

Interactive comment on “A Hybrid Stochastic Rainfall Model That Reproduces Rainfall Characteristics at Hourly through Yearly Time Scale” by Jeongha Park et al. (RC3)

Anonymous Referee #3

Received and published: 17 August 2018

Dear Anonymous Referee #3

The authors here propose and test a composite method which appears able to generate synthetic rainfall time series for a wide range of aggregation time scales. The issue of producing reliable synthetic rainfall time series is undoubtedly a central problem in hydrology, and the extension to a wide range of scales attempted in this study has the potential to be a relevant contribution. Therefore, the scope of this study makes it suitable for consideration in Hydrology and Earth System Sciences. However, there are a number of issues that needs clarification and/or correction, which I detail below. In particular, my main comment concerns the validation of the method (comment 6 below). I would recommend the paper for publication only after these issues have been addressed by the authors.

Authors' Response. We sincerely appreciate your constructive comments on our manuscript. All your comments tremendously helped us to improve the quality of the article. Our responses are as follows:

Major Comments

Comment 1. The so-called Module 2 uses only one quantity (mean hourly rainfall, as obtained from the 1st module) to generate finer scale rainfall statistics (mean, variance, correlation, and dry fraction). Is not clear to me how the regression analysis for module 2 is carried out (Page 10, line 10). For example, the variance is computed from the mean assuming the two quantities are linearly related. How is the slope $\alpha_{[6]}$ computed? Using the entire dataset available for each station and each month? If this is the case, is not true that the proposed model solely use monthly information (as stated e.g., in Figure 4) to produce fine scale rainfall statistics, and this should be clarified. I find that section 3.2 is overall not very clear, and could be improved. You use ‘functional relations’ between many quantities, but a relation is only shown for the mean vs standard deviation relationship. Perhaps this linear relation could be introduced more generally (e.g., $X = \alpha_i Y + \epsilon_i$), and then state that X and Y can be the variables of interest, for example mean, variance, ... ?

Authors' Response. We agree that the explanation about the second module is vague. We modified the manuscript as follows:

Revised Contents.

3.2 Generation of fine time scale rainfall statistics

The second module generates the fine time scale statistics corresponding to each monthly rainfall value generated through the SARIMA model. These synthetic fine time scale statistics will later be used for the calibration of the MBLRP model to downscale the monthly rainfall to the hourly level. In so doing we first consider the monthly rainfall, when divided by the number of days in the month times 24, as providing us with an estimate of the mean hourly rainfall for that particular month. Then, this estimated mean hourly rainfall is provided as the input variable of the module that generates the statistics needed to fit the MBLRP model, namely the mean, variance, auto-correlation coefficient, and the proportion of dry periods at 1-, 2-, 4-, 8-, and 16-hour aggregation intervals, as described in Figure 5. In this process, the module employs the information obtained from univariate regression analyses between the fine-scale statistics of the observed rainfall (Figure 6) and the mathematical formulae relating rainfall variance and auto-covariance at different time scales (Equation 4) as explained below.

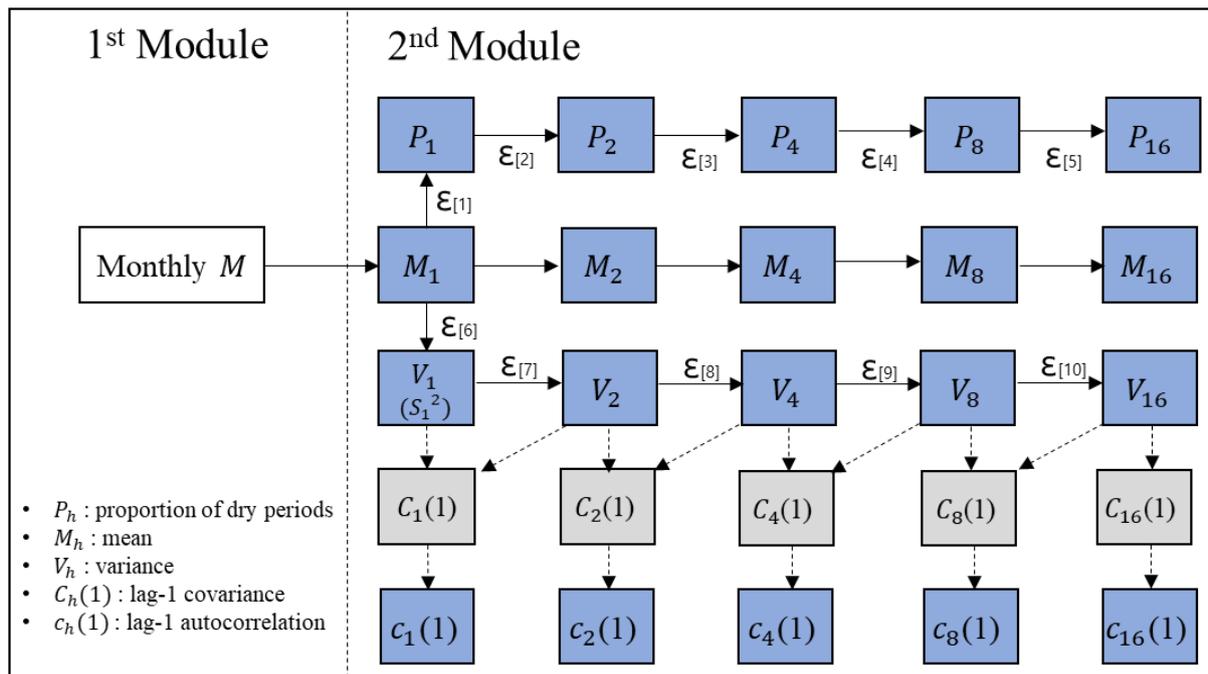


Figure 5: Schematic of the algorithm to generate fine time-scale rainfall statistics. The statistics in the blue boxes are used to calibrate the MBLRP model and the statistics in gray boxes are used to estimate the lag-1 autocorrelation.

