Review for manuscript "A Hybrid Stochastic Rainfall Model That Reproduces Rainfall Characteristics at Hourly through Yearly Time Scale"

The manuscript under review examines a mixed stochastic model using an autoregressive model (SARIMA) for the generation of monthly precipitation that is then disaggregated to hourly scale through a modified Bartlett-Lewis rectangular pulse (MBLRP) model. This method can capture the long-term behaviour of large scale (monthly) precipitation through the SARIMA model and, implicitly (through coupling), to preserve some statistical characteristics (e.g. marginal mean, variance and skewness, lag-1 autocorrelation, and dry proportions) of the small scale (hourly) precipitation through the MBLRP model, thus simulating some aspects of intermittency. The method is tested at 34 hourly stations located in USA where the monthly maxima were also well preserved there.

The paper is well organized and well written and it matches the of HESS journal. My only concerns are that some significant points (related to the innovations of this work) may still need some additional justification and discussion on other stochastic methodologies. Although my field of expertise is not on Bartlett-Lewis models (but rather on stochastic synthesis of processes from small to large scales in an explicit manner) it is highly relevant to some innovative points of the Authors' analysis. Therefore, I hope some of my comments and suggestions can contribute to the Authors' methodology and help them improve it and highlight it to Readers interested in stochastic modelling of precipitation in general.

Sincerely,

Panayiotis Dimitriadis

Major comments and suggestions

The main innovations of the presented model is the coupling of a large scale model (such as the SARIMA) that can reproduce some long-term properties (e.g. long-term behaviour of autocorrelation and monthly maxima), and a small scale one (such as the MBLRP) that can capture some statistical properties of the short-term behaviour of precipitation (e.g. marginal mean, variance and skewness, lag-1 autocorrelation, and dry proportions).

1) In other words, one of the innovations of this methodology is the coupling between an autoregressive model and a Bartlett-Lewis one, where often in literature is either a coupling of autoregressive models or of pulse models. In the first case, generalized linear models can reproduce the variability only at coarse time scales ("larger than one month"), whereas in the second case the (Poisson) cluster models cannot capture the large scale behaviour but can reproduce the small scale of storm events (P4L89 to P4L96). However, some of the approaches mentioned later in the text (P5L101 to P6L125) seem to propose how to tackle the above scale issues. For example, Koutsoyiannis (2001; see also Koutsoyiannis et al., 2003) suggested coupling several stochastic models of different scales and thus, preserving some small and large scale characteristics. Does the Authors' proposed methodology (a) resolve some of the limitations of these models (P5L101 to P6L125), and if yes what are those, or (b) is it a new model that can equally reproduce what all these models also reproduce? I would suggest a discussion to this question to be added to the Abstract and at the end of the Introduction (P6L125), since it seems important in my opinion.

2) Also, another innovation mentioned by the Authors is the modification of the Bartlett-Lewis model in module 2 (Fig. 4, 5, sect. 3.2). More specific, a modification is proposed to the Bartlett-Lewis Poisson pulse model, where a dependency is now introduced between the storm Poisson events and thus, the proposed model can now better represent the short (and medium) term autocorrelation of precipitation in contrast to other Bartlett-Lewis models in literature (e.g. the ones mentioned in P2L42 to P3L53).

The Authors may find interesting to discuss some works by Lombardo et al. (2012; 2017) where they also use an innovative downscaling method that can generate fine scale precipitation by preserving some aspects of intermittency. Additionally, the Authors may be interested in discussing a recent work by Dimitriadis and Koutsoyiannis (2018) where it is shown how the above "scaling issue" can be dealt by directly generating the fine scale process from small to large scale and thus, explicitly preserving the large scale behaviour and some aspects of intermittency (through the preservation of joint-statistics; see also Appendix D where a comparison is made to the copula method). This explicit generation is achieved based on a Moving Average scheme and not an AutoRegressive (AR) one since as also is stated by the Authors the AR models cannot capture the small scale intermittent behaviour (P4L94: "Models based on autoregressive properties of

rainfall are typically good at reproducing the observed rainfall variability only for a limited range of scales…”).

