

**Ref:** HESS\_2019\_427\_version2\_review

**Title:** The accuracy of weather radar in heavy rain: a comparative study for Denmark, the Netherlands, Finland and Sweden.

**Authors:** Marc Schleiss\*, Jonas Olsson, Peter Berg, Tero Niemi, Teemu Kokkonen, Søren Thorndahl, Rasmus Nielsen, Jesper Ellerbæk Nielsen, Denica Bozhinova, and Seppo Pulkkinen

### **General comments**

The authors have tried to perform a new bias estimation (section 2.3.1) and to present new analyses (section 3), which are nice ideas. However, the revised manuscript still lacks clarity on writing and the delivery of messages, which needs to be improved. In the following, I tried to point out some, but the authors should go through their manuscript more carefully and not to jump to conclusions without providing clear evidences.

I would not recommend for the publication without a major revision.

1. L368-372: The modelled bias by Eq. [5] for the Netherlands should be about 0.8. Where does 1.23 come from? The explanation on this value (that “the radar values even seem to be overestimated”) is neither clear (e.g., if the bias refers to rainfall intensity, mm/h as shown in Y-axis, or to total accumulation for this period shown in the upper left – “gauge 56.9 mm, radar 41.6 mm”) nor agree with the statement in L356 (the systematic pattern of “underestimation”), though the authors stated that such values should be carefully interpreted.
2. L420-421: This was a major conclusion (also appears in abstract). Does “the actual bias” mean the modelled one? If so, it cannot be “actual”. If not, it is not clear where and how such interpretation has been derived, given that i) the results based on eq.5 which require further analyses due to its assumption, and ii) the results (table1) were made only with highest aggregation time so what the authors mean by “after accounting for differences in scale”?
3. L429-455: i) Definition of the “conditional bias” with respect to the rainfall intensity is not clear. In L434, it seems that it refers to the “log ratio of ( $R_g/R_r$ )” with respect to the rainfall intensity, which is not the same as the modelled bias  $\beta$  shown in Eq 1, yet the results were described in terms of “bias”, ii) in Figure 7, was 0 mm (no rain) is also included and if so why? How would the error bar for the fit (or something that tells a significance level of the fit) look in this graph? iii) L438, this looks the slope of red line, does it make sense to use %? iv) It is not easy to follow L442-L444, can the authors explain better where we see such evidence; e.g., the differences in apparent biases shown in Figure 6 look small among Denmark (1.59), Sweden (1.66), Finland (1.56) and the NL (1.4).
4. Section 3.6: Although the authors introduced the comparison with X-band radar, in L248-250, “the X-band data can be used to provide valuable insight into the advantages and challenges associated with using high-resolution X-band radar measurement in times of heavy rain”, its analyses (L582-587) are poorly written and hardly provide insight as they aimed. If this radar is dual-pol, the results could have calculated with Kdp to prove their speculations (L583-585). It is also questionable in conclusion on the X-band data for Denmark (L640).
5. L547-550: where do we see that the “*multiplicative biases* between radar and gauges can *amplify* when data are aggregated to coarser time scales?” If the authors meant the “multiplicative bias” as the “peak intensity bias”, we still see that the bias values get smaller with coarser aggregation time (Figure 10). Similarly, it is not clear what

- this means in L550 that “single-radar products with daily (coarser than sub-hourly) rain gauge adjustments are *more vulnerable to error amplification*”. Can the authors provide a clearer example of error amplification and when this gets more vulnerable?
6. L551-559: Despite the explanation of Fig.11, it is not clearly written how significantly this can affect to the choice of hydrological models for flood prediction (L549). Also, in L558-559, how this should be interpreted for both bias correction and reduction to be applied for hydrological models?
  7. L582-587: Improve writing with clearer message.

### **Minor comments**

1. L112: Time period for KNMI gauges (2003-2017) does not match with those indicated in Table 1.
2. L278-279: does this “[-]” mean unitless? Is it necessary?
3. L421-422: “Moreover,..” What does this sentence mean?
4. L535-536 (Figure 10). Because the presented case is the same as shown in Figure 4, the peak intensity bias values and unit at the first interval (at the highest temporal resolution) are expected to be identical as those mentioned in L373-374 for each event. However, as seen right Y-axis in Fig. 10, the “peak intensity bias” is expressed in terms of in %, i.e.  $100*(1- 1/(\text{factor}))$ ”. Explain this value (technically it is not the “peak intensity bias”) better in the text, and revise the sentence.
5. Figure 11, revise the title of the X-axis.