Responses to referee#1

General comments

First of all we are grateful to anonymous referee for the great contribution to the manuscript coming from useful comments.

In this study we wanted to take advantage of the numerous Italian Intensive Observing Periods (IOPs) that affected the three Italian Target Areas (TAs) during the First Special Observation Period (SOP1) of the HyMeX campaign, but above all Central Italy (CI). Later, the choice fell on the IOP4 first of all because all the instruments activated was very successful (radar, sodar and microwave sensors were on alert in the Central Italy site from the evening Thursday 13 until Saturday 15 September 00UTC; extra operational soundings were performed on 13 September 18 UTC, 14 September 12 and 18 UTC in L'Aquila) and secondly it was a very interesting case with convective cells producing a remarkable amount of precipitation in a few hours (more than 150 mm) over Central Italy (Coastal Marche and Abruzzo) with precipitation peak of 300mm/24h. The event was quite well forecasted by all models operational during the campaign well in advance, but uncertainties remained until a few hours before the event regarding the exact location and amount of precipitation. On the other hand, we didn't find another Italian IOP, among those that have affected Central Italy, with so many radars activated simultaneously to enrich the analysis (for example during the IOP13 Monte Midia radar was out of service, whereas during the IOP16 Polar 55C was affected by some technical problems).

Concerning the novelty we claim in the paper, we know that many topics addressed in the manuscript have been already mentioned in previous studies, but except for Maiello et al. 2014, it is the first Italian experiment conducted on the Italian territory using the data of the Italian radars.

We are aware that constraining the analysis to maps of quantitative precipitation forecasts and relating scores could be a limit, but it was our choice to analyze the most important variable in a flash-flood event and to aim for the hearth of hydro-meteorological research. Nevertheless, we accept the advice to go deeply into the meteorology of the event to see which is its interaction with the data assimilation method.

We hope that the organization of the paper is now improved: section 2.2 has been moved after the presentation of the model configurations; section 4.1 has been shrink to few sentences and figures 6 and 7 have been removed; a table that summarizes the characteristics of the radars has been added. Moreover, several English mistakes have been corrected, the literature review has been updated and the quality of some figures has been improved. Also the title and the abstract have been modified.

Specific comments

Line 1: The word "Doppler" has been deleted and the title has been modified as follow: "Impact of Multiple Radar reflectivity data assimilation on the numerical simulation of a Flash Flood Event during the HyMeX campaign"

Line 16: The selected case study was tagged both as a Heavy Precipitation Event (HPE) and a Flash Flood Event (FFE). For this study we took advantage from all the instruments successfully activated during the event, with the aim of improving the forecast and alerting civil protection well in advance. In summary the objective here was to build a regionally-tuned numerical prediction model and decision-support system for civil prevention and protection within the central Italian regions. Moreover, the additional purpose is to find which type of observations (or a combination of several types) is more effective in improving the accuracy of the forecasted rainfall.

The sentence here has been modified as follows: "The main goal is to build a regionally-tuned numerical prediction model and decision-support system for civil prevention and protection within the central Italian regions, distinguishing which type of observations (or a combination of several types) is more effective in improving the accuracy of the forecasted rainfall."

Lines 31-34: We agree with the reviewer. The sentence has been modified as follows: " Nevertheless, the accuracy of the mesoscale NWP models is negatively affected by the "spin-up" effect (Daley 1991) and is mostly dependent on the errors in the initial and lateral boundary conditions (IC and BC), along with deficiencies in the numerical models themselves, and at the resolution of kilometers even more critical because of the lack of high resolution observations, beside for radar data."

Line 53: The references Ducrocq et al. 2014, Ferretti et al. 2014 and Davolio et al. 2015 have been added here.

Line 71: The sentence has been modified as follows " During the day of 14 September 2012 .... "

Line 78: A reference for DEWETRA has been added both in the text and in the references list.

