Manuscript prepared for Hydrol. Earth Syst. Sci. Discuss.
with version 4.1 of the LaTeX class copernicus_discussions.cls.
Date: 25 July 2014

Imperfect scaling in distributions of radar-derived rainfall fields

M.J. van den Berg[1], **L. Delobbe**[2], **and N.E.C. Verhoest**[1]

[1]Laboratory of Hydrology and Water Management, Ghent University, Coupure links 653, 9000 Ghent, Belgium
[2]Royal Meteorological Institute, Avenue Circulaire 3, 1180 Uccle, Belgium.

Correspondence to: M.J. van den Berg
(Martinus.vandenBerg@Ugent.be)

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Abstract

Fine scale rainfall observations for modeling 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-minute 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 behavior in rainfall can be exploited to yield a statistical estimate of the rainfall at a finer scale. At a very basic level, this behavior leads to a cascade of scales

$$\varepsilon_n = \varepsilon_0 \prod_{j=1}^{n} \mu\varepsilon_j \,, \tag{1}$$

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 figure 1. A variety of such cascades have been proposed, starting from Kolmogorov (1941) who described homogenous 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-beta generators (Menabde and Sivapalan, 2000) and a other simulatio methods including wavelets (Venugopal et al., 2006). More of these models can be found in Gupta and Waymire (1993); Menabde et al. (1997); 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 section 2). Generally, these models assume that $\mu\varepsilon$ is identically and independently distributed (*i.i.d.*). That is, the distribution of $\mu\varepsilon$ 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.* $\mu\varepsilon$. Empirical investigation of the scaling behavior 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 behavior 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 behavior was found to differ with the intensity of storms (Venugopal et al., 2006), and with that 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 $\mu\varepsilon \in [0, 1]$ 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 explainig the simulation of rainfall (section 2), followed by a description of the data and some investigation into its basic scaling behavior (section 3. Then, the α-stable distribution is described in some detail in section 4. This is followed by the methodology in section 5 and the result (section 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 $\mu\varepsilon_j$, cfr. equation 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)

$$\langle \varepsilon_\lambda^q \rangle = \lambda^{K(q)}, \tag{2}$$

with $K(q)$ a moment scaling function and $\lambda = L_{eff}/l$ 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 and infinity of parameters to describe the behavior. In practice, these cascades converge to a *universal* multifractal if the increments $\mu\varepsilon$ are from a log-stable distribution (Veneziano and Furcolo, 1999; Schertzer and Lovejoy, 1987), i.e. $\log \mu\varepsilon \sim S_\alpha(\cdot)$ where $S_\alpha(\cdot)$ is the (Levy) α-stable distribution (see section 4). If this is the case, $K(q)$ has the form

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

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

$$\varphi_\lambda = D^{-H}\varepsilon_\lambda\,, \tag{4}$$

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)

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

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

$$\zeta(q) = qH - K(q)\,, \tag{6}$$

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

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

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

$$\beta = 1 + 2H - K(2)\,. \tag{8}$$

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. Rebora 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 analyze 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 figure 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 multi-level scan every five minutes. 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-minute 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-minute 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 behavior 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 minutes. Measurements are collected up to 240 km with a resolution of 250 m in range and 1 degree in azimuth.

A time-domain Doppler filtering is applied for ground clutter removal. An additional treatment, based on a static clutter 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 dB 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 km $\times$ 0.6 km (Goudenhoofdt and Delobbe, 2009). Reflectivity values are then converted into precipitation rates using the Marhsall-Palmer relation

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

where Z_{dB} is the reflectivity in $[\mathrm{dB}]$ 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 ~ 0.1 mm/hr, 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/hr 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 figure 2.

3.1 Power Density Spectrum and Multiaffine analysis

We analyzed the rainfall fields both individually and for each of the storms (by averarging 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 Figure 3, together with a straight line fitted to the linear portion of the power spectrum. In each of these storms, linear behavior 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 lay be spurious.

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 coefficient H. These slopes are shown in figure 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 creedence 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 analyzed 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 analyzed 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.
- Each pixel has a weight associated with it the fraction of rainy pixels within the disjoint boxes at the finest scale level.

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

The results of this analysis are shown in figures 5 for the entire first storm. First, the resulting moments have been taking 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 figure 6. The moment scaling function appears to be well captured by the theoretical form of equation 3, suggesting that the field is indeed multifractal. The parameters α 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 section 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 behavior, including right- and left-skewed behavior, as well as symmetric behavior. 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 behavior, it includes several well-known distributions 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, 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).

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 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

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

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

When $\alpha = 2$, the α-stable distribution becomes the normal distribution. As a result, the effect of the parameter β diminishes as $\alpha \to 2$ 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 left-skewed, 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 of this paper, the method of McCulloch (1986) relies on a look up 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

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

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 behavior.

Subequently, 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

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

where $\#Rain$ denotes the number of active pixels in φ_{obs} within the region over which the sum is performed. Each point has a weight associated with it, equal to the fraction of rainy pixels within the area of averaging, i.e.

