Detailed Comments

1) Discrete scale ratio cascades versus continuous in scale cascades:

The authors consider only cascades built by iteratively dividing "parent" structures into λ_0 "daughter" eddies where λ_0 is an integer scale ratio ≥ 2, here taken = 4 (, in the authors' notation, $\lambda_0 = b = 1/s_i$; the most popular value is $\lambda_0 = 2$). The result is a highly artificial construct since for example, it is only scaling for integer powers of λ_0 (itself an integer). In contrast the real world presumably does not display such scale "quantisation" being on the contrary continuous in scale with structures at all intermediate scales. In addition, as mentioned above, these discrete in scale models are necessarily acausal (i.e. left-right symmetric on the time axis) whereas real rain sequences are not at all symmetric since the past influences the future but not visa versa. Causal cascade models have been around since (Marsan et al., 1996) so that in 2010 there is no need for this. In physicists' jargon, the authors' model is a "toy" model, introduced to illustrate the features of a full realistic continuous in scale model; it was never intended to be applied directly to the real world. Continuous in scale models have been available for over twenty years (Schertzer and Lovejoy, 1987) and numerically efficient implementations (including the treatment of finite size effects and causality) are now available (Lovejoy and Schertzer, 2010a), (Lovejoy and Schertzer, 2010b).

2 Weights versus fragmentation ratios:

Leaving aside the unrealism of the discrete in scale models, there is a fundamental theoretical "slippage" which occurs concerning "weights" and "fragmentation ratios". Indeed in one sentence (section 4.2), the expression "weight distribution of fragmentation ratios" is used implying that the two notions are indeed considered equivalent. The issue is quite fundamental and is worth clarifying since it is unfortunately widespread in the precipitation literature. To start, the authors express more or less correctly the construction of the cascade (eqs. 3, 4), where the *p*'s are conventionally called "weights". Before continuing, we need to issue a warning: the notation "*p*" dangerously leads one to think of "probabilities", which in general they cannot be (this is indeed the unfortunate origin of this borrowed notation). It would be better to use the standard "*w*" for "weights" or "$\mu\varepsilon$" for a "multiplicative increment" in the turbulent cascade flux ε). The real problem comes with the definition of the fragmentation ratio (the authors' eq. 5). Since the ratio is also denoted by p_j, it is clear that the authors already mix it up with the weights (eqs 3,4). Let us on the contrary denote the fragmentation ratio defined eq. 5 by f, and see under which conditions we obtain $f = p$ (note that in some publications the confusing term "multipliers" is used instead of "fragmentation ratio").

To illustrate why the key eq. 5 is in general wrong, it suffices to consider the simple dyadic case $\lambda_0 = 2$, and consider only the first two steps of the cascade. Denote the two parent weights at the first level by $p_0^{(1)}$ and $p_1^{(1)}$, (for the intervals (0,1/2), (1/2,1), respectively) and their four "daughter" weights at the second level by $p_{0,0}^{(2)}$, $p_{0,1}^{(2)}$, $p_{1,0}^{(2)}$, $p_{1,1}^{(2)}$ (for the intervals (0,1/4), (1/4,1/2), (1/2,3/4), (3/4,1) respectively). We see that after the two cascade steps, the actual values of the cascade process for the corresponding intervals is ($p_0^{(1)}p_{0,0}^{(2)}$, $p_0^{(1)} p_{0,1}^{(2)}$, $p_1^{(1)}p_{1,0}^{(2)}$, $p_1^{(1)}p_{1,1}^{(2)}$).

The problem is now how to recover the weights. If the details of the cascade construction are known, then we can indeed divide the second generation by the first generation i.e. $p_0^{(1)}p_{0,0}^{(2)} / p_0^{(1)} = p_{0,0}^{(2)}$ (the authors' eq. 5). The trouble is that in real world applications, we do not know the intermediate construction levels – only the final level - so that we are stuck with estimating the low resolution, weights $p_0^{(1)}$ (over the interval (0,1/2)) from the observed higher resolution values $p_0^{(1)}p_{0,0}^{(1)}$, $p_0^{(1)} p_{0,1}^{(1)}$ which are already products of weights over the two smaller intervals (0,1/4), (1/4,1/2). The only way to do this is to estimate the lower resolution cascade by spatial averaging:

