Lightning data assimilation in the WRF-ARW model for short-term rainfall forecasts of three severe storm cases in Italy

Albert Comellas Prat[1], Stefano Federico[2], Rosa Claudia Torcasio[2], Alex O. Fierro[3], Stefano Dietrich[2]

1. ISAC-CNR, Strada Prov.le Lecce-Monteroni, Lecce, Italy
2. ISAC-CNR, Via del Fosso del Cavaliere 100, Rome, Italy
3. Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, and NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma

Abstract

This study analyses the impact of total lightning data assimilation on cloud-resolving short-term (3 and 6 h) precipitation forecasts of three heavy rainfall events that occurred recently in Italy by providing an evaluation of forecast skill using statistical scores for 3-hourly thresholds against observational data from a dense rain gauge network. The experiments are performed with two initial and boundary conditions datasets as a sensitivity test. The three rainfall events have been chosen to better represent the convective regime spectrum: from a short-lived and localised thunderstorm to a long-lived and widespread event, along with a case that had elements of both.

This analysis illustrates the ability of the lightning data assimilation (LDA) to notably improve the short-term rainfall forecasts with respect to control simulations without LDA. The assimilation of lightning enhances the representation of convection in the model and translates into a better spatiotemporal positioning of the storm systems. The results of the statistical scores reveal that simulations with LDA always improve the probability of detection, particularly for rainfall thresholds exceeding 40 mm/3h. The false alarm ratio also improves but appears to be more sensitive to the model initial and boundary conditions. Overall, these results show a systematic advantage of the LDA with a 3-h forecast range over 6-h.

1. Introduction

The difficulty and challenges of accurate high impact weather forecast with state-of-the-art numerical weather prediction (NWP) models has been a longstanding problem (Stensrud et al. 2009). With the advent of ever-increasing computer power and the opportunity of assimilating data

from various observation platforms, there has been a notable improvement in the quality of limited-area NWP forecasts in the past two-three decades (e.g., Huang et al. 2009; Park and Xu 2013; Johnson et al. 2015).

Lightning data assimilation (LDA) can effectively improve NWP rainfall forecasts of high impact weather events. Lightning is typical of convective environments and permits to locate convectively active areas with a high degree of accuracy given that the geo-location errors for lightning flashes is often less than 1 km, which is smaller than typical NWP models horizontal resolution. Being a 2D dataset, lightning data requires minimal disk storage and, thus, alleviate the need for broad band network connections for transmission. Moreover, lightning data remain a reliable real-time product in terms of timeliness for delivery and feature generally few data gaps. Lightning can be detected in the long range, in mountainous regions and beneath high clouds with technologies observing lightning worldwide being readily available. All the aforementioned reasons render lightning data assimilation an attractive research venue for NWP, especially for Rapid Update Cycles (RUC). This explains why techniques to assimilate lightning have been the focus of active research in the recent years/decade.

In the context of LDA, the earliest attempts go back to the pioneer works of Alexander et al. (1999) and Chang et al. (2001). Infused by prior observational research establishing the well-known relationship between the presence of total lightning and deep convective updrafts and precipitation (Goodman et al. 1988, MacGorman et al. 1989), these modeling exercises demonstrated improvements of 12 to 24 h rainfall forecasts when an LDA converting lightning data into rain rates was utilized. Benjamin et al. (2004) also used rain rates as a proxy for lightning in the initial conditions of the Rapid Update Cycle (RUC) model.

Papadopoulos et al. (2005) nudged the relative humidity profiles to empirical profiles representative of observed convective regimes by making use of lightning flash rate data, while Mansell et al. (2007) used lightning data by modifying the Kain-Fritch cumulus scheme (Kain, 2004) to allow the flash rate to control the trigger function within the parameterization. A follow-on study by Pessi and Businger (2009) assimilated lightning data by adjusting the latent heating profile via a lightning-rainfall conversion relationship derived from Tropical Rainfall Measuring Mission (TRMM) data. Giannaros et al. (2016) applied an LDA method similar to Mansell et al. (2007) to eight convective events in Greece and documented overall statistically robust impact on the 24 h rainfall forecasts.

The aforementioned studies all made use of relatively coarser, convection permitting (> 5 km) horizontal grid spacing, which was not able to fully resolve convection and many of the associated cloud scale processes. Fierro et al. (2012) was the first to consider applying an LDA approach at convection-resolving scales (less than 3 km) using a methodology that imposed incremental increases in water vapor mass in the mixed-phase region (0°C to -20°C layer) based on observed

lightning flash density rates and simulated graupel mixing ratio. They showed general improvements of the representation of convection at the analysis time and at 1 h forecasts. Using the same nudging method, Fierro et al. (2014) improved the forecast of a derecho event up to 6 h and illustrated how the improvements incurred by the LDA could, in some cases, be comparable to that of three dimensional variational assimilation (3DVAR) of radar data (radial winds and reflectivity factor). Qie et al. (2014) adapted the nudging function of Fierro et al. (2012) to adjust graupel, snow and ice crystals mixing ratios and applied it to a quasi-linear convective system over Northern China. Because the system was generally outflow dominated, modulating ice mass (and, hence, cold pools) helped improve the representation of this system during the assimilation period and up to 1 h into the forecast.

Federico et al. (2017a) adapted the water vapor nudging methodology of Fierro et al. (2012) to evaluate the impact of LDA on very-short term (≤ 3 h), cloud-resolving forecasts using the Regional Atmospheric Modeling System (RAMS). They applied that LDA method to twenty widespread convective events that occurred over Italy during the fall of 2012, showing a general positive impact of the LDA on 3 h rainfall forecasts, as measured by standard statistical score metrics and 2x2 contingency table elements. Federico et al. (2017b) extended their prior analysis to longer forecast ranges of 3, 6, 12 and 24 h after the assimilation of lightning, showing that the main positive impacts were mainly seen during the 3 and 6 h forecast ranges, with predictive skill steadily and substantially decreasing after that with forecast time.

Using the WRF-ARW model in lieu of RAMS, the present work examines and assesses the impacts of the Fierro et al. (2012) LDA approach on the short-term (3 and 6 h) precipitation forecast for three high impact weather case studies that recently occurred in Italy. The three cases were selected to be representative of different convective modes, namely: short-lived and localised (12 August 2015 in Calabria), long-lived and widespread (9-10 September 2017 in Livorno and Central Italy), and the event in Sardinia (10 October 2018) that had traits from both. Real time operations revealed that these cases were challenging to forecast without data assimilation. Evaluation of the precipitation forecasts is performed against rainfall data from a dense raingauge network using the same statistical evaluation methods as in Federico et al. (2017a, b). This manuscript is structured as follows: section 2 describes the data, methodology and verification strategies followed; section 3 analyses the results of each case study and their overall statistical scores. Additionally, section 3 presents a sensitivity experiment of the LDA scheme for the first case study, and finally section 4 provides the discussion and concluding remarks.

2. Methodology

2.1. Data used

The initial (IC) and boundary conditions (BC) datasets were obtained from the Integrated Forecast System (IFS) global model from the European Centre for Medium-Range Weather Forecasts (ECMWF) at 0.25° horizontal resolution, and the Global Forecast System (GFS) forecasts from the US National Weather Service (NWS) at 0.5°. The lightning data used are from LINET (LIghtning detection NETwork, Betz et al. 2009). LINET exploits the VLF/LF electromagnetic bands to provide measurements of both intra-cloud (IC) and cloud to ground (CG) discharges. IC strokes are detected as long as lightning occurs within 120 km from the nearest sensor owing to an optimised hardware and advanced techniques of data processing (TOA-3D, Betz et al. 2004). LINET features more than 500 sensors worldwide with the largest density located over Central Europe. The antenna design and detection algorithms of LINET yield an average geo-location accuracy of 75 m as well as a homogeneous detection efficiency of peak currents as low as 2 kA (see www.nowcast.de). All discharges within a 1 s period and within a 10 km radius are grouped into a single flash (Federico et al. 2014) before data assimilation. For the model forecast verification, we made use of the hourly precipitation database from the DPC (Department of Civil Protection) network in Italy, consisting of more than 3000 rain gauges distributed across the country.

