Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-489-RC2, 2018

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Interactive comment on "A virtual hydrological framework for evaluation of stochastic rainfall models" *by* Bree Bennett et al.

Anonymous Referee #2

Received and published: 22 October 2018

The authors present a streamflow-based evaluation framework to assess the adequacy of hydrologic predictions from stochastic rainfall generators (SRGs). This is a "virtual framework" in that it benchmarks these predictions against streamflows produced by historical continuous simulations, rather than observed streamflow timeseries. The authors point out that this avoids the complicating issues of model structural errors. This is a useful approach to benchmarking SRGs, and could perhaps be applied to other fluxes of interest (not just streamflow) for which long-term observation records aren't just available.

I agree with the first reviewer, who stated that "the theoretical elements of the paper are very long". There seems to be a fair bit of repetition, or at least over-explanation, of the motivation, and I strongly recommend that the authors look closely at how Section

Printer-friendly version

Discussion paper

2 can be shortened.

Generally, I think that the demonstration would be more illuminating if the authors used it to compare two or more SRGs and/or hydrologic models.

The demonstration of these methods is provided at the monthly timescale. While this timescale might be useful for applications of water supply, it is not meaningful for flood processes in all but the very largest watersheds. It is easy to picture a hydrologic model that produces adequate performance in terms of monthly flows, but not daily or subdaily extremes, while the opposite is also possible. Similarly, it is also probably an easier task to create a stochastic rainfall generator that works well for producing monthly means and associated variability than fine-scale extremes. Thus, the virtual framework in this manuscript may not be as broadly useful for extremes as the authors claim (at least flood extremes, droughts might be a different story). I thus recommend that the authors acknowledge this shortcoming, and "tone down" the framework's purported usefulness for flood risk (e.g. page 9 lines 17), since this remains unproven.

It isn't clear how the boxplots (e.g. figure 3) are constructed. Is it the "13 errors" mentioned in page 8 line 24? Or is it somehow derived from the 10,000 synthetic rainfall years? Or the 73 years of observed data with synthetic rainfall "spliced in"? Either way, it isn't clear that the authors have avoided the proliferation of error metrics that they identify as a limitation of previous measures on page 3. If this method is applied to a large number of sites, it still seems like a not-entirely compact evaluation scheme. Perhaps the authors could clarify how this compares to other methods in this respect.

Relatedly, in Figure 6 and 7, shouldn't the "obs rain" and "virtual obs. flow" be a range, rather than a single value? There are 73 years of monthly rainfall and simulated flows. . . this variability would be valuable context for evaluating the variability of the stochastic realizations.

It is unclear how other meteorological forcings (temperature, etc.) are handled in this

framework. The authors focus on stochastic rainfall generators, as opposed to stochastic weather generators, meaning that the other forcings must be supplied independently of the rainfall. I would imagine that this could create some serious issues in some cases if synthetic rainfall is spliced together with inappropriate series of temperature or other forcings; one can imagine getting strange results in terms of precipitation vs. ET balances, with unclear consequences for the evaluation results.

I have never developed my own SRG, but I imagine that it might be hard to know exactly how to use the results from this analysis to refine that generator, despite the authors' claim that this is a valuable use of the framework. It identifies performance by month, rather than by "rainfall characteristics" (pg. 21 line 9). It is useful to know whether the SRG performs well for some months than others, but what next? If the authors plan to continue research on this topic, I would suggest that a method that "tracks" the propagation of rainfall through the model might be more effective. To me, the most clear way of doing this is to track how different rainfall statistical moments translate to different statistical moments in the streamflow, using both historical and synthetic rainfall. Such an approach would be amenable to changing the evaluation timescale. For these reasons, I recommend that the authors delete the statement that this framework "should be an essential step in the development and application of stochastic rainfall models" (page 21 line 22-23). On a related note, the authors should comment on how this technique would apply to distributed (i.e. high-resolution gridded) SRGs and hydrologic models.

I wonder if this framework should consider the autocorrelation in monthly rainfalls when doing this splicing. I don't know too much about the climate of South Australia, but I can imagine that autocorrelation at least in dry periods can be quite important, and this is likely not preserved during the splicing. It's not clear what the implications would be for the resulting evaluation.

Figure 3 and elsewhere: I don't understand what "(90% limits shown)" means.

Section 3: Mention basin size. Also, why are stations outside the watershed used? More importantly, is the rainfall hydroclimate stationary? If not, then it seems as though this whole issue of stochastic generation and comparison of resulting streamflows against a nonstationary continuous simulation would be more complicated. Please comment on this.

Minor comments: Page 1 line 9: change “is” to “has been” Page 1 line 10: change “is given” to “has been paid” Page 1 line 12: delete “whenever the simulated rainfall are poor” Page 1 line 15: change “months” to “seasons”-that is a more broadly relevant term. Hydrology varies seasonally, months are an arbitrary construct (this comment applies elsewhere in the paper, such as page 2 line 31) Page 1 line 19: change “catchment cycle” to “annual hydrologic cycle” Throughout paper: I recommend introducing an acronym for stochastic rainfall models and using it throughout. Page 1 line 24-25: I recommend deleting “risks” after “floods” and “droughts” and changing it to “hazards” after “hydrologic” Page 1 line 28: delete comma after “targeted” Page 2 line 10: “and/or” is not appropriate in technical writing. Use “or” Page 2 line 12: put “virtual experiments” in quotations when mentioned for the first time, for emphasis Page 3 line 16: add comma after “poor” Page 5 line 10: The goal is not to match streamflow observations. It is to match the statistics of streamflows Page 3 line 12: add “model” before parameters Page 3 line 13: Why would you call ET “extraneous”? It is generally very very important. Page 10 line 15: I think that “observed/virtual” is a strange term. Observations have very little usage in this study. . . Page 11 line 16: grammar problem “was fit good” Page 14 line 14: This sentence is a bit awkward. It isn’t perhaps so “common and obvious” to the reader. Page 19 line 3-4: streamflow arises from more than just rainfall integration over a catchment area-what about ET, etc.? Page 20 line 11-13: I don’t understand this sentence. Certainly model performance depends on the chosen model. Page 20 line 13: good place to mention that multiple SRGs could be used too, not just multiple hydrologic models. Page 20 line 18: put “memory” in quotations.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-489, 2018.

Printer-friendly version

Discussion paper