3) Additionally, please consider a couple of comments on the module 2 methodology:

A) In Eq. (5) the Authors estimate the lag-1 discrete autocorrelation, i.e. $\hat{c}(1)$, through an estimator $\hat{V}(2)/\hat{V}(1)/2 - 1$ that corresponds to the true value of the lag-1 autocorrelation (i.e. $V(2)/V(1)/2 - 1$). However, the statistical bias (for a discussion please see Dimitriadis, 2017, sect. 2.4.5) is not seem to be taken under consideration. I believe the use of this estimator is not adequately (but rather empirically) justified and it is based on the assumption that:

In case of an AR(1) model (Eq. 2) and for large timeseries samples, then $E[\hat{\underline{C}}(1)] \approx C(1) = V(2)/V(1)/2 - 1$. This can be derived based on the following analysis, where the expectation of autocovariance $C(h)$ as a function of the cumulative variance $V(h)$.is:

$$E[\hat{\underline{C}}(h)] = \frac{1}{\zeta(h)}\left((n-h)c(h) + \frac{V(h)}{n} - V(n)/h - \frac{V(n-h)}{n}\right) \tag{1}$$

where $\zeta(h)$ is related to the estimator of the autocovariance and is usually taken as n or $n - 1$ or n-h (Dimitriadis and Koutsoyiannis, 2015, Table 2, Eq. 9). Note that where $V(k) = k^2\gamma(k)$ is the variance of the cumulative process vs. scale, or else called cumulative climacogram, and $\gamma(k)$ is the variance of the averaged process vs. scale, or else climacogram (Koutsoyiannis, 2016, and references therein).

Therefore, for $h = 1$:

$$E[\hat{\underline{C}}(1)] = \frac{1}{\zeta(1)}\left((n-1)C(1) + \frac{V(1)}{n} - V(n) - \frac{V(n-1)}{n}\right) \tag{2}$$

We see that $E[\hat{\underline{C}}(1)] \neq C(1) = \frac{V(2)}{2} - V(1) + V(0) = \frac{V(2)}{2} - V(1)$. The above expressions may be used to correct the deviation between a straight line ($y = x$) as shown in Fig (5a) but also to help the Authors express the variances $V(h)$ of different lags h with autocorrelation $c(h) = C(h)/V(1)$, as shown in Eq. (5) of the manuscript.

Furthermore to the justification, since the Authors have chosen an AR(1) model (Eq. 2), then $\gamma(h) = V(h)/h^2$ would behave like a white noise process in large scales (Dimitriadis and Koutsoyiannis, 2015, sect. 2.3) and thus, $V(h) \approx ah$ (where a constant). Therefore, for large samples, one may assume that $\frac{(n-1)C(1)}{\zeta(1)} \approx C(1)$, $\frac{V(n)}{\zeta(1)} \approx a$, $\frac{V(n-1)}{n\zeta(1)} \approx 0$ and also $\frac{V(1)}{n(n-1)} \approx 0$, and thus, it can be assumed that $E[\hat{\underline{C}}(1)] \approx C(1) + a = \frac{V(2)}{2} - V(1) + a$, and for the autocorrelation, $E[\hat{c}(1)] \approx \frac{V(2)}{2V(1)} - 1 + a'$. In this case, P14L238 (“we therefore estimate the autocorrelation lag-1 of hourly rainfalls using $\frac{\hat{V}(2)}{2\hat{V}(1)} - 1 + \varepsilon''$”) can be now also analytically justified.

Alternatively, in case that the Authors wish to somehow take bias into consideration when estimating the autocorrelation $c(1)$, they could suggest an estimator of lag-1 autocovariance that is not only based on the cumulative variance but also takes into account the sample's length n, then (for e.g., $\zeta(h) = n - 1$), we have that $E[\hat{\underline{C}}(1)] = C(1) + \frac{V(1) - nV(n) - V(n-1)}{n(n-1)}$, and thus for an AR(1) process an estimator for the lag-1 autocorrelation could be:

$$\hat{c}(1) = \frac{\hat{V}(2)}{2\hat{V}(1)} - 1 + \frac{1 - a(n^2 + n - 1)}{n(n-1)} = \tag{3}$$

B) In Eq. (6) and (7) of the manuscript the Authors use an empirical expression for the mean and variances of the cumulative process. Based on the above analysis they could justify this linear approximation by the fact that for an AR(1) model (Eq. 2, P10L189) the variance of the cumulative process is (Dimitriadis and Koutsoyiannis, 2015, Table 4, Eq. 18):

$$V(h) = 2q^2 V(1)\left(\frac{h}{q} + e^{-h/q} - 1\right) \tag{4}$$

where $q = -1/\ln(c(1))$ is the AR(1) parameter (see also Koutsoyiannis, 2016, Table 1, Eq. T1.2 and T1.3, and references therein).

