Author's reply to interactive referee comments C1480:

We want to thank the reviewer for an encouraging introduction and thorough questions/comments. For detailed answers, see below.

Replies to Anonymous referee (Referee #2, RC C1480)

Q1. First, I disagree with the sentence on lines 20-22 (page 7). A large contribution of the differences between radar and raingauges observation is due to the different nature of the two measures. Problems with radar errors should be minimized by the radar data preprocessing carried on at FMI, following the discussion in section 2.2, and also the raingauges are quality controlled. Are there statistical studies on the radar error reduction algorithms used at FMI?

Looking at the station distribution, it seems that most of the time only one raingauge is present in a 3x3 km radar cell: the authors should discuss to what extent this raingauge can be representative of this large area (see as an example Kitchen and Blackall, 1992 J. Hydrol. 134, 13-33).

Here it is not clear to us what page and lines the reviewer refers to. Probably the reviewer refers to the sentence (page 2461, lines 16-18): "As described above, accumulation estimates based only on radar data usually differs from gauge observation values due to radar errors (see Sect. 2.2) or problems with the gauges (Sect. 2.3)"

A1. We agree that our attempt to summarize the previous chapters lead us to write a sentence which is not exactly accurate. This will be reformulated and the new version is:

"As described above, in addition to sampling differences (new section reference), accumulation estimates based only on radar data can differ from gauge observation values either due to radar errors (see Sect. 2.2) or problems with the gauges (Sect. 2.3)."

There are no separate statistical studies on the FMI's radar error reduction algorithms, but the resulting quality is described in two references already added as a reply to another question:

Koistinen, J. and Michelson, D. B.: BALTEX weather radar-based precipitation products and their accuracies, Boreal Environ. Res., 7, 253–263, 2002.

Koistinen, J., Michelson, D. B., Hohti, H., and Peura, M.: Operational measurement of precipitation in cold climates, in: Weather Radar Principles and Advanced Applications, edited by: Meischner, P., Springer, Germany, 337 pp., 2003.

Indeed, in Finland the density of rain gauges is much smaller than the density of radar pixels. This is illustrated in table 1 below, which we prepared as reply to another question (see Q3), and will probably be included as an annex in the next version of the manuscript. In a ring of 50 kilometres around selected stations, we have on average 11 stations, while we have there 7850 radar pixels.

The network of rain gauges in Finland is low. Results from Goudenhoofdt and Delobbe (2009) showed that simpler merging methods were less sensitive to a low network density of gauges. Therefore this is one reason to first consider simpler assimilation methods, such as used in this article, in areas with sparse gauge network.

Q2. Second, the authors select seven ground stations to provide "independent" measures for validation of the techniques ensuring they are "representative of a characteristic Finnish climatological or physiographical areas." This choice should be more substantially justified. If the authors are interested in validating the performances of the techniques over different background, the results for the seven stations should be separately analyzed and discussed. If they just want to give overall results, probably the best options would be to randomly select a variable number of stations and carry on several tests, providing averaged error values and their variances.

A2. We admit that the statement of "representative of a characteristic Finnish climatological or physiographical areas" is too bold. We will remove it from the next version of manuscript and instead try to describe how the distance between stations assures they are not only independent from the analysis, but independent of each other. In selecting them we tried to catch different parts of Finland (and radar stations). However, since all the runs are performed using the operational system (i.e. results are used in end-users applications), we could not set more stations aside without risking the quality of the end product. As the reviewer already pointed out, the total number of gauge stations is already low, compared to radar pixels.

Q3. Other points to be discussed: why seven (over 447) stations where selected? which is the station density in the neighborhood of the selected stations ? how would change the performance of the techniques in areas with more coarse (or more dense) station distribution? how much the hourly precipitation of the independent

stations is correlated to the rates measured in the neighboring stations?

A3. Seven independent stations were selected due to operational usage of this product, see answer above in Q2. The station density around the independent stations varies a lot (more stations near Finland's capital city, less in the country side). On average, there are 11 surrounding stations, within a radius of 50 kilometres from the independent station point. The average distance to the nearest station is 9.8 kilometres. See Table 1 for details.

The dependence of station distribution is briefly described in page 2462, lines 11-14 and page 2468, lines 9-14. Also, there is a reference to Goudenhoofdt and Delobbe (2009) where the topic of sensitivity to network density is described in more details.