$$p_0^{(1,d)} = (p_0^{(1)}p_{0,0}^{(2)} + p_0^{(1)} p_{0,1}^{(2)})/2 = p_0^{(1)}(p_{0,0}^{(2)} + p_{0,1}^{(2)})/2 \qquad (1)$$

where the "*d*" in the superscript indicates "dressed" or spatially averaged cascade quantity; the corresponding "bare" quantity is simply $p_0^{(1)}$ (we say "spatial" since as mentioned, this construction is adequate for 1-D spatial sections, but not (causal) time series). More generally, the "bare" cascade quantity is the result of developing the cascade over a total range of scales λ whereas the corresponding "dressed" quantity is for the cascade developed over a wider range of scales (possibly down to infinitely small scales) and then spatially averaged over the corresponding scale ratio λ. This is the fundmental "bare" versus "dressed" distinction elucidated in the 1980's and which is essential for understanding realistic cascades (particularly for their extremes, but that's another story!).

We can now see that the author's fragmentation ratio is:

$$f = \frac{p_0^{(2,b)}}{p_0^{(1,d)}} \qquad (2)$$

where we have added the superscript "*b*" to underlie the fact that the numerator requires a bare quantity while the denominator requires a dressed one. Since this argument is valid for any two consecutive levels, we can replace the "2" by "*j*" and the "1" by "*j*-1" and recover the authors' eq. 5 with the important difference that the numerator is a dressed and denominator a bare quantity.

We can now use the fact that the *p*'s all have the same distribution; let us take p_0, p_1 as simply two identically distributed random variables for the left and right daughter weights so that we obtain for the $\lambda_0 = 2$ case:

$$f_0 = \frac{2p_0}{p_0 + p_1}; \quad f_1 = \frac{2p_1}{p_0 + p_1} \qquad (3)$$

Here, we have expressed the dressed weights $p_0^{(1,d)}$ in terms of the bare weights (eq. 1) and f_0, f_1 are the left and right fragmentation ratios for the two daughters. The point is now clear: the author's eq. 5 is only valid if there are very strong correlations on all the daughter weights so that $p_0 + p_1 = 2$, leading to $f_j = p_j$ (the authors' eq. 5).

More generally for a division by an arbitrary positive integer λ_0, (and including the authors' case is $\lambda_0 = 4$) we have for the j^{th} fragmentation ratio f_j:

$$f_j = \lambda_0 \frac{p_j}{\sum_{i=1}^{\lambda_0} p_i} \tag{4}$$

where the p_j are identically distributed weights. For eq. 5 to be valid we require that they are strongly correlated so that $\sum_{i=1}^{\lambda_0} p_i = \lambda_0$. A cascade where this equation is satisfied by construction, it is called "microcanonical"; such cascades are clearly artificial constructs. In any case, it's one thing to use academic "microcanonical" models, it's another to postulate that $\lambda_0 = 4$ and that the (unobserved) empirical weights presumed to generate the data will somehow satisfy this constraint on the (more or less) arbitrarily partitioning involved in the empirical analysis.

In order to demonstrate the deleterious consequences of the microcanonical assumption, we made some simple numerical simulations using a Log-Levy discrete in scale model cascade which assumes - in addition to being identically distributed - that the weights are also statistically independent so that we only have the "canonical" conservation $\langle p \rangle = 1$. The bare statistics of this model are specified by the moment scaling exponent $K(q) = C_1(q^{\alpha}-q)/(\alpha-1)$. In standard notation, this function is defined by $K(q) = \log_{\lambda_0}\langle p^q \rangle$ (p is the weight) so that for example $K(q)>0$ for all $q>1$ (note that this is the opposite sign from the non-standard definition used by the authors).

Using parameters close to those found in satellite radar data (Lovejoy et al., 2008); $\alpha = 1.5$, $C_1 = 0.3$ (for the rainrate), fig. 1 shows the histograms for both the Log-Levy weights and the fragmentation ratios obtained by taking $\lambda_0 = 2$ and 4. We see that if we follow the authors and use the fragmentation ratios to estimate the weights then the all-important extremes are particularly biased. However, as λ_0 increases, the effect of the microcanonical constraint decreases. Fig. 2 shows the corresponding moment scaling functions $K(q) = \log_{\lambda_0}\langle p^q \rangle$ (weights) and $K(q) = \log_{\lambda_0}\langle f^q \rangle$ (fragmentation ratios). We see that for the higher moments, the difference is quite large. However, even for the important derivative $K'(1) = C_1$ which is the codimension characterizing the sparseness of the mean ($q = 1$), that there is still a strong bias ($C_1 = 0.3$ for the weights, but 0.15, 0.21 for the fragmentation ratios for $\lambda_0 = 2, 4$ respectively).