2.2. Methodology

To evaluate the LDA, the following modelling strategy was adopted. First, two control simulations (CNTRL) integrated for 36 h were performed – one for each IC and BC dataset – and were started at 12 UTC the day prior to the convective event of interest. Then, the two LDA experiments were conducted with each of the IC and BC datasets. The run labeled "ASSIM+3" assimilated lightning during the first 6 h of the simulation (assimilation phase) and only the following 3 h of forecast were considered, resulting in a total of 8 simulations to cover the targeted 24 h period. Additional experiments, denoted as "ASSIM+6", featured the same assimilation phase duration, but the following 6 h of forecast were considered instead for analysis. With this set up, a total of 4 simulations were needed to cover a full day. The first ASSIM+3 simulation was started at 18 UTC on the previous day and assimilated lightning until 00 UTC. Thus, the first forecast period was 00-03 UTC. The second simulation was initialized at 21 UTC so that the forecast period was between 03-06 UTC, and so on. In the case of ASSIM+6, the first forecast period was 00-06 UTC, the second 06-12 UTC, etc. In the subsections where each case study is analyzed, attention is focused on a particularly representative 3-h period. For the statistical scores, however, we took into account all 3-h periods within the simulated days (i.e., aggregate statistics).

To compare observed and predicted precipitation, we employed the nearest neighbourhood method. In particular, for each rain gauge, we considered all model gridpoints contained within a search radius of about 8.5 km (i.e., $2\Delta x\sqrt{2}$, with Δx model grid spacing) centered at the rain gauge

location, and then selected the gridpoint with the forecast rainfall closest in value to the observations. In this analysis, we also considered a larger search radius of 20 km to better visualize the score results using the Roebber performance diagram (Roebber, 2009).

The fully compressible, three-dimensional Weather Research and Forecasting Model (ARW-WRF version 4.1.2) was employed in this study (Skamarock et al. 2019). The simulation domain has horizontal dimensions of 530 x 530 grid points with 60 vertical eta levels in a hybrid sigma-pressure coordinate system (terrain-following in the low and middle levels, transitioning to constant pressure surfaces towards the upper model levels), with a lowest layer thickness of 50 m, a maximum thickness allowed of 1000 m, and an absolute top pressure of 40 hPa. The time step was set to 18 s and the horizontal grid spacing is 3 km.

The microphysics parameterization used is the Thompson 1.5-moment, which predicts rain, ice, snow and graupel mass mixing ratios, together with rain and ice number concentrations (Thompson et al. 2008). For cumulus processes the Kain-Fritsch convection parameterization scheme (CPS) (Kain, 2004) is employed; the Yonsei University scheme for the boundary layer (Hong, Noh and Dudhia, 2006), the Monin-Obukhov scheme for the surface layer (Monin and Obukhov, 1954), the Noah land surface model (Chen and Dudhia, 2001) for surface skin physics, and the Dudhia (1989) and RRTM (Mlawer et al. 1997) for short and longwave radiation schemes, respectively.

Although prior work showed that CPS schemes are generally not required for grid spacings below that of convection permitting scales (< 10 km, Kain et al. 2006; Kain et al. 2008; Skamarock et al. 2019), the 1-10 km range for horizontal grid spacing is often referred to as terra incognita or grey zone because a noteworthy portion of the turbulence spectrum still remains unresolved (e.g., Bryan et al. 2003; Wyngaard et al. 2004 and citations therein), and thus, no univocal conclusion on the ultimate benefits of CPS at these scales was reached (Gerard, 2007; Parodi and Tanelli, 2010; Li and Pu, 2009). Thus, because of this, simulations employing a CPS have often been applied at grid spacings ranging between 1 and 10 km, with their added value being largely case dependent. Preliminary tests for the three case studies considered herein showed that the activation of the CPS gave overall better forecast of the precipitation, which was deemed a reasonable rationale for its usage in this LDA exercise.

For the LDA, we have adopted the approach described in Fierro et al. (2012). First, the observed LINET lightning flashes were gridded onto the 3-km domain in 10-min intervals during the first 6 h of every simulation (therefore producing 36 periods for each assimilation window). Then, the parsed lightning data were perused to adjust (increase) in an incremental manner the simulated water vapor mass mixing ratio through the following relationship:

$$q_v = Aq_{vs} + Bq_{vs}tanh(CX) \cdot \left(1 - tanh\left(DQ_g^{\alpha}\right)\right) \quad (1)$$

in the mixed-phase region (defined here as the layer between 0°C and -20°C) and this only for grid points below a predefined relative humidity threshold of 95 % (set by A in eq. 1) and below a maximum graupel mixing ratio (Q_g) of 3 g/kg. Hence, water vapor adjustment (increase) is a function of the gridded lightning flash rate (X) and simulated graupel mixing ratio. q_{vs} is the water vapor saturation mixing ratio (g/kg), and the coefficients were set to A=0.95, B=0.07, C=D=0.25, and α=2.2. A was set to a larger value than in other studies (Federico et al. 2017a, Fierro et al. 2012) to maximize the number of gridpoints nudged, and therefore promote a faster onset of convection where lightning was detected, given the fact that some of the studied events were of rather short-lived nature. Additional sensitivity tests using A=0.85 were performed to evaluate the impact of introducing less water mass into the domain on the 3-6 h precipitation forecasts.

2.3. Verification

Statistical verification is performed through the use of 2 x 2 contingency tables for different precipitation thresholds, starting at 1 mm/3h and monotonically increasing up to 60 mm/3h. The hits *(a),* false alarms *(b)*, misses *(c)*, and correct negatives *(d)* of the contingency tables are computed. The probability of detection (POD; range [0, 1], where 1 is the perfect score, i.e., when no misses or false alarms occur), the false alarm ratio (FAR; range [0, 1], where 0 is the perfect score), the frequency bias (FBIAS; range [0, +∞), where 1 indicates no biases), and the equitable threat score (ETS; range [-1/3, 1], where 1 represents a perfect forecast skill and 0 no skill) were also used for forecast evaluation and are defined as follows (Wilks, 2006):

$$POD = \frac{a}{a+c} \quad (2)$$

$$FAR = \frac{b}{a+b} \quad (3)$$

$$FBIAS = \frac{a+b}{a+c} \quad (4)$$

$$ETS = \frac{a-a_r}{a+b+c-a_r} \text{ and } a_r = \frac{(a+b)(a+c)}{a+b+c+d}, \quad (5)$$

where a_r is the probability to have a correct forecast by random chance.

The elements of the contingency tables (a, b, c, d) are summed for all three storm events before computing the aggregate scores.

The POD gives the capability to the model of forecasting the exceedance of a certain rainfall threshold in the domain. Values close to one indicate that the experiment has a near-perfect performance, while in contrast, values near to zero show that the model could not correctly forecast the events. The FAR provides a measure of the fraction of rain forecast events that did not occur, therefore a value of one signifies that only “false alarms” and no correct events were forecast, while values of zero denote that only correct events and no “false alarms” were forecast. The FBIAS is

defined as the fraction of rain forecast events with respect to the rain observed events. The ETS measures the fraction of observed and/or forecast rain events that were correctly predicted, adjusted for hits associated with a random forecast, where the forecast occurrence/non-occurrence is independent of observation/non-observation.

All in all, the main hypothesis to be tested in this study is whether the water vapor-based LDA scheme is able to improve the model short-term precipitation forecasts via a better depiction of the intensity and locations of the main storm-scale objects in the domain.

3. Results

For every case study, a representative 3-h time period has been chosen for a more indepth examination of the forecast rainfall, which are discussed in the next subsections. Figure 2 shows the synoptic setup as portrayed by ECMWF's IFS at 00 UTC on the day of the case study. Figure 3 displays the lightning flash density projected onto the 3-km WRF grid during the target periods 09-12 UTC for the Sardinia (10 October 2018) and Livorno (10 September 2017) events, and 06-09 UTC for the Calabria (12 August 2015) case. Last, figure 4 shows the total daily observed precipitation amounts (mm) for each case study.

3.1. The 10 October 2018 case study

The intense rainfall event that impacted the southern portion of the island of Sardinia on 10 October 2018 produced a maximum 24-h amount of 452 mm in Capoterra just a few km west of the capital Cagliari, which barely registered 50 mm during that same day. The second maximum 24-h rainfall measurement was recorded only a few tens of km away in Tertenia over the southeastern Sardinian coast, with an amount of 360 mm (Fig. 4a). During the period 09-12 UTC, 182 mm of rain were recorded in Capoterra, while the second maximum amount was just 16 mm over the southeastern coastal region (Fig. 5g). At 00 UTC on 10 October, the synoptic setup was characterized by southeasterly flow at low levels that advected moist warm air towards the mountainous region of southern Sardinia. At midlevels, a cut-off low located over the Iberian peninsula favored larger-scale diffluence over Sardinia, along with relatively cold 500 hPa temperatures and appreciable vertical windshear (an approximately 90° clockwise turn from light breeze at surface to around 15 m/s at midlevels). These environmental factors led to a very unstable airmass triggering long-lived strong convection over that area (Fig. 2a). Figure 3a shows a circular and localised area of intense electrical activity over the central coast of the Gulf of Cagliari, with a peak 3-h lightning flash rate of about 100 for the period 09-12 UTC 10 October 2018 located over the Capoterra municipality.

