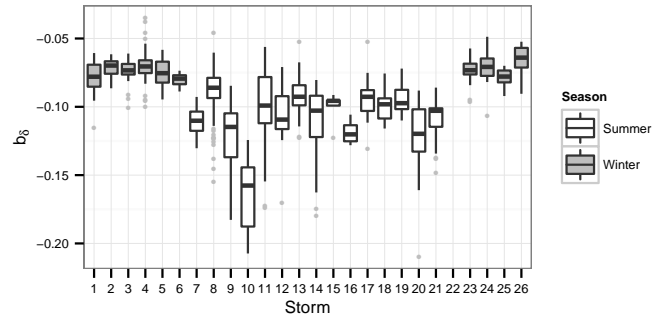


Figure 17. The parameter for the mean of the increment, shown as boxplots for each storm. Summer storms clearly have lower increments.

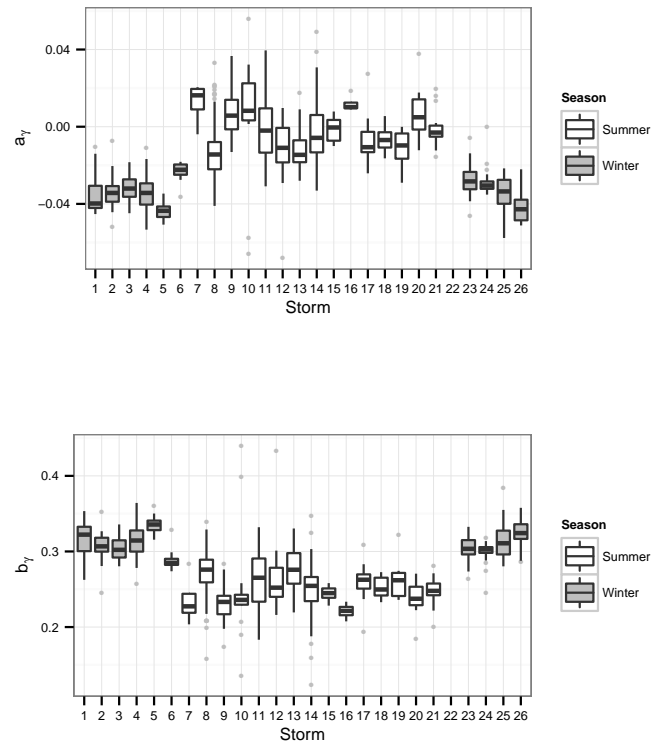


Figure 18. The parameters for the γ of the increment, shown as boxplots for each storm. The intercepts are higher for summer storms, suggesting more energy in the flux, but with smaller slopes; i.e. the scales are more similar. Winter storms show the exact opposite behaviour.

a correlated distribution, we use Eq. (17) where we substitute the Pearson correlation coefficients with the measure in Eq. (16). The effects of using this equation are investigated in Fig. 7 by comparing shape parameters fitted to the empirical distribution with shape parameters computed according to Eq. (17) over a single scale. Note that, in general, the errors appear to be mild, however, at lower values of α , several large errors can be observed. Fortunately, few rainfall images

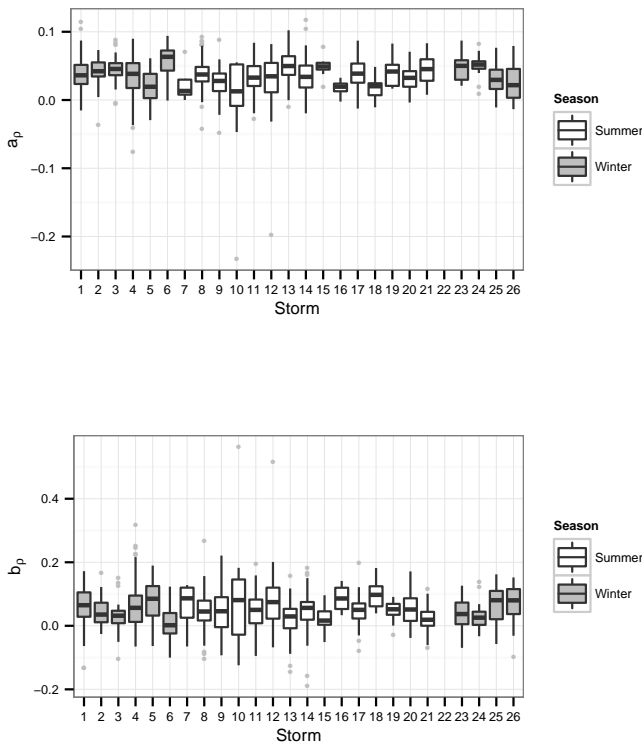


Figure 19. The parameters for the correlation of the increment, shown as boxplots for each storm. The functions for the correlation appear stable throughout winter and summer and have a high intercept, but a low slope suggesting that the correlation is somewhat the same for all scales.

have distributions with low α , making this a tenable approximation.

