Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-164-RC2, 2018

Hydrology and Earth System Sciences
Discussions

Interactive comment on "The potential of global re-analysis datasets in identifying flood events in Southern Africa" *by* Gaby J. Gründemann et al.

Anonymous Referee #2

Received and published: 20 June 2018

This study aims at evaluating the performances of several hydrological models in identifying flood events in South Africa. The models considered are members of the Water Resource Reanalysis (WRR) developed in the European Earth2Observe project. Models performances are evaluated using a frequency analysis and several skill scores related to flood detection. The authors also provide an interesting comparison with damaging events reported in three different disaster databases. Results convincingly show the ability of such models to capture the majority of flood events, despite their coarse resolution. Performances vary from one model to another, due to differences in model structure. The authors also pointed out the improvement due to the increase in spatial resolution (from 0.5° in WRR1 to 0.25° in WRR2). Before concluding, the authors discuss the main limitations of the method. The conclusions are consistent

with results presented all along the manuscript. The paper is well written and organized. This is a really interesting study and I think that the manuscript would be ready for publication provided that the authors address the following comments.

Major comment:

My major comment concerns the spatial resolution of the models. First, the authors often attribute the improvement in flood detection to the increases in spatial resolution. But in this case, the improvement can be due to three factors: (1) the improvement of models forcings: WFDEI (used in WRR1) and MSWEP (used in WRR2) rely on different methodologies (2) new model developments: each modeller involved in E2O included new developments in the models (e.g. multi-layer snow scheme in HTESSEL-CaMa, groundwater abstraction in LISFLOOD, reservoirs and water withdrawals in PCR-GLOBWB, aquifers and floodplains in SURFEX-CTRIP, etc.) (3) meteorological and hydrological processes better represented at higher resolution Many studies showed that simulated discharges are highly sensitive to meteorological forcing, especially precipitations, which are supposed to be of better quality in WRR2. Although points (1) and (3) are briefly mentioned in the discussion section (P13L2-5), I think they should be mentioned in the section presenting the models (section 2.2). Also all the conclusions on the differences in WRR1 and WRR2 performances should be put in this context. To have an idea of the impact of points (1) and (2) (without point (3)), the authors could consider the WRR2 version of the SURFEX-TRIP model which used the improved forcing (MSWEP) and new model developments but a spatial resolution of 0.5° for the routing scheme.

Another problem related to the spatial resolution is the selection of gauge stations (section 2.3.1). The authors mention that "the stations have upstream catchment areas that vary between 4 and 342,000 km2)". Is it realistic to compare observed and simulated discharges at stations with drainage area that small? Given that a model pixel has an area of approximately 2500 km2 for WRR1 and 650 km2 for WRR2, rivers with drainage area smaller than these thresholds are generally not represented in the models and the

correspondence between stations and model pixels is necessary wrong. This is consistent with authors results (P8L25-27, P12l30-32). This remark is mentioned P9L3-5, and in my opinion, stations with small drainage area should be excluded, even though there would remain a small number of stations. Also, could the authors give some details on the method used to select the model grid cell corresponding to each station? Was the river network of each model used to associate each station to a model grid cell?

Minor comments:

P3L31: “one of the largest basins”

P4L15-16: Have the impacts of such modifications on floods been studied already?

P5L11: Please provide a reference for “the Kolmogorov Smirnoff statistic for the Gumbel Extreme Value Distribution”.

P5L21: Please add a few words to explain what the criteria from NatCatSERVICE is.

P6L19: In my understanding, hydrological extremes include floods but also droughts. This study only focuses on floods. The authors should carefully revise the use of “extremes” throughout the manuscript (including the following subtitle).

P7L12: “. . . for both the observed and the modelled discharges...”

P8L14: The areas mentioned in L13 are mainly related to the models spatial resolution. This should be specified.

P8L25-27: My guess is that the difference between WRR1 and WRR2 performances mostly comes from the forcings (rather than from the resolution).

P8L27-29: This result is hardly visible from Figure 2. It is more evident from Table 4a.

P9L8-11: I think this kind of conclusion should be built on larger basins only. Also, a fair comparison would only consider WRR1 so that the influence of forcings are excluded.

P10L9: “. . . were established using the observed climatology.”

P10L26-27: These differences in performance are also (mainly?) the result of the improved forcings.

P12L5: “. . . and was evaluated by means of commonly used error statistics...”

P12L17-24: Would it be possible to get the dates of construction of major dams or reservoirs? It could be interesting to look at models performances before and after these dates.

P12L30: Improvements are also due to forcings improvements and new model development.

Figure 1: Please remove the marks within the inlet map. Also, please explain what the square and the circle represent (something like “stations used to illustrate the flood frequency analysis in section 3.2.1”).

Figure 2: Would the NSE results be better presented using a log scale?

Figure 3, in the caption: “The index value refer to models with 0.5 degree resolution (MM1 and MO1) and 0.25 degree resolution (MM2 and MO2).”

Figure 6: The figure is not clear (small points/lines, close colours). Would results be better presented using box plots with mean, median, 1st and 4th quartiles?

Table 2: What are the impacts of changing theses values on the results?

Table 4, in the caption: Change “Figure 3, upper/lower panel” to “Table 4, upper/lower panel”. Also, it seems that the selection of stations in each line of the table relies on a lower threshold, so the “lower or equal” sign should be “upper or equal”.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-164, 2018.