Figure 5 shows a schematic of the second module. In the figure, $M_h, S_h, V_h, c_h(1) = C_h(1)/V_h$ and P_h in each rectangle represent the rainfall mean, standard deviation, variance, lag-1 autocorrelation,

and proportion of dry periods at time-scale h hours, respectively. The statistic connected to each solid arrow head is stochastically generated based on its linear relationship to the one connected to the tail of the same arrow. In other words, the following equation is used:

$$Y = a_{[i]} X + b_{[i]} + \varepsilon_{[i]} \quad (2)$$

where Y is the variable being generated, and the X is the variable being used as the basis of the generation. Here, the variable X and Y can be substituted by any combination of two variables connected by the solid arrow; $a_{[i]}$ and $b_{[i]}$ are the parameters of the regression analysis, and $\varepsilon_{[i]}$ is a random number drawn from the normal distribution $\varepsilon_{[i]} \sim N(0, \sigma_{[i]}^2)$ fitted to the residuals of the regression analysis. Here, these three parameters are estimated from the univariate regression analysis relating the two variables observed during a given calendar month over multiple years as shown by black scatters in each plot of Figure 6, which shows the linear relationship between the rainfall statistics observed at gage NCDC-200164 (star mark in Figure 3) during the month of July of different years.

The statistic connected to the dashed arrow head is calculated based on the ones connected to the tail of the same arrow using the mathematical (deterministic) relationship connecting these variables (Equation 4). For instance, in Figure 5, V_1 and V_2 are connected to $C_1(1)$ through a dashed arrow, which means that $C_1(1)$ is derived from V_1 and V_2 . The following equations establish the relationship between the variances at time-scales h and $2h$ from which we shall obtain the relationship between V_1 and V_2 :

Comment 2. Also, it is not immediate to me how all these relations between rainfall statistics can be linearly related, especially rainfall mean and wet fraction. I think it would be helpful to show how these linear relations hold for all the stations in the study, not just a sample rain gauge. Is it possible they depend on season/rainfall regimes?

Authors' Response. Regarding the linearity, we prepared the plots for all gauge locations and all seasons, which can be accessed through the following website:

<http://www.letitrain.info/>

Here are some notes about our linearity assumptions:

- (i) We assumed that the hourly standard deviation (S_1), i.e. not the hourly variance (V_1), is linearly correlated to the hourly mean (M_1) as suggested by the black scatters in Figure 6(a).

After we generated S_1 from M_1 based on this relationship, we took the square of it to obtain the hourly variance (V_1). We believe that the linearity between M_1 and S_1 is not a bold assumption considering numerous previous studies that models the rainfall distribution as exponential (mean = λ^{-1} , standard deviation = λ^{-1}) or gamma (mean = $k\theta$, standard deviation = $k^{0.5}\theta$) distribution;

- (ii) The linearity between the variance at different aggregation intervals can be explained by the following equation given in the manuscript.

$$Var\left(Y_i^{(2h)}\right) = Var\left(Y_{2i-1}^{(h)}\right) + Var\left(Y_{2i}^{(h)}\right) + 2Cov\left(Y_{2i-1}^{(h)}, Y_{2i}^{(h)}\right)$$

$$V_{2h} = 2V_h + 2C_h(1)$$

We can consider two extreme cases. First, if $Y_{2i-1}^{(h)}$ and $Y_{2i}^{(h)}$ are independent, then we get a linear regression with the gradient of 2 ($V_{2h} = 2V_h$). Second, if $Y_{2i-1}^{(h)}$ and $Y_{2i}^{(h)}$ were identical, then the covariance is equal to the variance, so we would get $V_{2h} = 4V_h$. It is therefore reasonable to assume that a linear relationship with a gradient lying between 2 and 4 exists for the other cases.

- (iii) We could not find studies that explicitly deal with a relationship between hourly mean and hourly proportion of dry periods (M_1 vs P_1). However, our empirical analysis at all 34 stations suggests a strong linear relationship between the two variables. Please see the figures at:

<http://www.letitrain.info/>

- (iv) Regarding the relationship between the proportions of dry periods at different aggregation intervals, Onof et al. (1994) showed that the mean number of events at time scale h , is given by the following relation to the proportions dry period:

$$E(N_h) = \frac{P_h}{P_h - P_{2h}}$$

By rearranging the equation, we get:

$$P_{2h} = P_h \left(1 - \frac{1}{E(N_h)}\right).$$

This, therefore, suggests looking at whether the coefficient here is reasonably stable and therefore whether there is a linear relationship between these two proportions dry.

Reference. Onof, C., Wheater, H. S. and Isham, V.: Note on the analytical expression of the inter-event

time characteristics for Bartlett-Lewis type rainfall models, *J. Hydrol.*, 157, 197-210, 1994.

Comment 3. In general, figure captions should be improved and expanded throughout the manuscript, explaining more in depth what is in the figure. For example, the caption of figure 1 should state what the blue areas (cells) and shaded lines (total rainfall intensity I guess) are. Also, even if it is just a schematic figure, Figure 1 should have axes for time and rainfall depth. Figure 6 caption should specify that the results are for a single gauge, etc. The caption of figure 14 is probably swapped (blue and red lines – please check).

Authors' Response. We completely agree. Done as suggested as follow:

Revised Contents.