To investigate the behaviour of the scaling of the α -stable parameters through time, we first need to characterize this behaviour for each of the images. This is done by fitting a set of scale-dependent functions to the α -stable parameters for each image and its increments. The mean behaviour of the α -stable parameters for all images was used as a guideline for the function forms, shown in Figs. 8–10. These empirical functions all admit linear or stable behaviour, and thus we fitted

$$\delta_k = a_\delta + b_\delta \cdot \ln(\lambda), \tag{18}$$

$$\gamma_k = e^{a_\gamma + b_\gamma \cdot \ln(\lambda)}, \tag{19}$$

$$Q_k = a_Q + b_Q \cdot \frac{1}{\lambda}. \tag{20}$$

Note that the subscripts identifying that these parameters apply to $\log(\mu\varphi_l)$ have been dropped for notational convenience. The fit of the above functions is examined in Figs. 8–10, for each of the parameters respectively. First, the mean of the increments shows relatively erratic behaviour close to zero and was thus modelled as the mean over each of the scales. Secondly, γ showed a split behaviour where at smaller scales linear behaviour is observed, which flattens

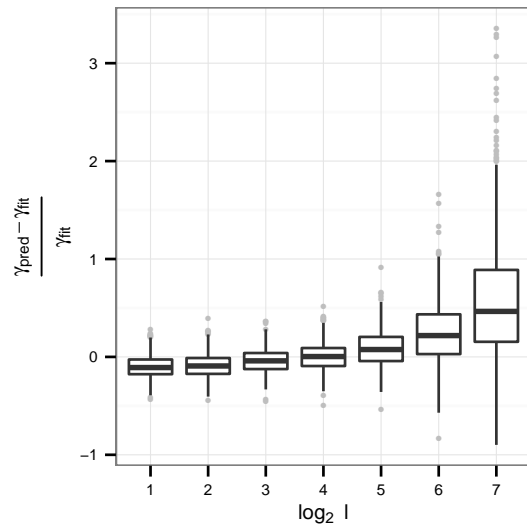


Figure 20. The difference between the relative error of the distribution without correlation, and that with correlation propagated from the largest scale. The inclusion of correlations leads to small, but consistent, improvements (i.e. negative values). It is immediately clear that the coarsest scale is not well captured by the functions, evidenced by the large relative error.

out at the larger scales and starts to behave somewhat erratically. Similar behaviour is observed in the correlations where the extreme scales are different in nature to the intermediate scales. It is more than likely that this is related to the scaling breaks observed in Sect. 3, nonetheless, we fit a linear function to both for the middle of their range.

Finally, the number of dry pixels are modelled based on the fractal box-counting dimension (Rupp et al., 2009). As the box-counting dimension is directly based on the number of dry pixels at each scale, it suffices to invert this relationship yielding

$$P(Y > 0)_l = \left(\frac{1}{l_k}\right)^{D_f} \cdot P(Y > 0)_{l_k=k_{\max}}, \tag{21}$$

where D_f is the fractal dimension and l_k is the side length of the pixel at scale k expressed in elementary pixels. This relationship functions nearly perfectly (Fig. 11). Note that this equation does not function on a pixel-by-pixel basis, but rather attempts the total fraction of zeros in the field.

6 Results

The assumption that the both the distribution of ε and $\mu\varepsilon$ (i.e. the conservative cascade) are of a log-stable family is common, as it is vital to the universal behaviour. However, empirical investigation of these fields is difficult, as removal of the effects of H can result in spurious scaling (Veneziano and Furcolo, 1999) and hence we operate on the non-conservative fields requiring some investigation as to whether this assumption is tenable.

After appropriate normalization, the distribution of the increments $\mu\varphi$ is well approximated by an α -stable distribution as shown in Fig. 12. The distribution of φ itself is somewhat more difficult to approximate due to the truncation of the lower tail at 0.1 mm h^{-1} . Despite this, a truncated distribution shows an excellent fit in Fig. 13, over all scales. The parameters found after free-fitting (i.e. without any preset parameters) shows that the α changes somewhat throughout the scaling range and between $\mu\varphi$ and φ , evidently, the α and β parameters fitted for the truncated distribution φ are unreliable, as both are strongly determined by the missing tail. It is likely that the changes in α are due to noise, and in part due to the truncation of the lower tail, causing both β and α to rise (i.e. the distribution to become more symmetric and normal). Moreover, the mean of the field is reasonably stable around zero, as would be expected of a normalized field.

