Climate change and uncertainty in high-resolution rainfall extremes

by B. Kianfar, S. Fatichi, A. Paschalis, M. Maurer and P. Molnar

Reviewer's comments

Summary

A stochastic weather simulator is used to assess the impact of climate change on rainfall extremes at sub-daily time scales. Results show that distributions of annual rainfall maxima for current and future climates largely overlap due to high internal stochasticity of rainfall at small-scales. The conclusions are that ... well, we can't say much about small-scale extremes in future climate within the proposed simulation framework. Whether this is because of model deficits or a general fact independently of the chosen simulation scheme remains unclear.

Personal Evaluation

The paper builds on an increasing body of literature devoted to the effects of climate change on precipitation and hydrological extremes. The topic is highly relevant and in line with previous research published in HESS. As with any emerging field, there are a number of challenges unique to this line of research, including the need to develop adequate statistical tools and conceptual frameworks. The authors propose one particular approach that builds on a previously developed simulation-downscaling scheme. The paper is well structured, brief and relatively easy to read. One notable exception is the description of the stochastic weather generator and multiplicative random cascade which, I assume, some readers may not be familiar with and therefore merits to be described in more detail.

Building mathematical models from rainfall observations is challenging, especially when there is a ton of noise (i.e., measurement uncertainty and internal stochasticity) or when the studied processes are characterized by highly scale-dependent and skewed distributions. There are many statistical models out there that can fit the data at hand. But how do you pick the right one? The most effective at making accurate predictions for future climate conditions? A good scientist knows not to approach data with a preferred result or preconceived notion in mind, because that can introduce bias and make people ignore information that goes against it. I understand the authors enthusiasm for their simulation-downscaling method and must admit it reproduces some of the observed features pretty well. However, some of the figures suggest (at least to me) that perhaps, the simulator does not perform as well as the authors think it does. In particular, there seems to be some issues with the autocorrelation, skewness and other higher-order moments. These are either overlooked or downplayed during the evaluation (see below for more details). I therefore think it is crucial the authors revise their paper and provide more solid tangible evidence that their NSRP-MRC simulation-downscaling scheme indeed reproduces realistic rainfall distributions and temporal structures at daily and sub-daily scales. Most importantly, I wish the authors were a little bit more self-critical of their approach and provided some additional discussion about the most important assumptions underlying it (random cascade weights, FoCs derived from imperfect model outputs, change of support, independence, stationarity etc...). Some of these issues are already mentioned in the text. But there is clearly more to say and the paper could benefit a lot from a more in-depth analysis and discussion.

My recommendation: **Major review**

Structure

1. **Number of MRC models:** I do not think you need two MRC models to make your point. MRCA is not used in the second part anyway so why go through all the trouble of describing it? Perhaps you could drop it and keep only MRCB. This would give you more space to describe the parametrization, advantages and shortcomings of NSRP and MRC models.

2. **Unnecessary Technicalities:** You can shorten the text by leaving out unnecessary details about macro-canonical cascades, as these are not used in this study.

Major Comments

3. **Changes in precipitation extremes at sub-daily scales:** One important point that is hardly made throughout the paper is that rainfall extremes at sub-daily scales are likely produced by different mechanisms than at larger scales. Small-scale extremes relate to highly localized systems of convective rainfall, which have very different dynamics and properties than the larger weather systems responsible for extremes at daily and weekly scales (i.e., stationary fronts, atmospheric blocking etc...). For example, one study by Berg et al. (2013) suggests convective precipitation may respond more sensitively to temperature increases than stratiform precipitation, and therefore play an increasingly important role in the future, dominating the events of extreme precipitation at smaller scales. Another study by Wasko et al. (2016) suggests heavy precipitation in a warming climate might become spatially more concentrated. It is no secret that GCM-RCM models and stochastic weather generators still struggle to reproduce realistic small-scale localized convective rainfall patterns. This is obviously a problem because it might undermine your ability to correctly predict changes in small-scale extremes. Please elaborate.

4. **Diurnal cycle:** In most regions of the world (including Switzerland), one can observe a strong link between small-scale rainfall extremes and the diurnal cycle. Your study does not even mention the diurnal cycle, nor is it clear how it is taken into account during the downscaling process. Please provide more details.

5. **Cascade weights:** As explained in the paper, empirical cascade weights are often scale and intensity dependent. In reality however, weights probably also depend on other factors such as the type of precipitation, temperature, wind speed and orography. The point I want to make is that there are different physical processes responsible for redistributing rainfall amounts in space and time across scales. One important assumption you make in this paper is that the rules of this redistribution remain unchanged between current and future climate conditions. But how sure are you exactly about this? For example, if cascade weights turned out to be temperature dependent, wouldn't that invalidate your approach and conclusions? Please provide detailed discussion of this issue and specify what reasons you have to believe the same cascade rules can be applied to downscale present and future rainfall events. MeteoSwiss stations provide co-located rainfall and temperature data that could be used to (partially) address this issue.

