We thank Referee #3 for her/his detailed comments on the paper.

Two important sets of issues are raised, and in both cases, we thank the referee for a careful reading of the paper which brings up issues that certainly need to be addressed if the message in the paper is to be conveyed clearly.

A. Derivations of the equations for moments of the rainfall depth.

A.1. The reviewer argues that no separation of integrals occurred in the original papers and that the condition on α remains valid (i.e. that it must be larger than 1):

In the model structure however, each storm has a different value of η. In Rodriguez-Iturbe et al. (1988), the derivation of the variance and covariances does not make use of separate integrals as claimed in the present paper: it just uses the expectation of $exp(-\eta\varphi\tau)/\eta$ where τ is a temporal lag; and (correctly) notes that this expectation exists only when $\alpha > 1$. It therefore looks to me as though the apparent problem noted by the present authors may be an artefact of an incorrect — or, perhaps, needlessly complicated — approach to the derivation.

Ans: In response, we would first like to point out that, if one looks at the first time these equations were derived, in the paper the referee refers to (Rodriguez-Iturbe et al., 1988), the derivation does in fact involve separating the integrals. This can be seen by looking at equation (2.2): the expectation that is calculated is obtained by separating into additive terms and taking the expectation of each of these terms (thus integrating each additive term separately).

This might seem fine, and indeed, one of the present authors rederived these equations in the past without noticing any problem. But in fact the integral of a sum of terms is only equal to the sum of the integrals of each additive term when the latter are finite. When the latter are infinite (which happens for $\alpha \leq 3$ not 1, because the pdf of the Gamma distribution is multiplied by this term as explained in the paper), this is not necessarily the case. That is, it is, in general, possible that the integral of the sum should be finite while the integrals of the additive terms are infinite.

Here is an example to illustrate this: consider the integral:

$$I = \int_0^x \frac{e^{\omega t} - e^{-\sigma t}}{t}\, dt$$

If we consider the sum of the two integrals:

$$I_1 = \int_0^x \frac{e^{\omega t}}{t}\, dt \text{ and } I_2 = \int_0^x \frac{-e^{-\sigma t}}{t}\, dt$$

we have two divergent integrals and we cannot say what this sum is ($+\infty - \infty$).

But $I \neq I_1 + I_2$ because, using Taylor expansions, we see that I is in fact finite:

$$I = \int_0^x \frac{1 + \omega t - 1 + \sigma t + o(t)}{t}\, dt = \int_0^x \big(\omega + \sigma + o(1)\big)\, dt$$

So, for x small:

$$I \cong (\omega + \sigma)x$$

This situation is analogous to that in the paper, with $x = \eta_0$. That is, we have shown that the integrand can be approximated by sums of Taylor series which, in the neighbourhood of $\eta = 0$ have terms that cancel out, so that the convergence of the integral as a whole is not defined by the convergence of the expectation of $exp(-\eta\varphi\tau)\ /\eta$ or $exp(-\eta\tau)\ /\eta$. As we show in the paper, in the vicinity of $\eta = 0$, the integrand of the variance of the RBL1 model is a term in $\eta^{\alpha-2}$ which converges as long as $\alpha > 1$.

To avoid this misunderstanding, we add some text to explain the key point that one cannot just treat integrals of additive terms separately when they are infinite.

A.2. The referee then raises a good point about the results found by previous authors and the claim that we make that they could be erroneous.

the authors' reporting of previous results with "non-valid" estimates for the α parameter (lines 178–181) should have made them stop and think more carefully. The reason is that the model fits are obtained by minimising an expression involving the theoretical model properties. Earlier authors must have calculated the properties for these values of α, therefore; but this wouldn't be possible if the integrals diverged (or the algebraic expressions would have produced results that are obviously wrong, such as negative values of $E(X^2)$).

Ans: There are two senses in which previous results might have been erroneous. The first, and most important one is that the domain of valid values of α was smaller than it need be in cases where the authors used the restrictions upon α required by the separate integration of the additive terms. This is the main reason for the work in this paper: it will allow a broader domain of values of α to be used.