In Fig. 14 all the correlations for each of the scales are shown, summarized as a boxplot. From this plot, it is evident that almost all storms exhibit a positive correlation between the increments and the rainfall field. This pattern is also seen in rank correlation measures (not shown), further corroborating that there is indeed a correlation. Taking this correlation into account according to Eq. (17) indeed results in a decrease in error, as is evidenced by the lower relative difference for the correlated than for the uncorrelated error (Fig. 15). The effects of this correction are evidently less at the higher scales, possibly due to the more erratic behaviour. Moreover, the significant changes in the shape parameter, γ further suggest that the i.i.d. assumption is, for these storms, incorrect.

The functions 18–20 are used to characterize the scaling behaviour. These functions exhibit a reasonably good fit for all storms, as determined through the relative error (Fig. 16). The resulting parameters are shown for each storm in Figs. 17–19. The mean of the increments is evidently below zero for the summer storm whereas it is generally positive for winter storms. The parameters for γ show a higher intercept for summer storms, but a lower slope. Moreover, some of these slopes are negative for the summer storms and their spread is higher. This behaviour is reversed for ρ , with lower intercepts for summer storms and higher slopes.

The analyses confirm the common finding that summer storms tend to be more energetic with higher variances and higher mean rainfall. Moreover, summer storms appear to exhibit a smaller decrease in correlations, resulting in a stronger correlation at the lower-scale levels.

Figure 20 shows the relative difference between the γ of the modelled distributions and the direct fitted distributions, propagated over the four scale levels. It can be seen that the model error increases as the number of scale levels increases over which we propagate the scaling. Nonetheless, the error remains relatively low, showing that the model captures the scaling behaviour quite well. The fractal model for the dry pixels works very well, as should be expected due to the direct relation with the actual number of dry pixels.

7 Conclusions

In this paper, we investigated the scaling behaviour of the distributions of rainfall. To this end, a novel scaling model was introduced that only relies on the basic assumptions regarding the cascade structure responsible for the fractal nature of rainfall. Furthermore, this framework is based on direct empirical comparison with the observed distributions. In contrast, most previous work relied on theoretical considerations and indirect use of the scaling distributions. Therefore, this framework allows for a more direct and empirical investigation into the scaling behaviour of rainfall, and provides a more adaptable framework to be used for practical purposes.

Rainfall was found not to be the result of an i.i.d. cascade, but rather of a cascade where the distribution changed and the increments are dependent on their coarse-scale parents. The changes in distribution, as described by the shape parameter of the α -stable distribution, were observed to change up to a scale of about 32 km. After this, the behaviour became erratic, possibly due to the large scale relative to the size of the images. Nonetheless, this lack of scaling at large scales has also been observed in similar studies on rainfall time series (Olsson, 1995). Furthermore, as the scales grow larger, the inclusion of non-rainy areas becomes unavoidable and more than likely affect the scaling behaviour (Olsson, 1998). Furthermore, the summer storms were observed to have increments with a higher variance, suggesting they are more energetic. This is in line with expectations, as well as the findings of other authors, e.g. (Venugopal et al., 2006) who suggested that the scaling behaviour is different dependent on the intensity of the storm.

The correlations found in the cascade were positive for almost all storms, and were shown to depend only on the large-scale values and not on the season. However, these correlations were clearly dependent on the scale of averaging, where larger scales resulted in larger correlations, up to the point where scaling became erratic. These dependencies have also been observed by other authors in time series (Rupp et al., 2009).

The inclusion of correlations into the distributional model showed only moderate improvements, in part due to the small magnitude of the scale parameters where the correlations were found. Nonetheless, the deviation from identical distributions, as evidenced through the change in γ , should be incorporated and give strong improvements.

In future research, the full dependence structure will need to be evaluated to allow for a more accurate representation of the dependence between scale levels and their increments. This will allow for a deeper investigation into this aspect of imperfect scaling and possibly a better way of representing the scaling behaviour. Moreover, it was observed that local trends were present in all rainfall images, not only in the mean of the field, but also in the correlations, this will need to be investigated further. Finally, the difference with respect to the scaling behaviour, between convective and stratiform

storms, will need further investigation, using a classification algorithm such as the Steiner algorithm (Steiner et al., 1995). A careful analysis of the behaviour of such algorithms will be required before using them to investigate the difference in scaling behaviour between stratiform and convective precipitation.

