Interactive comment on "Stochastic generation of multi-site daily precipitation for the assessment of extreme floods in Switzerland" *by* Guillaume Evin et al.

Guillaume Evin et al.

guillaume.evin@irstea.fr

Response to Interactive comment by Anonymous Referee #2

The authors propose extensions of a classical multisite daily rainfall generator initially proposed by Wilks in 1998. The framework of Wilks model is flexible enough to allow many adaptations, and the authors of this paper propose - to add more structure in the dynamics of the model by considering higher order Markov model for the occurrence process and an autoregressive component for the amounts - to use a hybrid distribution for the marginal distribution to deal with heavy tail distributions - to use a Student

copula for the spatial structure to catch upper tail dependence. I believe that all these extensions make sense and are interesting to try.

We thank the referee for this review and for these comments. Most of these suggestions will be incorporated in the modified manuscript.

1 General comments

1.1. Many extensions of the Wilks model have already been proposed in the literature. I think that a review of this literature must be included in the paper and that the authors should explain why the extension that they propose is original and useful with respect to this literature.
We agree that the differences between GWEX and the existing extensions of the Wilks models should be presented in the introduction. This discussion will be included in the revised version of the manuscript.

1.2. In my opinion, one weakness of the paper is that the model is formulated as a simulation tool rather than as a proper statistical model. It is also the case for the original Wilks model, but it has then been reformulated by other authors as a statistical model, see e.g. Thompson et al. (2007). I think that the paper would be easier to read for statisticians like me if a similar formalization was done in the paper. In particular, the various assumptions on the occurrence/amount processes should be written precisely using formulas and the definition of the model should be separated from the discussion on parameter estimation and simulation.
We thank the reviewer for this excellent suggestion. A more formal mathematical formulation of GWEX could certainly improve the presentation. We also agree that a specific section should be devoted to parameter estimation and model simulation. This will be done in the revised manuscript.

1.3. I believe that the validation part must also be improved. First, some usual validation criteria for rainfall generators, such as diagnostics based on the marginal distribution (e.g. qqplot) and the second order structure of the process (autocorrelation and crosscorrelation functions) are not shown and it makes it difficult to see the benefit of using a hybrid distribution and the autoregressive component. Also the chosen validation criteria does not permit to see the interest of using a student Copula (does it really improve the modeling of extremal dependence?).

These remarks have also been made by the referee #1 (comments #1.8, #2.5 and #2.6). QQ-plots will be provided to assess (visually) the quality of the fitting for the marginal distributions. Additional figures will also be provided to assess the performances concerning the autocorrelations at some stations and the reproduction of cross-correlations. Finally, an additional model version, a modified Wilks version with the EGPD, will be added to the current models, which will enable the assessment of the impact of the Student copula (versus a Gaussian spatial structure).

1.4. Finally, I find the simulation results generally disappointing. If I understand correctly the categorization, we should obtain about 90% of good if the model was able to reproduce the statistics of the observed rainfall? Is it satisfactory to obtain percentage around 50%?

Yes, we should obtain about 90% of good if the model is able to reproduce the observed statistics, and very few 'poor' cases. As indicated in the paper, our primary criteria to judge the overall performance of a model is the number of metrics for which 'poor' performances are obtained. We agree that these percentages are subjective (why 90%? Is 50% of good cases good enough?) but not more subjective, in our opinion, that the visual inspection of a QQ-plot. Furthermore, the purpose of the CASE framework, as presented in Bennett et al. (2017), is to enable a more systematic comparison of stochastic models. Our study also tries to promote this approach. A more systematic comparison of the models, which includes a consistent way to

compute the performance metrics, is important in order to obtain a fair assessment of the strengths/weaknesses of the different models. For this reason, this study applies the classification proposed by Bennett et al. (2017), without modifying the classification.

2 Specific comments

2.1. Keywords are missing?
In HESS, to the extent of our knowledge, keywords do not appear in the manuscript.

2.2. End of Page 1/top of page 2. I am not really satisfied by the proposed classification. For example weather type models are often used as multisite rainfall generators (without conditioning to large scale information). Also it would be useful to cite the review papers on rainfall generators here.

We agree that the terminology 'Multi-site models' is too vague here. A similar comment has been done by the referee #1 (see its comment #2.3.). We propose to replace 'multi-site models' at line 12 by 'statistical multi-site models'. Weather type models that are not conditioned by large scale information (for example using 'dry' and 'wet' states that are inferred) could belong to this class of models. We also agree that review papers could be included in the introduction and this will be done in the revised manuscript.

2.3. Section 2.1. The authors go directly from a Markov chain of order p=1 to a Markov chain of order p=4. I would expect that the best value of p is somewhere between these two values. The authors could try to find the optimal value of p, using for example standard model selection criteria.

We thank the reviewer for this suggestion. It is true that an optimal value might be found if there was an easy selection criteria. As this point is not central in our study, a direct comparison of Markov chains of order $p = 1$ and $p = 4$ is deemed sufficient.

2.4. Equation (5). I am surprised that the authors use a diagonal matrix for A. I would expect that it is useful to add some spatial structure here?

Initial version of GWEX were applying a full covariance matrix for A. However, it seems that large covariance matrices are often very close to a non positive definite. This is not really problematic during the estimation step, but leads to very unstable results during the simulation step. As applying a diagonal matrix for A does not degrade the performances of GWEX, this solution was retained.

2.5. Section 2.3 and 3.3 should be merged.

We thank the reviewer for this suggestion. Following comment #1.2., these sections will be re-organized with specific sections devoted to the estimation and the simulation steps. Sections 2.3 and 3.3 will thus be merged in the revised manuscript.

2.6. Section 3. Why is it called "Application"? I do not see any application here.

Following previous comments (comments #1.2. and 2.5.), the different parts of section 3 will be moved to other sections. Section 3.1 'Split-sampling procedure' is related to the evaluation framework. Section 3.2 'Regionalization of the ξ parameter' is related to the estimation of the parameters. Section 3.3 'Generation of scenarios' will be merged to section 2.3 (see previous comment).

References

Bennett, B., Thyer, M., Leonard, M., Lambert, M., and Bates, B. (2017). A comprehensive and systematic evaluation framework for a parsimonious daily rainfall field model. *Journal of Hydrology*. https://doi.org/10.1016/j.jhydrol.2016.12.043, In Press.