The above comments, by the reviewer, provide very interesting research topics. They would however lead to a different article and will therefore be considered in future research work and publications.

Station number	Station name	Road stations	Distance to nearest (km)	Weighing gauges	Distance to nearest (km)
2950	Fagerholm	4	8.5	3	36.9
2759	Porvoo	32	1.5	2	36.6
2770	Juupajoki	10	18.3	1	26.3
2778	Punkaharju	6	5.8	1	24.7
2915	Viitasaari	6	18.5	1	33.8
2852	Kokkola	8	7.8	0	56.5
2813	Rovaniemen	6	8	0	53

Table 1. Number of stations within 50 kilometres from the 7 independent stations. Porvoo is within 50 kilometres from Helsinki metropolitan area, hence the big number of road stations.

Q4. I also suggest a number of minor corrections:

Abstract. In the first line it is said that in this paper "four different methods used for combining radar data with precipitation gauge data

to produce while it seems the first method (LAPS-radar), is used as reference and does not use rain gauges data (line 4).

A4. This will be changed. For example:

"We investigate the appropriateness of four different methods to produce precipitation accumulation fields using radar data alone or combined with precipitation gauge data."

Q5. Section 2.2. I suggest to anticipate here the ground resolution of the radar maps.

A5. The "ground resolution" can be misleading, as LAPS is processing the radar data onto its own gridded coordinate system (which has the resolution of 3x3 km) thus benefiting from the larger resolution near radars and decreasing resolution with the distance from radars. However, the measurements in the FMI network have been designed to use the radar composite in Cartesian grid on 1x1 km. We will include this information in the text, in Sect. 2.2.

Q6. Page 6, line 14. What is the "standard Z-R relationship"? Please, specify.

Here is probably some confusion with page and line number. We assume the reviewer means the text (page 2460, lines 16-19) "This discrepancy is related to the use of the standard Z-R equation formula for all liquid precipitation cases, even though we know that drop size distributions vary from one precipitation case to another."

A6. This will be clarified in the revised article by including the reference (Marshall and Palmer, 1948) at page 2460, lines 16-19.

Q7. Page 7, lines 12-14. This is an important point, but the sentence is vague and the cited reference (Aaltonen et al., 2008) is difficult to reach. I suggest to quantify this "reasonable accuracy".

Again some confusion about page and lines. We assume the reviewer means page 2461, lines 5-10.

A7. Clarification to the text will be made by including following:

"In Finland the vertical profile of reflectivity (VPR) is the main source of bias in radar estimates of ground level accumulation at distances of more than 50 kilometres from the radar. Especially in winter, the underestimating bias of radars exceeds regularly 10 dB in the longest ranges. However, applying a range-dependent VPR correction or gauge adjustment, the bias of daily accumulation can be reduced, on average, to less than 1-2 dB at all ranges from the radar (Koistinen and Michelson 2002, Koistinen et al., 2003). Still after such corrections, due to major sampling differences between the two sensors, random errors remain at 2-3 dB, which is a typical, reasonable accurate figure in operational radar measurements (e.g. Collier 1986)."

Collier, C.G.: Accuracy of rainfall estimates by radar, Part I: Calibration by telemetering raingauges, J. Hydrol., 83, 207-223, 1986.

Koistinen, J. and Michelson, D. B.: BALTEX weather radar-based precipitation products and their accuracies, Boreal Environ. Res., 7, 253–263, 2002.

Koistinen, J., Michelson, D. B., Hohti, H., and Peura, M.: Operational measurement of precipitation in cold climates, in: Weather Radar Principles and Advanced Applications, edited by: Meischner, P., Springer, Germany, 337 pp., 2003.

Q8. Formula n. 8. In the definition of MAE, in the denominator, should be the absolute value of the difference.

A8. This will be corrected to be |Analysis-Gauge|

Q9. Figures. I suggest to use larger fonts for the labels.

A9. This will be corrected in revised article.

Q10. Figures 2, 3 and 4 are difficult to read, especially for low rainrates. The authors should try log-log scales or to use colors to resolve these values.

A10. We will improve the readability of these graphs by using another plotting tool, either by using shading or colors, and consider log-log scales.