Acknowledgements. We wish to thank the Special Research Fund (B.O.F.) of Ghent University and the Flanders Research Fund (FWO, grant number: G.0837.10) for funding this research. We would like to thank A. Langousis and A. Seed for a fruitful discussion and their insightful comments.

Edited by: A. Bárdossy

References

- Davis, A., Marshak, A., Wiscombe, W., and Cahalan, R.: Multifractal characterizations of nonstationarity and intermittency in geophysical fields: Observed, retrieved, or simulated, *J. Geophys. Res.*, 99, 8055–8072, 1994.
- De Jongh, I. L. M., Verhoest, N. E. C., and De Troch, F. P.: Analysis of a 105-year time series of precipitation observed at Uccle, Belgium, *Int. J. Climatol.*, 26, 2023–2039, 2006.
- Deidda, R.: Rainfall downscaling in a space-time multifractal framework, *Water Resour. Res.*, 36, 1779–1794, 2000.
- Delobbe, L., Dehem, D., Dierickx, P., Roulin, E., Thunus, M., and Tricot, C.: Combined use of radar and gauge observations for hydrological applications in the Walloon region of Belgium, *Proceedings of ERAD*, 418–421, 2006.
- Ferraris, L., Rudari, R., and Siccardi, F.: The uncertainty in the prediction of flash floods in the northern Mediterranean environment, *J. Hydrometeorol.*, 3, 714–727, 2002.
- Ferraris, L., Gabellani, S., Parodi, U., Rebora, N., Von Hardenberg, J., and Provenzale, A.: Revisiting multifractality in rainfall fields, *J. Hydrometeorol.*, 4, 544–551, 2003.
- Garel, B. and Kodja, B.: Signed symmetric covariation coefficient for alpha-stable dependence modeling, *CR Math.*, 347, 315–320, 2009.
- Gires, A., Onof, C., Maksimovic, C., Schertzer, D., Tchiguirinskaia, I., and Simoes, N.: Quantifying the impact of small scale unmeasured rainfall variability on urban runoff through multifractal downscaling: A case study, *J. Hydrol.*, 442, 117–128, 2012a.
- Gires, A., Tchiguirinskaia, I., Schertzer, D., and Lovejoy, S.: Influence of the zero-rainfall on the assessment of the multifractal parameters, *Adv. Water Res.*, 45, 13–25, 2012b.
- Goudenhoofd, E. and Delobbe, L.: Evaluation of radar-gauge merging methods for quantitative precipitation estimates, *Hydrol. Earth Syst. Sci.*, 13, 195–203, doi:10.5194/hess-13-195-2009, 2009.
- Gupta, V. K. and Waymire, E. C.: A statistical analysis of mesoscale rainfall as a random cascade, *J. Appl. Meteorol.*, 32, 251–251, 1993.
- Harris, D. and Foufoula-Georgiou, E.: Subgrid variability and stochastic downscaling of modeled clouds: Effects on radiative transfer computations for rainfall retrieval (1984–2012), *J. Geophys. Res.-Atmos.*, 106, 10349–10362, 2001.
- Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., and Stanley, H. E.: Multifractal detrended fluctuation analysis of nonstationary time series, *Physica A*, 316, 87–114, 2002.
- Kolmogorov, A. N.: The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers, in: *Dokl. Akad. Nauk SSSR*, Vol. 30, 299–303, 1941.
- Koutrouvelis, I.: Regression-type estimation of the parameters of stable laws, *J. Am. Stat. Assoc.*, 75, 918–928, 1980.
- Koutsoyiannis, D., Paschalis, A., and Theodoratos, N.: Two-dimensional Hurst–Kolmogorov process and its application to rainfall fields, *J. Hydrol.*, 398, 91–100, 2010.
- Lovejoy, S. and Schertzer, D.: *The weather and climate: emergent laws and multifractal cascades*, Cambridge University Press, 2013.
- McCulloch, J.: Simple consistent estimators of stable distribution parameters, *Communications in Statistics-Simulation and Computation*, 15, 1109–1136, 1986.
- Menabde, M. and Sivapalan, M.: Modeling of rainfall time series and extremes using bounded random cascades and Levy-stable distributions, *Water Resour. Res.*, 36, 3293–3300, 2000.
- Menabde, M., Harris, D., Seed, A., Austin, G., and Stow, D.: Multiscale properties of rainfall and bounded random cascades, *Water Resour. Res.*, 33, 2823–2830, 1997.
- Nolan, J. P.: Parameter estimation and data analysis for stable distributions, in: *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems & Computers*, 1997, vol. 1, IEEE, Pacific Grove(CA), USA, 443–447, 1997.
- Nolan, J. P.: Maximum likelihood estimation and diagnostics for stable distributions, in: *Lévy processes*, 379–400, Springer, 2001.
- Nolan, J. P.: *Stable Distributions – Models for Heavy Tailed Data*, Birkhauser, Boston, in progress, Chapter 1 available at: <http://academic2.american.edu/~jpnolan> (last access: 20 August 2013), 2012.
- Olsson, J.: Limits and characteristics of the multifractal behaviour of a high-resolution rainfall time series, *Nonl. Proc. Geophys.*, 2, 23–29, 1995.
- Olsson, J.: Evaluation of a scaling cascade model for temporal rainfall disaggregation, *Hydrol. Earth Syst. Sci.*, 2, 19–30, doi:10.5194/hess-2-19-1998, 1998.
- Over, T. and Gupta, V.: Statistical analysis of mesoscale rainfall: dependence of a random cascade generator on large-scale forcing, *J. Appl. Meteorol.*, 33, 1526–1542, 1994.
- Parisi, G. and Frisch, U.: A multifractal model of intermittency, *Turbulence and Predictability*, in: *Geophysical Fluid Dynamics*, edited by: Ghil, M., Benzi, R., and Parisi, G., 84–88, 1985.
- Paulson, K. S. and Baxter, P. D.: Downscaling of rain gauge time series by multiplicative beta cascade, *J. Geophys. Res.*, 112, D09105, doi:10.1029/2006JD007333, 2007.
- Rebora, N., Ferraris, L., von Hardenberg, J., and Provenzale, A.: Rainfall downscaling and flood forecasting: a case study in the Mediterranean area, *Nat. Hazards Earth Syst. Sci.*, 6, 611–619, doi:10.5194/nhess-6-611-2006, 2006.