Figure 5 shows the simulated precipitation for 09-12 UTC 10 October for all the experiments. For the simulations initialized with GFS data, the amount of rainfall is overall small over Capoterra and,

even, nonexistent in southeast Sardinia for CNTRL (a). ASSIM+3 (b), however, succeeds in producing up to 50 mm in Capoterra and 5-10 mm over the southeast. ASSIM+6 (c) also outperforms CNTRL, producing less precipitation over Capoterra than ASSIM+3 but in an area more confined along the east coast. When inspecting the ECMWF simulations, it becomes clear that CNTRL (d) is overall superior to CNTRL GFS, as evidenced by improved location and amounts of precipitation in Fig. 5d vs Fig. 5a. For the ECMWF simulations, the runs with LDA also outperform CNTRL (Fig. 5 e-f). The rainfall prediction from ASSIM+3 (Fig. 5e) is arguably the most accurate as accumulated precipitation fields not only match the ASSIM+3 GFS (b) rainfall forecast in the Capoterra region but also closely match observed values (15-20 mm) in the southeast (Fig. 5g). At the detriment of overestimating the rainfall peak there, ASSIM+6 ECMWF (Fig. 5f) also reproduces some of the observed precipitation in the southeast but underestimates the rainfall amounts in Capoterra relative to ASSIM+3.

Figure 6 shows vertical cross-sections of specific key variables at 10 UTC 10 October over the southern portion of Sardinia where the largest observed rainfall amounts were recorded. From this figure, it becomes apparent that all ASSIM experiments enhance vertical velocity and/or relative humidity fields with respect to CNTRL, with stronger updrafts and downdrafts and higher relative humidity. For the runs initialized with GFS data, ASSIM+3 (Fig. 6b) produced more vigorous and areally extensive vertical motions particularly over the areas with observed precipitation, while ASSIM+6 (Fig. 6c) shows weaker vertical velocities at midlevels with an areal coverage of near-saturation conditions that is notably larger. For the experiments initialized with ECMWF data (Fig. 6 d-f), both updrafts and downdrafts, at low and midlevels, are stronger and more widespread over the areas of observed precipitation compared to CNTRL (Fig. 6d), and also produce much larger values of relative humidity in the mixed-phase regions. As an example, the ASSIM+3 (ASSIM+6) experiment generates maximum vertical velocities of 1.9 (1.4) m/s, which is about three (two) fold those produced in CNTRL (0.6 m/s).

The above forecast evaluation analysis is complemented by an inspection focusing on the variation of some of the key variables impacting the forecast skill for the Sardinia event. Figure 7 shows the time evolution of the mean columnar contents of water vapor mass mixing ratio, rain water and ice, and maximum vertical windspeed on 10 October 2018 for the geographical domain shown in Fig. 7a. The mean values of water vapor mass mixing ratio for all simulations with the LDA are notably larger than their control counterparts for most of the event. This difference appears to be clearer for the simulations utilizing ECMWF data. All ASSIM+3 experiments consistently exhibit larger water vapor mass mixing ratio values than ASSIM+6. A similar pattern can be observed when examining in more detail rain water and ice mass columnar contents in Fig. 7b and c, respectively. Particularly for the time period 09-12 UTC of intense precipitation in Capoterra, it becomes clear that all

ASSIM experiments produce a relatively sharp decrease of these quantities compared to their CNTRL equivalents. The maximum vertical windspeeds displayed in Fig. 7d indicate that all experiments generate magnitudes that remained comparable for the most part of the day (until 18 UTC), with the peak of ASSIM+6 ECMWF at 12 UTC arguably being the result of a spurious convective cell developing over the southeast area (Fig. 5f). During the latter part of the day, a more active cold frontal system surged over the area to trigger the daily maxima of vertical velocity, water vapor mass mixing ratios, rain water and ice mean columnar contents.

3.2. The 12 August 2015 case study

During the morning hours and until around noon of 12 August 2015 (03-12 UTC), a severe storm cell developed and dwelled over the northern portion of the Calabrian Appennines, causing flash flooding in the Citrea Creek, which affected the nearby towns of Corigliano and Rossano, with notable damage to property and infrastructures. This storm generated a maximum rainfall amount of 185 mm from 00-09 UTC, and of 52.4 mm between 06-09 UTC (Fig. 8e). This was, however, an extremely isolated convective event, as evidenced by the fact that most rain gauges in the neighboring areas only registered very small amounts of precipitation (i.e., up to 30 mm for the whole day). Locally, the lightning activity was moderate and aerially confined, while an extensive area of more intense lightning activity located over the sea in the north of the Gulf of Taranto (Fig. 3b). The synoptic setup was marked by a well-defined midlevel cut-off low and a surface cyclonic circulation centered over the northern Sicilian coast, advecting warm and moist air from the east towards the Calabrian Appennines (up to 1500-2000 m ASL elevation). Together with cold temperatures at 500 hPa and moderate vertical windshear, these conditions favored the development of strong orography-driven thunderstorms (Fig. 2b).

Figure 8 is as figure 5 but for the period 06-09 UTC 12 August and with only one (the right) column for the ASSIM experiments since ASSIM+3 and ASSIM+6 coincide for this particular time period. When inspecting the two CNTRL runs, we notice that the rainfall forecast in the GFS simulation over land is, overall, quantitatively correct but suffers from a northwestern bias (of approximately 25 km) relative to the observations (Fig. 8a). This discrepancy is arguably due to inaccuracies (biases and errors) in the GFS initial and boundary conditions data. For the ECMWF run (Fig. 8c), the rainfall locations are more consistent with the observations, but their amount is generally underestimated. Turning our attention to the ASSIM simulations (Figs. 7b, d), it becomes clear that assimilating lightning yields precipitation forecasts in the correct areas over the Citrea Creek basin (where flash flooding took place), despite underestimating their amount by as much as 20 mm. Despite this improvement, the ASSIM simulations still induced too many false alarms in the Calabrian peninsula to the southeast of Citrea Creek, especially the ASSIM ECMWF run (d),

which forecasted a local maximum near Crotone, where lightning activity was also observed (Fig. 3b).

Figure 9 shows vertical cross-sections at 09 UTC 12 August along the areas affected by the flash flooding event. For CNTRL GFS (Fig. 9a), vertical ascent is rather weak and located mostly over the sea (consistent with predicted rainfall as indicated earlier) and areas of high relative humidity are confined to lower levels. For ASSIM GFS (Fig. 9b), however, stronger updrafts are evident (10-fold increase in maximum vertical velocity with respect to CNTRL) with a well-defined core centered at around 550 hPa over the Citrea Creek, together with near saturation conditions throughout the column. For the simulations with ECMWF, a similar pattern is seen. The convective core in the ASSIM experiment (Fig. 9d) is located over the same location as for the ASSIM GFS, but is overall weaker with a maximum value of 3.2 m/s (2.5-fold increase with respect to CNTRL (Fig. 9c)) and at a lower altitude (800 hPa). By design of the LDA, which adjusts (increases) water vapor mass mixing ratio, higher relative humidity values are also more vertically extensive in the ASSIM runs in comparison to CNTRL, with values exceeding 90 % reaching up to 400 (500) hPa in the case of ASSIM GFS (ECMWF).

3.3. The 10 September 2017 case study

This widespread and long-lived mesoscale convective system affected large portions of the Italian peninsula from northwest to southeast, as well as the island of Sardinia, from the evening of the 9 throughout 10 September 2017. This storm especially ravaged the Tuscan coastal city of Livorno first, and later impacted Rome and the Lazio region with severe-warned strength, while also affecting most of central Italy as illustrated in figure 4c. A maximum 3-h rainfall amount of 137.7 mm was measured near Rome between 09-12 UTC (Fig. 10g), with the maximum daily rainfall recorded in Bibione (Venetian coast) at 251.2 mm. In Livorno, nine casualties and severe damage were reported during the early hours of 10 September. The large-scale meteorological conditions at 00 UTC were dominated by an upper-level trough tilted from SW to NE (from southeastern Spain to northern Europe), and by a surface low pressure system over northwestern Italy that favored southwesterly advection of a moist airmass towards the peninsula over a deep layer. Large vorticity advection at 500 hPa over the Tuscany and Lazio coastlines further enhanced mesoscale vertical ascent there (Fig. 2c). As the day progressed, the system moved southeastward, and by 09-12 UTC intensified over the Lazio region and coast, as indicated by a marked uptake in lightning activity (Fig. 3c). From there, it slowly moved eastward and then weakened into scattered areas of largely stratiform precipitation along the Appennines, before intensifying again in the Venetian coast between 21-00 UTC to produce rainfall amounts up to 150 mm.