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

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

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

where $l+1$ is used to indicate the coarser scale. The resulting cascades can then be analyzed 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 section 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 values α 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 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

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

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, 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, modeling 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

$$\gamma^{\alpha}_{\varphi+\mu\varphi} = \gamma^{\alpha}_{\varphi} + \gamma^{\alpha}_{\mu\varphi} + \rho(2\sigma^{\alpha}_{\varphi})^{(1/\alpha)}(2\sigma^{\alpha}_{\mu\varphi})^{(1/\alpha)}, \tag{17}$$

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 a correlated distribution, we use equation 17 where we substitute the pearson correlation coefficients with the measure in equation 16. The effects of using this equation are investigated in figure 7 by comparing shape parameters fitted to the empirical distribution with shape parameters computed according to equation 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 have distributions with low α, making this a tenable approximation.

To investigate the behavior of the scaling of the α-stable parameters through time, we first need to characterize this behavior 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 behavior of the α-stable parameters for all images was used as a guideline for the function forms, shown in figures 8 to 10. These empirical functions all admit linear or stable behavior, and thus we fitted

$$\delta_\lambda = a_\delta \,, \tag{18}$$

$$\gamma_\lambda = a_\gamma + b_\gamma \cdot \log_2(l) \,, \tag{19}$$

$$\rho_\lambda = a_\rho + b_\rho \cdot \log_2(l) \,. \tag{20}$$

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 figures 8 to 10, for each of the parameters respectively. First, the mean of the increments shows relatively erratic behavior close to zero and was thus modelled as the mean over each of the scales. Secondly, γ showed a split behavior where at smaller scales linear behavior is observed, which flattens out at the larger scales and start to behave somewhat erraticaly. Similar behavior 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 section 3, nonetheless, we fit a linear function to both for the middle of their range.

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

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

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 performs nearly perfect (figure 11). Note that this equation does not perform 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 behavior. However, empirical investigation of these fields is difficult, as removal of the effects of H can result in suprious 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 figure 12. The distribution of φ itself is somewhat more difficult to approximate due to the truncation of the lower tail at 0.1 mm/hr. Despite this, a truncated distribution shows an excellent fit in figure 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's and β's 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 figure 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 equation 17 indeed results in a decrease in error, as is evidenced by the lower relative difference for the correlated than for the uncorrelated error (figure 15). The ef-

fects of this correction are evidently less at the higher scales, possibly due to the more erratic behavior. Moreover, the significant changes in the shape parameter, γ further suggest that the *iid* assumption is, for these storms, incorrect.

The functions 18 to 20 are used to characterize the scaling behavior. These functions exhibit a reasonably good fit for all storms, as determined through the relative error (figure 16). The resulting parameters are shown for each storm in figures 17 to 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 behavior 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 modeled 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 behavior 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 behavior 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 behavior 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 behavior 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 behavior (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 behavior is different dependent on the intensity of the storm.

The correl ations 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 were 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. None the less, the deviation from identical distributions, as evidenced through the change in γ should be incorporated and gives 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 possible a better way of representing the scaling behavior. 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 behavior, 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 behavior 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 Dr. A. Langousis and Dr. A. Seed for a fruitful discussion and their insightful comments.

References

Davis, A., Marshak, A., Wiscombe, W., and Cahalan, R.: Multifractal characterizations of nonstationarity and intermittency in geophysical fields: Observed, retrieved, or simulated, Journal of Geophysical Research, 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, International Journal of Climatology, 26, 2023–2039, 2006.

Deidda, R.: Rainfall downscaling in a space-time multifractal framework, Water Resources Research, 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, in: Proceedings of ERAD, pp. 418–421, 2006.

Ferraris, L., Rudari, R., and Siccardi, F.: The uncertainty in the prediction of flash floods in the northern Mediterranean environment, Journal of Hydrometeorology, 3, 714 – 727, 2002.

Ferraris, L., Gabellani, S., Parodi, U., Rebora, N., Von Hardenberg, J., and Provenzale, A.: Revisiting multifractality in rainfall fields, Journal of Hydrometeorology, 4, 544–551, 2003.

Garel, B. and Kodia, B.: Signed symmetric covariation coefficient for alpha-stable dependence modeling, Comptes Rendus Mathematique, 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, Journal of Hydrology, 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, Advances in Water Resources, 45, 13–25, 2012b.

Goudenhoofdt, E. and Delobbe, L.: Evaluation of radar-gauge merging methods for quantitative precipitation estimates, Hydrology and Earth System Sciences, 13, 195–203, 2009.

Gupta, V. K. and Waymire, E. C.: A statistical analysis of mesoscale rainfall as a random cascade, Journal of Applied Meteorology, 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, Journal of Geophysical Research: Atmospheres (1984–2012), 106, 10 349–10 362, 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: Statistical Mechanics and its Applications, 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, pp. 299–303, 1941.

Koutrouvelis, I.: Regression-type estimation of the parameters of stable laws, Journal of the American Statistical Association, 75, 918–928, 1980.

Koutsoyiannis, D., Paschalis, A., and Theodoratos, N.: Two-dimensional Hurst-Kolmogorov process and its application to rainfall fields, Journal of Hydrology, 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 Resources Research, 36, 3293–3300, 2000.

Menabde, M., Harris, D., Seed, A., Austin, G., and Stow, D.: Multiscaling properties of rainfall and bounded random cascades, Water Resources Research, 33, 2823–2830, 1997.

Nolan, J. P.: Parameter estimation and data analysis for stable distributions, in: Signals, Systems & Computers, 1997. Conference Record of the Thirty-First Asilomar Conference on, vol. 1, pp. 443–447, IEEE, 1997.

Nolan, J. P.: Maximum likelihood estimation and diagnostics for stable distributions, in: Lévy processes, pp. 379–400, Springer, 2001.

