Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-489-RC1, 2018

Hydrology and Earth System Sciences
Open Access
Discussions

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

Anonymous Referee #1

Received and published: 17 October 2018

General comments:

The manuscript introduces a new framework for the evaluation of generated rainfall time series in terms of their ability to reproduce runoff time series characteristics. This is done by two tests, an integrated test and a unit test. This topic is of broad interest for the hydrological scientific community and suitable for a publication in HESS.

However, I consider the integrated test not as a novelty, since it has been applied before in different studies, but the unit test is useful for rainfall model evaluation. Hence, I suggest to move the focus to the unit test and extend the validation by other runoff characteristics. Also, the theoretical elements of the paper are very long, the application and validation of the test should be extended and there is a lack of some crucial information reagrding the applied r-r model and its calibration procedure (for details

please see my specific comments). Due to the resulting workload I suggest a major revision of the manuscript.

Specific comments:

- P2l20-23 The so-called "virtual-observed streamflow"-approach and the integrated test is not new and a widely used evaluation method, especially in data-sparse regions or research fields. For example in urban hydrology, where measured runoff characteristics are not often available, the simulation of a reference streamflow is very common (e.g. Müller and Haberlandt, 2018). The authors even mention other studies using the integrated test (Li et al., 2014, 2016). However, the unit test is interesting and indeed provides useful insights into the rainfall-runoff (r-r) transformation process. It would be useful to move the focus on this test and proof it with additional runoff characteristics, e.g. flow duration curves, not using only the monthly runoff amount. Therefore, no new simulations are necessary, only additional analyses of the existing r-r simulation results.

P2l23-25 The sentence is not clear without the explanations given in section 2. Either here more information are provided or the sentence is left out.

P3l9-14 The idea behind the example provided by the authors is clear. Nevertheless, some of the rainfall characteristics mentioned are not clear and, since it is only an example, can be left out or can be replaced by other rainfall characteristics:

- rainfall on wet days - What does this characteristic represent (the daily total rainfall itself is mentioned later)?;

- Extreme value analysis on a monthly basis and autocorrelation on an annual basis are from my understanding rather uncommon rainfall characteristics for the evaluation of rainfall time series P3l20-21 The details provided in brackets can be left out, since without reading the reference there are no additional information for the reader.

P3l7-P4l20 The motivation for the introduction of the new evaluation strategy is quite long and can be shortened by the half. I think the majority of the community is quite aware of the issue with overlapping errors. Also Fig. 1 and Fig. 2 are quite clear from the text and could be left out. If kept, a box with "True rainfall" should be added in Fig. 1a) to be consistent with Fig. 1b ("True streamflow")

P5l10 "to match streamflow observations" -> "to match streamflow observations or statistics"

P6Table1 The authors should include a definition of the applied symbols in the caption, since the difference between "x" and "-" is not too intuitive (from only the table). Is in the last line, first column something missing (virtual hydrological. . .)?

P6Table 1 From my opinion the results from the virtual-observed streamflow approach can still be biased by the applied r-r model. For example, rainfall is generated in space and two rainfall generation methods show differences in terms of rainfall characteristics, but not in the simulated streamflow. After what I've read in the introduction and methods section, the conclusion is that the compared rainfall characteristics are then not practicable ("no impact") und useless (for the study region). But this also depends on i) the model choice (including e.g. spatial resolution, model type (fully / semi-distributed), several model approaches) and ii) the parameter identification. In a semi-distributed model differences in spatial rainfall could be dampened, while they are (maybe) not dampened in a fully-distributed model. The parameters have to be chosen a priori – a calibration on one of the rainfall data sets is not possible to avoid biases. Will the parameters be calibrated by an additional rainfall data set (the observed data) and if so, how can be avoided that this calibration introduces a bias (e.g. maybe the observed rainfall data is more similar to rainfall data set A under investigation than to B)? So all of the results depend on the chosen setup for the r-r simulations and drawn conclusions are only valid in context with the model setup and parameter set. This is of course always the case in hydrology, but it becomes more important if a virtual runoff time series is applied, since the "relation" between the model output and reality gets lost. However, the authors point these issues out later in their investigation (p20), but it

should be communicated earlier to the reader.