This analysis will focus on the time period of 09-12 UTC, during which the severe thunderstorm system passed over central Italy and produced considerable rainfall amounts in several locations over the northern regions of Lombardy and Friuli Venezia Giulia (Fig. 10g). Figure 10 shows the rainfall forecasts for all the experiments. The two CNTRL simulations do succeed in producing precipitation peaks of comparable magnitude, but remained too restricted/confined in areal coverage and positioned too far to the north with respect to the observations (Fig. 10g). The CNTRL GFS (Fig. 10a) experiment produces two narrow bands of precipitation over Tuscany, while CNTRL ECMWF (Fig. 10d) is able to predict a single, wider area of precipitation with values more aligned with the observations. In both cases, however, these rainfall bands are displaced tens of kilometres away from the observed locations (northwest bias). The lightning assimilation experiments ASSIM+3 with GFS (Fig. 10b) and ECMWF (Fig 9e) notably improve rainfall prediction over CNTRL by reproducing rainfall clusters very similar to raingauge measurements and of very similar amounts, extending all the way to the Adriatic coast. Both LDA cases also predicted reasonably well the rainfall amounts over the northeastern corner of Italy. The ASSIM+6 runs display the same tendency as ASSIM+3 in terms of forecasting the large area of precipitation in the correct locations, although ASSIM+6 GFS (Fig. 10c) falls short in terms of amounts but succeeds in capturing the rainfall event near Piombino in the Tuscan coast. The ASSIM+6 ECMWF (Fig. 10f) experiment does a reasonable job in forecasting the total amount of precipitation over Rome but its areal coverage is generally underestimated relative to the observations and relative to its ASSIM+3 counterpart.

Altogether, these results demonstrate how the assimilation of lightning flash density data helps to improve the skill of shorter-term precipitation forecasts by correctly forecasting the location, timing and intensity of the main convective band. This is also confirmed by examining the predicted fields of precipitable water and 850 hPa vertical velocity at 12 UTC, 10 September in figure 11. Both CNTRL runs (Fig. 11a, d) show the band of high precipitable water lagging behind to the northwest with respect to both ASSIM+3 simulations (Fig. 11b, e), which exhibit a well-defined low-level vertical ascent band just to the east of the area of predicted rainfall between 09-12 UTC (Fig. 10b, e). As expected, the results from ASSIM+6 (Fig. 11c, f) show a similar behaviour to ASSIM+3 in terms of the position of the rainfall band, although these experiments do not seem to predict a similarly well-defined band of 850 hPa vertical motion, explaining, in turn, why their precipitable water values are underestimated relative to the observations.

3.4. Statistical scores

The statistical scores were computed for rainfall thresholds of 1 mm/3h and from 2 to 60 mm/3h by 2 mm/3h increments. Figure 12 shows the four contingency elements for all simulations (FBIAS

(a), POD (b), ETS (c) and FAR (d)). It is clear that the two simulations without LDA (CNTRL GFS and CNTRL ECMWF) perform worse across all rainfall thresholds than the simulations that assimilated lightning. This difference becomes particularly notable from 30 mm/3h upwards, where the POD of the CNTRL runs is less than 0.2 and even asymptotes to 0 at higher rainfall thresholds, while the runs with LDA produce nearly constant POD values ranging between 0.3-0.4. This illustrates the ability of the LDA to reasonably capture severe rainfall events, in contrast to the CNTRL experiments. Among the simulations with LDA, it can be noted that the ASSIM+3 runs produce higher values of POD across all thresholds compared to their corresponding ASSIM+6 runs. Such result is explained by the decreasing impact of the LDA with forecasting time, as the inherent biases and errors contained in the initial conditions progressively govern the longer-term solution.

One apparent feature of these results is that all three sets of simulations (CNTRL, ASSIM+3 and +6) seem to exhibit FAR values (Fig. 12d) somewhat dependent to the global model used, with those using GFS data displaying better (i.e., lower) FARs. ASSIM+3 and ASSIM+6 with GFS exhibit markedly better FAR values than their corresponding ECMWF simulations. The differences in these values become already evident at the 10 mm/3h threshold. For the 20-45 mm/3h range, both ASSIM+3 and ASSIM+6 with GFS (ECMWF) produce FAR between 0-0.15 (0.20-0.40). Since the difference in hits for this rainfall threshold range is not significant between them (as seen in the PODs), it follows that the LDA simulations with GFS consistently produce fewer false alarms, as their lower FAR values are not the result of more hits. The LDA simulations with ECMWF show a steady reduction of the FAR with increasing precipitation threshold, until the differences between both groups become negligible at thresholds larger than 55 mm/3h. The CNTRL runs exhibit a generally opposite behavior for thresholds larger than 50 mm/3h: the FAR of CNTRL GFS equates 0 because this run produces no false alarms, but at the expense of also missing most rainfall events of intensity equal to or larger than 40 mm/3h. In contrast, the FAR values for CNTRL ECMWF hover near unity because this experiment solely generates false alarms but no hits for thresholds exceeding 48 mm/3h.

The FBIAS score results (Fig. 12a) show a similar trend to those seen for the POD. The CNTRL simulations performed poorly at all thresholds with biases approaching 0 for thresholds larger than 50 mm/3h, as only a very small number of intense rainfall events are forecast. For thresholds exceeding 15 mm/3h, the bias at 3-h forecast is always closer to 1 than at 6-h regardless of the global model used for model initialization (GFS or ECMWF), confirming the progressive decline of the LDA impact with forecast time. All LDA simulations produced a nearly constant bias ranging between 0.25-0.50 for rainfall thresholds exceeding 25 mm/3h.

In terms of ETS (Fig. 12c), ASSIM+3 GFS produces the best results for thresholds smaller than 35 mm/3h followed by ASSIM+6 GFS, which is traced back to low numbers of false alarms in these experiments. The ETS for ECMWF's ASSIM+6 are generally larger than ASSIM+3 for thresholds between 18 to 34 mm/3h. This counterintuitive result, however, likely is a consequence of the very small sample of cases considered in this study (i.e., only three), warranting the need to include a larger number of cases in future work.

It is important to highlight that the scores of figure 12 for the ASSIM+6 forecasts have been derived considering the forecasts between 3 and 6 h after the end of the assimilation phase. Because both ASSIM+6 ECMWF and ASSIM+6 GFS outperform their corresponding CNTRL forecast, this entails that the LDA still impacts the forecast after three hours.

Because model output for the current simulations are saved only in 3 h intervals, the assessment of the LDA impact cannot be performed for time ranges smaller than 3 h. However, a more detailed examination of the model results and precipitation forecasts can be conducted for longer time ranges; i.e., 6-h precipitation forecasts. This additional assessment was carried out for the simulations CNTRL ECMWF, CNTRL GFS, ASSIM+6 ECMWF and ASSIM+6 GFS. Results of this analysis (not shown) are similar to those depicted in Fig. 12 and can be summarized in two main take aways: a) the rainfall forecast is improved by the LDA regardless of the larger scale model used to derive the ICs and BCs for WRF (i.e., ECMWF or GFS); b) the forecasts using GFS as ICs and BCs, with or without LDA, perform better than the corresponding forecasts using ECMWF. This improvement is apparent for POD, FAR and ETS, while FBIAS shows mixed results with WRF driven by ECMWF showing better performance for most 6-hourly rainfall thresholds.

To complete the evaluation of NWP rainfall forecast, Roebber performance diagram for all simulations (Fig. 13) were constructed for the rainfall thresholds of 10 and 50 mm/3h and for nearest neighbourhood radii of 8.5 and 20 km. These diagrams highlight in a concise manner that ASSIM+3 GFS is the most skillful simulation regardless of the rainfall threshold or search radius selected, closely followed by ASSIM+6 GFS. This analysis offers additional evidence of the noteworthy sensitivity of the forecast rainfall to the initial and boundary conditions used to integrate the model with and without LDA. Another clear feature is the systematic underperformance of the control simulations when compared to LDA experiments independently of the search radii or threshold considered. As expected, increasing the neighborhood radius from 8.5 to 20 km improves model skill at all rainfall thresholds owing to larger hits and smaller FAR. However, the advantage of the LDA scheme come at the cost of exacerbating wet biases with respect to control experiments.

