Assessment of spatial uncertainty of heavy rainfall at catchment scale using a dense gauge network

by Sungmin O and Ulbrich Foelsche

This is the second time I review this paper. Most of my previous comments were taken into account and the authors have made substantial changes to the original manuscript. As a result, the quality has improved. That being said, there are still a few important issues that remain to be addressed before publication can be envisaged.

Major Comments:

1. Usefulness and novelty:

While the WegenerNet is an amazing and unique source of data, the paper itself is rather dull and empty, providing very few ideas and findings worth publishing. The conclusions can be summarized as follows:

- The spatial variability of rainfall varies with years, seasons and events (no big surprise there). The spatial autocorrelation structure is anisotropic and approximately decreases exponentially with distance (which has been known for a long time already).

- The small-scale variability during the cold season is higher while the decorrelation range is longer, in agreement with previous findings.

- The uncertainty in areal-rainfall estimates depends on the number of gauges available (obviously!). High-resolution gridded data provide more reliable information about extremes (also obvious).

Many of the numbers are specific to the WegenerNet and of little usefulness to others. The only new idea in my opinion is the power-law model for predicting the number of gauges needed to ensure the uncertainty affecting mean areal rainfall estimates is below a given threshold. But this part of the paper is not detailed enough and the model has not been properly validated (see comment 3). Otherwise, the paper remains very descriptive, with only a single equation in it. There's nothing wrong with it but I just don't think that it will be useful. A more useful approach in my opinion would have been to use the WegenerNet data to formulate predictive models for assessing uncertainties and validating findings on other networks or situations.

2. Too many inconsistencies:

The paper contains many inconsistencies. Often, the text says A while the figures show B.

- in the autocorrelation function, the text says that the largest considered distance is 15 km. Yet Figure 3 shows distances up to 20 km and more. This was already an issue in the previous version and has not been addressed properly.

- On page 4, the text says that a data transformation is necessary to deal with zeros and make the data more Gaussian distributed. Yet Figure 3 shows the autocorrelation without data transformation (or at least I assume, it's not 100% clear). If indeed the values change after transformation, why don't you show the ones matching the text?

- For the exponential model, the text mentions that another two-parameter model was tried with very similar results. Yet somehow the authors decided to go for the three parameter model anyway. The explanations to justify this choice are not convincing at all and proper validation is required to motivate this choice. Indeed, one of the parameters seems to be almost constant across scales, which means its probably not useful.

- Page 8, line 24: the text says that the study has confirmed that the WEGN provides very accurate areal-precipitation estimates. But actually, there is no evidence for this in the paper. Only comparisons where the average of the 150 stations is assumed to represent the truth.

- Page 8, lines 32-33: "More than 10 gauges guarantee that we can obtain constant results regardless of the number of gauges." This is not true! The only evidence you show is that the average error will be lower than 20%.

3. Validity of power-law model:
The power law model proposed in Figure 6 is potentially interesting as one could imagine situations in which people need to estimate the required gauge density for achieving a certain accuracy. However, it should be pointed out that (a) it would be better to formulate it in terms of gauge density (#gauges/km2) and (b) that such a model needs to be properly validated/assessed. Right now, it is just given without any further evaluation or critical discussion. One big question is how well does the power-law model generalize to other cases. For example, what if I work with a catchment of 100 km2, that is, 3x times smaller than the area of the WegenerNet? Does this mean that I can divide the number of gauges by 3? I assume not since the autocorrelation varies faster over the first 10 km compared with the 10-20 km range. There is a lot of potential here for formulating a model that can be used by others. But this requires additional work.

4 Data transformation:
On page 4 line 26 it is said that a transformation $x \rightarrow \log(x+1)$ is applied to the data to avoid issues due to zeros. However, there are crucial details missing. For example, you need to be aware of the fact that due to the non-linearity of the log transform, the results of the correlation analysis will depend on the units in which x is expressed (e.g., using mm/h or mm). This is a common statistical fallacy and needs to be addressed more carefully. Moreover, it is still unclear to me how zeros are actually handled by the authors. Do you take all time steps (including the ones where all gauges record zeros) or only a subset? I asked for more details about this in the previous round of review but there hasn't been much progress on this aspect. Same for the anisotropy maps in Figure 5. You need to specify how zeros were handled and whether the maps show the values obtained after data transformation or before.

5 Selection of events:
It's not 100% clear from the text how the events in Section 4 were selected. The corresponding sentence on Page 6, lines-6-7 is not well formulated and open to interpretation. Please revise. Also, the selection seems to be done based on total rainfall accumulation, which tends to favor certain types of events (i.e., persistent and widespread), potentially providing biased results. Please comment on this in the paper.

Minor issues:

- Page 1, abstract: "these dense networks are only available at sub-pixel scales and over short periods of time". Too vague. Please explicitly state what pixels we are talking about. There are large differences in weather radar products and some of the latest X-band products have resolutions as high as as 30 meters.

- In Equation 1, I suggest to replace c1 by 1-c1. This would make the interpretation easier, making c1 the drop in autocorrelation (= the nugget) instead of the intercept. Higher c1 would then mean higher small-scale variability which makes more sense in my opinion.

- Figure 4: the units are missing

- Figure 4, the RMSE values (<0.01) you provide are misleading. They give the impression to the reader that the exponential fit is very good whereas the RMSE in the plot is the one that you calculated by taking the average value of the yearly-averaged autocorrelation values. However, individual fits (for a given year or a season) are not that good and the spread remains large, as shown in Figure 3.

- Figure 4: Most decorrelation distances shown here are much larger than the maximum observable size of the network (20 km). So how much do you trust these values? The text provides a small warning and a reference to a paper but I believe this warrants more attention than that. Having worked on spatial structures myself, I know that such large range estimates are often the result of bad fits or the choice of the model rather than physical. In any case, I think it's illusory to think that you can infer a 200 km range from data extending over 20 km. A better approach would be to limit the range of scales.

- There seems to be a confusion between the e-folding distance (which is a distance and should be in units of km) and the autocorrelation corresponding to the e-folding distance (which is unitless).

- Figure 9: it's almost impossible to see any difference in color here. Please adapt the scales.

- Section 4. It would be more useful to talk about gauges/km2 to avoid introducing too much dependence on the total area. In fact, the entire discussion should be centered around gauge density rather than the number of gauges. Also, other thresholds than 20% should be considered, as this is rather large in my opinion and is more characteristic of the accuracy one would like to achieve at the point scale.

- Page 5, lines 10-14: You mainly attribute the higher small-scale variability to solid or mixed precipitation here. Another explanation is that winter-type events are more heterogeneous and spatially disorganized than convective cases. Moreover, since they are lower intensity, the uncertainty affecting the measurements of two neighboring stations plays a bigger role. Bottom line: there are more explanations that can be given here and it's not clear from the evidence that you present that the higher small-scale variability is indeed attributed to snow and ice. Please rephrase.

Typos:

- Page 2, line 7: "On the contrary, gridded rainfall data are nowadays available..." The expression "On the contrary" does not appear adequate here.

- Page 2, line 34, "in order to contribute to the effort better and more broadly assessing the uncertainty". Bad English, please rephrase.

- Page 7, line 25: "decreases" instead of "deceases"

- Figure 3: there is a typo in the caption (separation)

- Figure 5: there is a typo in the caption (north-south)