Nolan, J. P.: Stable Distributions - Models for Heavy Tailed Data, Birkhauser, Boston, in progress, Chapter 1 online at academic2.american.edu/∼jpnolan, 2012.

Olsson, J.: Limits and characteristics of the multifractal behaviour of a high-resolution rainfall time series, Nonlinear processes in Geophysics, 2, 23–29, 1995.

Olsson, J.: Evaluation of a scaling cascade model for temporal rain-fall disaggregation, Hydrology and Earth System Sciences Discussions, 2, 19–30, 1998.

Over, T. and Gupta, V.: Statistical analysis of mesoscale rainfall: Dependence of a random cascade generator on large-scale forcing, Journal of Applied Meteorology, 33, 1526–1542, 1994.

Parisi, G. and Frisch, U.: A multifractal model of intermittency, Turbulence and Predictability in Geophysical Fluid Dynamics M. Ghil, R. Benzi, G. Parisi, 84–88, 1985.

Paulson, K. S. and Baxter, P. D.: Downscaling of rain gauge time series by multiplicative beta cascade, Journal of Geophysical Research, 112, D09 105, 2007.

Rebora, N., Ferraris, L., Von Hardenberg, J., and Provenzale, A.: Rainfall downscaling and flood forecasting: a case study in the Mediterranean area, Natural Hazards and Earth System Science, 6, 611–619, 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 Resources Research, 45, 14 PP., 2009.

Schertzer, D. and Lovejoy, S.: Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes, Journal of Geophysical Research, 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", Journal of Applied Meteorology, 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, pp. 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, Nonlinear Processes in Geophysics, 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, Journal of Applied Meteorology, 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, Journal of Hydrology, 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 resources research, 42, 2006.

Verrier, S., Mallet, C., and Barthès, L.: Multiscaling properties of rain in the time domain, taking into account rain support biases, Journal of Geophysical Research: Atmospheres (1984–2012), 116, 2011.

Wuertz, D., Maechler, M., and core team members., R.: stabledist: Stable Distribution Functions, http://CRAN.R-project.org/package=stabledist, r package version 0.6-5, 2012.

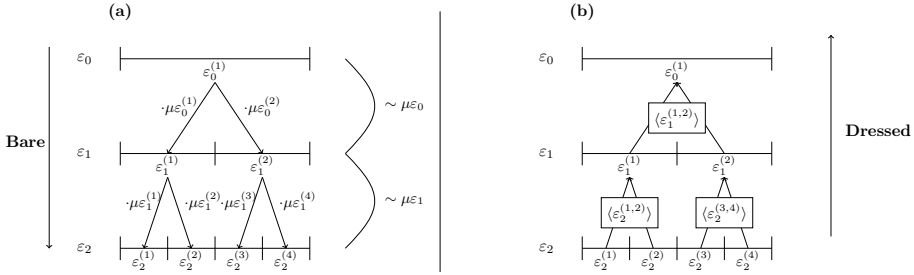

Fig. 1. A basic rainfall model, graphically illustrated. The left hand side of the image is the dressing procedure, whereas the right hand side is the generation.

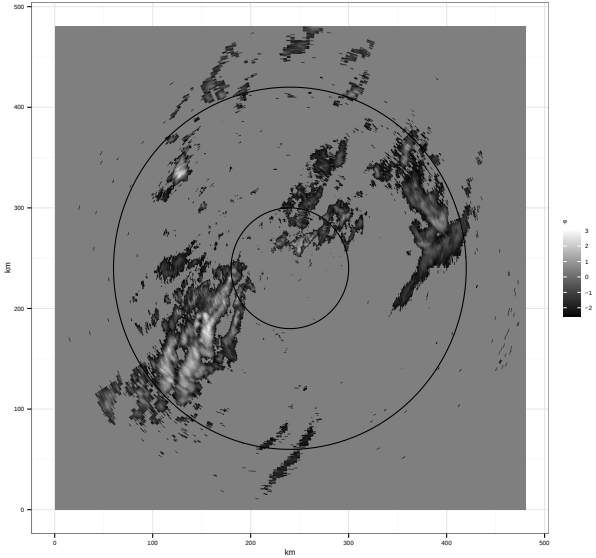

Fig. 2. A log-transformed rainfall field, together with the radius of reliable observations (circles) at 60 km and 180 km.

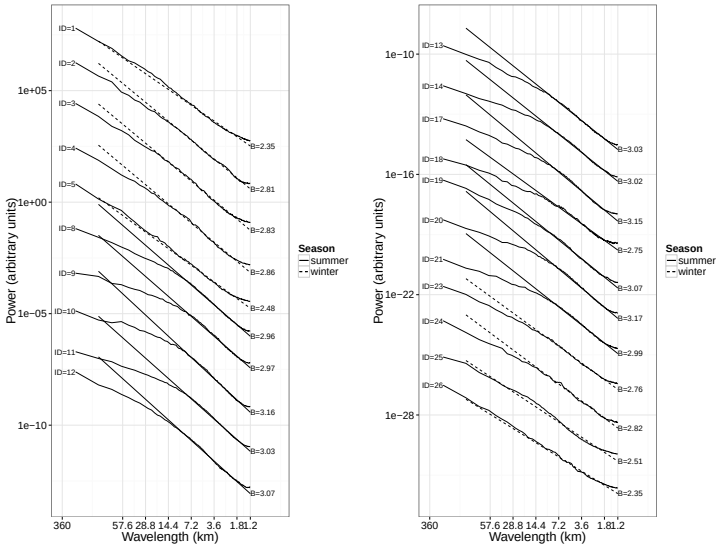

Fig. 3. The power spectra of all rainstorms, up to a range of 180 km from the radar, for images which have more than 10% active pixels and storms with at least 10 valid images. The storm spectra are found by averaging together the spectra of each of the images. The number at the end of each line is the slope β, whereas the number at the beginnging 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.