- 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.
- Schertzer, D. and Lovejoy, S.: Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes, *J. Geophys. Res.*, 92, 9693–9714, 1987.
- Schertzer, D. and Lovejoy, S.: Universal multifractals do exist!: Comments on – A statistical analysis of mesoscale rainfall as a random cascade, *J. Appl. Meteorol.*, 36, 1296–1303, 1997.
- Schertzer, D., Lovejoy, S., and Hubert, P.: An introduction to stochastic multifractal fields, in: *ISFMA Symposium on Environmental Science and Engineering with related Mathematical Problems*, 106–179, Higher Education Press, Beijing, 2002.
- Serinaldi, F.: Multifractality, imperfect scaling and hydrological properties of rainfall time series simulated by continuous universal multifractal and discrete random cascade models, *Nonl. Proc. Geophys.*, 17, 697–714, 2010.
- Steiner, M., Houze Jr, R. A., and Yuter, S. E.: Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data, *J. Appl. Meteorol.*, 34, 1978–2007, 1995.
- Tchiguirinskaia, I., Lu, S., Molz, F., Williams, T., and Lavallée, D.: Multifractal versus monofractal analysis of wetland topography, *Stochastic Environmental Research and Risk Assessment*, 14, 8–32, 2000.
- Veneziano, D. and Furcolo, P.: A modified double trace moment method of multifractal analysis, *Fractals*, 7, 181–196, 1999.
- Veneziano, D., Furcolo, P., and Iacobellis, V.: Imperfect scaling of time and space-time rainfall, *J. Hydrol.*, 322, 105–119, 2006.
- Venugopal, V., Roux, S. G., Foufoula-Georgiou, E., and Arneodo, A.: Revisiting multifractality of high-resolution temporal rainfall using a wavelet-based formalism, *Water Resour. Res.*, 42, W06D14, doi:10.1029/2005WR004489, 2006.
- Verrier, S., Mallet, C., and Barthès, L.: Multiscaling properties of rain in the time domain, taking into account rain support biases, *J. Geophys. Res.*, 116, D20119, doi:10.1029/2011JD015719, 2011.
- Wuertz, D., Maechler, M., and R metrics core team members: *stabledist: Stable Distribution Functions*, <http://CRAN.R-project.org/package=stabledist>, r package version 0.6-5, 2012.