Answer to the reviewer #2 comments on NHESS-2018-319.

We acknowledge the reviewer for the useful comments on the paper, both in the general comments section and in the pdf file. Our answers are in red.

Before starting the discussion, we note that, in reviewing the paper, we found two errors: a) the length scale of the background error matrix in the x and y directions varies between 14 and 25 km and not, as stated into the paper, between 20 and 30 km; second the lightning number for each day written into the initial manuscript are wrong. The correct numbers are 82 331 for the 9 September, 291 164 for the 10 September (170 000 is written into the manuscript) and 105 467 for the 16 September (60 000 written in the manuscript). We apologize for these errors. However, the results shown in the paper were obtained using the correct number of flashes and the correct length scales in the background error matrix.

Extracted from the general comments:

This manuscript addresses an interesting and challenging topic, moreover represents a substantial contribution to the understanding of natural hazards and their consequences matching the scope of NHESS. However, the scientific and presentation quality are poor, above all because the results are presented in a "repetitive and heavy" manner and the English language needs a deep revision.

Specific comments: please see the notes on the pdf for each section and also for each figure's caption, moreover please deeply motivate the reason why: a) radial velocity is not assimilated: the operator is not implemented or the data are not available?; b) data assimilation is performed in the domain D02 only; c) you used a background error matrix of fall 2012 for case studies of late summer 2017. Technical corrections: please see the notes on the pdf for each section and also for each figure's caption.

Comment "…in a repetitive and heavy manner". In the first submission of the paper we stressed the important improvement given by the data assimilation at the local scale on the precipitation VSF (Very Short term Forecast, 0-3h). To highlight this point, we showed the many ways in which the forecast is improved by the assimilation of lightning, radar or both. For example, the two stages of the Serano case show that the radar (first phase 03-06 UTC on 16 September) or lightning (second phase of the event, 18-21 UTC on 16 September) were the key observation to assimilate in order to improve the precipitation VSF. Also other stages had some specific aspects that we discussed.

Our attempt, however, was not successful, given the comments of both reviewers and the results section (Section 4) underwent a substantial rewriting.

In particular, in the revised version of the paper, we will delete the Section 4.1.2 (second phase of the Serano case) and Section 4.2.1 (first case of the Livorno case). The results Section 4.2.1 will be shortly commented in Section 5 (Discussion and conclusions) to highlight that there is space for improvement. Following the comments of the reviewer #2, the scores of the phases discussed into the paper will be put in three tables (Tables: 4-6) for specific thresholds (1, 6, 10, 20, 30, 40 mm/3h and, for Livorno, also 50 mm/3h) and, following the remarks of the reviewer #1, different neighborhood radii will be used to compute ETS and POD scores.

The space gained by deleting the two sub-sections stated above will be used to extend the discussion about the methods of assimilating lightning and radar in the RAMS@ISAC model and to add two short sections to the result section. In particular, we will extend the section "Lightning data assimilation" to include a discussion of the useful comments raised by the Reviewer#1, we will extend the section "Radar data assimilation" to show an example of 3D-Var assimilation of reflectivity factor (this will also answer to few comments of the Reviewer #1). A draft of these revised sections will be reported at end of the answer to the Reviewer #1 comments.

Finally, we will add a section (Section 4.3) to show how the lightning and radar data assimilation works together, presenting the evolution of the total water averaged for all VFS of the two cases and including in this discussion the assimilation stage, as well as sensitivity tests of the precipitation scores (POD and ETS) to the nudging formulation (Section 4.4). Section 4.4 requested new simulations with different model settings (see Table 3 at the end of this answer).

A draft of the new Results section (Section 4), in clouding the new sections 4.3 and 4.4, is shown at the end of this answer. This could not be the final form because minor changes are still possible.

Comment "…the English language needs a deep revision". We will revise the English of the paper, also according to the suggestions of the reviewer #2 in the PDF file. Also, the copy-editing service of the journal will improve the quality of the English before the eventual publication of the paper.

Specific question a) We are working on the assimilation of the radial velocity but the operator is not yet implemented in the 3D-Var. Also, while the reflectivity factor measured by the radar network is operationally available, the product of radial velocity is under development. At the moment, it needs further research to solve some issues (complex orography, operations of the radars not optimal for the Doppler retrieval and others). For these reasons, the attention was on the assimilation of radar reflectivity factor. These motivations will be discussed in the revised version of the paper in Section 3.3 by writing:

"Radial velocity is not assimilated in the RAMS@ISAC model because the operational product of radial velocity needs research to solve issues (complex

orography, operation of radars non optimal for Doppler retrieval, not homogeneous coverage of the country), and it is not available for assimilation. Also, the implementation of radial velocity data assimilation is under development in RAMS-3DVar and it is not available for testing. For these reasons, we didn't consider the assimilation of radial velocity in this work. "

Specific question b) Data assimilation is not performed on domain D3 (R1) because we don't have background error statistics for this grid . The background error statistics are important in 3D-Var and, while radar data (we apply 3D-Var for radar only) are cloud-scale observations and could be used for the 3D-Var assimilation in domain D3, we cannot apply it on this domain.

Background error statistics for the domain D2 are computed by the NMC method, which, for this paper, is based on HyMeX-SOP1 simulations. The Appendix A and B of Federico (2013) shows the details. of the application of the method, which requires a number of simulations (see Barker et al., 2004 for the general discussion). These simulations are not available for the innermost grid of the Livorno case, which was introduced to better resolve the precipitation at the local scale and to show how precise can be the impact of the lightning and radar data assimilation on the VSF. These motivations will be clarified in the revised version of the paper.

Of course, this limitation is only for radar reflectivity factor because flashes are assimilated by nudging. Nevertheless, we could not compare simulations with or without data assimilation for a specific domain assimilating lightning in the innermost domain and for this reason we assimilated flashes over the D2 only.

In the paper we will specify better the role of the domain D3 and the reason for not assimilation lightning and radar reflectivity factor over the domain D3.

We will write in section 3.1 (RAMS@ISAC and simulations set-up)

"The third domain covers the Tuscany Region, has 4/3 km horizontal resolution (R1), and it is used for Livorno to represent with higher spatial detail the precipitation field over Tuscany and to show better the precision of the rainfall VSF using data assimilation at the local scale. The fine structures of the precipitation field are smeared out over Tuscany using only domains D1 and D2. The operational implementation of the RAMS@ISAC model uses the domains D1 and D2 and no refinement for specific areas of Italy are used because Italy is a complex orography country and grid refinements for a specific event can be done only a-posteriori, i.e. after the occurrence of the event."

And few lines below:

"It is noted that data assimilation is performed in the domain D2 (R4) only, and the innovations are transferred to the domain D3 (R1), for the Livorno case, by the two way-nesting. The domain D3 is used for the Livorno case to refine the resolution of the precipitation field over Tuscany and to show the spatial and temporal precision of the precipitation forecast over Tuscany using data assimilation. However, its usage is exceptional because, as stated above, Italy is a complex orography country and grid refinements over specific areas are used only after the occurrence of an event. For these reasons the domain D3 is usually not used in RAMS@ISAC and statistics about the background error aren't available for this grid. The background error in RAMS-3DVar is computed by the NMC method (Parrish and Derber, 1992), which requires a number of simulations (at least two-weeks) verifying at the same time but starting with a lag of 12 h. These simulations are not performed in this paper and background error statistics for the domain D3 are not available. Of course, being lightning assimilated by nudging, they could be assimilated over the domain D3. Nevertheless, to preserve the rationale of the paper, i.e. comparing simulation with or without data assimilation for specific domains, we didn't assimilate lightning over the domain D3.

