Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2017-226-AC1, 2017

Hydrology and Earth System Sciences
Discussions

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

Received and published: 14 June 2017

Response to Interactive comment by Anonymous Referee #1

We thank the referee for this thorough review and for the numerous constructive suggestions. We agree that the general presentation can be improved and these suggestions will be incorporated in the modified manuscript.

1 General comments

1.1. The title of the paper is a bit misleading. The three models may be used for the spatial assessment of floods and hydrological modelling is mentioned not only in the title, but also throughout the manuscript. However, the precipitation models are not applied in an impact assessment in this study and for this reason in my eyes the title should solely contain the comparison of three precipitation models. It is a bit irritating that the authors refer to the importance of several aspects of the precipitation model performance whose importance is not really demonstrated.
We agree that the title can be misleading and could be replaced by 'Stochastic generation of multi-site daily precipitation focusing on extreme events'. We think that it is important to indicate the emphasis on the reproduction of very large precipitation events, in terms of intensity, duration, and spatial extent.

1.2. The names of the new precipitation models are a bit misleading. First, "1D" and "3D" give the impression of any type of one- and three-dimensional simulation methodology. However, they represent days ("D"). I would rename the models into something more suitable.
This is a good suggestion and the names will be replaced by:

1. *Wilks: the current 'Wilks' model,*
2. *Wilks_EGPD: A modified Wilks version, with the EGPD and a Markov chain of order 4, as suggested by the referee (see comment #1.6),*
3. *GWEX: the current GWEX-1D,*
4. *GWEX_Disag: the current GWEX-3D. It would indicate more clearly the disaggregation step which follows up the simulation at a 3-day scale.*

1.3. As far as I understand from the paper, the new GWEX models are actually "Wilks models" but with a new method to simulate the precipitation amounts (using temporally and spatially correlated random numbers from an autoregressive process and using a Student copula for the spatial component). I think this should be stated as such in the paper as the manuscript presents the new models more as a revolution rather than an evolution. So one of the first sentences could be that the paper deals with two modifications of the Wilks approach.
We agree. The fact that GWEX are evolutions of the Wilks model must be clearly stated. In fact, it is already indicated at p.2/l.25 and p.4/l.10 and throughout the presentation of the models. As suggested by the referee, we will also indicate this point directly in the abstract. However, it must also be underlined that GWEX is a significant evolution of the Wilks model. First, as underlined by the referee, the methodology applied to simulate the precipitation amounts is considerably modified. We consider different temporal and spatial dependences, and we also discuss the choice of the marginal distribution in details, which is currently overlooked in the literature of precipitation stochastic models. Second, GWEX-3D (which will be named GWEX_Disag) combines simulations at a 3-day scale and a disaggregation approach, which represents a further step in the complexity of the model. In our opinion, GWEX cannot only be considered as a slight modification/evolution of the Wilks model.

1.4. The motivation behind the study and for the new model developments is the impact assessment. However, without the same the reader will not be able to really understand the sensitivity of certain statistics in regard to the assessment of extreme floods. I think the importance of some of the statistical metrics should be explained in more detail referring to the area of their application, and proof must be given of their relevance. Other literature in such a study (complex mountain region) is not very convincing to me.
We thank the referee fo this suggestion and additional details regarding the importance of the statistical metrics will be provided in the modified version. In particular,

Froidevaux (2014) analyze meteorological events triggering floods in Switzerland. These studies have been very briefly mentioned at the beginning of p.12 and these results must be discussed in more details, as will be done in the revised manuscript. However providing a proof of the relevance seems complicated without the hydrological application (which is clearly beyond the scope of this paper, as discussed in comment #2.14.). If the referee has more specific metrics that could be presented, we would be glad to include them in our study.

1.5. The abstract is incomplete and must be much more detailed and specific. What is an "event"? What is "large"? What are "recent advances"? The abstract should mention the Wilks model, the two new models (maybe also a short sentence how they work) and the basic outcomes of the study.
We thank the referee for this constructive suggestion. Additional details will be added to the abstract.

1.6. The Wilks model could likewise be applied with E-GPD distributions for precipitation intensities and Markov chains of the order 4. That is, revealed weaknesses of the Wilks model can easily be addressed. I recommend adapting the Wilks approach for a more objective comparison. The original Wilks approach is not a given and was just one application for a specific dataset in the US and in my eyes it should always be revised for other study areas and climates.
We thank the referee for this suggestion. An additional version of the Wilks model, with the EGPD instead of a mixture of exponential distributions for the marginal distributions, and a Markov chain of order 4 instead of order 1, will be presented in the modified manuscript.