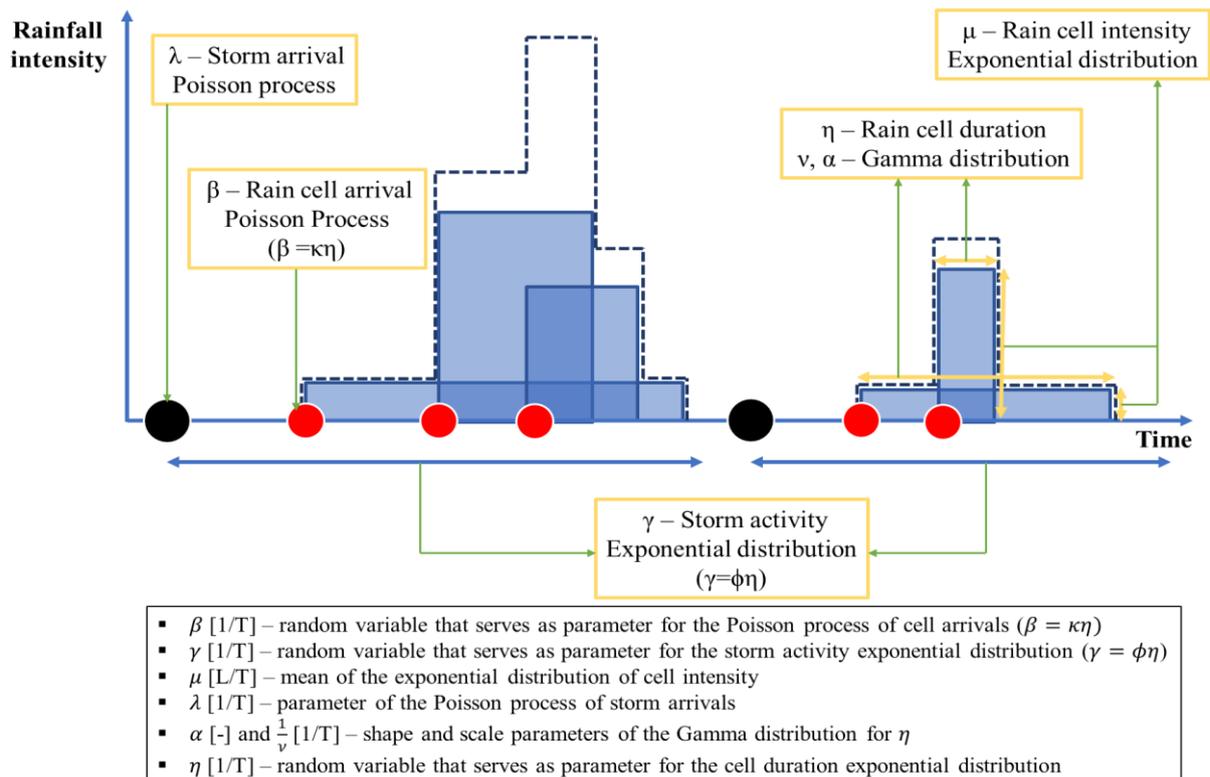


Figure 1: Schematic of the Modified Bartlett-Lewis Rectangular Pulse Model. The blue area represents duration (width) and intensity (height) of each rain cell, respectively. The dashed line represents superposed sum of the rain cell intensities.

Figure 6: Linear relationship between various fine time-scale statistics of the rainfall observed for the month of July of different years at gage NCDC-200164 (black dots). The solid black line represents the

least squares regression line. Based on this regression relationship, a set of the 20 fine-time scale statistics are generated, which are immediately used as the basis of the MBLRP model parameter calibration. If the objective function of the parameter calibration corresponding to the generated set is greater than a given threshold, the set is rejected (blue squares), and the set with the objective function lower the threshold value is only chosen (red squares).

Figure 7: (a) Comparison of estimator $\widehat{c(\mathbf{1})}$ (horizontal axis) with estimator $\frac{\widehat{V}_2}{2\widehat{V}_1} - \mathbf{1}$ (vertical axis) of the autocorrelation lag-1 of hourly rainfall, (b) The histogram of the discrepancies between these two estimators *at gage NCDC-200164*.

Figure 8: (a) Relationship between $\epsilon_{[7]}$ and $\epsilon_{[8]}$ and the fitted bivariate distribution. (b) Color map of the correlation coefficient between different $\epsilon_{[i]}$ s *at gage NCDC-200164 on September*.

Figure 14: Degree of over/underestimation of extreme values by our model (red) and the traditional MBLRP model (blue). ER_{syn} and ER_{obs} are extreme rainfalls estimated from synthetic rainfall and observed rainfall, respectively.

Comment 4. The authors present results for a particular station in some of the figures, and this station is not the same throughout the paper. I think it would be better to be consistent and present the results for the same station.

Authors' Response. We agree with your concern. We replaced the figures in the methodology section (Figure 6, 7, and 8) such that they represent the result of the same station (NCDC-200164). In the results section, Figure 10 is the only one corresponding to an individual gauge, but Figure 11 conveys the information of Figure 10 for all gauge locations. All other figures in the results section (Figure 9, 11, 12, 13, and 14) display the analysis of all gauge locations. The discussion section shows the analysis on gage NCDC-460582 (Figure 15 and 16), which clearly showed our point of discussion.

Comment 5. Page 15, line 17: it is stated that the Module 2 may fail in generating realistic fine scale rainfall statistics. The Authors should include a bit more explanation for this. How often does it happen? Given that Module 2 is based on linear relations, is it possible for some of this relations not to be linear in some cases, and cause these failures? (This may vary with precipitation types/rainfall regimes). This could be assessed checking if, in a cases of 'failure', any of the relations between rainfall statistics in Module 2 exhibit a divergence from linearity more marked than in other cases.

Authors' Response. We appreciate this thorough review. The cases of failure in which the parameter

set satisfying the objective function criteria could not be generated even after the 50 iterations were identified for 11 months over 8 gage locations. (NCDC-85663(Jun and Dec), 111549(Mar), 138706(Jan and Dec), 360106(May and Dec), 401094(Nov), 431081(Dec), 447201(May), and 460582(Jul)), which comprises 0.01 percent of the entire months simulated ($200 \text{ years} \times 12 \text{ months per gage} \times 34 \text{ gages} = 81,600 \text{ months}$). For these months, the set with the lowest objective function was selected among the 50 generated statistics set. It should be noted that the case of failure was not repeated for the same month at the same gage. If the linearity assumption was not correct, we would expect to see a repetitive (or systematic) behavior of failure. Therefore, we have no grounds for viewing the linearity assumption as the primary reason of the failure.

Comment 6. Figure 12 summarizes the performance of the rainfall generation method, comparing observed and simulated rainfall statistics. I think the method should be validated using an independent set of observations, not the same used for calibration (i.e., used for computing the regression coefficients). For example, the authors could divide the gauge time-series in two independent samples, using one of them for calibration and the second for validation. In the end, this is what we want to achieve with a synthetic rainfall generator: match as much as possible statistics of time series that we do not have available. This analysis could also resolve a second issue: Since the proposed model is more complex/has more parameters than the original MBLRP, using independent samples for validation would show to what extent the additional model complexity can improve method performance. Also for figures 9, 10 and 11 it would be helpful to show results obtained using independent samples.

Authors' Response. We completely agree with this view. We used the period of the year 1951 to 1980 as the validation period, which is the set of 30 years prior to the calibration period. Figure 9 to Figure 12 were redrawn, and the following paragraphs were added in the manuscript.

Revised Contents.

3.5 Validation for Ungaged Periods

One of the primary purposes of the stochastic rainfall model is to provide synthetic rainfall for the un-gaged periods, which can be the periods of missing data or future periods. For this reason, we separated the period of model calibration and validation at some gage locations (square marks in Figure 2) where record length of each period is sufficiently long (60+ years). Then, we tested our model not only based on the statistics of the calibration period (1981-2010) but also based on the validation period (1951-1980).