Therefore, the link of cumulative variances for different lags h is not always be close to linear but rather depends on the q parameter. For large q, or else small $c(1)$, we have that $V(h) \approx 2q^2 V(1)(h/q - 1)$ and so, $V(h)/V(h') \approx \left(\frac{h-q}{h'-q}\right)$, which is not a linear expression.

Maybe the above configurations could explain the deviation from linearity ($y=x$) in Fig. 6

C) Finally, in Eq. (3) the Authors linearly connect the mean of the process to the standard deviation and to the dry proportions. This is a very good result that I would recommend the Authors to further highlight as one of the empirical results of this work (for example in the Introduction), since for a larger scale such as daily the link between mean and standard deviation seems to be a rather power-type expression (e.g. Sotiriadou et al., 2016, sect. 7).

Also, since this a link based on the marginal distribution of precipitation rather than on its dependence structure, a more analytical justification could be that all the distributions applied to the Bartlett Lewis model (i.e. exponential, gamma and Poison) have a linear combination between their mean and standard deviation. For example, for the gamma distribution (i.e. $f(x) = x^{k-1} e^{-\frac{x}{\theta}}/\Gamma(k)\theta^k$) we have that $\mu/\sigma = \sqrt{k}$, and similar results can be drawn for the other two distributions. Therefore, this could be another evidence that the proposed modified model MBLRP can describe well some properties of the fine scale precipitation.

Minor comments and suggestions

1) P1L27: "the observed rainfall record is oftentimes not long enough (Koutsoyiannis and Onof, 2001).".
   The Authors could also add some of the drawbacks of the limited timeseries with large length n, such as for example the statistical bias (as discussed on the above 2nd comment).

2) P2L41: "...so they are good at reproducing the first through the third order statistics of the observed rainfall..."
   What are "the first through third order statistics"? Do you mean first to third marginal statistics (i.e., mean, standard deviation and skewness)? Also, the title of the paper gives the impression that all the rainfall characteristics can be well reproduced. Maybe the title could be altered to "A Hybrid Stochastic Rainfall Model That Reproduces some important Rainfall Characteristics at Hourly through Yearly Time Scale". The preservation of the first three (or even four with kurtosis) statistics is very important and sometimes preserving more is unnecessary. For example in Dimitriadis and Koutsoyiannis (2018, sect. 3.1 and sect. 4) a discussion is made for the impracticality of estimating high-order moments in geophysical processes and, in all applications there, it is exhibited that beyond the first four moments there is a negligible increase in accuracy of the representation of the marginal distribution.

3) P3L57: "These model assumptions deprive the model of the ability to reproduce the long-term memory of rainfall that is often observed in reality (Marani, 2003)."
   Do the Authors mean "short or medium memory"? Since their improvement (module 2) deals with the fact that Poisson events are considered independent and thus, by introducing a dependency among the rainfall events, the model's short and medium term preservation is enhanced (see also P3L73: "...the Poisson cluster rainfall model because it can only reproduce short-term memory in the rainfall signal through its model structure..."). The long term behaviour is achieved by the SARIMA model that generates the large (monthly) scale precipitation.

4) P4L81: "Here, the MBLRP model used the parameter set that was calibrated to reproduce the observed rainfall mean, variance, lag-1 auto-covariance, and proportion of dry periods at sub-daily aggregation intervals (1, 2, 4, 8, and 16-hour)…".
   Why not adding the preservation of monthly skewness as shown in Fig. 9?

5) P4L85: "that the variability of the observed rainfall is systematically greater than that of the synthetic rainfall."
   P4L86: "In addition, the monthly extreme values shown as star marks are also underestimated by synthetic rainfall."
   P4L87: "This is, in particular, caused by the aforementioned limitations of the Poisson cluster rainfall models."
   Maybe this is also due to the bias effect as shown in the major comments above.