1.7. For flood modelling, the lagged cross correlations (see Wilks 1998, page 183) can be very important as they represent the progression of weather systems across the study area. Especially at larger scales the progression of weather events

may be important. I strongly recommend plotting these statistics for all three models.
We appreciate this judicious suggestion and lagged cross correlation will be added to the set of statistics presented in the paper.

1.8. The autocorrelation of precipitation is addressed by MAR(1) models in the GWEX models. I would recommend plots for the autocorrelation of the precipitation intensities for some sites to see potential differences in their performance.
We thank the referee for this suggestion. Plots for the autocorrelation of the precipitation amounts will be shown for some of the stations.

2 Specific comments

2.1. Line 8. I think there is a language issue.
This sentence will be reformulated in the revised version.

2.2. Line 10. Not only conceptual models. There are more recent studies for coupling WGs with impact models.
Thanks for this remark. References about these recent studies can be included in the revised version.

2.3. Page 1 bottom/ Page 2 top: In my eyes the classification is not fully correct. All these models are multi-site models. Also resampling methods are multi-site models. I recommend a more suitable classification even though I admit that the variety of the existing multi-site models makes a clear classification more and more difficult (also the authors combined parametric and non-parametric techniques).
Here, multi-site models refer to models that target the reproduction of statistics at

specific sites. They can be opposed to random fields that mainly intend to reproduce spatial properties (e.g. the variogram). We agree that resampling methods are also referred as multi-site models in the literature (see, e.g. Mehrotra et al., 2006). Hence, we propose to replace 'multi-site models' at line 12 by 'statistical multi-site models' in order to clarify the distinction between resampling methods and the set of models cited in the paragraph 'multi-site models' which applies various statistical structure (copulas, truncated Gaussian distributions, V-copula transform, etc.).

2.4. Page 4, Line 8. Are Thiessen polygons suitable for such a complex mountainous study region?
The computation of areal precipitation values is a difficult task considering the spatial and temporal variability of precipitation events, the complex topography of the sudy area, and the limited number of pluviographs. In Switzerland, Schäppi (2013) shows that the topography impacts rainfall amounts differently according to the type of meteorological event. In a preliminary study, the impact of different interpolation methods (inverse distance, ordinary kriging, kriging with external drift, Thiessen polygons) and different sets of stations (399, 211, 129, 47 and 22 stations) on extreme areal precipitation amounts has been analyzed. The main conclusion was that the number of stations was a much more important factor than the interpolation method. This was the main motivation for the application of the stochastic models to a high number (105) of stations. Furthermore, it is important to notice that applying more complex interpolation methods (e.g. kriging methods) increase significantly the computational cost, which can be prohibitive for the production of long meteorological scenarios.

2.5. Page 7 “Marginal distributions”. Can any proof be given that the more complex fitting of a combined distribution is really significantly better for the simulation of the extremes in this region? Also here, the most prominent argument is other literature.
QQ-plots will be provided in the revised version in order to assess the quality of the

fitting of these marginal distributions. However, it is very important to note that local applications give limited proof regarding the performance of a distribution for the fitting of extreme values. As indicated in "Papalexiou, S. M. and Koutsoyiannis, D. (2013) Battle of extreme value distributions: A global survey on extreme daily rainfall, Water Resources Research, 49, 187201", most studies of extreme rainfall are inconclusive because they are too specific to particular areas or stations. The main explanation for these failures is that fitting and inferring the distribution tails is subject to high uncertainties in the estimation of the parameters, even for long time series (this point is also discussed and illustrated in Evin et al., 2016). The references given in the paper (Papalexiou and Koutsoyiannis, 2013; Serinaldi and Kilsby, 2014) are conclusive precisely because they are the result of a very large number of applications, and give strong arguments in favor of the application of heavy-tailed distributions. Furthermore, Figures 2 and 3 prove that low tail-distributions (like a mixture of exponentials) would lead to an under-estimation of extreme precipitations in some regions (regions where ξ is different from 0, in green, yellow and red).

2.6. Page 9, top of the page. If the Gaussian copula is not suitable for simulating spatially dependent extremes but the Student copula is, this could be demonstrated. I am thinking of readers who want to build the code but are not experts in copulas and want to understand the significance.
With an additional version of the Wilks model, with the EGPD instead of a mixture of exponential distributions for the marginal distributions (see comment #1.6.), we will be able to assess the difference between a Gaussian copula and a Student copula for the reproduction of daily precipitation extremes.

2.7. Page 9 bottom. Why are Markov chains of the order 4 used? Have there been statistical tests or sensitivity studies to underline this decision? Later on, some remarks are given on the simulation of short dry spells, but I think this should be addressed in a more structured way.