3.5. LDA sensitivity experiment

Given the fact that the LDA scheme used a relatively high (A=95 %) relative humidity threshold below which the water vapor mass is adjusted following the relationship in Eq. 1, we opted to perform a sensitivity analysis for the 10 October 2018 case in which the constant A was lowered down to 85 %. As indicated in Fierro et al. (2012), reducing A has for primary effect to reduce the integrated impact of the LDA over the domain since fewer grid points will satisfy RH < 85 % (relative to 95 %) for water vapor mass adjustments.

With these settings, the ASSIM+3 simulations predict notably less rainfall over the Capoterra area than the original LDA simulation with A=95 % (Fig. 14). These differences remain overall negligible over the southeastern portions of Sardinia. The differences in forecast precipitation for the ASSIM+6 simulations remain minimal in the Capoterra area. Over the southeast, however, the ASSIM+6 ECMWF run using RH < 85 % does not produce the spurious maximum precipitation that was forecast by its counterpart at RH < 95 %. This highlights an obvious consequence of wet bias exacerbation by the LDA scheme when more grid points are activated at higher thresholds for A (or RH). Nevertheless, ASSIM+6 GFS with RH < 95 % is still able to forecast – overall – more consistent precipitation patterns and amounts relative to the raingauge observations compared to the new simulation with RH < 85 %.

Inspecting now the time evolution of the mean columnar variables and vertical velocity for this experiment in Fig. 15, it can be seen that all ASSIM experiments produce somewhat lower values of water vapor mass than their RH < 95 % counterparts (Fig. 7), but become virtually undistinguishable from the control experiments for the rain water and ice variables. The same holds for the maximum vertical windspeed.

All in all these results suggest that, while there is some value in attempting to minimize water vapor mass adjustments in the model during the DA, the performance in terms of precipitation forecast for this case is generally worse than when using the higher moisture threshold for the LDA activation.

4. Discussion and conclusions

In this work, it is illustrated and confirmed how a simple nudging scheme published in the literature for total lightning data implemented in the WRF-ARW model is able to notably improve short-term (≤ 6 h) convection-resolving (3 km) rainfall forecasts for three severe events over Italy characterized by different convective regimes. This work also explores forecast sensitivity to different initial and boundary conditions datasets. Using observations from a dense rain gauge network as ground truth, the statistical scores for 3-hourly precipitation thresholds confirm the overall advantage of this LDA method over control simulations not assimilating any data. The LDA scheme imposed incremental increases of water vapor mass mixing ratio at every model timestep during the first 6 h of simulation (DA phase) within gridpoints that meet a specific, user-defined

fixed set of conditions, which are: a) at least one lightning flash has been detected in the last 10-min interval, b) the water vapor adjustment is only applied in the mixed-phase region defined as the layer between 0°C to -20°C, c) the grid-point relative humidity must be less than 95 % and the simulated graupel mixing ratio less than 3 g/kg for this water vapor adjustment to take effect.

A closer inspection of each case study helped us identify how the simulations are improved by the 6-h LDA phase. The simulations with LDA for the event of 10 October 2018 in Sardinia predict stronger updrafts and wider downdraft areas, and humidity profiles closer to saturation and higher values of columnar water vapor, rain water and ice content than the control simulations. In this case study, such effects from the LDA appear more noticeable in the runs using ECMWF data. For the rain event that affected Calabria on 12 August 2015, LDA simulations are able to advect the storm inland and over the mountains where it was observed, in contrast to the control runs which misplaced the heaviest rainfall areas mostly over the sea in addition to exhibiting large displacement bias for rainfall over land. In stark contrast to the control simulations, the vertical structure of the thunderstorms in the LDA experiments shows notable improvements in terms of their updraft cores and humidity profile. Nonetheless, as indicated in the original work describing this LDA scheme, false alarms can often arise from the addition of water vapor, such as the rainfall maxima produced over the Calabrian peninsula where lightning was detected but observed precipitation was negligible. The case of widespread flooding rainfall that occurred on 9-10 September 2017 in Livorno and Central Italy, provided perhaps the most convincing evidence of the LDA's ability to substantially improve the 3 and 6-h rainfall forecast. While the control experiments only generated relatively narrow bands of precipitation displaced hundreds of km away relative to the observations, the 3-h LDA simulations succeeded in predicting both the magnitude and areal extent/coverage of the intense and large precipitation region observed in Central Italy. The 6-h forecasts do not perform as well as for 3-h, but still depict successfully the area of precipitation in comparison to the control experiments. Although it is not feasible to reach robust conclusions with only three case studies, the overall results – supported by aggregate statistics – seem to suggest that the positive impact of the LDA is more appreciable when widespread and longer-lived convective events (especially MCSs) are forecasted (versus small-scale and shorter-lived), a finding in line with Fierro et al. (2015) for this LDA scheme evaluated over a 70 day period. This feature of our results makes sense not only from the perspective of the higher inherent predictability of mesoscale-dominated convective events over small-scale and shorter-lived phenomena, but also from the relatively larger amount of lightning activity assimilated into the model for this MCS case in comparison to the other localised events.

After a brief sensitivity test performed with the goal of assessing if the LDA base moisture threshold of 95 % was exacerbating wet biases, one of the case studies was reforecast but with a

lower threshold for DA activation (A=85 %). This additional experiment illustrated that, for that particular case, it was preferable to use the initially *high* threshold for LDA activation (A=95%) as this simulation was able to capture more realistically the primary high-intensity rainfall event over Capoterra in southern Sardinia (despite underestimating the overall amounts), while the lower moisture threshold experiment fell substantially short of forecasting the observed event. More systematic tests covering a broader range of values for A will be performed in the future to better address this relevant aspect of the LDA, which has been shown to also prevail with more complex DA methods such as 3DVAR (e.g., Fierro et al. 2016; 2019).

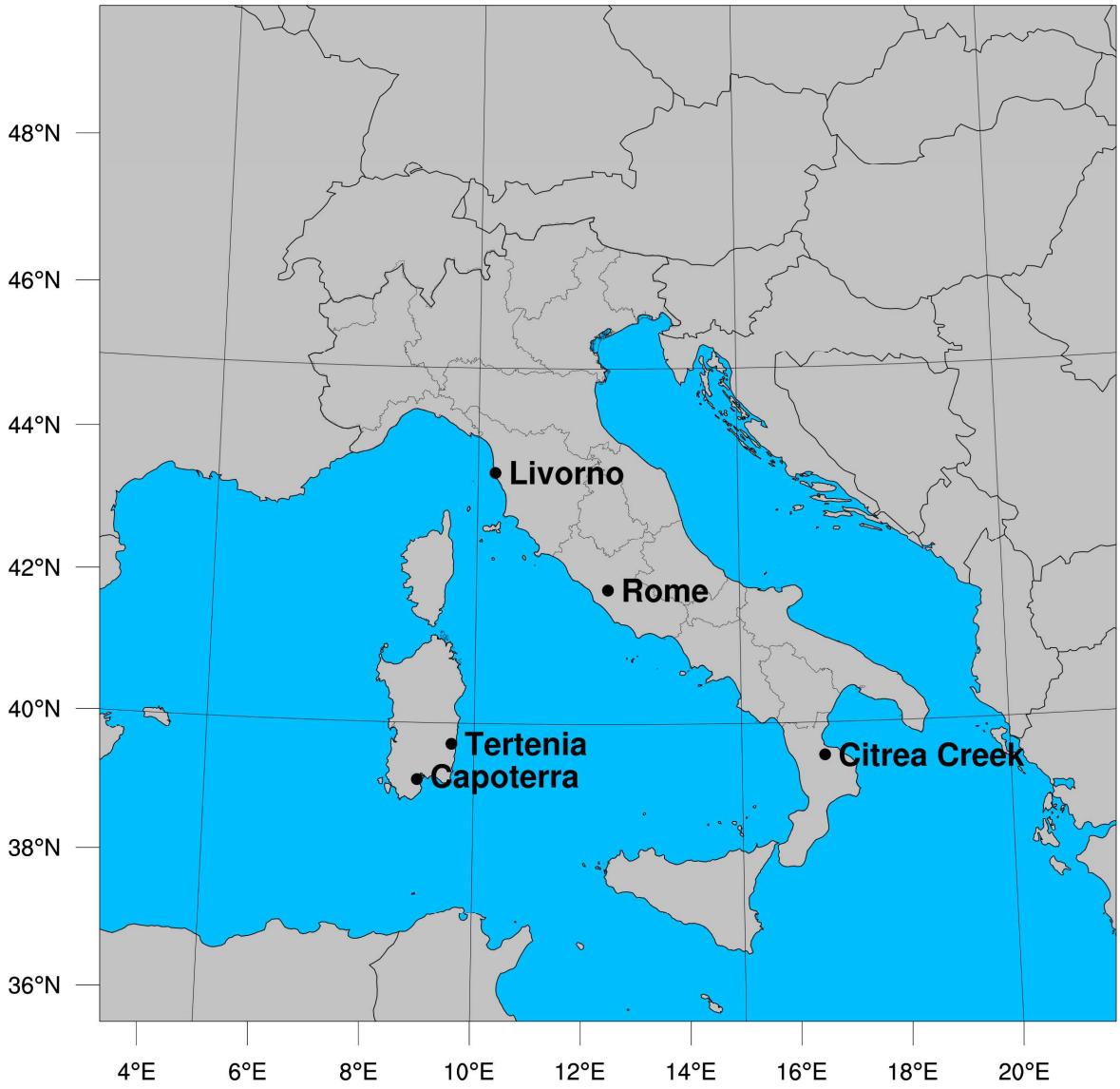
The 3 and 6-h statistical scores are generally improved with the LDA, which is traced back to notable improvements in POD. For these cases, LDA simulations are able to particularly improve rainfall intensities higher than 40 mm/3h. The FAR show some noteworthy degree of sensitivity to the initial conditions used (downscaled from the GFS or ECMWF) with the GFS experiments showing lower FAR values across most rainfall thresholds. Regardless of that, all LDA simulations exhibit a notable advantage over the control experiments especially for high rainfall thresholds, wherein false alarms decrease but hits remain nearly constant (i.e., there's a net improvement in the ability to not overpredict high rainfall events). These results are encouraging, given that increases in false alarms are expected by design of the LDA (i.e., increase in water vapor mass). The frequency bias scores also highlight the superiority of the LDA runs with a 3-h forecast range over that of 6-h. At high rainfall thresholds, biases remain roughly constant while the control runs exhibit negligible to no skill. The same general pattern is observed in the ETS scores. It is also noted that LDA runs initialized by downscaling GFS data produce the best scores overall. While this result cannot be generalized with only three cases, the results highlight the impact of initial and boundary conditions on the rainfall forecast.