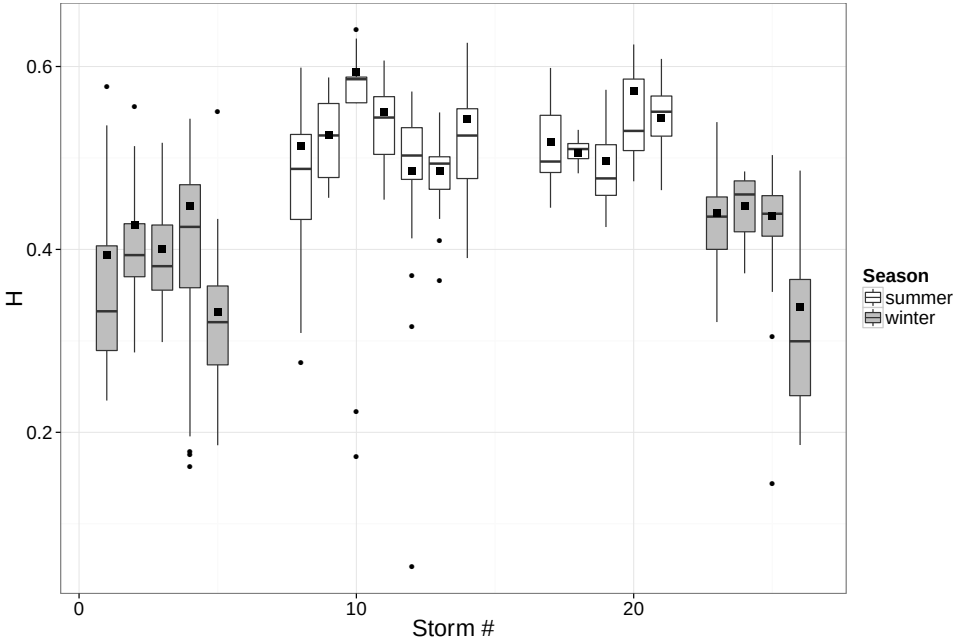

Fig. 4. A boxplot of the non-conservation parameter H for all images. The overall parameter for each storm is marked as a black square. Summer storms show a higher non-conservation parameter than winter storms, suggesting a smoother behavior within the storm (as structure functions are computed only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

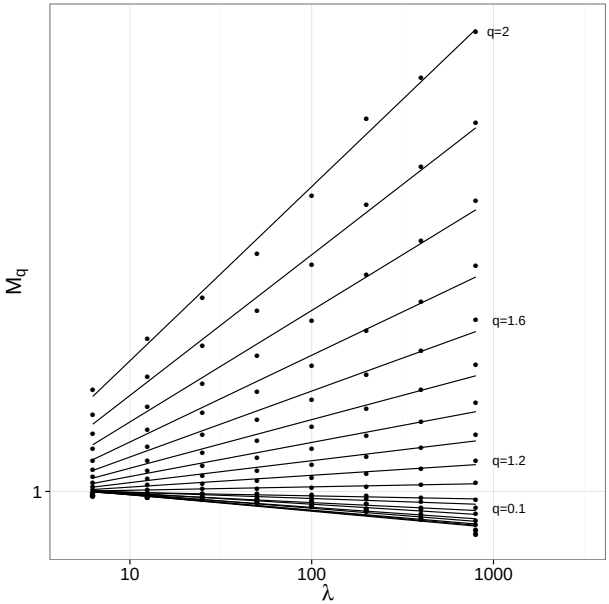

Fig. 5. The empirical moments M_q as a function of scale for storm 1 in its entirity. The moments show good scaling characteristics as evidenced by the good fit of the linear function for each of the moments q.

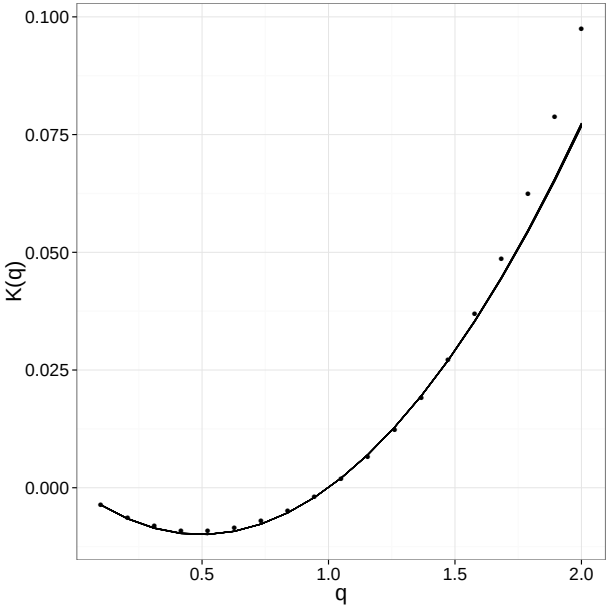

Fig. 6. The empirical function $K(q)$, together with the fitted function using the WMA method. The empirical function crosses zero for $K(0)$ because of the use of the WMA method. Above $q \sim 1.5$ the fit starts to diverge from the empirical function, suggesting that above $q = 1.5$ the scaling breaks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