At p.5/l.10, we indicate that Srikanthan and Pegram (2009) apply a 4-order Markov chain and show that it improves the reproduction of dry/wet period lengths. This point will be reminded at p.9.

2.8. Page 11, Table 11 (and figures). Red and green are not suitable for figures, please change the colours as some people cannot read them otherwise (https://www.nature.com/nature/journal/v510/n7505/full/510340e.html).
We thank the referee for this comment.These colors will be modified to be more suitable for most color-blind people, (https:// www.nature.com/ nmeth/ journal/ v8/ n6/ full/ nmeth.1618.html). As we understand this issue, it seems that types of green (bluish green) and red (vermillion) are more adapted to color-blind individuals.

2.9. Page 12, Line 28. I guess it is very difficult to say if an extreme precipitation amount is unrealistic or not as long as they are physically possible?
It is true that an extreme precipitation amount cannot be considered as unrealistic if the amount is physically possible. However, it is difficult to define what amount can be considered as impossible. Here, we indicate that large ξ parameters (> 0.25) lead to extremely heavy-tail distributions. In practice, very large daily amounts (e.g. > 1000 mm) will often be obtained when long runs are performed (e.g. 1000 years). The reference given in the paper (Serinaldi and Kilsby, 2014) indicates that these large ξ parameters are often spurious as they usually are the results of high parameter uncertainties. Furthermore, for $\xi > 1/3$, the Generalized Pareto distribution has an infinite variance, which is not a desirable properties. Note that this constraint has a limited impact in our case since we always obtain $\xi < 0.25$ in our study area (see dark red areas in Fig. 2 and 3). However, we think that it is important to inform potential users of GWEX (and more generally anyone who applies a GPD to extreme precipitation amounts) that they need to be very careful if they obtain very high ξ estimates.

2.10. Page 16, Line 18-20. If the order of the Markov chain is the issue for

short dry spells, this can be easily adapted by using the same order in the original Wilks approach. What was the argument for using the first order Markov chains in the Wilks model? (see comment above)
See comment #2.7.

2.11. Page 21, Line 8-9. Please explain the seasonal differences with explicit reference to the study area and its climatology for better understanding.
More explanations about these seasonal differences, specific to Switzerland, wil be provided in the revised version.

2.12. Page 22, section 4.4. and figure 10. To me, the performance looks fair for all three models. The main difference is the simulation of higher extremes with the GEWX models. The authors mention the difference but it needs further discussion. Also, how can we know that the extremes of one method are more realistic than from another? While we know little about the validity of the simulated extremes, they may have a big impact on simulated floods, especially in small catchments (but as mentioned before, this is not examined in the paper).
We agree with the referee, the performance looks fair for all three models. However, this figure only points out differences of behavior between the three models. As mentioned in the comment #2.5., these illustrative examples cannot be used to test the performances of the different models for the simulations of extreme daily precipitation amounts. The only way to perform such a validation is to apply metrics on a large set of applications (here, for example, at all the stations), which is done at Figures 14 and 15. This remark will be added to the revised version of the manuscript.

2.13. Page 26 Line 10-13. It is not surprising that the non-parametric disaggregation leads to a better performance. I understand its strengths but it may likewise be a limiting factor in generating extremes.
In our opinion, GWEX-3D (which will be named GWEX_Disag) represents the best

combination between a purely statistical approach and a nonparametric approach. The results presented in this study confirm this statement. As we know little about very extreme precipitations, it is, in our opinion, impossible to know if it is a limiting factor or not.

2.14. Page 29, first line 2-9. As already mentioned, I see the motivation behind the study (and it is generally a good one). But without any proof that the differences in the performance of the three precipitation models really have a significant impact on the simulation results of hydrological extremes (also considering all the uncertainties in hydrological models), the significance of the research outcomes remain questionable.
The rational behind the study is the following: we want to develop a stochastic model for precipitation which preserve the most critical properties of precipitation at different spatial and temporal scales. Froidevaux (2014) shows that extreme 3-day precipitation amounts often trigger important floods. In this context, it seems natural to target the reproduction of 3-day precipitation amounts. However, we agree that hydrological applications would validate the importance of such properties. Actually, hydrological applications are currently undertaken by the University of Zürich. A conceptual hydrological model (HBV) is applied to 87 sub-basins partitioning the whole study area, using precipitation scenarios produced by GWEX as inputs. Numerous technical issues still need to be resolved. Some basins are ungauged, or with very short streamflow series. The hydrological system of the Aare-Rhine river needs to be treated as a whole since floods at larger spatial scales need also to be investigated. Rating curves have very high uncertainties in some basins and need to be re-evaluated. All these aspects have to be treated in details. It is important to note that the hydrological study (as well as our study) is particularly challenging considering the large spatial extent of the Aare river catchment. These studies stand out from similar studies which are usually limited to e few precipitation stations and one "small" catchment (see, e.g., Keller et al., 2015, recently published in HESS, with an application to 8 precipitation stations located in a catchment with a size of 1700 km^2, to be compared with our