Of course, being lightning and radar cloud scale observations, their assimilation at higher horizontal resolution is foreseeable in future works. "

Specific questions c) We chose the background error matrix computed for HyMeX-SOP1 because the period was characterized by several convective events over Italy, as documented in Ferretti et al., (2014), while the period preceding the convective events of this paper was characterized by fair weather, typical of the summer Mediterranean season. For this reason, we believe that the matrix for the HyMeX-SOP1 is more representative of convective events compared to the matrix computed for the period of the storms occurrence. We will write:

"The background error matrix is computed using the NMC method (Parrish and Derber, 1992; Barker et al. 2004) applied to the HyMeX-SOP1 (Hydrological cycle in the Mediterranean Experiment – First Special Observing Period occurred in the period 6 September-6 November 2012; Ducroq et al., 2014). This choice is motivated by the fact that HyMeX-SOP1 contains several heavy precipitation events over Italy and the background error matrix is representative of the convective environment of the cases considered in this paper. In particular, 10 out of 20 declared IOP (Intense Observing Period) of HyMeX-SOP1 occurred in Italy (Ferretti et al., 2014). On the contrary, the period of September 2017, before and after the events selected in this study was characterized by fair and stable weather conditions over Italy and the background error matrix for September 2017 is less representative of the convective environments that characterise the events of this paper."

PDF file with technical corrections:

Considering the pdf file, all corrections will be accepted. There are, however, few points that need a short discussion. They are listed below.

Elements of novelty of the paper: we will highlight better the elements of novelty of this paper in the introduction. We will write:
"This paper presents for the first time the assimilation of the radar reflectivity factor in the RAMS@ISAC model and shows how the assimilation of the radar reflectivity factor works together with lighting data assimilation. Also, this paper shows how accurate in space and time can be the forecast of the precipitation field using cloud scale observations over complex terrain, contributing in this way to a number of works on the same subject."

Comment on Line 276: The frequency bias was not shown to keep the discussion concise. However, important to point out was that the model had a wet bias, especially when assimilating radar reflectivity factor. For this purpose, we introduced the score discussing it to highlight that the model has a wet bias, nevertheless we didn't show any graphs to keep the discussion more concise. In the revised version of the paper the wet bias of the model will be highlighted better and a discussion on how to reduce the wet bias will be added in the "Discussion and conclusions" section (Section 5).

In particular, at the end of Section 4.4 we will write:
"It is finally noted that RAD and RADLI have high POD values for all thresholds, nevertheless their ETS is below that of LIGHT and SAT up to 32 mm/3 h (RADLI) and 42 mm/3h (RAD). This behaviour is caused by the larger number of false alarms given by assimilating radar reflectivity factor compared to those assimilating lightning. This result shows again that the RAD and RADLI configurations have a wet frequency bias. In particular, the frequency bias of RAD and RADLI configuration is about 3 between 20 and 40 mm/3h."

and in the "Discussion and Conclusions " section:

"The wet bias of RAD and RADLI forecast is the main drawback of the results of this paper. To reduce the moisture added by radar and lightning data assimilation further research is needed and different approaches are possible (Fierro et al., 2016). In particular: a) assimilating for a shorter time period (0-6h in this paper); b) reducing the length-scales of the 3D-Var in the horizontal directions to limit the spreading of the innovations (or assuming an innovation equal to zero for grid points without lightning and with zero reflectivity factor observed and simulated); c) reducing the amount of water vapour added to the model (for example reducing the values of A and B constants for lightning data assimilation or relaxing the request od saturation when radar reflectivity is observed in areas where the model has zero reflectivity); d) adding moisture to a shallower vertical level.

It is also noted that a combination of heating and moistening could provide the same buoyancy with less water vapour addition (Marchand and Fulberg, 2014)."

Line 306: There was a typing error into equation 2. The correct equation is:

$$q_v = Aq_s + Bq_s \tanh(CX)(1 - \tanh(Dq_g^{\alpha}))$$

and q_g is in the last term. We also note that the number of lightning assimilated by the model is larger than that reported in the first submission of the paper for both cases for a mistake we did in typing the numbers. The correct number of assimilated flashes is reported at the start of this answer. The results, however, were obtained using the correct number of flashes.

Line 314: "The check and eventual substitution of the water vapor is performed every five minutes and it is made only in the charging zone (0 °C, -25°C)." To better explain this choice, we will reformulate it adding references to the charging zone:

"The check and eventual substitution of the water vapor is performed every five minutes and it is made within the mixed phase layer zone (0 °C, -25°C), wherein electrification processes are the most active (Takahashi 1978, Emersic and Sounders, 2010; Fierro et al., 2015). It is also noted that some authors use the layer (0 °C, -20°C) (Fierro et al., 2012; 2015), however specific tests for the case studies of this paper did not show significant changes of the simulated precipitation to this choice."

Line 314: the comment is: could you add some more details about this quality control?
We will add two references. We will write:
"The processing chain aims at identifying most of the uncertainty sources as clutter, partial beam blocking and beam broadening. The radar observations are processed according to nine steps detailed in Vulpiani et al. (2014), Petracca et al. (2018) and references therein."

In the following we show the new section Results (Section 4). Note that sections 4.3 and 4.4 are completely new. Note also that minor changes are possible before the eventual submission of the revised version of the paper.

4. Results

In this section, we discuss the most intense phase of the Serano case, 03-06 UTC on 16 September, and two VSF forecasts, 00-03 UTC and 06-09 UTC on 10 September, for the Livorno case. The two VSF for Livorno correspond to the most intense phase of the storm in Livorno and to a very intense phase over Lazio region, Central Italy. The aim of the section is to show the notable improvement given to the very short term forecast by the lightning and radar reflectivity factor data assimilation. In the discussion paper two additional VSF are discussed, one for Serano and one for Livorno; also,

the discussion paper shows the behaviour of the scores for a number of rainfall thresholds larger than those shown in this section.

In particular, we consider four types of VSF: a) CTRL, without radar reflectivity factor and lightning data assimilation; b) LIGHT, assimilating lightning but not radar reflectivity factor; c) RAD, assimilating radar reflectivity factor but not lightning; d) RADLI, assimilating both lightning and radar reflectivity factor. A_76 and SAT show the sensitivity of the results to the nudging formulation. Table 3 shows the types of simulations considered in this paper.

4.1 Serano: 03-06 UTC 16 September 2017

In this period, an intense and localised storm hit the central Italy, while light precipitation occurred over northern Italy (Figure 15a). Considering the storm over central Italy, 10 raingauges observed more than 30 mm/3h, 6 more than 40 mm/3h, 3 more than 50 mm/3h and 1 more than 60 mm/3h, the maximum observed value being 63 mm/3h.

The CTRL forecast, Figure 15b, misses the storm over central Italy and considerably underestimates the precipitation over Northern Italy, giving unsatisfactory results.

The assimilation of the radar reflectivity factor improves the forecast, as shown by Figure 15c. In particular, RAD forecast shows localized precipitation (30-35 mm/3h) close to the area were the most abundant precipitation was observed. However, the maximum precipitation is underestimated. Also, the RAD forecast better represents the precipitation over Northern Italy compared to CTRL.