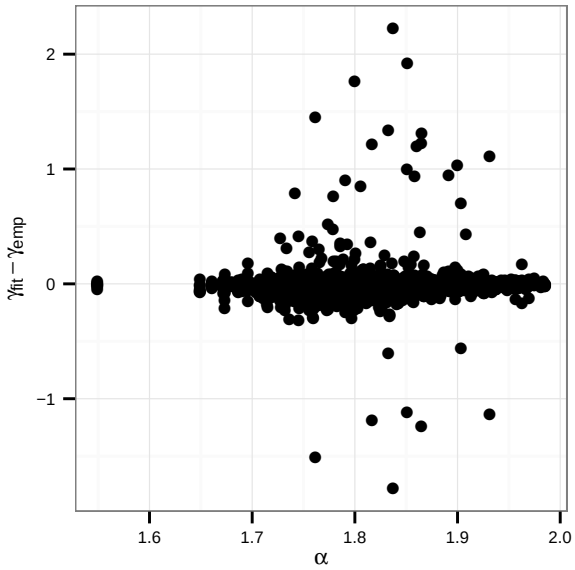

Fig. 7. The difference between the parameter γ as fitted, and as predicted with correlations versus the tail parameter α. The few very large values are due to the largest scales, but generally the approximations appears to behave without bias or response to deviations from normal when $\alpha \ll 2$.

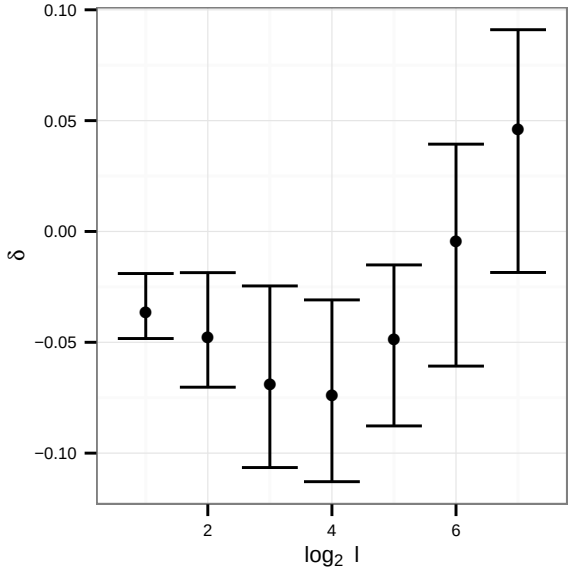

Fig. 8. The empirical means of the increments, averaged over all images, and its fit. The errorbars denote the 25^{th} and 75^{th} percentiles. There does appear to be some steady behavior, but it appears highly complex, but with relatively small values suggesting that the mean might be sufficient to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

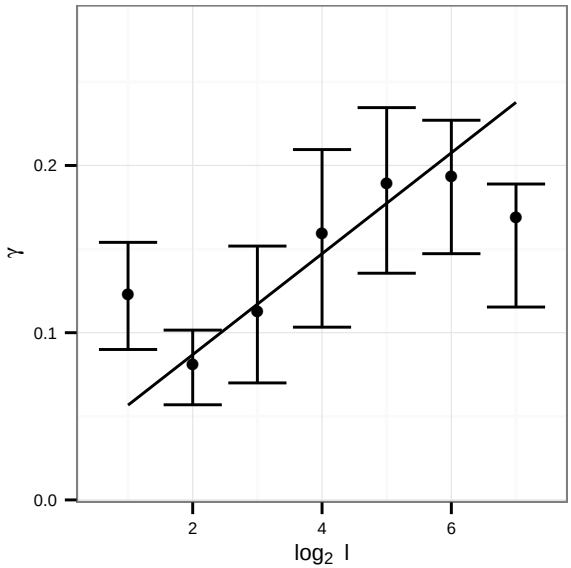

Fig. 9. The empirically fitted γ of the increments, averaged over all images, and its fit. The errorbars denote the 25^{th} and 75^{th} percentiles. The behavior of the function is linear in the middle of the scaling range, but breaks above $\log_2 l \approx 6$ or 30 km, and for the first scale, which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

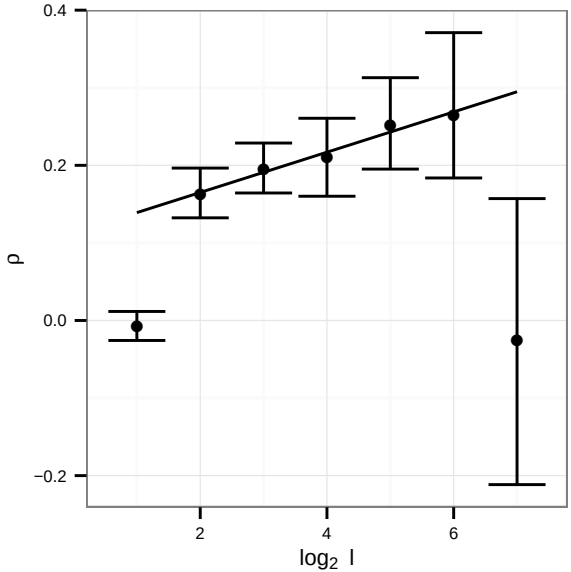

Fig. 10. The empirical ρ of the increments, averaged over all images, and its fit.The errorbars denote the 25^{th} and 75^{th} percentiles. Almost the exact same pattern is observed as for γ.