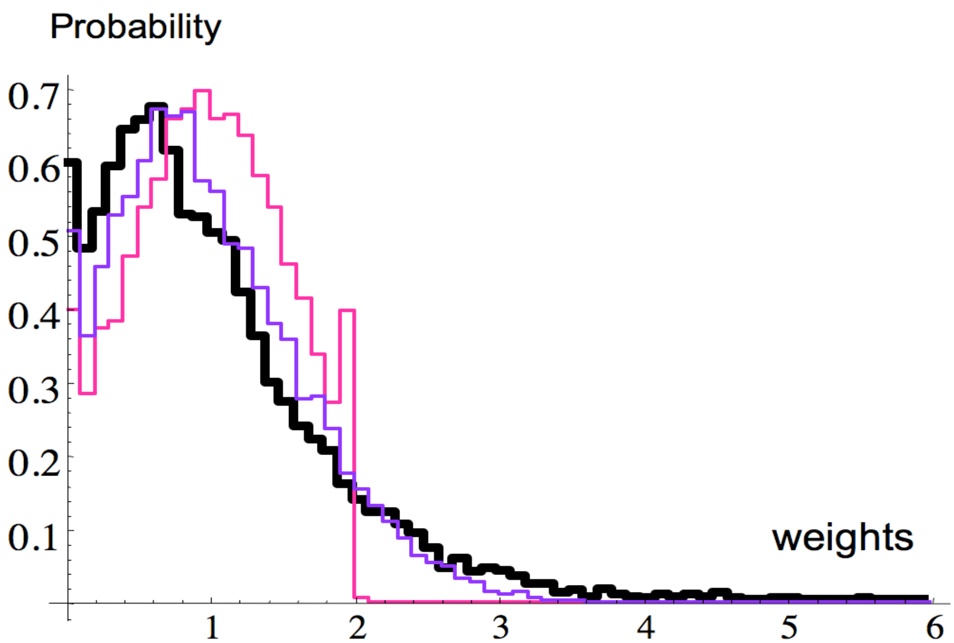

Fig. 1: Histograms showing the probability density of the weights (black) and the fragmentation ratios (red, purple) for λ_0 =2, 4 respectively. The weights were from a Log-Levy distribution with α =1.5, C_1 =0.3; there were 10,000 simulated ratios.

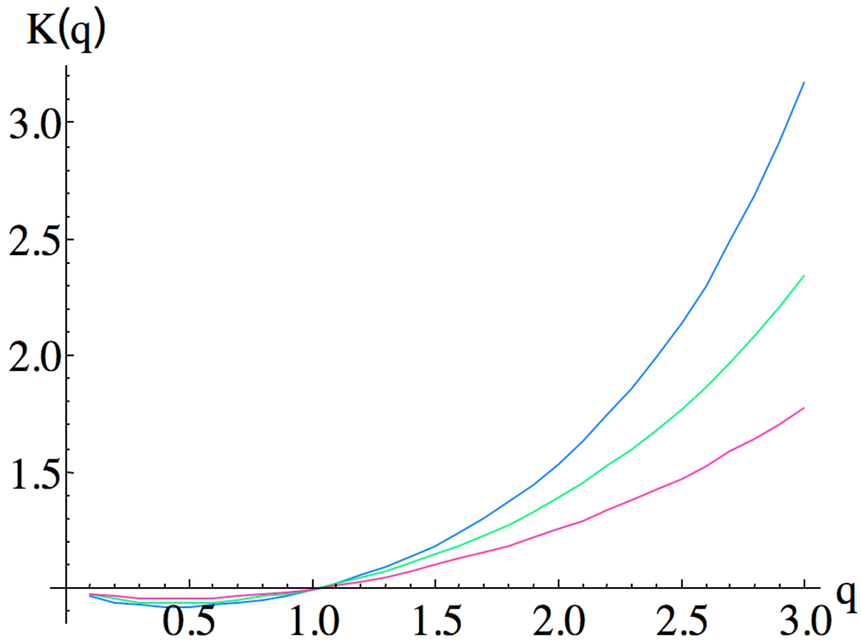

Fig. 2: The scaling moment function $K(q)$ for the weights (top right), the fragmentation ratios for λ_0 =4 (right, middle), and for λ_0 = 2 (right, bottom). The corresponding $C_1 = K'(1)$ are 0.3, 0.21, 0.15.