The precipitation forecast of LIGHT, Figure 15d, shows some improvements compared to CTRL because the precipitation over central Italy has a maximum of 25-30 mm/3h, close to the area where the maximum precipitation was observed. LIGHT, however, has a worse performance compared to RAD because it misses the small precipitation amount over northern Italy. Also, similarly to RAD, LIGHT underestimates the maximum precipitation.

RADLI forecast, Figure 15e, shows the best performance. The precipitation over central Italy is well represented because the maximum rainfall (40-45 mm/3h) is in reasonable agreement with observations, and also because the area with intense precipitation (> 25 mm/3h) is elongated in the SW-NE direction in agreement with raingauge measurements, giving a much better idea of the real storm intensity compared to RAD and LIGHT, as well as CTRL. The precipitation over northern Italy is well represented by RADLI.

Table 4 shows the ETS and POD scores for selected rainfall thresholds for different neighbourhood radii. Different radii are considered to account for the well-known double penalty error (Mass et al., 1996; Mittermaier et al., 2013) caused by displacement errors of the detailed precipitation forecast in convection allowing grids.

CTRL was unable to predict rainfall larger than 6 mm/3h. The comparison between RAD and LIGHT shows that assimilating radar reflectivity factor performs better than assimilating lightning.

This behaviour, however, is not general and sometimes the assimilation of lightning has a better performance than assimilating radar reflectivity factor (see section 4.2.1).
RADLI forecast has the best performance among all model configurations. In particular, it is the only forecast having positive scores for thresholds larger than 30 mm/3h.
In conclusion, for this VSF, the assimilation of lightning and radar reflectivity factor acted synergistically to improve the precipitation VSF and the simulation assimilating both data performs considerably better than simulations assimilating either lightning or radar reflectivity factor.

4.2 Livorno

The Livorno case lasted for several hours starting at 18 UTC on 9 September 2017 and ending more than a day later. The most intense phase in Livorno and its surroundings was observed during the night between 9 and 10 September. In the following, we will show two representative VSF (3h), including the most intense phase in Livorno.

4.2.1 Livorno: 00-03 UTC 10 September 2017

This period represents the most intense phase of the storm in Livorno. In particular, the raingauge close to the label A (Figure 16a) reported 151 mm/3h (Collesalvetti), while the one close to the label B measured 82 mm/3h. Among the 518 raingauges reporting valid data, 75 observed more than 10 mm/3h, 31 more than 20 mm/3h, 17 more than 30 mm/3h, 9 more than 40 mm/3h, and 6 more than 50 mm/3h.
The CTRL precipitation forecast is shown in Figure 16b. The forecast is poor because it misses the precipitation swath from the coast towards NE. A precipitation swath is forecasted about 50 km to the North of the real occurrence, but it is less wide compared to the observations.
The forecast of RAD, Figure 16c, shows that the assimilation of radar reflectivity factor gives a clear improvement to the forecast. The largest precipitation in the coastal part of the swath (we searched the maximum value in the area with longitudes between 10.20E and 10.70E and latitudes between 43.10N and 43.60N) is 94 mm/3h. Another local maximum is in the southern part of the domain (label B). The maximum location is well represented, but the forecast value is 55 mm/3h while the observed maximum is 82 mm/3h.
An improvement, compared to both CTRL and RAD, is given by the assimilation of lightning (Figure 16d). Also for this simulation there is a precipitation swath from coastal Tuscany to the Apennines, but the shape of the swath better resembles that observed. The maximum value close to Livorno, i.e. in the coastal part of the swath, is 158 mm/3h.
The LIGHT simulation also shows the local maximum in the southern part of the domain (about 50 mm/3h), but the amount is underestimated.

Figure 16e shows the rainfall forecast by RADLI. The precipitation swath from coastal Tuscany towards NE is more apparent compared to LIGHT and RAD. The maximum rainfall accumulated close to Livorno is 186 mm/3h. Also, the second precipitation maximum in the southern part of the domain reaches 70 mm/3h in good agreement with observations (82 mm/3h). RADLI is the only run giving a satisfactory precipitation field over the south-eastern Emilia Romagna (north-eastern part of the domain), on the lee of the Apennines.

It is also noted that the main precipitation swath forecasted by RADLI is too broad in the direction crossing the swath compared to the observations. This is confirmed by the FBIAS of RADLI (not shown), which is more than 3 for thresholds larger than 42 mm/3h.

The analysis of the scores (Table 5) confirms the results outlined above. CTRL has the lowest performance and the improvement given by the data assimilation to the VSF is apparent for POD and ETS for all thresholds and neighbourhood radii considered. For this specific VSF, lightning data assimilation gives a better improvement to rainfall forecast compared to RAD. RADLI has the best performance, especially for 25 km and 50 km neighbourhood radii, nevertheless it over forecast the precipitation field (Figure 16). Because ETS penalizes false alarms, the value of this score for RADLI is sometimes lower than that for LIGHT.

4.2.3 Livorno: 06-09 UTC 10 September 2017

In this period, the most intense phase of the precipitation occurred over central Italy, over the coastal part of Lazio (Figure 17a). More in detail, among the 2695 raingauges reporting valid data over the domain of Figure 17a, 307 reported more than 10 mm/3h, 132 more than 20 mm/3h, 86 more than 30 mm/3h, 66 more than 40 mm/3h, 49 more than 50 mm/3h and 35 more than 60 mm/3h. Among the 35 raingauges measuring more than 60 mm/3h, 33 were over Lazio, showing the heavy rainfall occurred over the Region.

Some precipitation persisted over Tuscany but the rainfall is much lower compared to previous 6h (the rainfall over Tuscany between 03 and 06 UTC was very intense, not shown). Other notable precipitation areas are over the NE of Italy (moderate to low amounts), over Central Alps (moderate values) and over the whole Sardinia (small amounts).

Figure 17b shows the rainfall simulated by CTRL. The forecast is unsatisfactory, mainly for the following two reasons: a) heavy precipitation is simulated over Tuscany (> 75 mm/3h), also close to the Livorno area; b) very small precipitation is forecasted over central Italy. The rainfall over NE Italy is well represented in space, but overestimated.

Considering the evolution of CTRL rainfall forecast for the two VSF of Livorno, we conclude that CTRL was able to predict abundant rain over Livorno, but this was delayed compared to the real event.

The rainfall simulated by RAD (Figure 17c) clearly improves the forecast compared to CTRL. First, the precipitation over Lazio is very well predicted and the rainfall values are higher than 40 mm/3h (up to 65 mm/3h), so the RAD forecast well represents the main precipitation spot over Italy for this VSF. Second, the precipitation over Tuscany is less than for CTRL, showing the ability of radar reflectivity factor data assimilation to dry the model when it predicts rain that is not observed. Third, the precipitation over central Alps is represented, even if located about 30 km to the East. It is noted, however, that the area of intense rainfall (>60 mm/3h) is overestimated by RAD, showing a wet forecast. This is confirmed by the wet frequency bias of the RAD simulation, which is greater than 3 between 14 and 44 mm/3h.

LIGHT forecast, Figure 17d, shows a worse performance compared to RAD for this time period. The precipitation forecast is mainly over Tuscany, where it is overestimated, with a small precipitation spot over Lazio.

The precipitation forecast of RADLI, Figure 17e, represents very well the precipitation over Lazio, and the rainfall amount is better predicted compared to RAD. The precipitation over Sardinia is well represented by RADLI as well as the precipitation over Central Alps, giving the best results among all forecasts.