Lines 102-111: A table that summerizes the characteristics of the three radars has been added and lines 102-111 have been rewrited as follows: " Volumetric reflectivity taken from three C-band Doppler radars operational during the IOP4 have been assimilated to improve IC. Radars have different technical characteristics and were operated with different scanning strategies and operational settings as shown in Table 1. Monte Midia (MM) and San Pietro Capofiume (SPC) radars are included in the Italian radar network, while Polar 55C (P55C) radar is a research radar working on demand which was operational during HyMeX IOPs (Roberto et al., 2016)."

We consciously decided to assimilate only reflectivity data, probably the term "Doppler" in the title was misleading (it has been dropped). A high quality of Doppler velocity is required for assimilation. However, quality of available data, especially due to the need of correct for aliasing was not suitable for assimilation in the case of the considered event. Therefore we have preferred assimilating only reflectivity.

Lines 112-119: Reflectivity data were quality controlled before ingested into the 3DVAR. However, an observation thinning before the minimization to avoid as much as possible error correlations between adjacent pixels is not performed because this procedure is not yet developed into WRFDA 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! (Cardinali et al. 2004, "Influence matrix diagnostic of a data assimilation system", Q. J. R. Meteorol. Soc., 130, 2827-2849). Indeed, the experiments performed using different numbers of outer loops allowed to compare the impact of a small sub-group of very influential data (i.e. radar observations, experiments with 3OL) on the forecast as the full amount of data. As future development, a thinning of radar data has to be undertaken either to reduce the observation-error spatial correlation or the computational cost of the assimilation (Montmerle and Faccani, 2009).

Concerning the data conversion to the model format, conventional and radar observations are treated in a different way. Conventional observational data are converted in LITTLE_R format using the Observation Preprocessor (OBSPROC) program provided by WRFDA system. The purposes of OBSPROC are to:

- Remove observations outside the specified temporal and spatial domains
- Re-order and merge duplicate (in time and location) data reports
- Retrieve pressure or height based on observed information using the hydrostatic assumption
- Check multi-level observations for vertical consistency and super adiabats
- Assign observational errors based on a pre-specified error file

- Write out the observation file to be used by WRFDA in ASCII or BUFR format

For what concern radar data, an ad hoc shell script in Fortran language has been written and adapted to each radar characteristics to perform conversion to the model format (more details about this have been added in the text).

Line 130: We agree with the reviewer. The sentence has been modified as follows: "a one-way nested configuration using *ndown* program is used"

Lines 150-152: We agree with the reviewer. The sentence has been modified as follows: " Data assimilation (DA), which applications arise in many fields of geosciences perhaps most importantly in weather forecasting and hydrology, in this context is the procedure by which observations are combined with the product (*first guess* or *background forecast*) of a NWP model and their corresponding error statistics to produce a bettered estimate (the *analysis*) of the true state of the atmosphere (Skamarock et al., 2008).

Line 162: The word "fonts" has been replaced by "sources".

Line 165: Pseudo relative humidity and total water mixing ratio are both control variables for the analysis of moisture observations in a global atmospheric data assimilation system. In a variational framework, the choice of control variable is important because the notion of “distance” between model and observations depends on it. A pseudo-relative humidity can be defined by scaling the mixing ratio by the background saturation mixing ratio. A pseudo-relative humidity analysis is shown to be equivalent to a mixing ratio analysis with flow-dependent variance specifications. The “pseudo” relative humidity is the water vapor mixing ratio divided by its saturated value in the background state.

Line 171: The microphysics scheme used is the New Thompson (Thompson et al., 2004). This scheme adopted a generalized gamma distribution shape for each hydrometeor species. The observational operator, on the other hand, uses the more simple Marshall and Palmer DSD which is an exponential one. This is a simplified gamma distribution, assuming 0 as exponent for the drop diameter. The main differences between the two DSDs are bounded where coalescence and evaporation processes and break-up process are active; these are the smallest and biggest drops region, i.e. the tails of the DSD. The difference introduced using these two DSDs plays a minor role respect to other errors like for example time and position shift.

Lines 200-205: The experiments names in the text and in table 2 are now consistent. The acronyms "LR" and "HR"mean low and high resolution respectively, in the sense that in the first case D01 is showed, D02 in the second case.