3. The Log-Poisson model:

As indicated in the introduction, this model is a poor choice since it has a built in *a priori* assumption that the rain process is somehow incapable of generating events stronger than a critical value; a maximum order of singularity denoted by γ_+ (for details see (Schertzer et al., 1995)). Defining $c = C/\log 2$, the maximum singularity is: $\gamma_+ = c(1-\beta)$ where C and β are the authors' parameters (eq. 7). Using the author's table 2, we find that the monthly γ_+ is in the range $\approx$ 0.5 to 0.8 which is not so high – an unsurprising result given the large bias (underestimates) of the extreme weights that result when the extreme fragmentation ratios are used instead. Also, rather than being an insurmountable maximum, the observed γ_+ varies as expected as though it were indeed no more than a random fluctuation from a process without any intrinsic maximum value (or from one with a true maximum much higher than were observed).

Even ignoring the bias from the fragmentation ratios, surely it doesn't make sense to use a finite sample with a finite maximum in order to estimate the maximum that the model can ever produce on an infinite ensemble? It's like taking the maximum of 100 sample values from a process and then making a model constrained *a priori* to never exceed that particular empirical value. Why not use a model (such as the Log-Levy one) that is capable of producing arbitrarily large singularities and then study the actual sample to sample variation in its maximum? Even better: see if the random variation of the maxima are just as the model predicts!

4. Zeroes and low rain rates

The authors mention the zero rain rate problem. Actually, there are two problems here. The first is empirical: the extreme difficulty of all known instrumental methods in objectively, and accurately measuring low and zero rain rates. Recall that for multiplicative cascades, it is the even more problematic logarithm that is most relevant. The second problem is theoretical: how to introduce the zeroes in the model (assuming that they are indeed real). There are two basic methods, either multiply the rain field by a "beta" model (i.e. an indicator function for a fractal set), thus –following (Gupta and Waymire, 1993) -making the rainrate zero everywhere except on a fractal support (i.e. it almost surely never rains anywhere!), alternatively, put a threshold on the low rain rates and set them to zero. In this case, the threshold breaks the scaling – at least for the low rain rates – but the authors already note that the empirical scaling is indeed poor for these. Indeed, using such a thresholded model (Lovejoy et al., 2008) showed that the observed poor scaling at low rainrates can be accurately reproduced as an artefact of the thresholding (presuming that the threshold is an instrumental effect; if it is real, the mechanism still works, but the poor scaling is no longer an artefact).

Are the authors really wise to attempt to reproduce the detailed poor scaling of the low and zero rates even though these may be no more than artefacts of the instruments?

-Shaun Lovejoy

References

V.K. Gupta and E. Waymire, A Statistica l Analysis of Mesoscale Rainfall as a Random Cascade, *Journal of Applied Meteorology* **32**(1993), pp. 251-267.

S. Lovejoy and D. Schertzer, On the simulation of continuous in scale universal multifractals, part I: spatially continuous processes, *Computers and Geoscience*(2010a), p. 10.1016/j.cageo.2010.1004.1010.

S. Lovejoy and D. Schertzer, On the simulation of continuous in scale universal multifractals, part II: space-time processes and finite size corrections,, *Computers in Geosciences*(2010b), p. 10.1016/j.cageo.2010.1007.1001.

S. Lovejoy, D. Schertzer and V. Allaire, The remarkable wide range scaling of TRMM precipitation, *Atmos. Res.* **10.1016/j.atmosres.2008.02.016**(2008).

D. Marsan, D. Schertzer and S. Lovejoy, Causal space-time multifractal processes: predictability and forecasting of rain fields, *J. Geophy. Res.* **31D**(1996), pp. 26,333-326,346.

D. Schertzer and S. Lovejoy, Physical modeling and Analysis of Rain and Clouds by Anisotropic Scaling of Multiplicative Processes, *Journal of Geophysical Research* **92**(1987), pp. 9693-9714.

D. Schertzer, S. Lovejoy and F. Schmitt, Structures in turbulence and multifractal universality. In: M. Meneguzzi, A. Pouquet and P.L. Sulem, Editors, *Small-scale structures in 3D and MHD turbulence*, Springer-Verlag, New York (1995), pp. 137-144.