But second, and this is the case that the referee is alluding to, the issue of convergence of the integrals was often not addressed by the authors (including some published with one of the present authors as co-author). This would however not necessarily be picked up when carrying out the optimisation, because all it means is that the expression for the corresponding moment would not have had a numerical value that was that of the model skewness. This would not have involved any negative numbers (only in some exponents). So, for instance, if we look at the claim we made on lines 173-179, we are drawing attention to the fact that, for the convergence of variance and covariance of the RBL1, the condition $\alpha > 3$ is required. The papers listed have some parameters smaller than 3. How does that translate in terms of the kind of quantities these authors would have found for, say, the variance. Well, if we look at the equation for the variance they would have used (see lines 203-205), it involves some values of the function T: $T(2, \dots, \dots)$ and $T(3, \dots, \dots)$. Going back to the definition of this function T (lines 157-158), we see that when α is less than 3, the term in the denominator, i.e. $(\nu + u)^{\alpha-k}$ has an exponent $\alpha - k$ which is $\alpha - 2$ for $k = 2$ and $\alpha - 3$ for $k = 3$. These exponents are less than 1 when $\alpha \leq 3$. But we will, as the referee points out, get some negative terms because of the negative $\alpha - 3$ but overall, we find positive variances. For instance, looking at the parameters A_1 and A_2 found by Rodriguez-Iturbe et al. (1988),

$$A_1 = \frac{\lambda\mu_C\nu^\alpha}{(\alpha-1)(\alpha-2)(\alpha-3)}\left[E(X^2) + \frac{\kappa\varphi\mu_X^2}{\varphi^2-1}\right]$$

$$A_2 = \frac{\lambda\mu_C\kappa\mu_X^2\nu^\alpha}{\varphi^2(\varphi^2-1)(\alpha-1)(\alpha-2)(\alpha-3)}$$

we find that term A_1 on p.288 of that paper is negative when $1 < 2 < \alpha < 3$, while A_2 is positive. This will in fact generally be the case as can be seen from the following reasoning. Recall that φ is the ratio of the mean cell duration in the storm and the mean duration of storm activity. In practice (by which I mean, for a physically realistic set of parameters), we will therefore have $\varphi < 1$. From this we can conclude that $A_2 > 0$ when $1 < 2 < \alpha < 3$.

Next, consider κ: we will in practice have $\kappa > \varphi$, because κ is the ratio of the mean cell duration in the storm and the mean cell interarrival time in the storm, whereby the latter is less than the mean duration of storm activity. With these relations, we can show that the terms in square brackets is positive, and therefore $A_1 < 0$. Starting with the fact that $E(X^2) > \mu_X^2$ since the difference between left- and right-hand side is the variance of the cell depth X, we can then write the following for the terms in square brackets:

$$E(X^2) + \frac{\kappa\varphi\mu_X^2}{\varphi^2-1} = E(X^2) - \frac{\kappa\varphi}{1-\varphi^2}\mu_X^2 > E(X^2)\left(1 - \frac{\kappa\varphi}{1-\varphi^2}\right)$$

since the coefficient of μ_X^2 is negative. Further, we get:

$$E(X^2)\left(1 - \frac{\kappa\varphi}{1-\varphi^2}\right) = E(X^2)\left(\frac{1-\varphi^2-\kappa\varphi}{1-\varphi^2}\right) > E(X^2)\left(\frac{1-\varphi^2-\kappa^2}{1-\varphi^2}\right) > E(X^2) > 0$$

The expression for the variance is:

$$\begin{aligned}\operatorname{var}\left(Y_i^{(h)}\right) = {} & 2A_1\{(\alpha-3)h\nu^{2-\alpha} - \nu^{3-\alpha} + (\nu+h)^{3-\alpha}\} \\ & - 2A_2\{\varphi(\alpha-3)h\nu^{2-\alpha} - \nu^{3-\alpha} + (\nu+\varphi h)^{3-\alpha}\}\end{aligned}$$

in which we have just shown that $A_1 < 0$ and $A_2 > 0$.