6) P4L95: "Models based on autoregressive properties of rainfall are typically good at reproducing the observed rainfall variability only for a limited range of scales...".
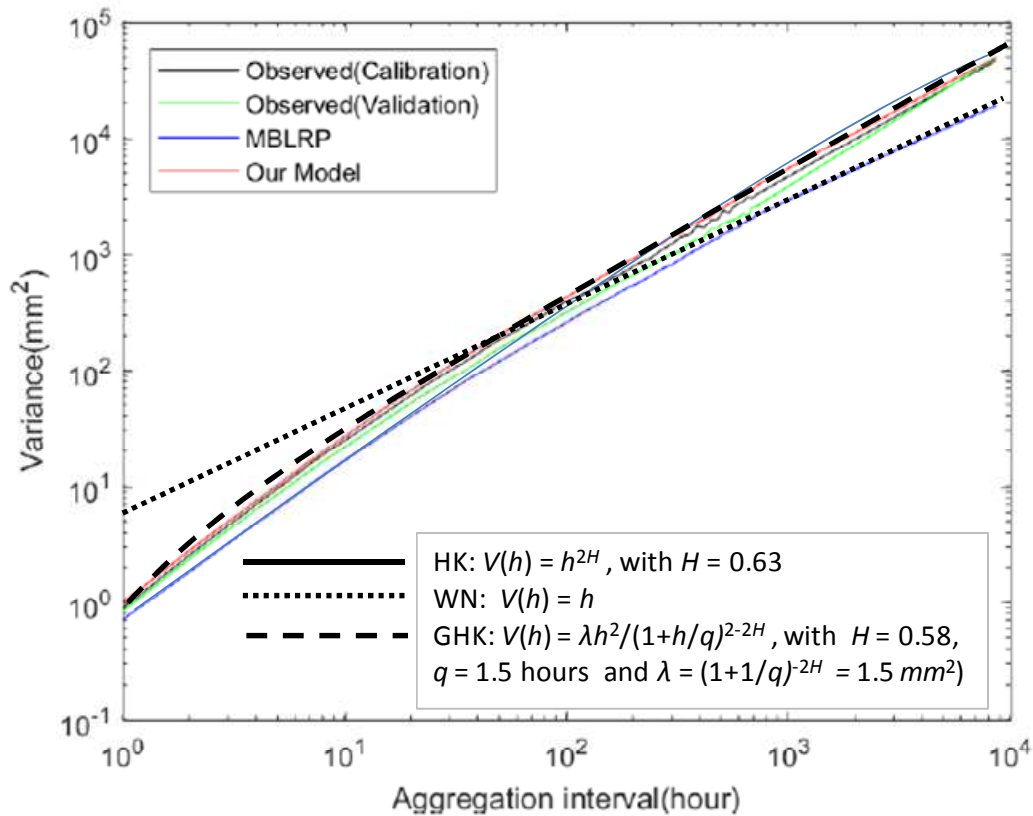
Do you mean that mixed Autoregressive (AR) and Moving-Average (MA) models (such as the SARIMA model used here) are required to reproduce long-term behaviour, and that solely the AR models cannot achieve this? In fact, a Sum of arbitrary many AR models (SAR) can also preserve the long-term behaviour as recommended by Mandelbrot (1971). Also, a SAR algorithm with the parameters analytically estimated is introduced in Koutsoyiannis (2010) with 3 AR(1) models, where the long-term (or else called Hurst-Kolmogorov –HK-) behaviour is preserved for 1000 scales, and in Dimitriadis and Koutsoyiannis (2015, sect. 3 in suppl. mat.) with arbitrarily many AR(1) models, where the HK behaviour is preserved for as many scales as needed.

7) P6L120: The Authors may find useful discussing other models (like SARIMA) that can reproduce the long-term behaviour (such as the SAR one mentioned above) as well as the monthly seasonality such as the Langousis and Koutsoyiannis (2006) and Efstratiadis, et al. (2014).

8) Figure 5: "ε_i is a random number drawn from the normal distribution" is missing from the Figure's legend at the lower left corner.