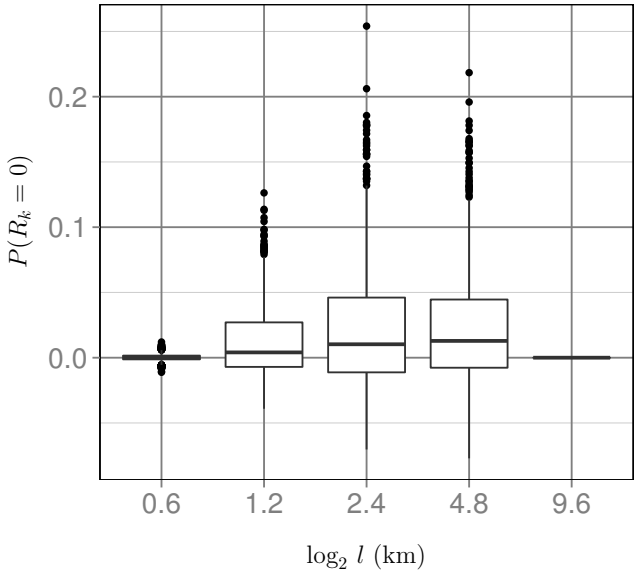

Fig. 11. The difference between the probability of dry pixels as predicted, and as observed.

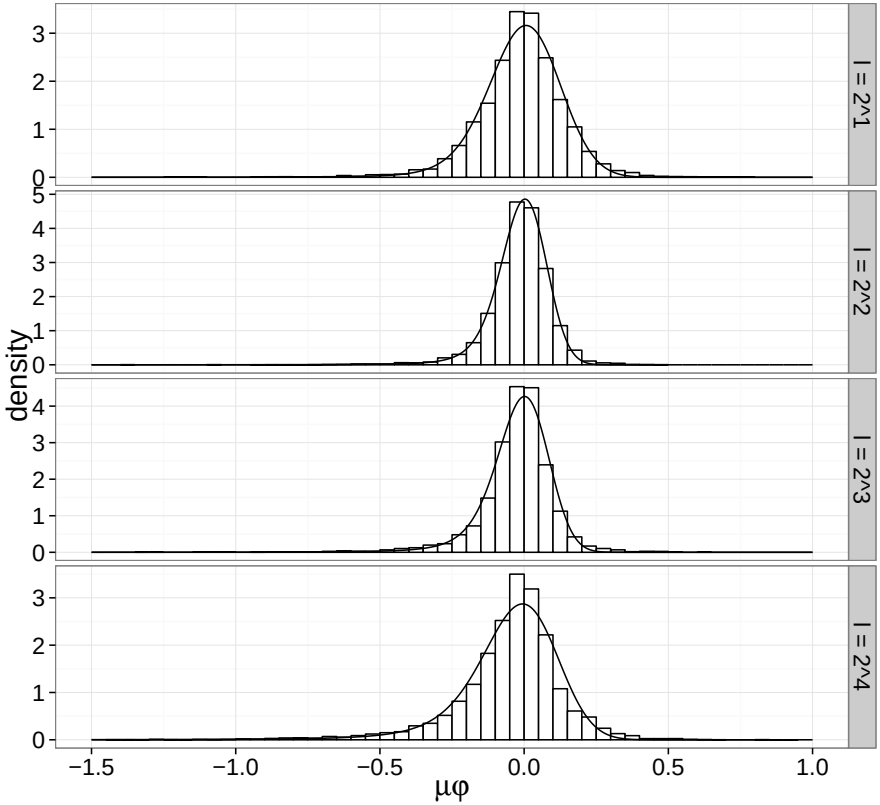

Fig. 12. The weighted histograms of a rainfall field increments at a range of scales. The lines are the fitted distributions without any preset parameters.

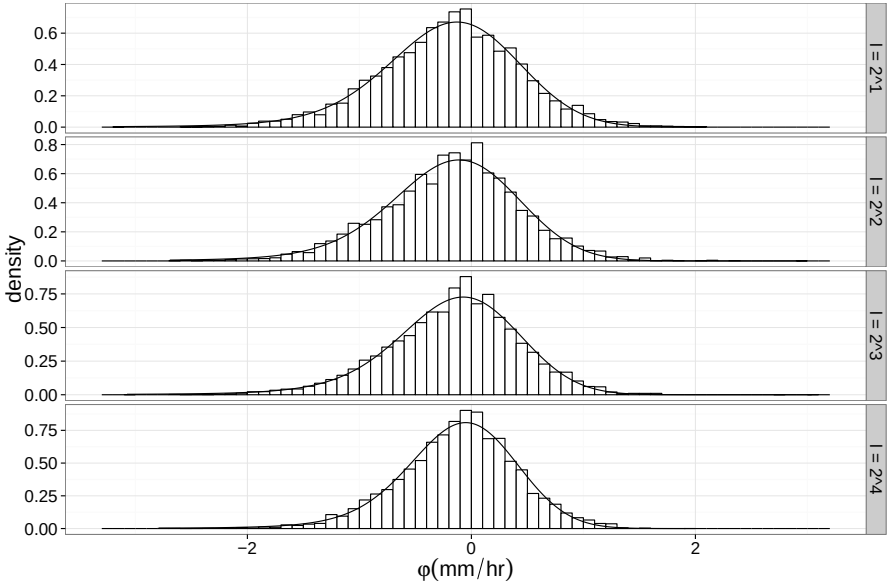

Fig. 13. The weighted histograms of a rainfall field at a range of scales. The lines are the fitted distributions without any preset parameters. Note that the observed tail is a result of the local normalization, not a natural feature of the rainfall field.

