Ref: HESS_2019_427_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

This manuscript presents rainfall comparisons estimated from radar and rain gauge observations (in Denmark, the Netherlands, Finland, and Sweden for 50 top heavy rainfall events occurred in each country). Biases and some discrepancies were analyzed with respect to different temporal and spatial scales and rainfall (and peak) intensity to assess the performance of radar products capturing heavy rainfalls. Because of the differences in radar hardware, radar data processing, and the collected events from each country, the interpretation of the comparison is not straightforward and challenging but the authors made a worthwhile multinational effort to document the differences.

The topic fits to the scope of the journal's special issue, but the manuscript needs to be better clarified for its publication. Please see the comments below. Line numbers are indicated with "L"

Recommendation: Major revision required

Major comments

1. (Abstract) L15-19: Throughout the manuscript, supporting materials for urban hydrology and mitigations of attenuation are not presented. Revise this part and reflect what has been presented.
2. The link with hydrology or urban flooding/forecast:
    a. One of the objectives of this study is explicitly written to better understand the link between rainfall and urban flooding (L7-9) or/and the use of radar in hydrology and flood forecasting (L84-85). However, very few discussions were presented in this aspect. Add either more supporting martials for flooding parts (link with the presented work) or clarify better the objective of the presented work.
    b. Hydrological model (L171, L205, L397, L472, L490) has been mentioned in several sections without reference cited and the statements are rather generally made, which requires improvement in either writing or strengthening the explanation with more supporting materials (particularly for the statement made in the conclusion).
3. Better clarification and more supporting materials are required in results and conclusions (see the minor comments 16-37).

Minor comments

1. L10-L11: Clarify better "the top 50 events", "overall agreement", "the peaks" of what.
2. L44: need clarification of "accuracy" (of what).
3. L46-47: This term "higher-level" composite is less objective and vague. Rephrase it.

4. L59-60: ", the longest…15-20 years at best." Is it the case for world-wide or those countries presented in the manuscript?
5. L76-78: "Often…the results" This is not clearly written in the context. Specify better. Also, adding more backgrounds/references to support strong needs in multinational assessment and comparisons will be necessary. At least, in Europe, there has been an effort made with BALTRAD products (Michelson et al. 2018, referenced already in the manuscript but in later chapter) and with the OPERA products (e.g., Saltikoff et al. 2019, Park et al 2019), which can be referred in the introduction.

   Saltikoff, E.; Haase, G.; Delobbe, L.; Gaussiat, N.; Martet, M.; Idziorek, D.; Leijnse, H.; Novák, P.; Lukach, M.; Stephan, K. OPERA the Radar Project. Atmosphere 2019, 10, 320

   Park, S., Berenguer, M., Semper-Torres, D., 2019: Long-term analysis of gauge-adjusted radar rainfall accumulations at European scale. J. Hydrol., 573, 768-777. doi:10.1016/j.jhydrol.2019.03.093

6. Table 2: Clarify the data resolution original vs. used for the comparison, e.g., in the text Line 128, Danish data has been interpolated to 1 min. In Table 3, is the comparison also done 5 km not 1 km?
7. L153-154: reference missing for the operational product.
8. L164: "Polar radar measurements". Describe better, it seems a jargon, meaning radar measurement done at polar grid.
9. L170: After applying HIPRAD, the temporal/spatial resolution of the data remains the same shown in Table 2?
10. L178, "Aalborg" add country name and indicate the coverage of this radar in Fig 1.
11. L188: what is "tas BALTRAD"?
12. L206-208: Add reference
13. L290: "the HIPRAD" here, isn't it BALTRAD?
14. L249: "the highest available temporal" This term is used several times later, but isn't it the same as gauge sampling resolution (shown in table 1)? Is there any reason for such term? If so, explain better.
15. L 249: "Top event" → Event 1 (fig. 2), where these gauges located in Fig 1?
16. L253-254: Some results presented were already gauge adjusted and one (Finland) not. It is not clear to compare these numbers from literature examples (which is not clearly mentioned either if they (literature examples) were also derived before the adjustment or after?). Is it necessary?
17. L258: "The third rainfall peak" indicate here figure 4 (perhaps better with 4a indicating Denmark).
18. L264-265: "the relatively large peak intensity biases of 2.17, 2.09, 1.98 and 1.73 for Denmark, Finland, the Netherlands and Sweden…confirms this hypothesis" if the hypothesis refers the previous sentence, the bias for Netherlands should be larger than that of Finland because the peak intensity is higher for NL than for Finland (L256), isn't it?
19. L272 "at these scales" and L275 "such small scales". What does it mean? Is it related to storm scale? Or do you mean that the comparison was done with the instantaneous and point estimates (that affects representativeness error)?
20. L283: This is redundantly written (merge with L280-282)
21. L300-301: Are these numbers MB after the ARFs reduction applied? is it also shown in Table 3?
22. L302-302: Is the statement made before applying the ARFs? Clarify better. After ARFs, Swedish result shows the best, doesn't it?

23. L306-307: This does not support any argument and redundantly written in L300. Rephrase or remove it.
24. L324, L405: “deeper analysis” Avoid “deeper” (somewhat subjective word) and revise the sentence.
25. L325: “temporal aggregation time scale” → aggregation time scale (isn’t it the same as shown in Figures 8-10?)
26. L338-339: “Furthermore, the quality….an important role”. Add supporting explanation.
27. L359: It is not clear in Table 3 that the Danish products are the best in terms of RRMSE and CC. Revise this part.
28. L363-364: “However, a closer analysis….only 0.2”, what does it mean?
29. L375-376: Clarify what is “viewpoints”. Apart from the statement, how the attenuation and VPR correction applied to the group 2 data (Yes for Danish C band data, not explicitly indicated for the Swedish) were performed?
30. L379: “a coarser scale” in time or/and space?
31. L397-399: add reference. Is there any example run for the presented event?
32. L418: “the same order…than for…” → the same order…as for
33. L421-L422: This statement needs better supporting explanation, e.g., what dual-polarization capabilities was used in the processing of the data?
34. L469-470: “Bias correction…on peak intensity bias”. Is this conclusion derived from all the presented cases for four countries? There are some explanations for the Dutch product (L348-349), but not easy to find for the others. For Finland, the presented examples are not even bias corrected, so it is not clear what the authors mean.
35. L471-472: Throughout the manuscript, “the importance of high-resolution radar observations in hydrological study” is hardly demonstrated/literature-reviewed with respect to the high-resolution radar products, which makes such conclusive statements weak. Add more solid outputs or references.
36. L488-489: Add references.
37. L489-490: Add references or strengthen supporting material for the referred rainfall uncertainties in hydrological models (e.g., some examples among any of the events -50 events*4 countries as a part of discussion or more explanation in L397-399).