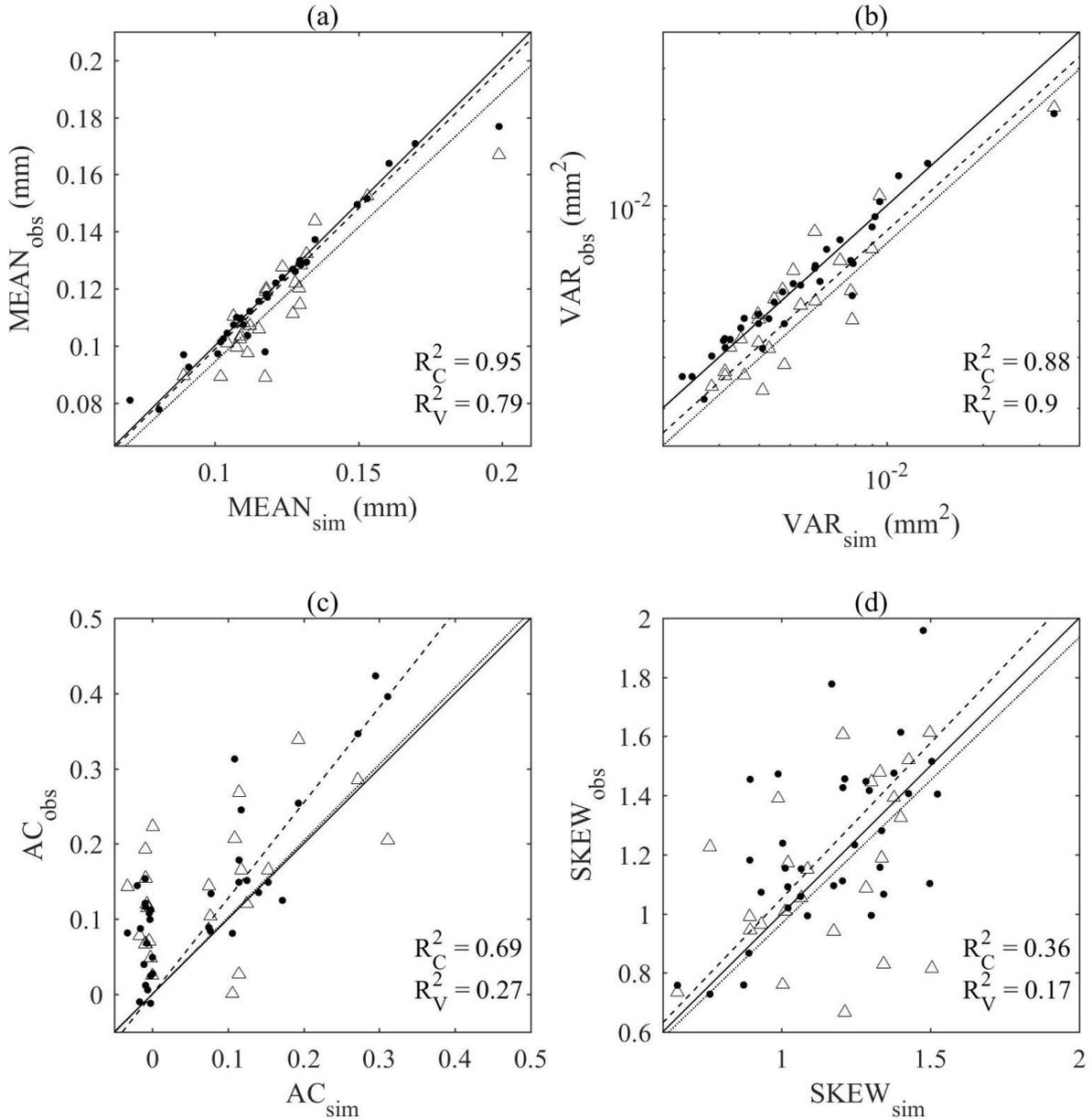


Figure 9: Comparison of (a) mean, (b) variance, (c) lag-1 autocorrelation, and (d) skewness of the synthetic (x) and observed (y) monthly rainfall. Filled circles (dashed line) and hollow triangles (dotted line) correspond to the calibration (1981-2010) and validation period (1951-1980) respectively.

4.1 Monthly Rainfall Statistics Reproduction

Figure 9 compares the mean, variance, lag-1 autocorrelation, and skewness of the monthly rainfall time series generated by the SARIMA model (x axis) to those of the observed monthly rainfall time series (y axis). Each scatter represents one rainfall gage. For the calibration period (1981-2010), the first and the second-order moments were reproduced accurately with the coefficient of determination ranging from 0.69 to 0.95. Skewness was reproduced fairly well with the coefficient value of 0.36. For the validation period (1951-1980), mean and variance were reproduced, but not lag-1 autocorrelation and skewness. However, this discrepancy cannot be attributed solely to limitations in the model because the

discrepancy in each plot of Figure 9 directly results from the differences between the statistics of the calibration and validation periods. In other words, had the statistics of the calibration period been similar to those of the validation period, we would have expected similar performance for both periods, and vice versa.

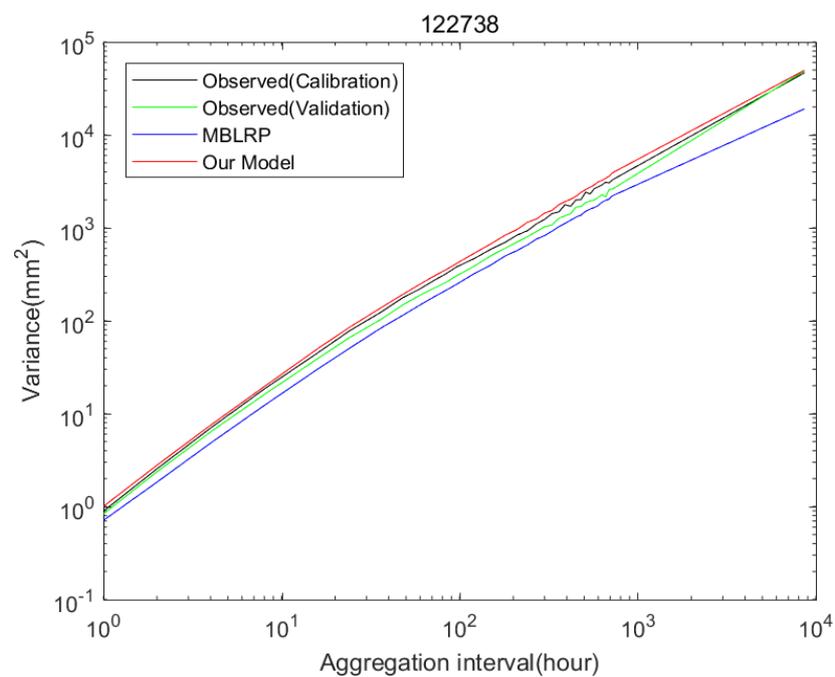


Figure 10: Behaviour of the rainfall variance with regard to the aggregation interval of rainfall time series at gage NCDC-122738. The behaviour corresponding to the observed-calibration (black, 1981-2010), observed-validation (green, 1951-1980), MBLRP (blue) and our hybrid model (red) are shown together.