9) Please consider moving Figures 6, 12, 13, 15, 16 and 19 (or some parts of them) to an Appendix or a supplementary material since, in my opinion, they are quite large for the main text.

10) Also, I would recommend the Authors to add a Table with all the statistical characteristics of the 34 hourly stations (mean, stdev etc.) as well as the fitted parameters of the SARIMA and MBLRP models.

11) P16L275: "Here, it should be noted that a time step with rainfall less than 0.5 mm was considered dry when the proportion of non-rainy period was calculated because small rainfall values are known to distort the "true" proportion of non-rainy period exerting an adverse effect on calibration process (Kim et al, 2016, Cross et al., 2018)."
The value 0.5 mm may seem somehow arbitrary. Why not estimating the average of the lowest positive values (> 0) observed at the 34 timeseries and set this value as the dry threshold?

12) Some comments on Fig. 10, which seems very interesting and in my opinion should be further discussed since it highlights (additionally to the other results) the strength of the applied method in terms of other methods existing in literature. Specifically:
a) From this Fig. 10 one could estimate the Hurst parameter that is related to the significance of the long-term behaviour of the process. In fact, I made a rough estimation of the Hurst parameter (based on the log-log slope of the cumulative variance shown in the Figure below; for this method see also Koutsoyiannis, 2016), where $H \approx 0.6$. This is also consistent to what the Authors mention on Marani (2003) about the long-term behaviour in P3L58, and to the global analysis of Iliopoulou et al. (2016) and Tyralis et al. (2018).
Also, it can be shown at the Figure below that the MBLRP model can well preserve the short-term behaviour but not the long-term behaviour as the Authors mention, since the MBLRP model exhibits a white noise (WN) slope (as shown from the fitted dashed line).

b) There is an evident smooth behaviour of the cumulative variance (or else cumulative climacogram, please see above comments) vs. scale as an estimator of the long-term behaviour as compared to the autocovariance (or autocorrelation) vs. lag, where a larger variability of the sample statistics at large lags could prevent depicting this behaviour. This diversity is discussed and thoroughly analyzed in Dimitriadis and Koutsoyiannis (2015), where the power spectrum is also compared to the other two estimators of autocovariance and climacogram, and it is found that the variability of the latter is smaller in large lags and thus, enabling a more accurate estimation of the long-term behaviour and of the Hurst parameter.

c) Also, one could fit a more generalized process based on the Figure below and describe within a single expression of $V(h)$ how the cumulative climacogram is increasing from the finer to the larger scales h. In fact, a one-parameter power-type model (named HK, i.e., $V(h) = V(1)h^{2H}$) or a more generalized two-parameters one (named GHK, i.e. $V(h) = \lambda h^2/(1 + h/q)^{2-2H}$, where coefficient $\lambda = V(1)(1 + 1/q)^{2-2H}$) seem to work also very well. Note that if one is focused to finer and finer scales, eventually the dependence structure of the cumulative process will also have to drop down to zero in a quicker manner than power-type (or in other words, it will have to stabilize at some point in terms of the dependence structure of the averaged process). It is interesting to mention that a similar HK and GHK behaviour has been detected to a daily precipitation timeseries (Dimitriadis and Koutsoyiannis, 2018, sect. 4.2, Fig. 4). The Authors may well use these results to further highlight their work in the sense that their proposed methodology seem to can very well preserve the expected dependence structure for a large range of scales and thus, it is equally strong to other methods that are based on the expected value of the cumulative climacogram.

The figure legend contains:

- Observed(Calibration)
- Observed(Validation)
- MBLRP
- Our Model

HK: $V(h) = h^{2H}$, with $H = 0.63$

WN: $V(h) = h$

GHK: $V(h) = \lambda h^2/(1+h/q)^{2-2H}$, with $H = 0.58$, $q = 1.5$ hours and $\lambda = (1+1/q)^{-2H} = 1.5\ mm^2$)

Axis labels: Variance(mm²) versus Aggregation interval(hour)

13) P29L443: "While significant variability is observed for all six parameters, the parameter μ, which represents the average rain cell intensity, showed the greatest variability, ranging over two orders of magnitudes."