It is useful to document some of the main differences between the results herein and those in a similar study using the RAMS model (Federico et al. 2017a). In this work, we found that both the bias and FAR decrease with rainfall threshold, whereas in their study the opposite behavior was documented. A possible explanation for that is the usage of a single-moment (WSM6) microphysics parameterization in Federico et al. (2017a), compared to a 1.5-moment (Thompson) scheme here. Setting aside the different NWP model used, this work also employed slightly finer horizontal grid spacing (3 km vs 4 km). More work is needed to determine the primary causes and factors for these differences.

As a final note, possible future research routes would include: (i) reducing the horizontal grid spacing further down to cloud-resolving scales (1-2 km), (ii) using a different definition for the mixed-phase layer, and instead of using the 0°C level as a lower limit, employing the lifting condensation level to potentially increase the vertical extent of the water vapor adjustments (e.g., as

in Fierro et al. 2016, 2019) and/or (iii) implementing a radar reflectivity factor assimilation method using 3DVAR in addition to lightning, as demonstrated in Federico et al. (2019) or Fierro et al. (2016; 2019; Hu et al. 2020).

561 **Figures**



594

595 Figure 1. Illustration of the WRF model domain used in this study, which consist of a 530 x 530
596 grid with horizontal spacing of 3 km, spanning about 1600 km in both horizonal directions.
597

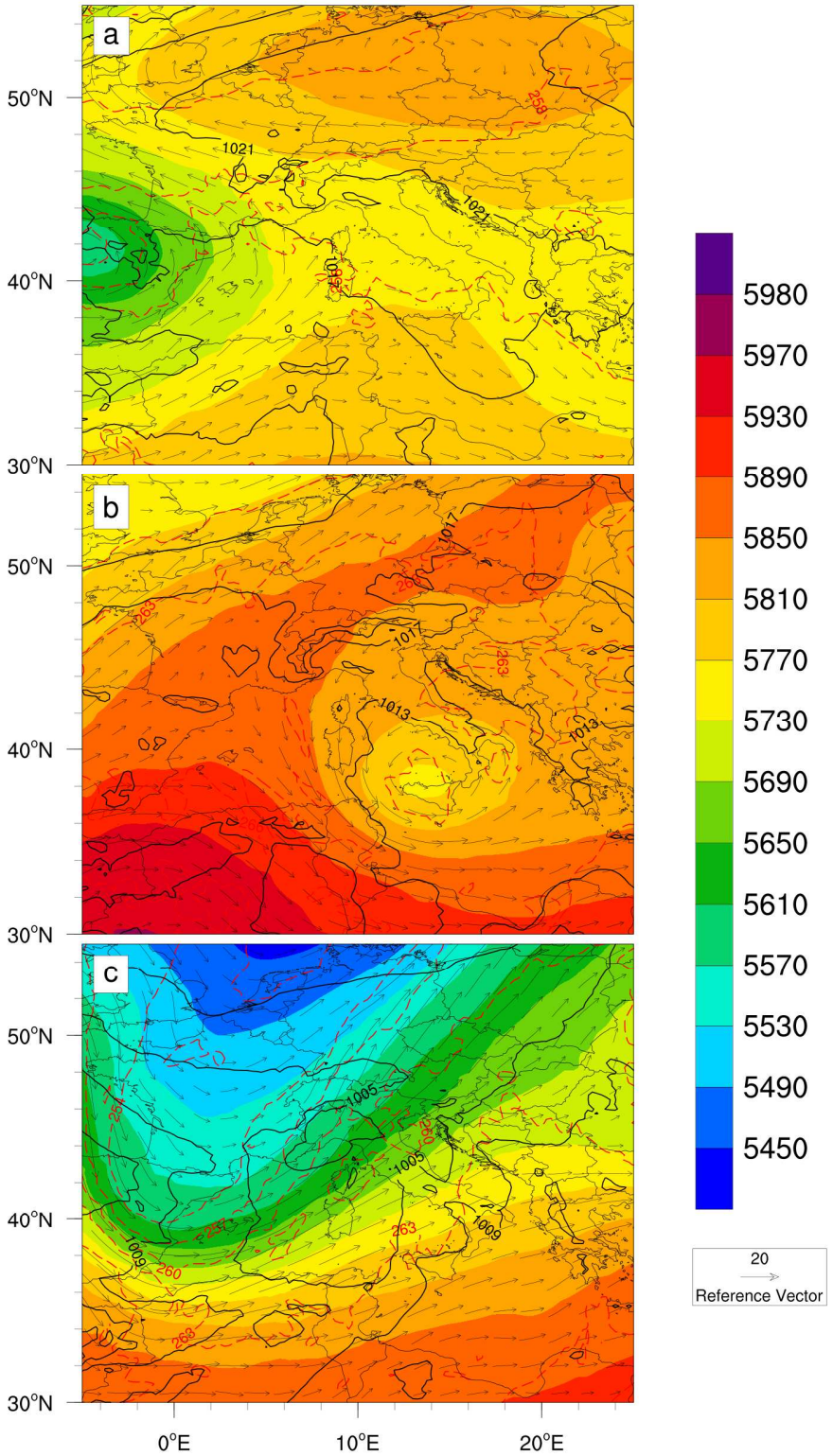


Figure 2. The ECMWF synoptic setups at 00UTC for each case study; on the 10 October 2018 for the Sardinia event (a), on the 12 August 2015 for the Calabria event (b), and on the 10 September 2017 for the Livorno event (c). The 500 hPa geopotential height filled contours are in m, the 500 hPa wind field's reference vector in m/s, the 500 hPa temperature contours (dashed red lines) are drawn every 3 K, and the mean sea level pressure contours (solid black lines) are shown every 4 hPa.

