**General comments and recommendation**

The manuscript by Hohmann et al. presents an analysis of the effect of different meteorological network densities, as well as different interpolation schemes on the simulated runoff for a meso-scale catchment in southeastern Austria. While indeed many questions about the optimal meteorological network density, about the adequacy and representativeness of stations' distribution and about the suitability of different interpolation schemes for hydrological modelling have not yet been fully answered, I do not see right now in this study an adequate and robust assessment, offering an added value to what we already know. The setup and design of the experiment is not thorough enough to support reliable statements with evidence.

The manuscript has per se a clear structure, the methods are generally described in a comprehensible way or supported by relevant sources, however, some explanations could be more clear and concise (e.g. the calibration procedure). Even though the discussion provides some good points, generally there are quite a few redundant paragraphs, while more interesting and critical points are not examined closely enough.The manuscript generally features high-quality and interesting figures; some of the tables and their captions should be reorganized in a more meaningful and efficient way. I would also suggest a native speaker to read it and correct it, some sentences definitely need to be rephrased.

Because of these considerations, I think the manuscript requires and extension of the experimental design, a more critical discussion and further work, before it can be possibly recommended for publication.

Please find my specific and technical comments here following.

**Specific comments**

- Experimental design:
  - One important drawback of your setup is the fact that despite the stations' density is high, it is not covering your whole study area, but only its central part.
  - You have "fix" a priori chosen subnetwork configurations, but for a fair evaluation of the effect of stations' subnetwork density you should try different (random?) configurations for the same number stations. Tentatively, you could also consider smaller jumps between one configuration and the next one.
  - You chose 2 interpolation methods (I wouldn't refer to three methods, as you simply changed a parameter of the second method), I think it would be more appropriate to use more interpolation methods, such as kriging, etc.. for a more robust evaluation of the effect given by the interpolation method, and also definitely to use a smaller searching radius (second spurious peaks in your hydrological simulations are not surprising, given how you interpolate precipitation).
  - The number and sample of events you are analyzing is simply too limited to allow you to make any meaningful assessment, now your analysis might show only very specific and localized effects, and might not hold for a larger ensemble of events. Is there any way to increase the number of events you analyze? You say you selected first heavy precipitation events among the top 10% heaviest rainfall days during summer, and out of these you selected through visual inspection your 6 events. I would suggest you rather select the top 100 events for 1d and 3d accumulation periods, for instance, and further analyze these?
  - (You are only using discharge gauging stations on the main river trunk, but you analyze also the contributing subcatchments. Sure, this is fine for looking at the effect of the different precipitation inputs, but not enough to disentangle the effects possibly stemming from the parameters of the hydrological model. You implicitly assume the hydrological model is working equally well on much smaller subcatchments, with the same parameters. As the

model is process oriented I am fine with this assumption, but you might want to spend a few words on this?)

- Data:
  - You have a 10 year period to choose from resp. analyze, is there some way to extend it?
  - You only report the return period for discharge, but actually it would be interesting – and relevant? - to know the return period associated with your rainfall events too.
- Model setup:
  - P7-L128-130: Lumped doesn't necessarily mean that a model is not process-oriented. I guess you mean conceptual?(that are often lumped, but not always)
- Calibration of the hydrological model
  - Please describe better and provide more details on your calibration procedure.
  - Why do you calibrate the model only basing on one summer, and only for one interpolation method?
  - Why do you use "only" NSE and KGE to define your objective function, instead of including further goodness-of-fit measures and criteria, perhaps more specifically tailored for floods, their volume and/or their timing?
- Runoff analysis approach:
  - I would recommend you expand your analysis approach by including also the peak timing, as this also is an important aspect in your modelling exercise, e.g. considering the effect of superposition of peaks simulated in the subcatchments.
- Results:
  - P16-L271: I wouldn't say that, e.g. Kornbach is rather systematically overestimated.
  - P16-L275-277: this is not true, see for example the short-2 event or the long-3 event.
- Discussion:
  - P 20: it is confusing you mentioning first Lopez et al.2015 report no increase in performance after including 24 gauges per 1000 km2, and later on saying they actually used 12 stations per 1000km2. They used gridded products to be able to increase the number of stations with "hypothetical" gauging stations. I think you should either specify this, or not go so much into detail.
  - P20-L360-363: Rephrase please this last paragraph.
  - P21-L380-381: Isn't this possibly case specific? Refer also to Lobligeois et al. 2014: In all regions, natural variability allows for contradictory examples to be found, showing that analyzing a large number of events over varied catchments is warranted.
  - I think you should expand more on the representativeness of stations' location.

**Technical corrections**

- P6-L124: developed..by who et al.?
- P7-L128: BAFU is FOEN in english
- Figure 2: shouldn't you remove the box with snow accumulation/melt, as far as I understood you are not using it?
- Figure 4: You should be consistent, and either use the min-max range in both cases, or the 5$^{th}$ and 95$^{th}$ percentile range. And why don't you also show the same kind of information for the 3$^{rd}$ source of precipitation data?
- Figure 6: Why don't you show observations in the last column?(i.e. for the Neumarkt/Raab gauging station)
- P17-L301 ..in these subcatchment should be in *this* subatchment
- P21-L373: ..more event-depended should be event-*dependent*