study area of 17,000 km^2). Clearly, the hydrological application should be presented in future publications, considering the complexity of this work and the amount of results. However, we agree that the hydrological application would emphasize the significance of this study, and this point must be discussed in the manuscript.

2.15. Page 29, Line 21-22. Please explain why, see comments above.
See Comment # 3.11.

2.16. Page 29, Line 27-28. The issue of larger spatial scales could be addressed by running more analyses at smaller scales. So the key motivation of the study is probably to examine large flood events and their spatial dependences? If so, this should be better explained. But again, without really simulating the floods throughout different scales the arguments for a particular precipitation model choice is questionable.
The key motivation is to develop a stochastic model for precipitation which preserve the most critical properties of precipitation at different spatial and temporal scales, and especially for extreme precipitation amounts. This motivation is rather general and not specific to some characteristics of flood events (e.g. their spatial dependence).

2.17. Page 30. Is the underestimation of the inter-annual variability such a big issue in Switzerland and for flood modelling? I would assume it is more an issue in more arid regions and for example agricultural studies? Some more remarks on the relevance in Switzerland and floods in general would be useful.
Thanks for this remark. We agree that more comments about the relevance of these metrics should be provided, which need to be more specific to applications in Switzerland

3 Summary of review

3.1. The abstract needs revision and must be more detailed (see general comments).
See comment # 1.5.

3.2. The introduction is not very well structured. The arguments for the construction of the new precipitation methodologies are mainly based on other literature and reasoning. The context of the paper should (i) either be revised (comparison of precipitation models) or (ii) proof must be given of the advantages using the new models by really coupling them with a hydrological model and examining the estimated flood events in the study region. I think it is the key weak point of the paper: reference is given to an application, which is not really done. Also, the title and abstract are a bit misleading and the reader may expect a flood modelling study and thus more than what has been presented.
We agree that the introduction can be misleading and it will be modified in order to clearly indicate that this study mainly aims at comparing precipitation models, the hydrological context being the motivation for the thorough assessment of areal precipitation extremes.

3.2. For the three different precipitation models, I would recommend a flow chart with the Wilks model as the central component and then the adaptations that have been done. This makes it easier for the reader to understand all models and what has been changed.
This is an excellent suggestion and a flow chart will be added to the revised version of the manuscript.

3.3. Although the level of English is very good, some (minor) mistakes can be found in the manuscript and a native speaker should probably have a final look before resubmission.

The revised version will be proofread by a native speaker.

References

Evin, G., Blanchet, J., Paquet, E., Garavaglia, F., and Penot, D. (2016). A regional model for extreme rainfall based on weather patterns subsampling. *Journal of Hydrology*, 541, Part B:1185–1198.

Froidevaux, P. (2014). *Meteorological characterisation of floods in Switzerland*. PhD thesis, Geographisches Institut, University of Bern.

Keller, D. E., Fischer, A. M., Frei, C., Liniger, M. A., Appenzeller, C., and Knutti, R. (2015). Implementation and validation of a Wilks-type multi-site daily precipitation generator over a typical Alpine river catchment. *Hydrology and Earth System Sciences*, 19(5):2163–2177.

Mehrotra, R., Srikanthan, R., and Sharma, A. (2006). A comparison of three stochastic multi-site precipitation occurrence generators. *Journal of Hydrology*, 331(1–2):280–292.

Papalexiou, S. M. and Koutsoyiannis, D. (2013). Battle of extreme value distributions: A global survey on extreme daily rainfall. *Water Resources Research*, 49(1):187–201.

Schäppi, B. (2013). *Measurement and analysis of rainfall gradients along a hillslope transect in the Swiss Alps*. PhD thesis, ETH Zürich, Zürich, Switzerland.

Serinaldi, F. and Kilsby, C. G. (2014). Rainfall extremes: Toward reconciliation after the battle of distributions. *Water Resources Research*, 50(1):336–352.

Srikanthan, R. and Pegram, G. G. S. (2009). A nested multisite daily rainfall stochastic generation model. *Journal of Hydrology*, 371(1–4):142–153.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-226, 2017.