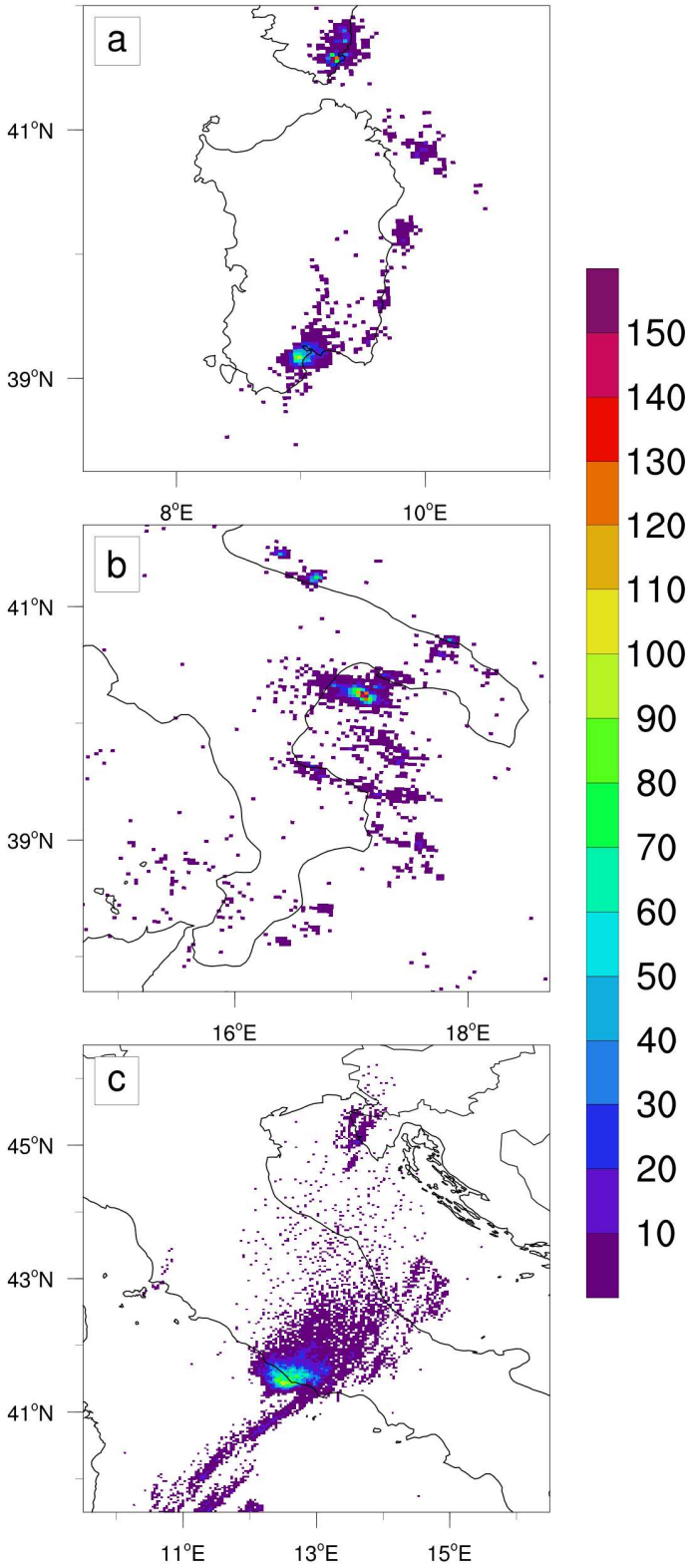
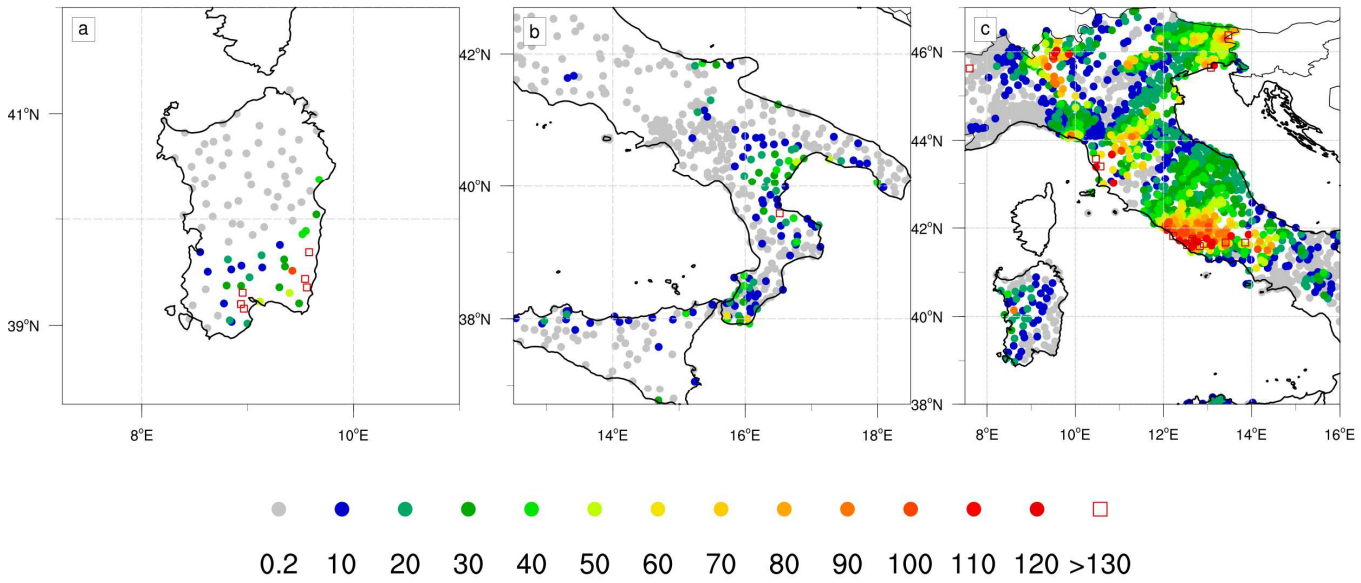


Figure 3. Lightning flash density (per 9 km^2) during the period 09-12UTC for the Sardinia (a) and Livorno (c) events, and 06-09UTC for the Calabria event (b). These have been obtained from the LINET network and remapped onto the WRF grid for said periods, although the model uses these data in 10 minute intervals as explained in Section 2.



611
612 Figure 4. Total daily observed precipitation (in mm) for the Sardinia (10 October 2018, left),
613 Calabria (12 August 2015, middle) and Livorno (10 September 2017, right) events.
614

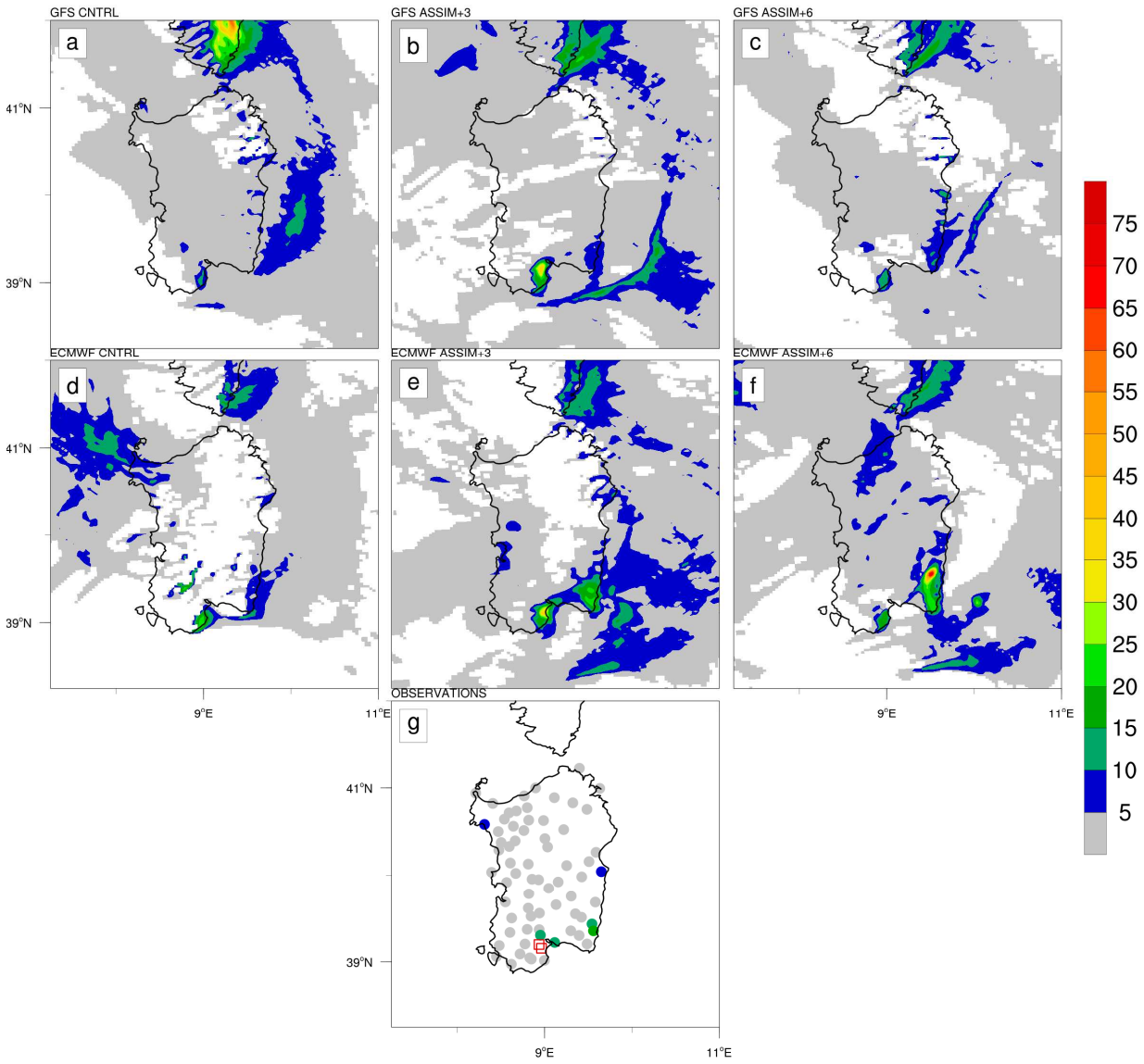


Figure 5. Simulated precipitation (mm) for 09-12UTC 10 October 2018 for runs with GFS (top) and ECMWF (middle row), and for the control (left), assimilation +3 (middle) and assimilation +6 (right) simulations. The panel on the bottom row shows the observed precipitation for the same time period (red square marks indicate values ≥ 75 mm).