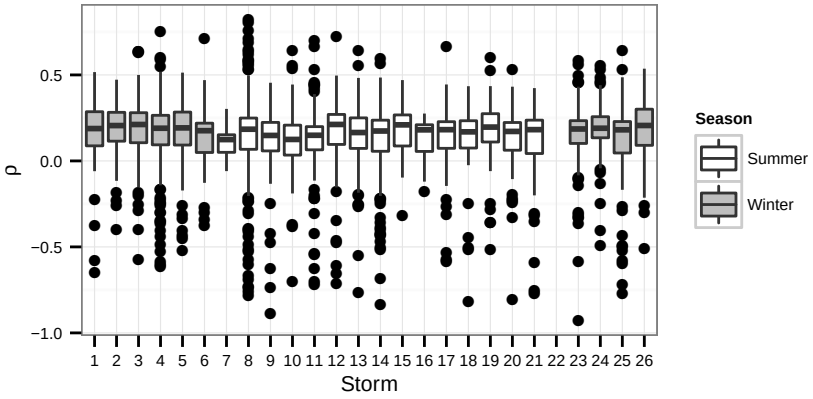

Fig. 14. The correlations of all scales for each of the storms. Note that almost all rainfall fields exhibit correlations, and that almost all of them are positive.

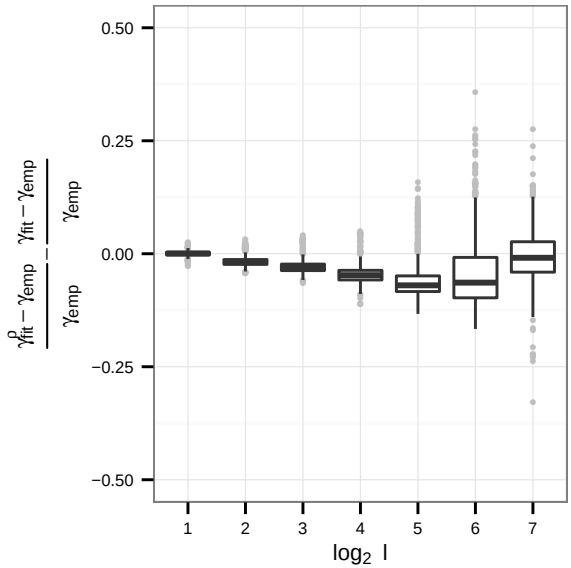

Fig. 15. The difference between the relative error of the distribution without correlation, and that with correlation propagated from the largest scale. The inclusion of correlations leads to small, but consistent, improvements (i.e. negative values). It is immediately clear that the coarsest scale is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

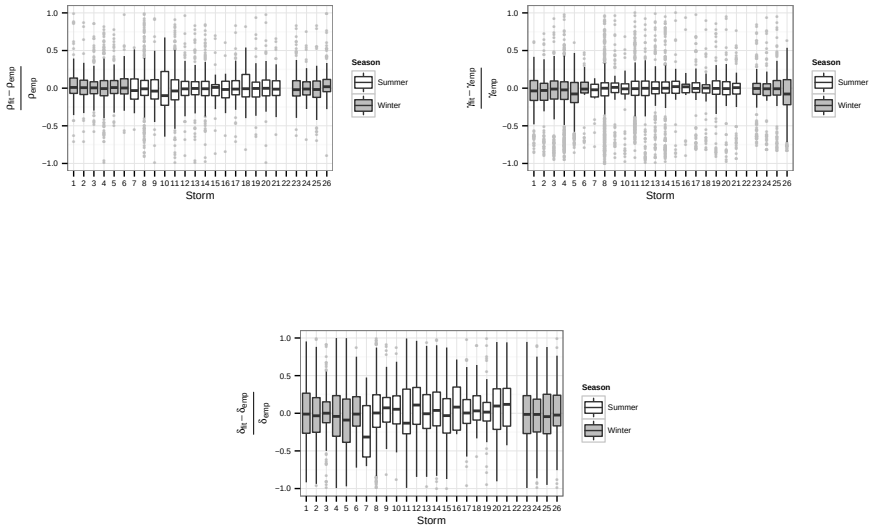

Fig. 16. The relative errors of the mean, shape and correlation functions. All functions appear to behave relatively stable troughout winter and summer, with the exception of γ where the summer storms have a smaller error, but more outliers.

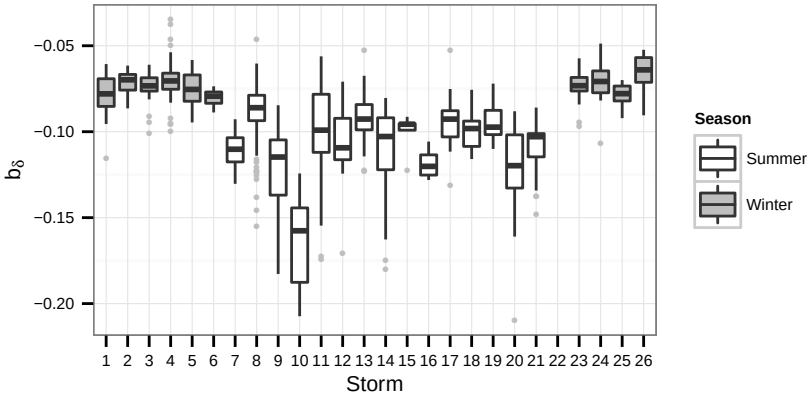

Fig. 17. The parameter for the mean of the increment, shown as boxplots for each storm. Summer storms clearly have lower increments.