The terms in curly brackets are both negative for the parameters we have looked at. The key to the sign of the variance will therefore be the relative sizes of $|A_1|$, $|A_2|$ and of the associated curly bracket terms (denote $|C_1|$ and $|C_2|$ here). Values for these expressions with a typical parameter set for which $2 < \alpha < 3$ are given in Table 1. As can be seen, with typical parameter values: $|A_1| < |A_2|$ and $|C_1| \gg |C_2|$. Therefore, we still get a positive expression for the variance. This means that this would not be picked up as anomalous in calibrating the model.

To clarify these issues, we introduce a sentence explaining the two types of problems and emphasising the one about non-optimality because of an unnecessarily narrow parameter space (which is currently not sufficiently clear in the paper). Since we cannot estimate the impact of the second problem without looking at the particular data sets used in past papers, we rephrase the comment we make about previous studies so that it indicates that one cannot be sure that the parameters found in these previous studies are optimal.

Table 1: Calculations of the variance expression, given in Rodriguez-Iturbe et al. (1988), and the associated parameters at 1-h timescale ($h = 1$) for $2 < \alpha < 3$. Other parameters used are $\lambda = 0.025$, $\mu_x = 1.3$, $\nu = 0.28$, $\kappa = 0.65$ and $\varphi = 0.04$.

α	C_1	A_1	$C_1 \times A_1$	C_2	A_2	$C_2 \times A_2$	$\mathrm{var}\left(Y_i^{(h)}\right)$
2.1	-1.0871	-1.0031	1.0904	-0.0003	206.7770	-0.0574	2.2958
2.2	-1.1085	-0.4554	0.5048	-0.0006	93.8752	-0.0524	1.1145
2.3	-1.1159	-0.2820	0.3147	-0.0008	58.1307	-0.0482	0.7257
2.4	-1.1048	-0.2017	0.2229	-0.0011	41.5858	-0.0445	0.5348
2.5	-1.0703	-0.1592	0.1703	-0.0013	32.8071	-0.0414	0.4234
2.6	-1.0060	-0.1368	0.1377	-0.0014	28.2088	-0.0386	0.3526
2.7	-0.9046	-0.1296	0.1172	-0.0014	26.7155	-0.0362	0.3069
2.8	-0.7573	-0.1415	0.1071	-0.0012	29.1578	-0.0340	0.2823
2.9	-0.5533	-0.2098	0.1161	-0.0007	43.2381	-0.0321	0.2963

A.3. The referee proposes a simplification of the equations by using the fact that ratios of Gamma functions can lead to simpler expressions.

The authors' concerns about convergence are all focused on the situation where $l = 0$, because this is where the integrand can become infinite. In this case however, the final numerator in the expression above is a complete gamma function so that the expression can be written as

$$\Gamma(k,u,l) = \frac{\nu^\alpha}{(\nu+u)^\alpha}\frac{\Gamma(\alpha-k)}{\Gamma(\alpha)}$$

But if $k > 0$ is an integer (which I think it is throughout the paper), we have $\Gamma(\alpha)/\Gamma(\alpha-k) = (\alpha-1)(\alpha-2)\ldots(\alpha-k)$ providing $\alpha - k$ isn't a negative integer (if it is, then $\Gamma(\alpha-k)$ is undefined). Thus

$$\Gamma(k,u,l) = \frac{\nu^\alpha}{(\nu+u)^\alpha(\alpha-1)\ldots(\alpha-k)}$$

which is obviously finite providing none of the terms in the denominator is zero. Unless I've missed something, this seems to resolve the convergence issue much more simply.

Ans: This simplification is indeed well known and indeed, going back again to the Rodriguez-Iturbe et al., (1988), it was used for instance in equation (2.4): this is why no Gamma functions appear in the expressions of coefficients A_1 and A_2, but rather products such as $(\alpha-1)(\alpha-2)(\alpha-3)$.

As the referee notes we can only use the Gamma function (and the simplifications that follow) when the variable in the Gamma function is non-negative (there is actually an extension of the Gamma function to negative non-integer numbers, but there is no obvious justification for using it: the Gamma function is introduced in the calculations because of the Gamma distribution, which requires positive parameters). This means, as stated by the referee that $\Gamma(\alpha-k)$ is defined as long as $\alpha > k$. This is precisely the issue that our paper is dealing with: this condition arises when calculating the integrals separately, and unnecessarily restricts the domain of feasibility for the minimisation of the objective function. By not treating the integrals separately which leads to using these (complete) Gamma functions, we show that the domain of possible values of α is broader.