Lines 221-241: We agree with the reviewer. Section 4.1 has been rearranged as follows and figures 6 and 7 have been removed: "From the sensitivity test to different cumulus parameterization scheme (Table 2) the best performance is obtained by Grell3D scheme which is able to simulate the peak precipitation cumulated in 24 hours over Campo Imperatore, whereas KAIN-FRITSCH completely misses it (not shown here). The MET statistical analysis support the previous finding and the simulation with *cugd_avedx* activated shows higher performances in terms of accuracy, equitable threat score and false alarm ratio than the other two simulations. Here after GRELL3D_MYJ_CUGD is referred as the control (CTL) experiment performed without any data assimilation. Therefore, a new set of simulations are performed following the previous strategies already mentioned in Section 4."

Lines 251-253: The statistical indexes have been calculated using the pointstat tool of MET (as reported in the lines 210-214). The MET Guide (Developmental Testbed Center, 2013: MET: Version 4.1 Model Evaluation Tools Users Guide. Available at http://www.dtcenter.org/met/users/docs/overview.php. 226 pp.) reports more details about the calculation of the statistical indexes. The reference will be added also in lines 251-253. The Fig. 9 (now Fig.7) reports the statistical indexes for the 12 hours accumulated precipitation. The 12 hours accumulations have been calculated from the 2012-09-14 12:00:00 to 2012-09-16 00:00 every 6 hours, i.e. the scores of 2012-09-14 12:00:00 refers to the accumulated precipitation from 2012-09-14 00:00:00 to 2012-09-14 12:00:00. The word MEAN(9) on the title refers to the interpolation method used to match the gridded model output to the point observation. In details for this study the distance weighted mean in a 3 x 3 square has been used. The scores reported in Fig.9 have been averaged all over the data points belonging to the same threshold in the simulated time range.

Lines 251-253: The results will be tempered to make these clearer.
Fig. 1-2 report the histograms for the 12 hours accumulated precipitation. As you can see the first bin, including the precipitation lower than 10 mm/12h, is the most populated with approximately 20000 data points (Fig.1).

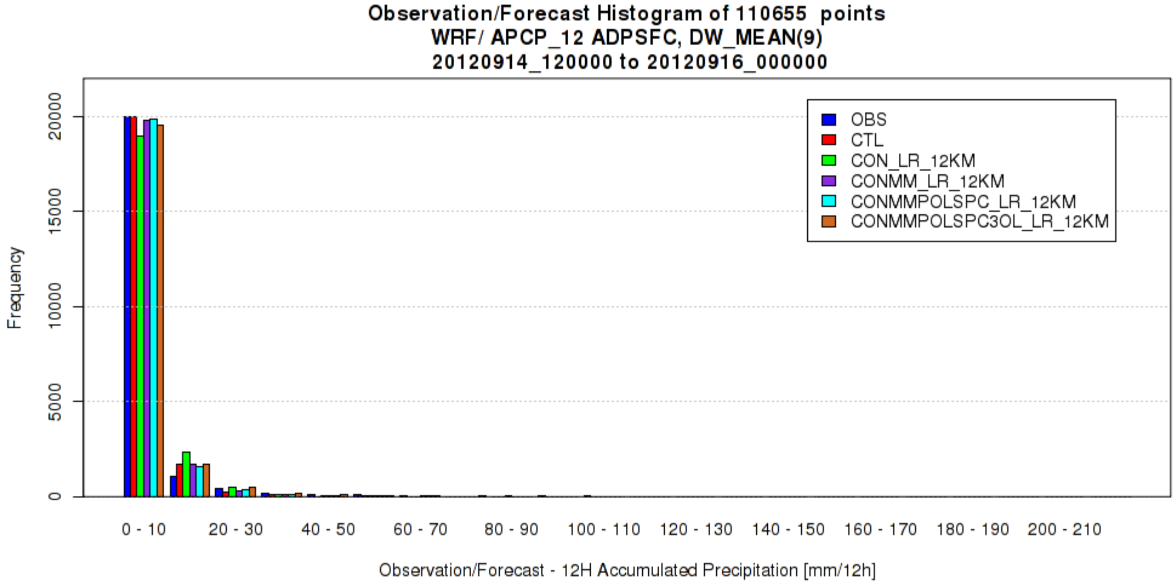

Fig 1. The histogram for the accumulated precipitation.