This could be also justified by the fact that in long-term processes (such as the one examined in this paper) there seems to be a larger variability of the mean of the process rather than of some higher-order moments. The Authors may be interested in the analysis of Dimitriadis and Koutsoyiannis (2018), where in sect. 3.1, they present a benchmark case with a N(0,1) distribution and in Fig. 1 they show how the variability of the mean (in terms of its standard deviation) is changing as a function of the Hurst parameter, and in Fig. 2 how the variability of the mean is larger than that of the first four moments for a large range of scales.

References

Dimitriadis, P., Hurst-Kolmogorov dynamics in hydrometeorological processes and in the microscale of turbulence, PhD thesis, National Technical University of Athens, 2017. [for online version see https://www.itia.ntua.gr/en/docinfo/1767/]

Dimitriadis, D., and D. Koutsoyiannis, Climacogram versus autocovariance and power spectrum in stochastic modelling for Markovian and Hurst–Kolmogorov processes, Stochastic Environmental Research & Risk Assessment, 29 (6), 1649–1669, doi:10.1007/s00477-015-1023-7, 2015. [Supplementary Material also in https://www.itia.ntua.gr/en/docinfo/1514/]

Dimitriadis, P., and D. Koutsoyiannis, Stochastic synthesis approximating any process dependence and distribution, Stochastic Environmental Research & Risk Assessment, doi:10.1007/s00477-018-1540-2, 2018.

Efstratiadis, A., Y. Dialynas, S. Kozanis, and D. Koutsoyiannis, A multivariate stochastic model for the generation of synthetic time series at multiple time scales reproducing long-term persistence, Environmental Modelling and Software, 62, 139–152, doi:10.1016/j.envsoft.2014.08.017, 2014.

Iliopoulou, T., S.M. Papalexiou, Y. Markonis, and D. Koutsoyiannis, Revisiting long-range dependence in annual precipitation, Journal of Hydrology, 556, 891–900, doi:10.1016/j.jhydrol.2016.04.015, 2016.

Koutsoyiannis, D., The Hurst phenomenon and fractional Gaussian noise made easy, Hydrological Sciences Journal, 47 (4), 573–595, doi:10.1080/02626660209492961, 2002.

Koutsoyiannis, D., Generic and parsimonious stochastic modelling for hydrology and beyond, Hydrological Sciences Journal, 61 (2), 225–244, doi:10.1080/02626667.2015.1016950, 2016.

Koutsoyiannis, D., C. Onof, and H. S. Wheater, Multivariate rainfall disaggregation at a fine timescale, Water Resources Research, 39 (7), 1173, doi:10.1029/2002WR001600, 2003.

Langousis, A., and D. Koutsoyiannis, A stochastic methodology for generation of seasonal time series reproducing overyear scaling behaviour, Journal of Hydrology, 322, 138–154, 2006.

Lombardo, F., E. Volpi, and D. Koutsoyiannis, Rainfall downscaling in time: Theoretical and empirical comparison between multifractal and Hurst-Kolmogorov discrete random cascades, Hydrological Sciences Journal, 57 (6), 1052–1066, 2012.

Lombardo, F., E. Volpi, D. Koutsoyiannis, and F. Serinaldi, A theoretically consistent stochastic cascade for temporal disaggregation of intermittent rainfall, Water Resources Research, 53 (6), 4586–4605, doi:10.1002/2017WR020529, 2017.

Mandelbrot, B. B., A fast fractional Gaussian noise generator, Water Resources Research, 7(3), 543–553, 1971.

Sotiriadou, A., A. Petsiou, E. Feloni, P. Kastis, T. Iliopoulou, Y. Markonis, H. Tyralis, P. Dimitriadis, and D. Koutsoyiannis, Stochastic investigation of precipitation process for climatic variability identification, European Geosciences Union General Assembly, Vol. 18, Vienna, EGU2016-15137-5, doi:10.13140/RG.2.2.28955.46881, European Geosciences Union, 2016. [for a printed version see https://www.itia.ntua.gr/en/docinfo/1658/]

Tyralis, H., P. Dimitriadis, D. Koutsoyiannis, P.E. O'Connell, K. Tzouka, and T. Iliopoulou, On the long-range dependence properties of annual precipitation using a global network of instrumental measurements, Advances in Water Resources, doi:10.1016/j.advwatres, 2017.