Review of 'Impact of Multiple Radar reflectivity data assimilation on the numerical simulation of a Flash Flood Event during the HyMeX campaign' by Maiello *et al.*

General comments

I do acknowledge the efforts made by the authors to address my comments. However, some new elements have been introduced in the manuscript that call for clarifications:

- In my previous review, I suggested using confidence intervals to evaluate the statistical significance of the results. The authors have computed these confidence intervals, which add value to the manuscript. However, the results are presented in a cumbersome way. The authors provide 9 (!) tables with skill scores and associated confidence intervals. The authors should pick up salient features in these tables and explicitly refer to them in the text to help the reader (and convince him/her!; see some related comments here below, *which are not exhaustive*). In particular, the most striking (and not much discussed) feature is that nearly all figures have overlapping confidence intervals, which definitely call for cautious interpretations and justify backing up any conclusion carefully.
- The use of a dynamical thinning in relationship with the outer loop technique needs to be clarified (see my comment below). In the end, are there more or less radar data ingested with this technique?

Specific comments

- Section 2.1: It should be mentioned in the text that Figure 2 was produced with DEWETRA. Otherwise, the reader who overlooks Figure 2's caption does not understand why DEWETRA is introduced here.
- Section 3.1: The authors explain that 'volume reflectivity radar data, for each elevation, are projected onto the Cartesian plane in order to find the closest radar bin for each Cartesian grid point and then they are interpolated by the 3D-Var code of WRF'. This is still unclear to me. Does it mean that there is a radar observation assimilated at every model grid point (that of 'the closest radar bin')? What kind of interpolation is done by the 3D-Var code? In other words, the interesting (and missing) piece of information here is the *spatial resolution* of the observations.
- Section 3.1: The authors write: 'Moreover, no observation thinning is performed because this procedure is not yet developed into the 3D-Var system

for radar data. Nevertheless, a dynamical thinning has been devised that selects, for every assimilation cycle, the most influential partition of a particular measurement, from information based on the previous cycle: this is the multiple outer loops technique explained later in Section 4.' I have a different understanding of the outer loop technique. I understand that it is meant to update linearised operators (such as the observation operator) during the minimization process. As a consequence, more observations are assimilated with each iteration and the quality of the analysis is improved.

I do not see the relationship between the outer loop technique and thinning. The purpose of the latter is to counterbalance the use of an overly simplistic (ie, diagonal) observation error covariance matrix or to reduce the computational cost of the assimilation. Thinning actually results in reducing the amount of observations.

So what is the 'dynamical thinning [that] has been devised that selects, for every assimilation cycle, the most influential partition of a particular measurement, from information based on the previous cycle'? In the cited literature, Rizvi et al. (2008) pertains to the outer loop technique (in passing, it may be more appropriate to cite peer-reviewed articles such as Hsiao et al. 2012), and Liu and Rabier (2002) pertains to thinning, but no citation refers to both thinning and outer loop.

The sentence in question is almost a verbatim excerpt from Cardinali (2013, 2014)[1]. Does it mean that the authors used a dynamical thinning based on the influence matrix, which is the topic dealt with by Cardinali (2013, 2014)? In that case, they should add a reference to the technique they used or give more details about how it works. If this is related to the outer loop technique, the authors should formulate this relationship more explicitly.

- Section 5, comments on Table 4: Table 4 contains a lot of figures and the conclusions which are drawn from it are that the values are 'good' for ACC and FAR (which is expected when the considered events are rare) and that the experiments overestimate light precipitation. Is Table 4 really needed?
- Section 5: What message do the authors want to convey with the following sentence: 'MET indices in Table 5 suggest that CTL and CON_HR_12KM

[1] Compare (common terms are highlighted in bold face):

> Nevertheless, **a dynamical thinning** has been devised **that selects,** for **every assimilation cycle, the most influential partition of a particular measurement, from information based on the previous cycle**: this is the multiple outer loops technique explained later in Section 4. (Maiello et al. 2016)

and:

> In this case, **a dynamical thinning** can be thought/considered **that selects,** at **every assimilation cycle, the most influential measurement partition of a particular** remote sensing instrument, **from information based on the previous cycle** (see also Rabier et al., 2002). (Cardinali 2013, 2014, p 158)

have the widest spread between the CIs limits for higher thresholds'?

- Section 5, l 327: I do not understand how the conclusion that 'CONMMPOLSPC3OL_HR_12KM is the simulation with the best response' is reached. The score values for all experiments are quite close to each other and within the uncertainty intervals, and CONMMPOLSPC3OL_HR_12KM even scores lower than CTL for ACC(1 mm), FBIAS and ETS(1 mm).

- Section 5, ll 338-339: I do not understand that 'the frequency of rainfall overestimation for higher thresholds has been reduced by radar reflectivity assimilation performed only on D01'. For higher thresholds, FBIAS is systematically below 1, which means that the experiments *under*estimate the frequencies of large rainfall accumulations. The underestimations are even worse when radar reflectivity data are assimilated in D01 only: all FBIAS score values lie below .31 when radar reflectivity is assimilated, vs .47 and .49 for CTL and CON_HR_12KM, respectively.

- Section 5, ll 342-344: 'The assimilation, operated on both 12 km and 3 km, gives better results than the ones on column 1, but a worse response than the others on column 2 is given for higher thresholds.' Could the authors please back this up? It is far from straightforward to see it.

- Section 5, ll 378-380: How can shielded radar data lead to underestimating precipitation forecasts? I understood that they had been filtered out (see ll 184-185 'all the data that are affected by partial beam blocking and clutter have been filtered out').

References

Cardinali, Claudia (2013). *Observation influence diagnostic of a data assimilation system*. URL: http://www.ecmwf.int/sites/default/files/elibrary/2013/16938-observation-influence-diagnostic-data-assimilation-system.pdf.

— (2014). 'Observation influence diagnostic of a data assimilation system'. In: *Advanced Data Assimilation for Geosciences: Lecture Notes of the Les Houches School of Physics: Special Issue, June 2012*. Ed. by É. Blayo et al. Lecture Notes of the Les Houches Summer School. Oxford University Press. Chap. 5, pp. 137–163. ISBN: 9780198723844. URL: https://books.google.it/books?id=RSXSBAAAQBAJ.

Hsiao, Ling-Feng et al. (2012). 'Application of WRF 3DVAR to Operational Typhoon Prediction in Taiwan: Impact of Outer Loop and Partial Cycling Approaches'. In: *Wea. Forecasting* 27.5, pp. 1249–1263. ISSN: 1520-0434. DOI: 10.1175/waf-d-11-00131.1. URL: http://dx.doi.org/10.1175/WAF-D-11-00131.1.

Liu, Z.-Q. and F. Rabier (2002). 'The interaction between model resolution, observation resolution and observation density in data assimilation: A one-dimensional study'. In: *Quart. J. Roy. Meteor. Soc.* 128.582, pp. 1367–1386. DOI: 10.1256/003590002320373337.

Maiello, Ida et al. (2016). 'Effects of Multiple Doppler Radar data assimilation on the numerical simulation of a Flash Flood Event during the HyMeX campaign'. In: *Hydrology and Earth System Sciences Discussions*, pp. 1–25. DOI: 10.5194/hess-2016-320.

Rizvi, Syed R. H. et al. (2008). 'Impact of outer loop for WRF data assimilation system (WRFDA)'. In: *9th WRF Users' Workshop*. (Boulder, CO, 23rd–27th June 2008).