Review of “Signature and sensitivity-based comparison of conceptual and process oriented models, GR4H, MARINE and SMASH, on French Mediterranean flash floods” by Abubakar Haruna et al.

General comments:

The study explores the parameter sensitivity and event-based performance of three different models in aspect of French Mediterranean flash floods. The authors did a lot work resulting in 18 figures, 11 tables and abundance appendix materials, however, which makes the main point neither condensed or concise. Overall, the novelty of this study is not strong and has not been articulated.

Here are some major comments:

For distributed models, why using different spatial resolution for MARINE (0.5km) and SMASH (1km)? Is it fair in terms of inter-model comparison? I suggest adding a table to describe the resolutions of data and models, in order to avoid confusing.

I also wonder how do you consider the influence of hydrograph shape (one peak or multi-
peak, thin or flat hydrograph) to model performance?

How robust it is to set SIM2 as benchmark for evaluating model performance of soil moisture?

What’s the point to soil moisture comparing? Is there any possible that the peak flow is related to better soil moisture simulation? It would be better if the result sections can be related to each other.

For the text throughout this study, it would be better if you use the same order when you describe the results of the tree models (e.g. always in the order of GR4H, MARINE and SMASH). Please also put figures and tables in the correct position. It will also be my pleasure if there is more analyzation and discussion, but less repeated figures or tables.

Specific comments:

P8. L212. 1. Can you please introduce more about Mediterranean climate of these two catchments to support the selection of flood events? For example, is there snowmelt in spring to cause flood events of return period higher than 2 years?

2. What are the difference and similarity of these two catchments? Why do you choose these two catchments? Is there any influence for model performance since they don’t share similar catchment shape?

P12. L295-302. Did you use GR4H and SMASH to simulate long-term hourly runoff process and use event based MARINE model to simulate selected events? If so, the performance evaluation is present in event based scale. How do you consider the advantage of event-based model MARINE?

P19. L406-407. This is an interesting finding. Is it because the flood of these two catchments is mainly relied on transferring or production components? Is it possible that it can be related to certain catchments attributes?
Figure 2. The longitude and latitude grid is needed. It would be better if you can show Mediterranean climate regions in this figure.

Figure 12. Please also include P in the legend.

Table 3. Is *No of classes* represents number of soil classes? Why *Cp* differs a lot between these two catchments?

Table 4. Is Gardon in the first row correct? Should it be replaced by range?

Table 5. 1. Please uniform the flood events name for two catchments.

2. As you introduced before: 1) heavy rainfall always appears in autumn (line 27) and, 2) intense rainfalls are in the autumn and winter (line 223) for Mediterranean climate. So please give a brief explanation about why spring events can be selected under this kind of climate.