The analysis of the scores confirms the above results (Table 6). CTRL has a poor performance as shown by the POD and ETS values, close to zero, for all thresholds above 30 mm/3h and for all neighbourhood radii. The simulations assimilating radar reflectivity factor performs better than LIGHT, the difference being larger for higher rainfall thresholds and for smaller neighbourhood radii.

It is also notable the good performance of RADLI forecast for the nearest neighbourhood radii (ETS=0.43, POD=0.92) for the 50 mm/3h threshold.

4.3 Evolution of total water

Because lightning data assimilation and radar reflectivity factor data assimilation both adjust the water vapour mixing ratio (q_v), it is interesting to evaluate the contribution of each data source to the q_v adjustment including in that evaluation the assimilation phase (0-6 h).

For the 3D-Var approach the impact of the contribution of data assimilation on q_v can be done using maps similar to Figure 14b. For example, Fierro et al. (2016), using a 3D-Var approach to assimilate lightning, used the layer averaged q_v between 3 and 10 km to quantify the water vapour added to the WRF model by lightning data assimilation. However, because in this paper lightning are assimilated by nudging, this kind of representation is not practicable because it is difficult to separate the contribution of the nudging from other processes in the evolution of q_v.

Fierro et al. (2015) used the total water substance mass (forecasted accumulated precipitation + total hydrometeors and water vapour mass) to quantify the impact of lightning data assimilation by

nudging. In this paper, a similar approach is used. More specifically, we consider the forecasted accumulated precipitation and the total hydrometeors and water vapour mass in the atmosphere averaged over the grid columns. Also, we averaged all VSFs for Serano and Livorno. The evolution of the forecasted accumulated precipitation is shown in Figure 18a, while the evolution of the total hydrometeors and water vapour mass in the atmosphere is shown in Figure 18b.

Considering the Figures 18a and 18b it is apparent that flashes add less water vapour compared to radar reflectivity factor data assimilation and, of course, RADLI has the largest impact. In particular, the total water mass added to the background is 2.5%, 5.7% and 7.4% for LIGHT, RAD and RADLI, respectively. Importantly, the total water substance mass added by RADLI to the background is less than the sum of the total water substance mass added by RAD and LIGHT. This happens because 3D-Var adds water to the background limiting the impact of nudging during the simulation. For example, in an already saturated atmosphere the nudging of Eqn. (2) doesn't have any impact.

Accumulated precipitation accounts for the largest part of the water vapour added to the simulation, similarly to Fierro et al. (2015). At the end the assimilation phase (6h), the evolution of the total water vapour and hydrometeors mass in the atmosphere converges towards the background as boundary conditions propagates into the domain.

4.4 Sensitivity to nudging formulation

As stated in Section 3.2, there are limitations when applying the nudging method of Fierro et al. (2012) to RAMS@ISAC. Also, the optimal setting of the coefficients of Eqn. (2) depends on the case study. For these reasons, it is interesting to evaluate the sensitivity of the results to changes in nudging formulation. For this purpose, we show the variability of ETS and POD scores to changes in the A and B coefficients of Eqn. (1). The scores are computed considering all the VSF for the two case studies for different configurations: A_76 has the coefficients A=0.76 and B=0.25; LIGHT has A=0.86 and B=0.15 (default setting), SAT has A=1.01 and B=0; RADLI has A=0.86 and B=0.15 (default setting); CTRL, and RAD are as defined in Table 3.

The scores are computed for the second RAMS@ISAC domains and are shown for the nearest neighbourhood. ETS score (Figure 19a) shows that all configurations assimilating either lightning or radar reflectivity factor alone or a combination of lightning and radar reflectivity factor improves the forecast for all thresholds. RADLI has the best ETS for rainfall intensity larger than 32 mm/3h in line with the results of the three VSF discussed above.

For rainfall lower than 32 mm/3h, the simulations assimilating lightning perform better, because they have less false alarms compared to those assimilating radar reflectivity factor (not shown). From the comparison of LIGHT and SAT with A_76, it is apparent that the latter has the worst score. The comparison between LIGHT and SAT shows mixed results: SAT performs better up to

38 mm/3h, while LIGHT is better for higher thresholds. This result is confirmed by POD, Figure 19b, which shows that SAT performs better up to 32 mm/3h, while LIGHT is better for higher thresholds. A visual inspection of the model output reveals that SAT can generate spurious convection in some areas while missing convection in other areas that are correctly forecast by LIGHT or even A_76, i.e. adding less water vapour to the model because of the different trajectories in the phase space followed by the model using different settings.

Lynn et al. (2015) implemented a method suggested by Fierro et al. (2012) to suppress spurious convection in WRF model. The method compares the lightning forecast during the assimilation period with lightning observations to filter out spurious convection. The application of the methodology on 10 July 2013 improved the forecast of the squall line from Texas to Iowa, which was the focus of the forecast on that day; however, the application of the method to 19 and 21 March 2012 over the US gave mixed results, improving the forecast in the first 6h and worsening it after 6h.

The implementation of this method in the RAMS@ISAC could be used to suppress spurious convection in simulations assimilating lightning, especially SAT.

It is finally noted that RAD and RADLI have high POD values for all thresholds, nevertheless their ETS is below that of LIGHT and SAT up to 32 mm/3 h (RADLI) and 42 mm/3h (RAD). This behaviour is caused by the larger number of false alarms given by assimilating radar reflectivity factor compared to those assimilating lightning. This result shows again that the RAD and RADLI configurations have a wet frequency bias. In particular, the frequency bias of RAD and RADLI configuration is greater than 3 between 20 and 40 mm/3h.

Table 3: Types of simulations performed.

Experiment	Description	Data assimilated	Model variable impacted
CTRL	Control run	None	None
RAD	RADAR data assimilation	Reflectivity factor CAPPI (RAMS-3DVar)	Water vapour mixing ratio
LIGHT	Lightning data assimilation (A=0.85; B=0.16 in Eqn (2))	Lightning density (nudging)	Water vapour mixing ratio
RADLI	RADAR + lightning data assimilation (A=0.86; B=0.15 in Eqn	Reflectivity factor CAPPI (RAMS-3DVar) + Lightning density	Water vapour mixing ratio

	(2))	(nudging)	
A_76	Lightning data assimilation (A=0.76; B=0.25 in Eqn (2))	Lightning density (nudging)	Water vapour mixing ratio
SAT	Lightning data assimilation (A=1.01; B=0. in Eqn (2))	Lightning density (nudging)	Water vapour mixing ratio

Table 4: ETS and POD scores for three different neighbourhood radii. Scores are computed over the domain D2.

