**Interactive comment on “Runoff sensitivity to spatial rainfall variability: A hydrological modeling study with dense rain gauge observations” by Clara Hohmann et al.**

**Anonymous Referee #2**

Received and published: 23 October 2020

This study tried to explore the effect of interpolation methods and station density on runoff simulation. However, this work has some severe problems and does not reach the standard of HESS in the current form. My comments are as follows.

Major comments:

1. It can hardly say that this study is novel or contributes much to our knowledge. The authors used two simple and widely used interpolation methods, TP and IDW. Results and conclusions are generally similar to previously published work. Besides, as a study focusing on the effect of station density and interpolation methods on hydrological modeling, the design is too simple. It is better and not difficult to (1) design station networks with random numbers and locations of stations instead of using pre-defined networks (Table 3), and (2) conduct a sensitivity test of the weighting power instead of using 2 and 3. A negative example caused by (1) is that the authors state “32 stations would be enough...”, but this resulted from a large jump from 16 to 32 to 64 stations.

2. A large part of the study area is not covered by 158 rain gauges. Both IDW and TP cannot achieve reasonable estimates outside the WEGN network. As a result, only rainfall in the middle reach is largely affected by different density and interpolation methods, while rainfall in the upper and lower reach could always contain much larger errors. Meanwhile, the hydrological model is built over the whole domain, and runoff at catchment and sub-catchment outlets is compared. The effect of biased estimates in upper and lower reach on the runoff simulation could be very large.

3. Please perform the analysis based on a larger collection of precipitation events. Currently, only three small-scale short-duration and three large-scale long-duration events are selected. Although the six events may be representative, a comprehensive view in a long historical period is useful and necessary to demonstrate the all-aspect effect of interpolation and station density. The current results and analyses are all based on those limited events, making the results more “casual” than “causal”. For example, the interpolation based on 5 stations could be largely affected by the location of storm centers, and thus the results based on 100 events could be different with results based on 6 events.

4. Figure 5: The IDW precipitation map based on 158 stations looks quite unrealistic. An actual precipitation event should be spatially continuous like that in Figure 4. However, the bull eye effect in Figure 5 is too obvious. The authors stated that all stations within a 50 km searching radius would be used in IDW. 50 km is quite large based on the measuring scale in Figure 1. This makes the bull eye effect even weirder.

5. The authors strengthened “spatial rainfall variability” in the title. But the spatial variability is not explicitly analyzed in this study. The title does not reflect what this
study actually did, i.e., station density and interpolation methods. Besides, the effect of station density and interpolation methods cannot be simply represented using "spatial variability".

Specific comments:
6. Please adjust the font of units which is often different from that of texts.
7. Line 25: It is better to state the “measurement uncertainties of rain gauges” because other approaches of rainfall measuring are not mentioned here.
8. Line 32-34: The pros and cons of radars and satellite sensors are similar in many cases and should be stated together.
9. Figure 1: It will be helpful to add latitude and longitude.
10. Line 124: Please complete the reference for WaSiM.
11. Line 150: Although I understand what you mean by 50% NSE and 50% KGE, please rephrase to be more formal.
12. Line 155-162: Given the authors state that manual recalibration is necessary, please add some descriptions on the benefits of manual recalibration. For example, what’s the KGE and NSE before manual recalibration? Besides, the NSE decreased from calibration to validation periods but KGE increased. Please add some explanations.
13. Line 167: Strictly speaking, you used two, not three methods. Two different parameters do not make IDW two different methods.
14. Figure 4: Please use the shapefile of the catchment to replace the black box, which can help identify whether storm centers are located within or outside the river basin.


C3