4.2 Reproduction of Large Scale Rainfall Variability

Figure 10 shows the behaviour of the rainfall variance varying with temporal aggregation interval between 1 hour and 1 year at gage NCDC-122738. The behaviour corresponding to the observed-calibration (black, 1981-2010), observed-validation (green, 1951-1980), MBLRP (blue) and our hybrid model (red) are shown together. ...

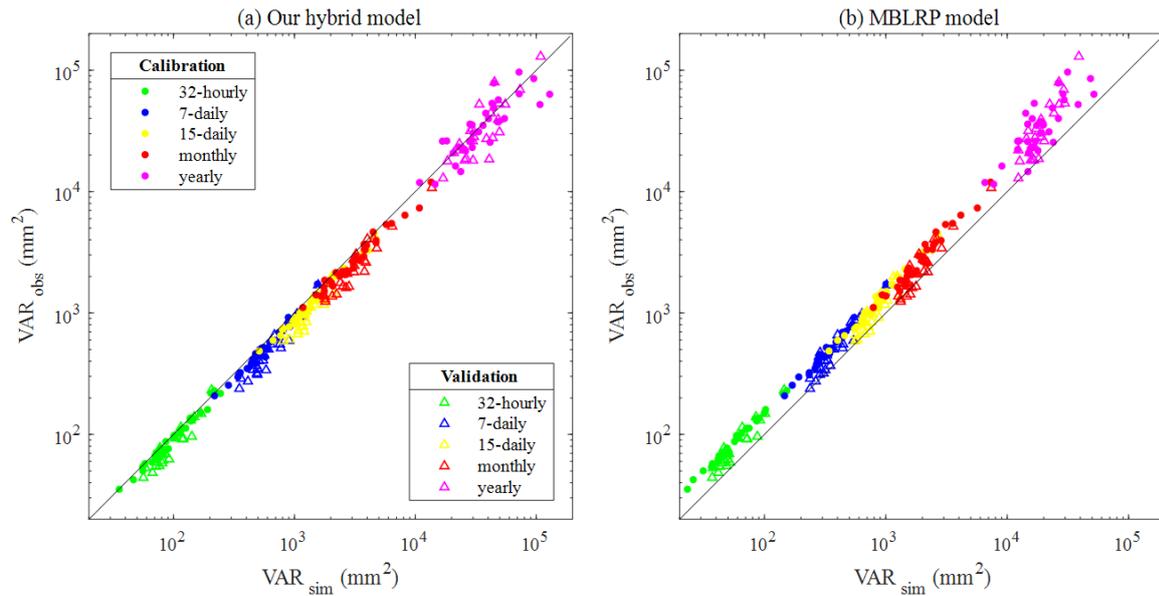
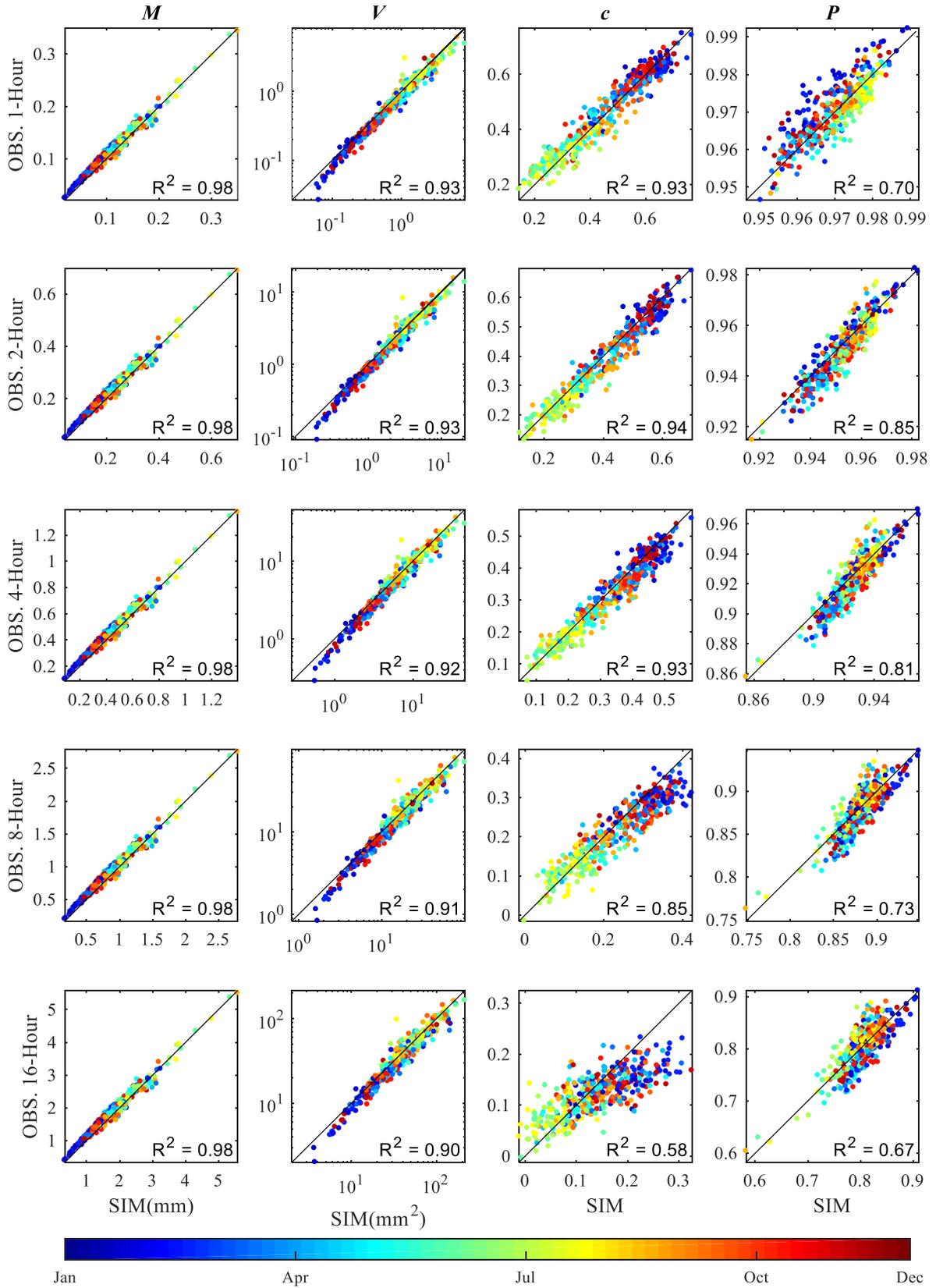