Threshold (mm/3 h)	ETS nearest neighboorhood (CTRL, RAD, LIGHT, RADLI)	POD nearest neighbourhood (CTRL, RAD, LIGHT, RADLI)	ETS 25 km (CTRL, RAD, LIGHT, RADLI)	POD 25 km (CTRL, RAD, LIGHT, RADLI)	ETS 50 km (CTRL, RAD, LIGHT, RADLI)	POD 50 km (CTRL, RAD, LIGHT, RADLI)
1	(0.42,0.36,0.44, 0.33)	(0.57,0.87,0.60, 0.81)	(0.68,0.73,0.68, 0.73)	(0.77,0.93,0.75, 0.89)	(0.79,0.89,0.82, 0.87)	(0.84,0.92,0.84, 0.90)
6	(0.06,0.10,0.14, 0.13)	(0.0,0.5,0.20,0. 72)	(0.11,0.44,0.72, 0.41)	(0.11,0.86,0.72, 0.83)	(0.19,0.86,0.86, 0.92)	(0.19,0.86,0.86, 0.92)
10	(0.,0.05,0.,0.15)	(0.,0.26,0.,0.79)	(0.,0.66,0.58,0. 74)	(0.0,0.84,0.58,0 .89)	(0.,0.95,0.74,0. 90)	(0.,0.95,0.74,0. 90)
20	(0.,0.,0.,0.41)	(0.,0.,0.,0.8)	(0.0,0.41,0.33,0 .87)	(0.,0.47,0.3,0.9)	(0.,0.73,0.80,1. 0)	(0.,0.73,0.80,1. 0)
30	(0.,0.,0.,0.31)	(0.,0.,0.,0.5)	(0.,0.,0.,0.90)	(0.,0.,0.,0.9)	(0.,0.,0.,1.0)	(0.,0.,0.,1.0)
40	(0.,0.,0.,0.)	(0.,0.,0.,0.)	(0.,0.,0.,0.33)	(0.,0.,0.,0.33)	(0.,0.,0.,0.50)	(0.,0.,0.,0.50)

Table 5: ETS and POD scores for three different neighbourhood radii. Scores are computed over the domain D3.

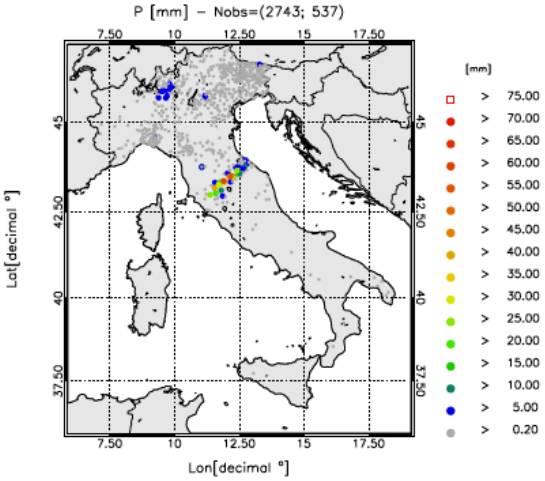
Threshold (mm/3h)	ETS nearest neighboorhood (CTRL, RAD, LIGHT, RADLI)	POD nearest neighbourhood (CTRL, RAD, LIGHT, RADLI)	ETS 25 km (CTRL, RAD, LIGHT, RADLI)	POD 25 km (CTRL, RAD, LIGHT, RADLI)	ETS 50 km (CTRL, RAD, LIGHT, RADLI)	POD 50 km (CTRL, RAD, LIGHT, RADLI)
1	(0.43,0.64,0.70, 0.56)	(0.67,0.86,0.98, 0.99)	(0.68,0.80,0.82, 0.71)	(0.83,0.92,0.98, 0.99)	(0.68,0.80,0.82, 0.71)	(0.83,0.92,0.98, 0.99)
6	(0.1,0.31,0.60,0 .49)	(0.24,0.58,0.89, 0.95)	(0.49,0.70,0.91, 0.96)	(0.55,0.76,0.96, 0.97)	(0.49,0.70,0.91, 0.96)	(0.55,0.76,0.96, 0.97)
10	(0.11,0.33,0.56, 0.54)	(0.19,0.56,0.75, 0.80)	(0.48,0.76,0.91, 0.97)	(0.52,0.79,0.92, 0.97)	(0.48,0.76,0.91, 0.97)	(0.52,0.79,0.92, 0.97)
20	(0.02,0.30,0.52, 0.59)	(0.03,0.39,0.74, 0.81)	(0.18,0.73,0.97, 0.93)	(0.19,0.74,0.97, 0.97)	(0.18,0.73,0.96, 0.93)	(0.19,0.74,0.97, 0.97)
30	(0.,0.27,0.51,0. 47)	(0.,0.29,0.76,0. 76)	(0.,0.64,0.94,1.)	(0.,0.65,1.,1.)	(0.,0.64,0.94,1.)	(0.,0.65,1.,1.)
40	(0.,0.44,0.27,0. 27)	(0.,0.44,0.56,0. 67)	(0.,0.89,1.,1.)	(0.,0.89,1.,1.)	(0.,0.89,1.,1.)	(0.,0.89,1.,1.)
50	(0.,0.33,0.66,0. 50)	(0.,0.33,0.67,0. 67)	(0.,0.67,1.,1.)	(0.,0.67,1.,1.)	(0.,0.66,1.,1.)	(0.,0.67,1.,1.)

Table 6 ETS and POD scores for three different neighbourhood radii. Scores are computed over the domain D2.

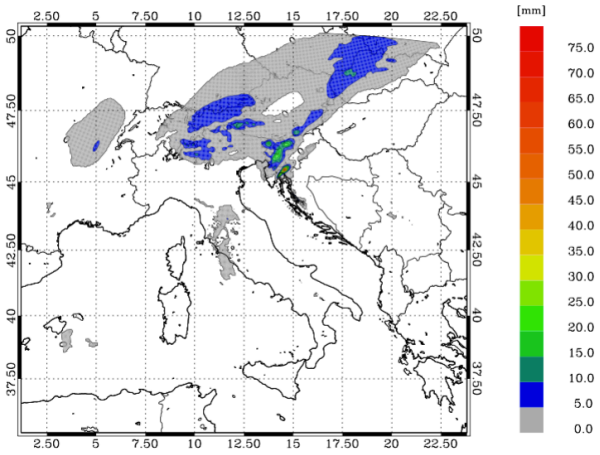
Threshold (mm/3h)	ETS nearest neighboorhood (CTRL, RAD, LIGHT, RADLI)	POD nearest neighbourhood (CTRL, RAD, LIGHT, RADLI)	ETS 25 km (CTRL, RAD, LIGHT, RADLI)	POD 25 km (CTRL, RAD, LIGHT, RADLI)	ETS 50 km (CTRL, RAD, LIGHT, RADLI)	POD 50 km (CTRL, RAD, LIGHT, RADLI)
1	(0.41,0.63,0.61, 0.65)	(0.66,0.89,0.89, 0.93)	(0.79,0.83,0.82, 0.83)	(0.89,0.95,0.95, 0.96)	(0.88,0.92,0.93, 0.94)	(0.93,0.97,0.98, 0.98)
6	(0.2,0.4,0.39,0. 47)	(0.43,0.82,0.77, 0.88)	(0.45,0.63,0.71, 0.76)	(0.63,0.90,0.95, 0.96)	(0.72,0.86,0.88, 0.92)	(0.82,0.96,0.97, 0.96)
10	(0.,0.24,0.18,0. 28)	(0.14,0.78,0.55, 0.80)	(0.14,0.47,0.58, 0.62)	(0.24,0.86,0.82, 0.93)	(0.32,0.91,0.96, 0.95)	(0.35,0.95,0.97, 0.97)
20	(-0.03,0.18,0.13, 0.22)	(0.01,0.81,0.30, 0.80)	(0.09,0.46,0.57, 0.61)	(0.11,0.86,0.59, 0.90)	(0.15,0.84,0.91, 0.96)	(0.15,0.90,0.92, 0.97)
30	(-	(0.,0.90,0.23,0.	(0.01,0.79,0.46,	(0.01,0.93,0.47,	(0.02,0.95,0.93,	(0.02,0.95,0.93,