P9Fig3b Maybe the authors can spend a more detailed explanation of the two different indices k and t. For me the difference was not quite clear at the beginning. Also, it is clear that rainfall in June can affect the runoff in July (or from April by filling storages and hence affecting runoff in July). But how can rainfall in July affect runoff in June, although the months August to January obviously don't? Is the rainfall information transformed into runoff over such a long period in the model? Since there is no rain in the summer half year, shouldn't the storages run empty?

P9Section 2.4 It would be useful for the reader to illustrate the implementation of the framework with a flow chart, since the authors use step 1, step 6 and so on throughout the section (and the manuscript).

P10l14-17 What is the 90 % limit of the simulated statistic? If m=10 mm, everything between 1 mm and 19 mm is considered as good? Here an additional explanation is required.

P11l9-10 From Table 2 it cannot be seen, how long the time series used for the calibration of the rainfall generator are. It would be useful to the reader to characterize the time series more in detail (wet spell durations and amount, dry spell durations and maybe even on a monthly basis, since further investigations are carried out on a monthly basis). At least a hint to Fig. 6 and Fig. 7, which include some monthly observations, would be useful.

P11l16 For the calibration of the model the reader is referred to Westra et al. (2014), which is a non-reviewed technical report with 100+ pages, as far as I can see. In context with my former specific comment it is necessary to provide information in the actual manuscript, how the model has been calibrated. Which rainfall data was used for the calibration? If all 22 stations have been applied, how was the areal rainfall estimated as input for the lumped r-r model?

P11Section 3 Although the observed discharge time series is not used in the investigation, it would be useful for the reader to provide some runoff characteristics (e.g. mean discharge) to get a feeling for the catchment.

P11l14-18 On p9l15-17 you mention "The hydrological model should be selected on the basis that it is capable of simulating streamflow for the timescales, magnitudes and physical processes of interest to the intended application.“ Is the lumped model able to simulate the physical processes of a catchment with a few 100 km^2 catchment area (I could not find the catchment area in the manuscript).

P12l8-10 Which result is analyzed? Integrated test or unit test?

P19l19 In Fig. 5 the results for rainfall are worse than for runoff (for mean values).

P19 Results-section Before it was mentioned that also the influence of spatial rainfall patterns can be evaluated. Since this is not done in the manuscript, it can be moved to the outlooks of the manuscript. Otherwise a spatial analyses can be implemented in the manuscript (what I would recommend), to show further advantages of the unit test.

P20l20-21 With the introduced framework it is still not possible to identify, which rainfall characteristics are important for streamflow prediction. Based on the high non-linearity of the rainfall-runoff transformation process, a single rainfall characteristic cannot be sufficient to draw conclusions about the impact on the resulting runoff. If this would be the case, r-r models wouldn‘t have to be used anymore. However, could the authors identify, based on their analysis, which rainfall characteristics are important for the resulting runoff behaviour? (of course, the results depend on the study site, model choice and so on, but nevertheless. . .)

P21l32-p22l1 This example is hard to follow, maybe the authors can extend it. From my understanding it depends on the calibration of the storage coefficients. If storage coefficients are small, the results from the monthly rainfall will be transfered to runoff immediately. This would be possible with the "traditional“ approach.

Technical corrections:

P23 There is a reference of Li et al. (2015b), but no Li et al. (2015a). Also Li et al. (2016) is mentioned before Li et al. (2015b)

References:

Müller, H., Haberlandt, U. (2018): Temporal rainfall disaggregation using a multiplicative cascade model for spatial application in urban hydrology, Journal of Hydrology 556, 847-864

Sikorska, A. E., Vivrioli, D., Seibert, J. (2018): Effective precipitation duration for runoff peaks based on catchment modelling, J. Hydrol., 556, 510–522

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