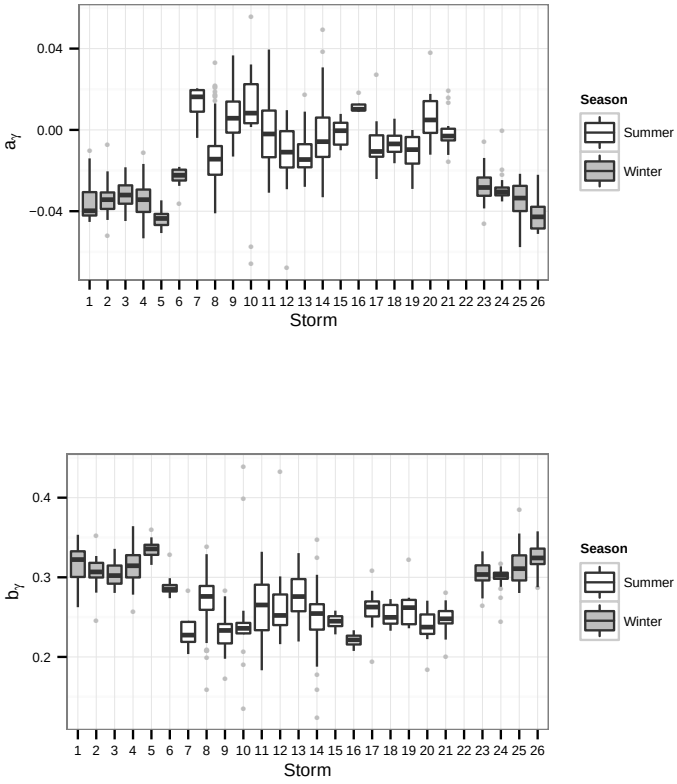

Fig. 18. The parameters for the γ of the increment, shown as boxplots for each storm. The intercepts are higher for summer storms, suggesting more energy in the flux, but with smaller slopes, i.e. the scales are more similar. Winter storms show the exact opposite behavior.

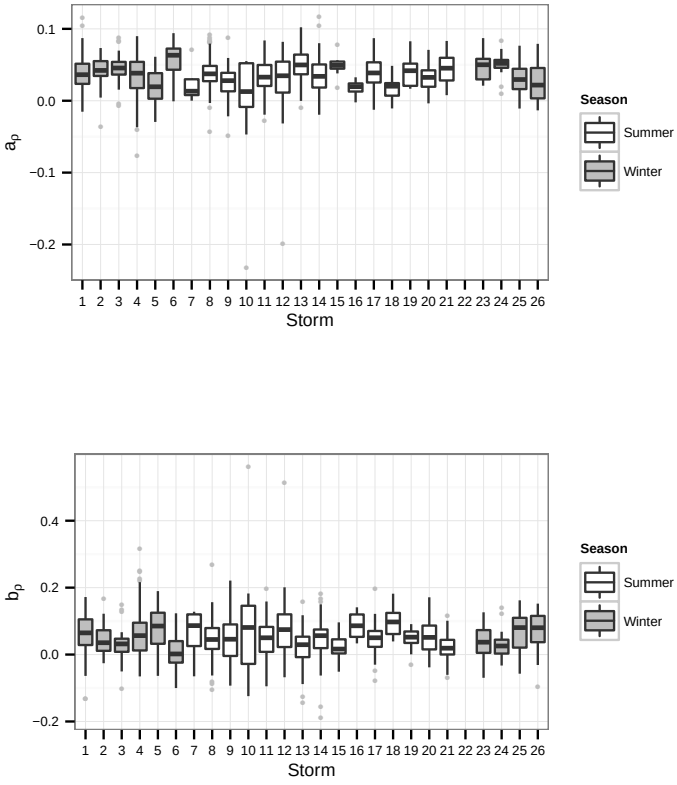

Fig. 19. The parameters for the correlation of the increment, shown as boxplots for each storm. The functions for the correlation appear stable throughout winter and summer, and have a high intercept, but a low slope suggesting that the correlation are somewhat the same for all scales.

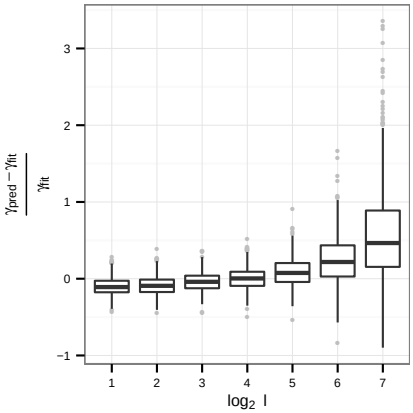

Fig. 20. The difference between the relative error of the distribution without correlation, and that with correlation propagated from the largest scale. The inclusion of correlations leads to small, but consistent, improvements (i.e. negative values). It is immediately clear that the coarsest scale is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