	0.02,0.22,0.13, 0.28)	88)	0.80)	0.94)	0.99)	0.99)
40	(- 0.1,0.24,0.08,0. 36)	(0.,0.83,0.12,0. 89)	(0.01,0.83,0.37, 0.83)	(0.02,0.97,0.38, 0.97)	(0.1,0.97,0.95,0 .98)	(0.02,0.98,0.95, 0.98)
50	(- 0.01,0.27,0.,0.4 3)	(0.,0.67,0.,0.92)	(0.,0.90,0.,0.90)	(0.,0.94,0.,0.96)	(0.,0.96,0.,0.96)	(0.,0.96,0.,0.96)

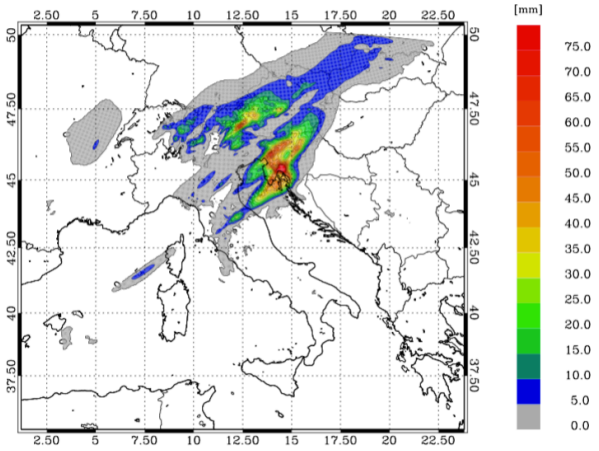
a)



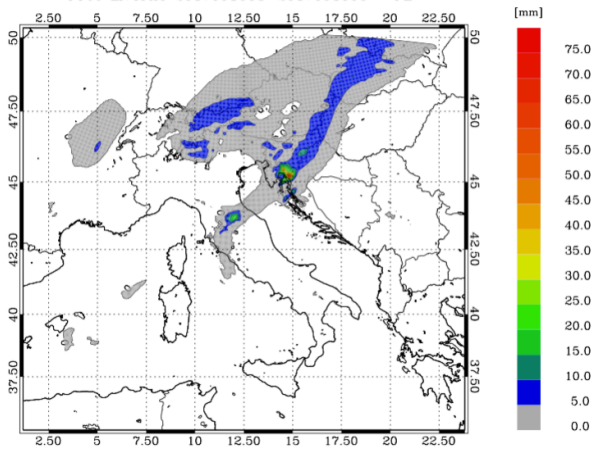
b)



c)



d)



e)

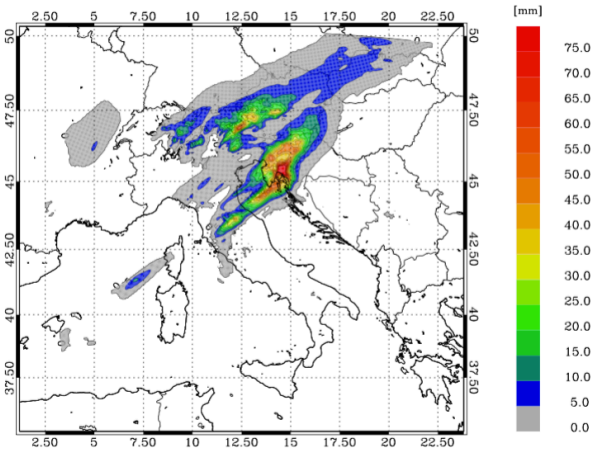
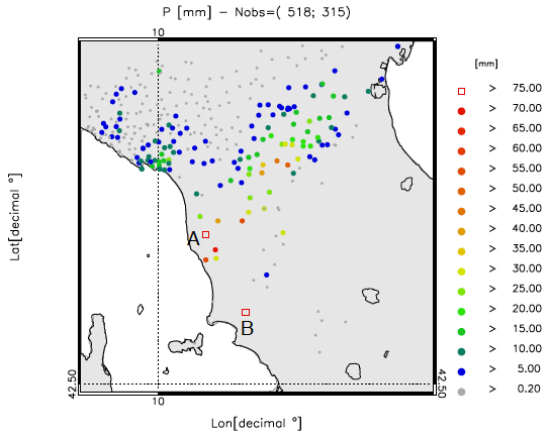
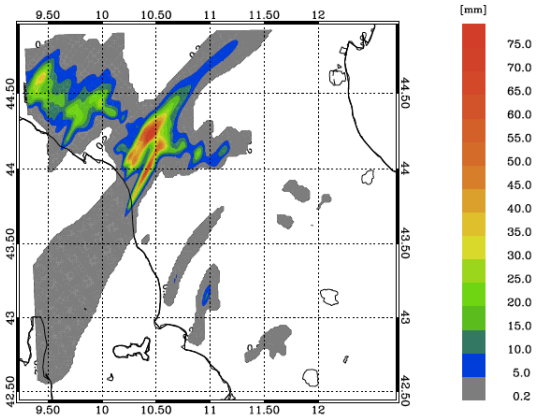


Figure 15: a) rainfall reported by raingauges between 03 and 06 UTC on 16 September 2017. Only raingauges observing at least 0.2 mm/day are shown. The first number in the title within brackets represents the available raingauges, while the second number represents those observing at least 0.2 mm/3h; b) as in a) for the CTRL forecast; c) as in a) for the RAD forecast; d) as in a) for the LIGHT forecast; e) as in a) for the RADLI forecast.

a)

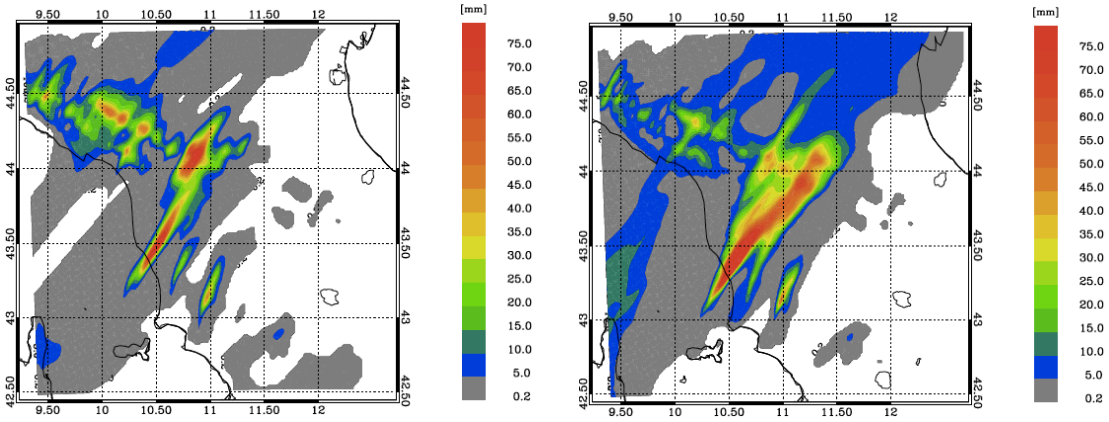
b)





c)

d)



e)