The Fig. 2 reports the same histogram removing the first bin to show how is crowded the following bins. The bin including the precipitations from 40 to 50 mm/12h has approximately 200 data points.

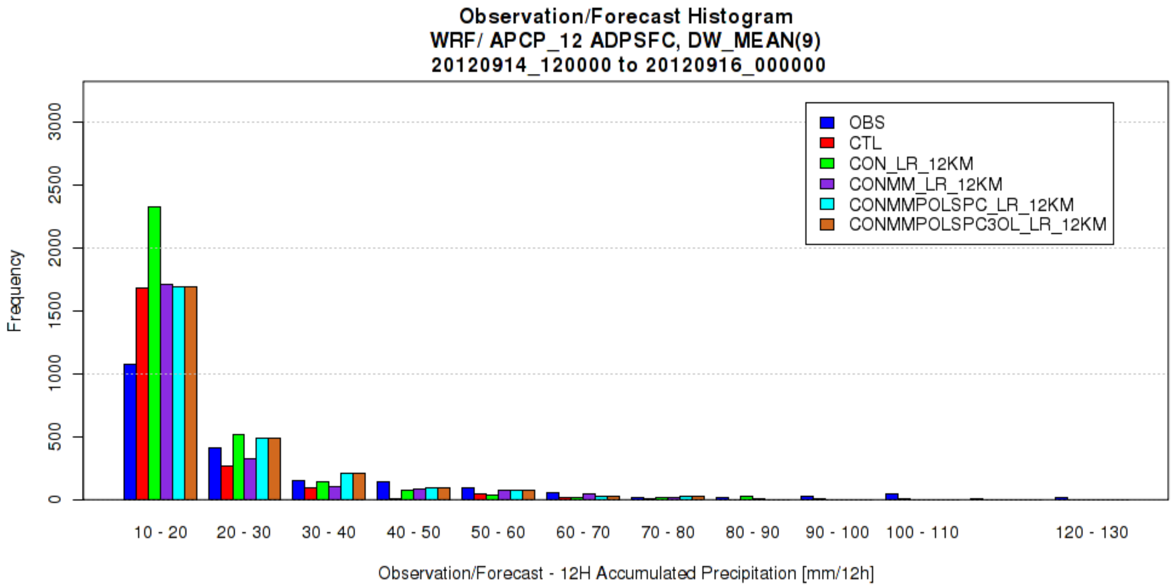

Fig 2. Zoom of the histogram for the accumulated precipitation (the first bin has been removed).

Lines 277-280: We found that when the assimilation is performed on the highest resolution domain only few SYNOP and even less TEMP fell down in the 3km domain at the analysis time of the assimilation procedure. For example after applying the WRFDA Observation Preprocessing procedure only a total of 338 observations (331 SYNOP and 7 TEMP) have been ingested into the D02 (Italy), compared to a total of 989 (967 SYNOP and 22 TEMP) into the D01 (Europe).

Lines 307-309: Since the three radars are managed by different organizations, a different radar data preprocessing procedure is followed and it depends on the case study.
Reflectivity is not corrected neither for total nor for partial beam blocking; nevertheless, all the data that are affected by partial beam blocking and clutter have been filtered out. In a future operational context, we could think to harmonize the processing of the three radars in order to achieve a spatially uniform quality.

Lines 336-337: We are aware that the assimilation of radar data is already operational at several meteorological services, but not in Italy. The Center of Excellence Cetemps (Abruzzo, Italy) is the only meteorological center in Italy that has radar data assimilation in operational mode since 2013, together with SYNOP and TEMP.

Lines 392-393: The reference has been updated.

Line 397: The reference has been corrected. Moreover, all the references have been checked both in the text and in the list.

Fig.1: The quality of figure 1 has been updated; a description of the meaning of isolines and colour shades has been added in the caption. The model used is WRF and the graphical tool GRADS.

Fig.2: The coloured circles represent the warning pluviometric thresholds as follows: 

The legend has been added in figure 2.

Fig.3: Figure 3 has been updated with units and scale.