Since it is likely that some readers will also wonder why we did not use complete Gamma functions and the simplifications which ensue, we add a sentence to explain this.

B. The block estimators

The referee points to the fact that we have not used the standard unbiased estimators of the variance and that, if the problem that we are flagging is that some previous authors have not done this, then it is not worth being discussed extensively, if at all.

The authors' second main point relates to the calculation of "block" statistics used for model calibration with uncertainty. They claim that the block estimators of variances and other quantities are biased (e.g. lines 257-260). However, the expressions that they give for these estimators are incorrect because there is no adjustment for degrees of freedom in the denominator in either case: the denominator in the first expression should be $N_y N_{m,h} - 1$ and that in the second expression should be $N_y(N_{m,h} - 1)$. In fact, Section 5.1 of their Jesus and Chandler (2011) reference (cited on line 244) discusses the need for careful treatment of small-sample biases: that discussion would probably be relevant to quantities such as the skewness coefficient, discussed by the present authors at lines 291-294. Jesus & Chandler did not discuss the variance specifically: I assume that this is because the form of an unbiased variance estimator is well-known so they didn't think it needed mentioning. If the variance expressions given by the authors are indeed in standard use, this is worrying: a decent journal is probably not the best place to draw attention to such a basic error, however. The bottom line is that there isn't necessarily a problem with block estimators per se; but (as with any other sophisticated technique) if you're going to use them then you need to do it carefully.

Ans: We fully agree that, indeed, if the issue were the bias that is introduced by using $N_g N_{m,h}$ rather than $N_g N_{m,h} - 1$ and $N_g(N_{m,h} - 1)$, then this would not be worth discussing. As the referee points out, the differences between estimators using the first rather than the second coefficients are of relevance to small samples.

However, we are not concerned with small samples here. The smallest sample that might be involved would be of size 30 (in the case of the daily time-scale, for the block estimation method, and for all other time-scales used in the fitting, they are considerably larger). The estimators which do not use the '-1' adjustments are biased (and we indicate that, at least for one of them, after equation (10)). We would not agree that they are 'wrong': they just have a bias that can be corrected by using the '-1' adjustments pointed out by the referee, but they are asymptotically unbiased. And for samples that are, at worst, of size 30, the bias is not of much practical relevance (e.g. typically a few percent).

Still, one might ask why we did not use the unbiased estimators in any case, to avoid introducing even such small biases. That is a good point which we should have indicated in the paper. The reason is that it leads to simpler expressions for the comparison of the standard and block estimators. So, we get an equation (10) in which one can easily interpret the difference between the block estimator and standard estimator in terms of an additional term that is the variance of the monthly averages. In effect, this result is just the well-known result expressed in terms of sums of squares in

ANOVA (e.g. see Kottegoda and Rosso (2008), p. 285). Since the reason for putting in this equation is just to give some understanding of where the differences between the estimators comes from, this seems sufficient for that purpose. Ultimately, significant differences between block and standard estimators only arise when dealing with ratios, as we explain in the paper, and for these, there are no useful such equations to write down.

In fact, the numerical estimates that we have shown are for the unbiased estimators, so that removes any concern about the impact of the biases in question. However, this points to the fact that we need to clarify these issues in the paper which, as currently presented, can be confusing. We therefore add some text to indicate that for the biased (but asymptotically unbiased) estimators of the variances, we get the relations shown in the equations, while the relations are a little more complex when using unbiased estimators. We also add a sentence to indicate that the numerical values are for unbiased estimators. We also indicate that this brief investigation is not aimed at making any general point about block estimators, but simply indicating the problems that arise in the case of the estimator of a ratio like the coefficient of skewness.

References

Kottegoda, N.T. and Rosso, R. (2008) *Applied Statistics for Civil and Environmental Engineers*, 2nd ed., London: Blackwell

Rodriguez-Iturbe, I., Cox, D.R., Isham, V. (1988) A point process model for rainfall: further developments*, Proc. Roy. Soc. Lond.*, A417, 283-298