6. **Model validation:** The validation of the stochastic simulation-downscaling scheme absolutely needs to be improved. Please provide hard evidence (e.g., goodness-of-fit metrics) and detailed exploratory data analyses (e.g., histograms, box-plots, etc...) showing the simulator indeed reproduces key rainfall characteristics a sub-daily scales such as autocorrelation, skewness, seasonal and diurnal cycles. Your conclusions completely depend on this! For now, arguments are mostly subjective and qualitative and frankly, do not convince me.

7. For example, I strongly disagree with the authors' claim that *"the main statistics of precipitation are reproduced very well"*. On the contrary, when you look at Fig 3-4, it is quite clear that both MRCA and MRCB failed to correctly reproduce higher-order rainfall statistics at small-scales like lag-1 autocorrelation and skewness at 10-min. Discrepancies between observations and

simulations are quite large and therefore merit further attention. I am particularly concerned by the fact that differences are at their highest during Summer, which is critical for capturing small-scale extremes caused by localized convective processes. The fact that differences increase as we progress from large to small scales is also a sign that the downscaling scheme might not be performing as well as you think. Please revise accordingly.

8. **Factor of Change:** One of the key assumptions in this study is that the FoC of spatially-averaged rainfall statistics changes in the same way as the FoC of point rainfall statistics. The authors claim this is a reasonable assumption. But their argument, i.e., that the spatial variability in FoC in a single climate model is likely to be much smaller than FoC variability between models, does nos sound very convincing to me. Please specify exactly what you mean by that and, if possible, how this could affect the presented conclusions. In particular, it would be worth testing whether the FoC is sensitive to the size of the grid cells in the numerical model. This could be done by aggregating the model outputs and computing FoCs at coarser resolutions.

9. **Extrapolation of FoCs:** As GCM-RCM outputs are only available at daily resolutions, FoCs must be extrapolated (i.e., downscaled) to sub-daily resolutions. This is a critical step for which the authors provide very little detail. Clearly, this passage merits more attention, especially with respect to higher-order moments such as variance and skewness. Many studies have shown that rainfall distributions and key statistics like skewness and tail properties strongly depend on spatial and temporal aggregation scales. FoCs are no exception and I found it surprising that the authors assumed FoCs for skewness is scale independent. Please state the motivations behind this assumption and what evidence you have to support it.

10. **Extremes and Climate Change:** On page 7, ll.12-14, you say that *"We assume that if climate projections lie outside this range (i.e., the 10-90% uncertainty bounds on the mean), a practically meaningful signal of changes in extremes which goes beyond historical expectation is present."*

    This threshold seems totally arbitrary. Actually, it means there will be a rather large probability (under current climate conditions) of seeing an extreme outside the observed uncertainty range. Please carefully explain why you tolerate such a large margin of error and what exactly you think can be conclude based on such a weak statistical test.

11. page 7, ll.28-30: *"This makes statements about future climate change impacts on rainfall extremes very uncertain. In fact in the absence of considerable departures, it may be concluded that the climate change impact on high-resolution rainfall extremes may be difficult or impossible to characterise"*

    Difficult or impossible to characterise using your simulation-downscaling scheme yes. You forget to mention this might be due to model deficits and does not necessarily imply there won't be any significant changes.

12. A study in nearby northern Italy by Boni et al. (2006) has shown that rainfall extremes exhibit strong seasonal components. Even worse, the likelihood of an extreme occurring over a given region and month might depend on the chosen aggregation time scale. The paper provides very little information about the seasonality of extremes in Switzerland and the timing of the observed/simulated extremes. Maybe this could be discussed in more detail.

Literature Review

13. In the Introduction, you could mention the study by Allan and Soden (2008) who investigated precipitation extremes (at larger scales) with respect to temperature using model and satellite data and concluded that models might underestimate changes in extremes.

14. In a recent study, Anderson et al. (2015) proposed a method to detect secular trends in station-based precipitation variations, including trends in occurrences of rainfall extremes. Their approach is different but also based on stochastic weather generators and worth mentioning. In particular, they address the important issue of internal variability.
15. You could mention the work by Bechler et al. (2015) who proposed a two-step spatial hybrid downscaling method (SHD) to study small-scale extremes in the south of France.
16. You should absolutely cite the work of Langousis et al. (2016) who used climate simulations from the ENSEMBLES project together with a (different) statistical downscaling scheme to assess precipitation extremes at regional scales over the Flumendosa catchment in Italy.
17. You may want to have a look at the work of Loriaux et al. (2016). They performed an interesting statistical analysis of the relation between large-scale conditions and hourly precipitation extremes in the Netherlands.
18. Other studies that could be of interest: Lenderink and van Meijgaard (2008), Muschinski and Katz (2013), Harding and Snyder (2014), Scherrer et al. (2016), Wasko et al. (2016).