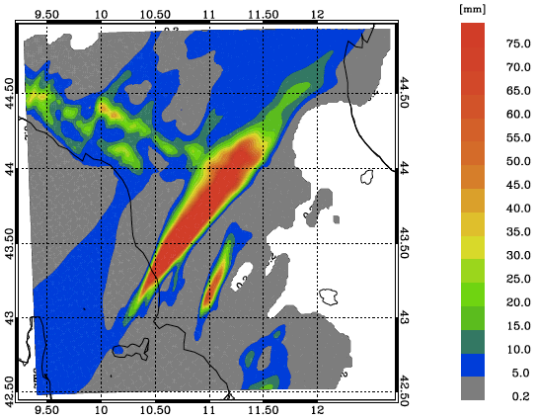


Figure 16: a) rainfall reported by raingauges between 00 and 03 UTC on 10 September 2017. Only stations reporting at least 0.2 mm/3h are shown. The first number in the title within brackets represents the number of raingauges available over the domain, while the second number shows those observing at least 0. 2 mm/3h; b) as in a) for the CTRL forecast; c) as in a) for the RAD forecast; d) as in a) for the LIGHT forecast; e) as in a) for the RADLI forecast; f) POD score for the period 00-03 UTC on 10 September; g) as in f) for the ETS score; h) as in f) for the HR score.

a) b)

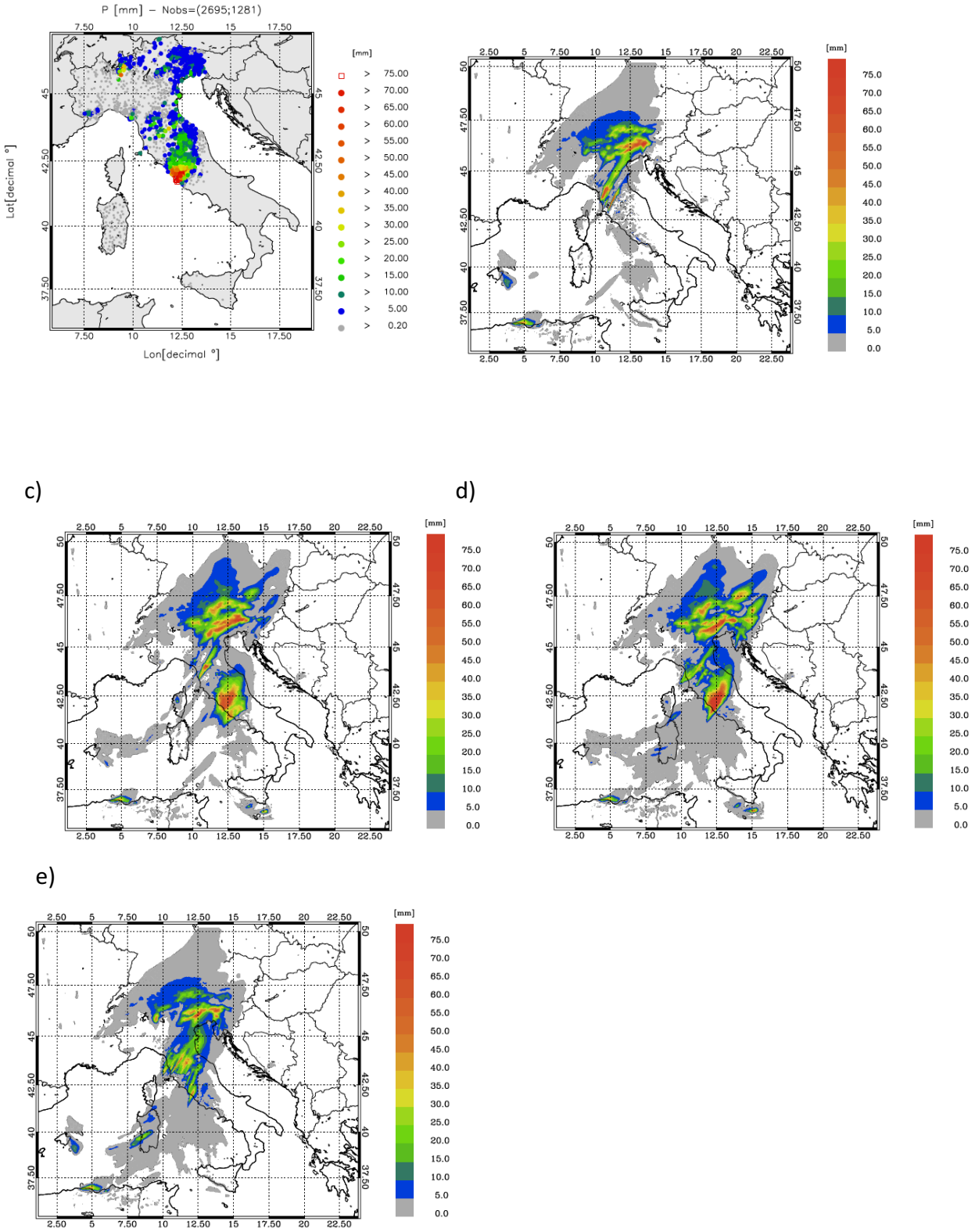


Figure 17: a) rainfall reported by raingauges between 06 - 09 UTC on 10 September 2017. For this time period 2695 raingauges reported valid observations in the domain, however only stations reporting at least 0.2 mm/3h are shown The first number in the title within brackets represents the number of raingauges available over the domain, while the second number shows those observing at least 0.2 mm/3h; b) as in a) for the CTRL forecast; c) as in a) for the RAD forecast; d) as in a) for the LIGHT forecast; g) as in a) for the RADLI forecast.

a) b)

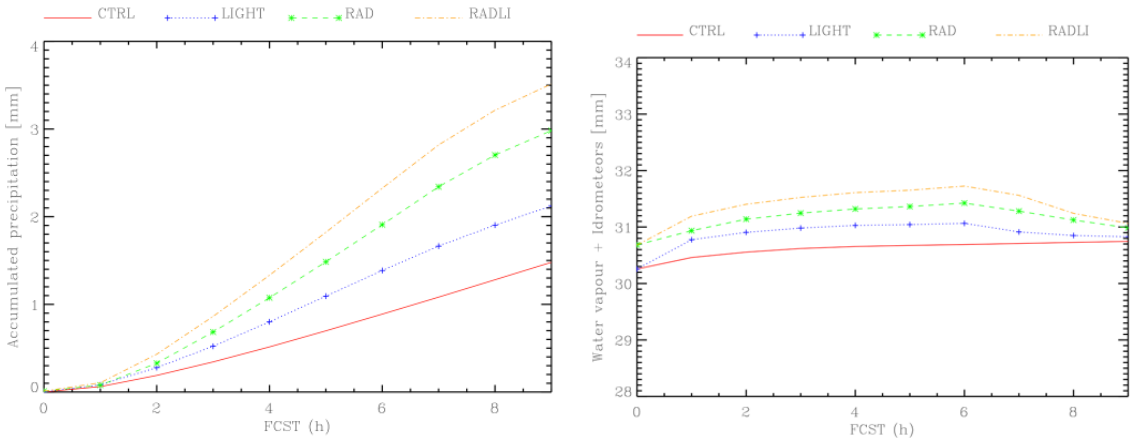


Figure 18: a) Evolution of the accumulated precipitation for different model configurations and for all forecast hours; b) as in a for the hydrometeor mass plus the water vapour equivalent mass per unit area. All quantities are expressed in [mm] and are averaged over the number of grid columns.