Figure 11: (a) Comparison of the large scale rainfall variance of the rainfall generated by our hybrid model (x) and the observed rainfall (y); (b) Comparison of the large scale rainfall variance of the rainfall generated by the traditional MBLRP model (x) and the observed rainfall (y). The different colours of the scatter correspond to the different aggregation interval of rainfall time series. Filled circles and hollow triangles correspond to the calibration and validation periods respectively.

...Figure 11 compares the variance of the synthetic (x) and observed (y) rainfall time series at yearly (purple), monthly (red), 15-daily (yellow), weekly (blue), and 32-hourly (green) aggregation levels. The comparison of the variance at the finer time scale is carried out in the following section.

As indicated by the concentration of the scatters above the 1:1 line in Figure 11b, the traditional MBLRP model systematically underestimates the variability at time scales greater than 32 hours. Our model did not show any bias in this range of large time-scales as shown in Figure 11a.

(a) Calibration



(b) Validation

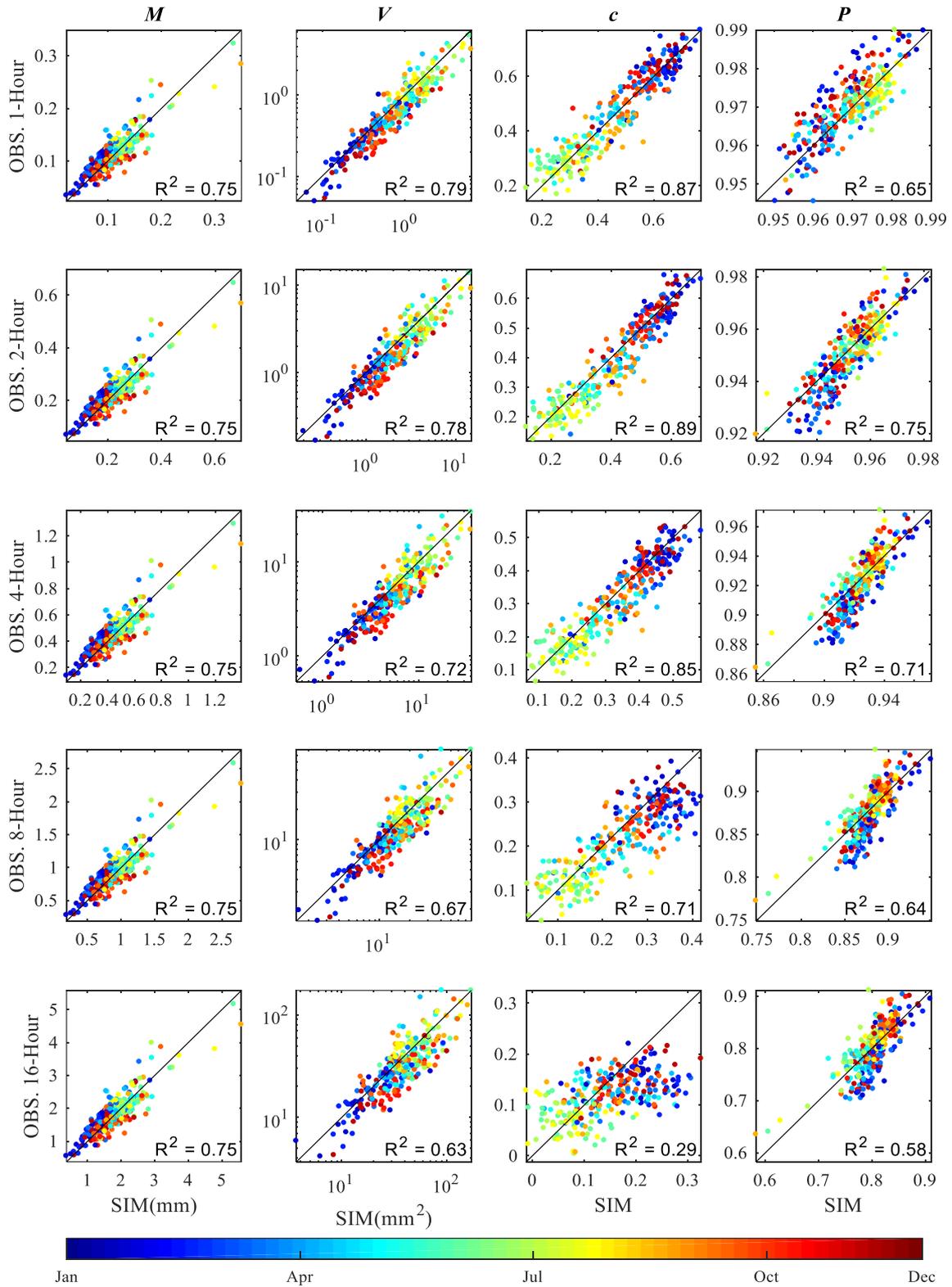


Figure 12: Comparison of the statistics of the synthetic (x) and observed (y) rainfall time series at sub-daily time scale. The colour of the dots represents the statistics of each calendar month. The results of (a) the calibration period (1981-2010) and (b) the validation period (1951-1980) are shown.