Minor Comments

19. Please provide more details about the cascade weights, their probability distribution and the way they depend on intensity and scale. How well does the theoretical model of cascade weights fit the observations?
20. In the captions of Fig 3-4, the word "distributions" is misspelled.
21. *"The problem of the underestimation of serial correlation in MRC models is known and has been ascribed to the temporal independence in cascade weights assumed in such models (Paschalis et al., 2012) and to the discrete nature of the cascade (Lombardo et al., 2012)."*

    This argument does not really hold for MRCB which uses intensity dependent weights and should therefore be able to better reproduce the temporal dependence structure. Yet this does not seem to be the case. Please explain.
22. *"The good performance of the stochastic simulator (including correlation and intermittency) at all temporal scales reflects in a good performance for extremes as well. In most cases the 10-90% confidence bounds on the distributions of simulated annual maxima for all analysed durations contain the observations, such as in the Lugano example (Figs 3f and 4f). This result was reproduced at the other stations as well. This is not a trivial outcome since the stochastic simulator is not explicitly calibrated to reproduce annual maxima."*

    First of all, results for Lugano are not as good as claimed by the authors. Correlation and skewness in particular seem quite far off. Second, it should come as no surprise that annual maxima are well reproduced by the MRC (given the huge uncertainty and large confidence intervals). Indeed, simulated annual maxima at 1h and 10 min will be strongly conditioned by annual maxima at daily resolutions. In fact, any decently parameterized conservative cascade model - even MRCA - will yield similar results. Please be more self-critical of your results and provide a more detailed and fair evaluation of the models' performances.
23. *"Simulated mean extremes fit observations very well at all stations, for all three studied time resolutions and return periods"*

    Not true. At best, the fit is good or fair. Certainly not excellent. Please rephrase or provide solid evidence supporting your claims.
24. What about solid or mixed-phase precipitation? Switzerland gets a fair amount of snow and hail. Yet I could not find any information about how these are treated in the analyses.

References

Allan, R. P. and B. J. Soden, 2008: Atmospheric Warming and the Amplification of Precipitation Extremes. *Science*, **321 (5895)**, 1481–1484, doi:10.2307/20144800.

Anderson, B. T., D. J. Gianotti, and G. D. Salvucci, 2015: Detectability of historical trends in station-based precipitation characteristics over the continental United States. *J. Geophys. Res. Atmos.*, **120 (10)**, 4842–4859, doi:10.1002/2014JD022960.

Bechler, A., M. Vrac, and L. Bel, 2015: A spatial hybrid approach for downscaling of extreme precipitation fields. *J. Geophys. Res. Atmos.*, **120 (10)**, 4534–4550, doi:10.1002/2014JD022558.

Berg, P., C. Moseley, and J. O. Haerter, 2013: Strong increase in convective precipitation in response to higher temperatures. *Nature Geosci.*, **6 (3)**, 181–185, doi:10.1038/ngeo1731.

Boni, G., A. Parodi, and R. Rudari, 2006: Extreme rainfall events: Learning from raingauge time series. *J. Hydrol.*, **327 (3-4)**, 304–314, doi:10.1016/j.jhydrol.2005.11.050.

Harding, K. J. and P. K. Snyder, 2014: Examining future changes in the character of Central U.S. warm-season precipitation using dynamical downscaling. *J. Geophys. Res.*, **119 (23)**, 13 116–13 136, doi:10.1002/2014JD022575.

Langousis, A., A. Mamalakis, R. Deidda, and M. Marrocu, 2016: Assessing the relative effectiveness of statistical downscaling and distribution mapping in reproducing rainfall statistics based on climate model results. *Water Resour. Res.*, **52 (1)**, 471–494, doi:10.1002/2015WR017556.

Lenderink, G. and E. van Meijgaard, 2008: Increase in hourly precipitation extremes beyond expectations from temperature changes. *Nature Geosci.*, **1 (8)**, 511–514, doi:10.1038/ngeo262.

Loriaux, J. M., G. Lenderink, and A. P. Siebesma, 2016: Peak precipitation intensity in relation to atmospheric conditions and large-scale forcing at midlatitudes. *J. Geophys. Res. Atmos.*, **121 (10)**, 5471–5487, doi:10.1002/2015JD024274.

Muschinski, T. and J. I. Katz, 2013: Trends in hourly rainfall statistics in the United States under a warming climate. *Nature Clim. Change*, **3 (6)**, 577–580, doi:10.1038/nclimate1828.

Scherrer, S. C., E. M. Fischer, R. Posselt, M. A. Liniger, M. Croci-Maspoli, and R. Knutti, 2016: Emerging trends in heavy precipitation and hot temperature extremes in Switzerland. *J. Geophys. Res. Atmos.*, **121 (6)**, 2626–2637, doi:10.1002/2015JD024634.

Wasko, C., A. Sharma, and S. Westra, 2016: Reduced spatial extent of extreme storms at higher temperatures. *Geophys. Res. Lett.*, **43 (8)**, 4026–4032, doi:10.1002/2016GL068509.