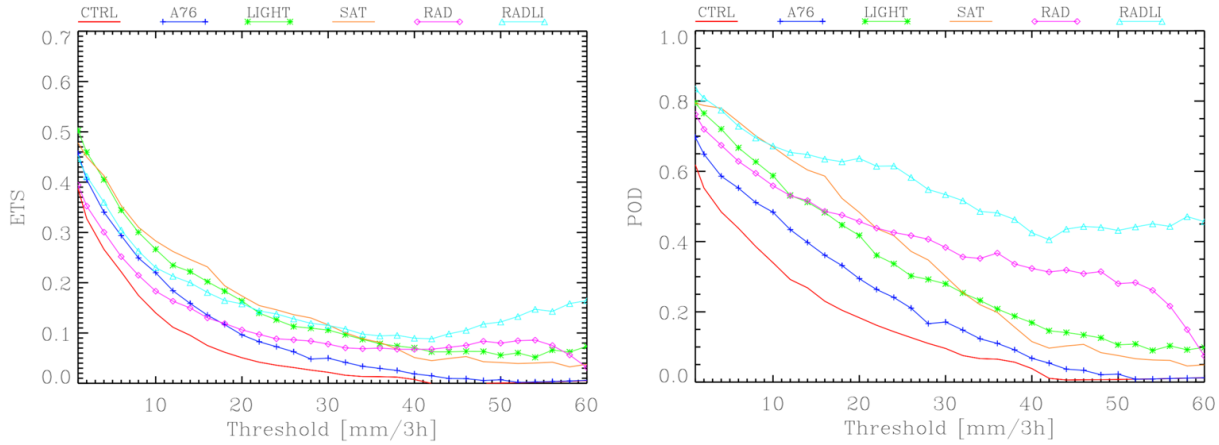


Figure 19: a) ETS score for all VSF considered in this paper; b) as in a) for the POD score.

References

Barker, D.M., Huang, W., Guo, Y.-R., and Xiao, Q.N.: A Three-Dimensional Variational Data Assimilation System For MM5: Implementation And Initial Results, Monthly Weather Review, 132, 897-914, 2004.

Ducrocq, V., Braud, I., Davolio, S., Ferretti, R., Flamant, C., Jansa, A., Kalthoff, N., Richard, E., Taupier-Letage, I., Ayral, P.-A., Belamari, S., Berne, A., Borga, M.,

Boudevillain, B., Bock, O., Boichard, J.-L., Bouin, M.-N., Bousquet, O., Bouvier, C., Chiggiato, J., Cimini, D., Corsmeier, U., Coppola, L., Cocquerez, P., Defer, E., Delanoë, J., Di Girolamo, P., Doerenbecher, A., Drobinski, P., Dufournet, Y., Fourrié, N., Gourley, J.J., Labatut, L., Lambert, D., Le Coz, J., Marzano, F.S., Molinié, G., Montani, A., Nord, G., Nuret, M., Ramage, K., Rison, W., Roussot, O., Said, F., Schwarzenboeck, A., Testor, P., Van Baelen, J., Vincendon, B., Aran, M., and Tamayo, J.: HYMEX-SOP1 The Field Campaign Dedicated to Heavy Precipitation and Flash Flooding in the Northwestern Mediterranean. Bull. Amer. Meteor. Soc., 95, 1083-1100, https://doi.org/10.1175/BAMS-D-12-00244.1, 2014.

Emersic, C., and C. P. R. Saunders, 2010: Further laboratory investigations into the relative diffusional growth rate theory of thunderstorm electrification. Atmos. Res., 98, 327–340, doi:https://doi.org/10.1016/j.atmosres.2010.07.011.

Federico, S.: Implementation of a 3D-Var system for atmospheric profiling data assimilation into the RAMS model: Initial results, Atmospheric Measurement Techniques, 6(12), 3563-3576, 2013.

Ferretti, R., Pichelli, E., Gentile, S., Maiello, I., Cimini, D., Davolio, S., Miglietta, M. M., Panegrossi, G., Baldini, L., Pasi, F., Marzano, F. S., Zinzi, A., Mariani, S., Casaioli, M., Bartolini, G., Loglisci, N., Montani, A., Marsigli, C., Manzato, A., Pucillo, A., Ferrario, M. E., Colaiuda, V., and Rotunno, R.: Overview of the first HyMeX Special Observation Period over Italy: observations and model results, Hydrol. Earth Syst. Sci., 18, 1953-1977, https://doi.org/10.5194/hess-18-1953-2014, 2014.

Fierro, A. O., Mansell, E., Ziegler, C., and MacGorman, D.: Application of a lightning data assimilation technique in the WRFARW model at cloud-resolving scales for the tornado outbreak of 24 May 2011, Mon. Weather Rev., 140, 2609–2627, 2012.

Fierro, A. O., A. J. Clark, E. R. Mansell, D. R. MacGorman, S. Dembek, and C. Ziegler, 2015: Impact of storm-scale lightning data assimilation on WRF-ARW precipitation forecasts during the 2013 warm season over the contiguous United States. Mon. Wea. Rev., 143, 757–777, doi:https://doi.org/10.1175/MWR-D-14-00183.1.

Fierro, A.O., Gao, l., Ziegler, C. L., Calhoun, K. M., Mansell, E. R., and MacGorman, D. R.: Assimilation of Flash Extent Data in the Variational Framework at Convection-Allowing Scales: Proof-of-Concept and Evaluation for the Short-Term Forecast of the 24 May 2011 Tornado Outbreak. Mon. Wea. Rev., 144, 4373-4393,https://doi.org/10.1175/MWR-D-16-0053.1, 2016.

Mass, C. F., Ovens, D., Westrick, K., and Colle, B. A.: Does increasing horizontal resolution produce more skilful forecasts?, B. Am. Meteorol. Soc., 83, 407–430, 2002.

Marchand, M., and H. Fuelberg, 2014: Assimilation of lightning data using a nudging method involving low-level warming. Mon. Wea. Rev., 142, 4850–4871, doi:10.1175/MWR-D-14-00076.1.

Mittermaier, M., N. Roberts, and S. A. Thompson, 2013: A long-term assessment of precipitation forecast skill using the Fractions Skill Score. Meteor. Appl., 20, 176–186, doi:https://doi.org/10.1002/met.296.

Parrish, D.F., and Derber, J.C.: The National Meteorological Center's Spectral Statistical Interpolation analysis system, Monthly Weather Review, 120, 1747-1763, 1992.

Petracca, M., L. P. D'Adderio, F. Porcù, G. Vulpiani, S. Sebastianelli, and S. Puca, 2018: Validation of GPM Dual-Frequency Precipitation Radar (DPR) rainfall products over Italy. J. Hydrometeor., 19, 907–925, https://doi.org/10.1175/JHM-D-17-0144.1

Takahashi, T., 1978: Riming electrification as a charge generation mechanism in thunderstorms. J. Atmos. Sci., 35, 1536–1548, doi:https://doi.org/10.1175/1520 0469(1978)035<1536:REAACG>2.0.CO;2.

Vulpiani, G., A. Rinollo, S. Puca, and M. Montopoli, 2014: A quality-based approach for radar rain field reconstruction and the H-SAF precipitation products validation. Proc. Eighth European Radar Conf., Garmish-Partenkirchen, Germany, ERAD, Abstract 220, 6 pp., http://www.pa.op.dlr.de/erad2014/programme/ ExtendedAbstracts/220_Vulpiani.pdf (last access January 2019).