Dear Authors,

The referees have raised several issues in their reports. You addressed most of them in a satisfactory manner in your replies. So I ask you to compile a revised version of your manuscript with the required changes made. Please also provide a version where the changes made are marked for easier checking and a single document where all of your responses (including those for the questions below) in a single document.

I would like to draw your attention especially to the following points which require a major revision. Some of them were already mentioned by the referees:

- The issue was raised about whether the raingauge observations are the best benchmark to compare the other methods against. Although the raingauge density in your study is quite high (~ 1 per 5 km²), this is true due to the points Dr. Mazzetti raised. Actually, you can reduce the problem of not knowing the true rainfall field by cross-validation: Split your gauge data set (even one station per 10 km² is still a good coverage), use one part to fuel the gauge-based estimation techniques, evaluate all methods on the remaining gauges or products derived thereof. This way you can rank all methods against a non-used data set. This does not remove the problem of a general bias of the raingauges (undercatch in case of wind etc.), but does not pre-favor gauge-based methods over the others. I realize that this requires substantial additional work, but this is important to make your study technically sound.
- Generally, you explore and present an interesting new avenue for multi-sensor based estimation of a true rainfall field for small spatial scales and high spatio-temporal resolution by aiming to overcome major limitations of existing merging techniques, which generally rely on Gaussian data distributions and smoothing/interpolation approaches. Your singularity-based approach takes singularities into account, but is relatively 'complex'. What I would really like to see, in the sense of a parsimonious approach to solve a problem, is an evaluation of how a standard merging technique (here the original Bayesian merging) performs with data that have been brought closer to a Gaussian distribution beforehand, e.g. by log-transform. This way, the major problem of smoothing away singularities should be mitigated without much effort.
- In your approach, you consider singularities in the radar image. However, singularities will also be present in your raingauge observations, e.g. if one gauge is exactly below a local convective cell or in a rainfall 'hole'. The singularity is of course hard to detect due to the sparse spatial sampling, but can still be there. Would it make sense to apply the singularity-removal also to the gauge data beforehand?
- The quality of the radar data play a major role in your merging approaches. It is therefore necessary to include a much more in-depth description of the radar data processing chain, operational corrections performed etc. For example, some of the low radar quality during high rainfall intensities could be attributed to path attenuation, if no attenuation correction is made.
- Having only four storms and using three of them for the calibration of the hydrological/hydrodynamic model, and doing calibration such that it favors all raingauge-derived rainfall estimation methods is a problem, but I can see the reason for this limitation from the limited duration of the observation period. Still I think things can be learned from the hydrograph-based discussion. I suggest to keep it in, but stress its limitations, and also reduce its extent in the paper and impact on the final conclusions. Please also include some performance statistics of the calibrated models (Nash etc) in 3.1.1. Also, please state in which temporal resolution the rainfall data were fed into the model. 2-min? 5-min? Was it different for the raingauge- and radar-based rainfall estimates?
- Some technical aspects

- Like Dr. Sinclair, I found it hard to fully understand the singularity approach (extracting and imposing the singularties) from the text. So I encourage you to go into more detail here as indicated in your response to Dr. Sinclair.
- Figure 7 is indeed too full. Consider distributing the timeseries over several plots (e.g. all SIN plots of an event into a separate plot), and/or to zoom into the most interesting parts of the hydrographs (peaks).
- Throughout the paper I found it hard to understand in which temporal aggregations the data were used and why: How were the 2-min raingauge and 5-min radar data attuned for combination or comparison of the instantaneous rain rates?
- Please include the location of the radar in your overview map in Fig. 2.

Yours sincerely,

Uwe Ehret