4.3 Reproduction of Sub-Daily Rainfall Statistics

Figure 12 compares the mean, variance, lag-1 autocorrelation, and the proportion of dry periods of the synthetic (x) and observed (y) rainfall time-series at hourly through 16 hourly aggregation levels. Each scatter represents the statistics at a given gage for a given calendar month. The colours of the scatters represent the calendar months. In each plot, the coefficient of determination (R^2) of the linear regression between the two variables is shown. All four statistics were accurately reproduced across various sub-daily time scales with R^2 equal to 0.98 (mean), and varying between the following limits for the other statistics: 0.90 and 0.93 (variance), 0.58 and 0.93 (lag-1 autocorrelation), and 0.67 and 0.85 (proportion of dry periods) on the calibration period (Figure 12a). Similar ranges of coefficient of determinations were obtained for the validation period (Figure 12b).

5.3 Calibration versus validation

Our approach of separating the period of calibration and validation adopted to some gage locations, may seem surprising because most stochastic rainfall generators are calibrated based upon statistics under an assumption of temporal stationarity of the rainfall process. According to this assumption, the statistics of the periods of calibration and validation should be the same, which obviates the need for validating the model for separate periods. However, this assumption often does not obtain, for example, in case the observation period is too short (e.g. a few extreme events are included in only one part of the time period) or in when the time series is indeed non-stationary. For this reason, the discrepancy of the model performance between the calibration and the validation period should not only attributed to the model limitations but also to the differences between statistics from the two periods. In view of these considerations, our primary purpose of separating the period of calibration and validation should be understood as an assessment of the model's applicability to rainfall generation for a future period. From the point of view of the calibration period, the validation period is an ungauged period just as any future period, and our model can be easily extended to the future period by adding a term accounting for long-term rainfall non-stationarity to the SARIMA model (first module). Our hybrid model assumes not only the stationarity of the typical rainfall statistics such as mean, variance, covariance and proportion of dry periods but also the relationship between them (See Figure 6). The latter has not been explicitly discussed by previous studies, so it will be very interesting to see whether such relationships between the statistics hold over different temporal periods and how the discrepancy affects the final model performance, if there is any. Figure 19 compares the slope of the regression analysis between the statistics shown in Figure 6 for the calibration (x axis) and validation (y axis) periods. The plots corresponding to the variances at different scales are not shown because there are theoretical reasons for having a solid slope close to 2 (See Equation 5 and the preceding equations). There is no significant discrepancy between slopes estimated using statistics on calibration and

validation period implying that relationships between the fine time scale statistics are stationary and that our model can be used for future or unged periods.

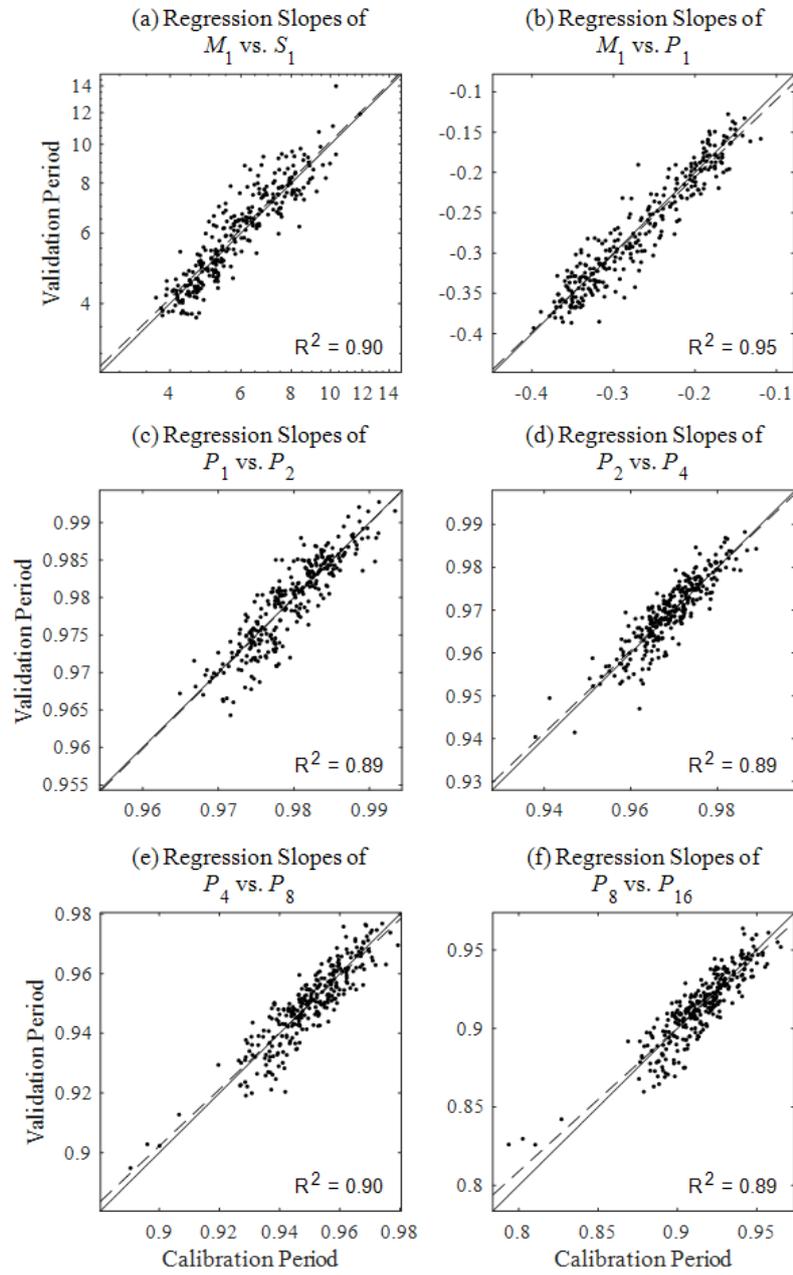


Figure 19: Comparison of the slope of regression analysis between the statistics shown in Figure 6 for the calibration (x) and validation (y) period. The slopes of regression analysis (a) between mean and standard deviation and (b) between mean and proportion of dry periods and (c)-(f) between proportion of dry periods at the different time scale were compared. Solid lines are 1:1 line and dashed lines represent the regression lines.

Comment 7. In the MBLRP model, cell durations are represented by a double stochastic process which produces a long-tail matching observation as stated at line 14. However, rainfall intensities are also known to be of sub-exponential nature (Wilson and Toumi 2005). While MBLRP uses an exponential distribution for rainfall intensities, the method proposed here appear to improve the performance in reproducing rainfall extremes. You state that this results on rainfall extremes (better performance compared to MBLRP) is ‘surprising’: However, this is likely due to the inter-annual variation of the parameters as you also discuss in section 5.1. On this effect, see Zorzetto et al (2016), where it is shown that the interannual variation of exponential-type rainfall distributions, as is the case with your findings, can explain the emergence of fattailed extreme value distributions. The fact that your model generates time series with improved extreme value properties is very appealing, as this is a traditional shortcoming of Poisson-cells based rainfall models.

Authors’ Response. This is the best compliment made about this article. Indeed, the heavier tail could be reproduced after introducing stochasticity to the parameters of the exponential distribution representing the cell intensity. Please see the following figure that we added to the revised manuscript.

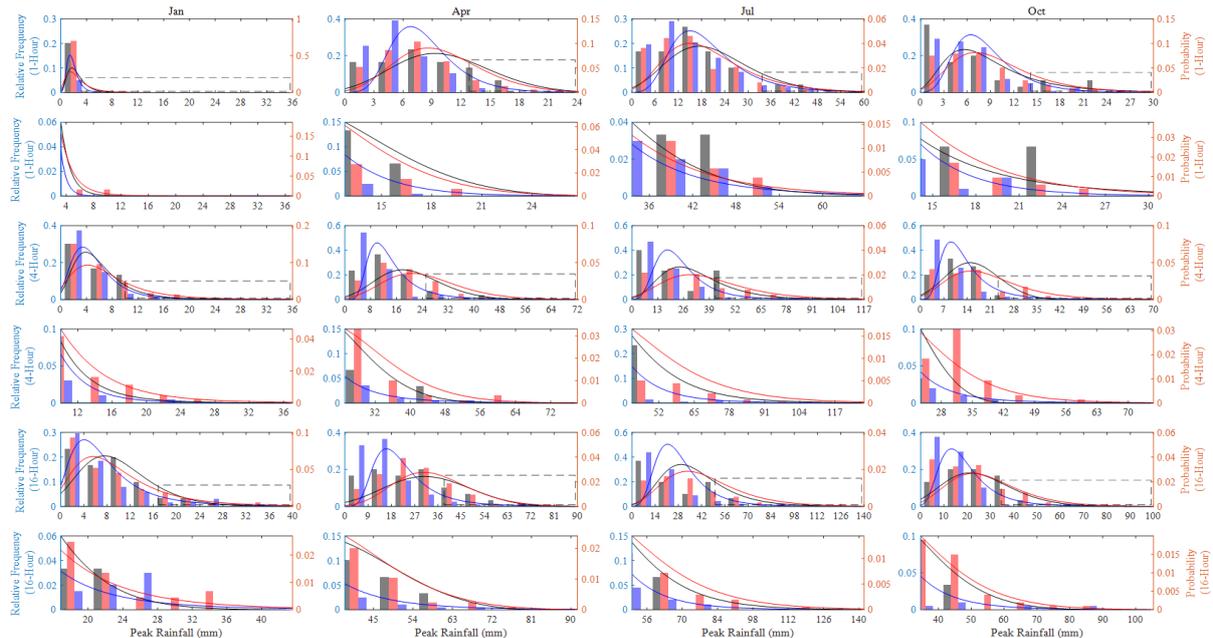


Figure 15. Relative frequency and the fitted GEV distribution of the 1, 4, and 16-hour monthly maxima of January, April, July, and October rainfall at gage NCDC-132203. Results of Observed rainfall (black), our hybrid model (red), and the traditional MBLRP model (blue) are shown. The upper 10 percentile part of the distribution (dashed box in the plots in the first, third, and fifth row) is magnified in the lower rows (plots in the second, fourth, and sixth row).

In addition, we added the following sentences in the manuscript:

Revised Contents.

Zorzetto et al. (2016) also briefly discussed this matter. They introduced a novel framework of meta-

statistical extreme value (MEV) analysis. In this MEV formulation, one can show that interannual-variation of exponential-type rainfall process leads to a fat-tail for its extreme values.

Reference. Zorzetto, E., Botter, G. and Marani, M.: On the emergence of rainfall extremes from ordinary events, *Geophys. Res. Lett.*, 43, 8076-8082, 2016.

Minor Comments

1) Page 4, line 29 – Rain gauge with ID 85663, please correct

Authors' Response. Done as suggested.

2) Page 5, line 2 – ‘the discrepancy between their quartiles/ their range..’. Please also specify in the caption what the whiskers of the boxplots are (quartiles?)

Authors' Response. Done as suggested.

Revised Contents.

Figure 2: Box plots of the observed monthly rainfall at gage NCDC-85663 in Florida, USA (red). The box plots of the synthetic monthly rainfall generated by the Modified Bartlett-Lewis Rectangular Pulse model at the same gage are shown for reference (blue). Whiskers reach to minimum and maximum values of monthly rainfall during the period between 1961 and 2010 and gray shaded boxes represent the discrepancy of the variability of the two monthly rainfalls.

3) The sentence at lines 27-28 is not clear, please clarify. The variance is not represented by the vertical axis in Figure 2, even though the boxplots do give an idea of the variability of the seasonal monthly rainfall distributions.

Authors' Response. Done as suggested.

Revised Contents.

In Figure 2, the red box plots represent the distribution of the monthly rainfall observed at gage NCDC-85663 located in Florida, USA during the period between 1961 and 2010.

4) Figure 3: please state in the legend what the values in the map are (SARIMA parameters) and what the coloring is, even if it is already explained in the text.

Authors' Response. Done as suggested.

Revised Contents.

Figure 3: Study area and 34 NCDC hourly rainfall gages. The label of the markers is presented in the following format: aaaaaa(i,j,k)(x,y,z)₁₂, where aaaaaa represents the NCDC gage ID, (i, j,k) represent the orders of the autoregressive, differencing, and moving average terms of the SARIMA model, and (x,y,z) represent the orders of the seasonal autoregressive, differencing, and moving average terms of SARIMA model. The colour of the markers represent the Bayesian Information Criterion (BIC) value of the SARIMA model. The lower BIC indicates either more parsimonious parameterization, larger likelihood, or both. Model description of SARIMA is detailed in Section 3.1.

5) Page 22, line 10: 'It is important TO NOTE THAT..' or something on this line.

Authors' Response. This line was already pointed from RC2. It was changed as follows:

Revised Contents.

A good rainfall model should reproduce not only the extreme values but also the distribution of the maxima from which extreme values are derived.

6) Page 26, line 1: 'and the subsequent EFFECTS ON human..'

Authors' Response. Done as suggested.

7) Page 26, line 3; 'time scaleS'

Authors' Response. Done as suggested.