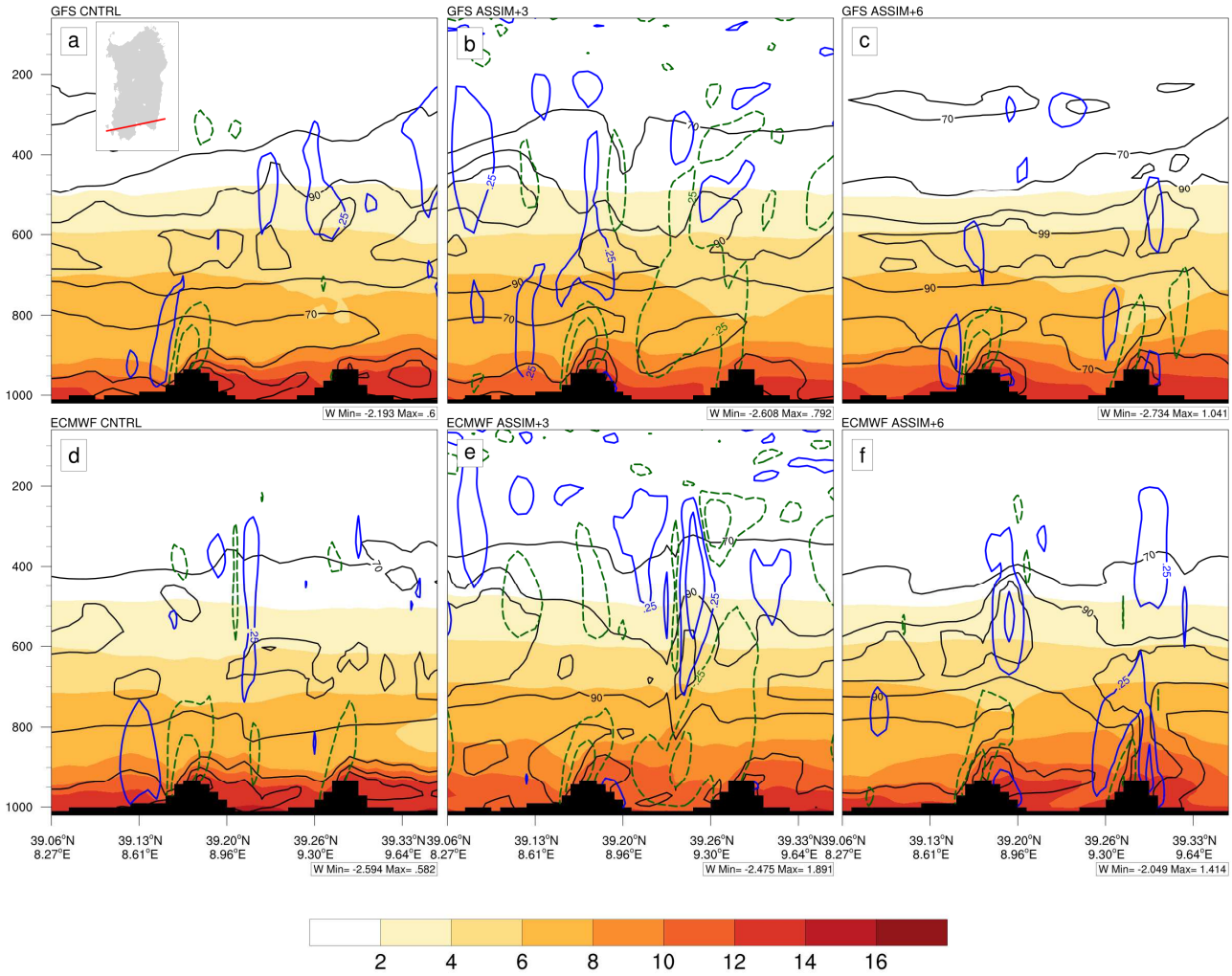
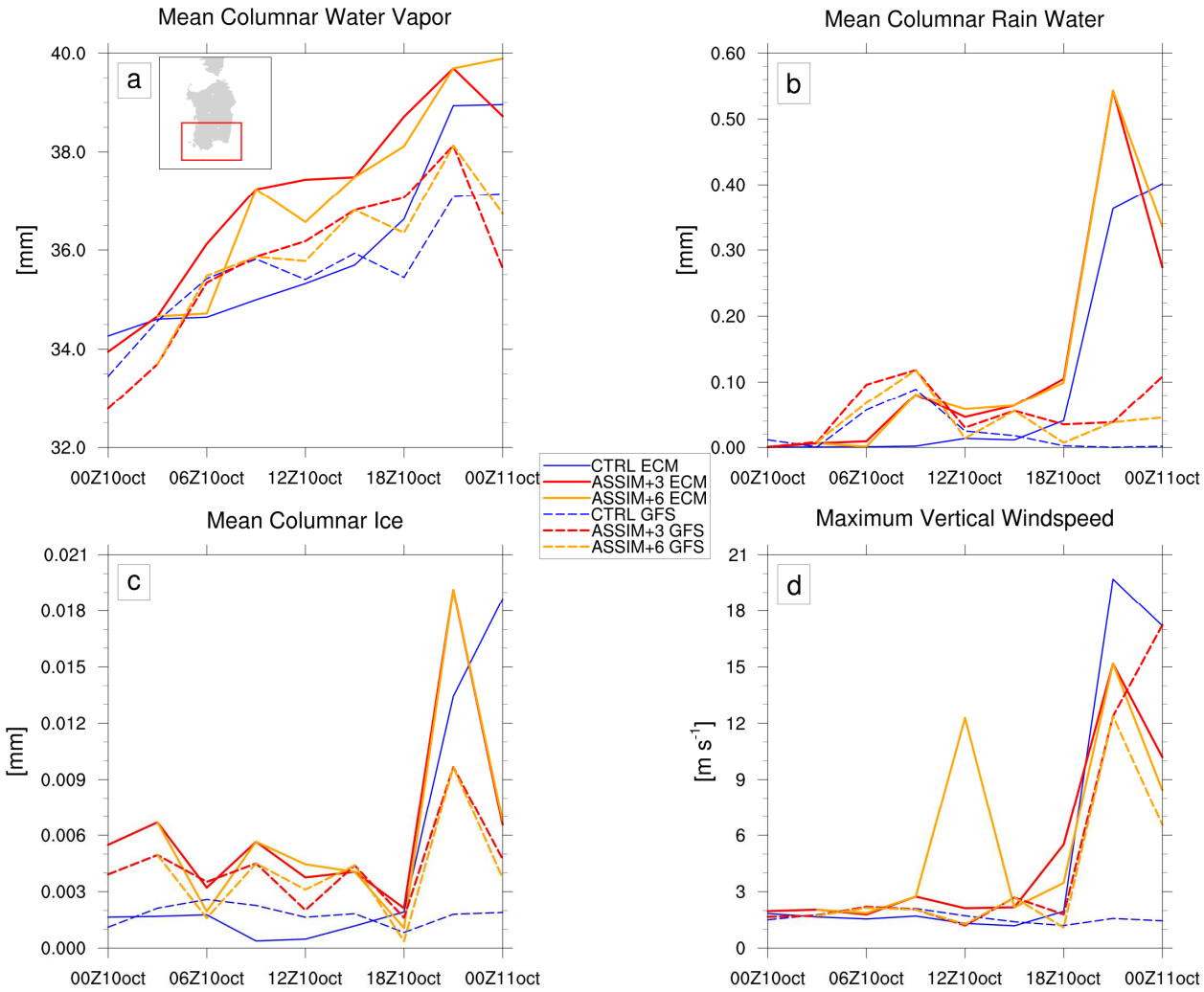
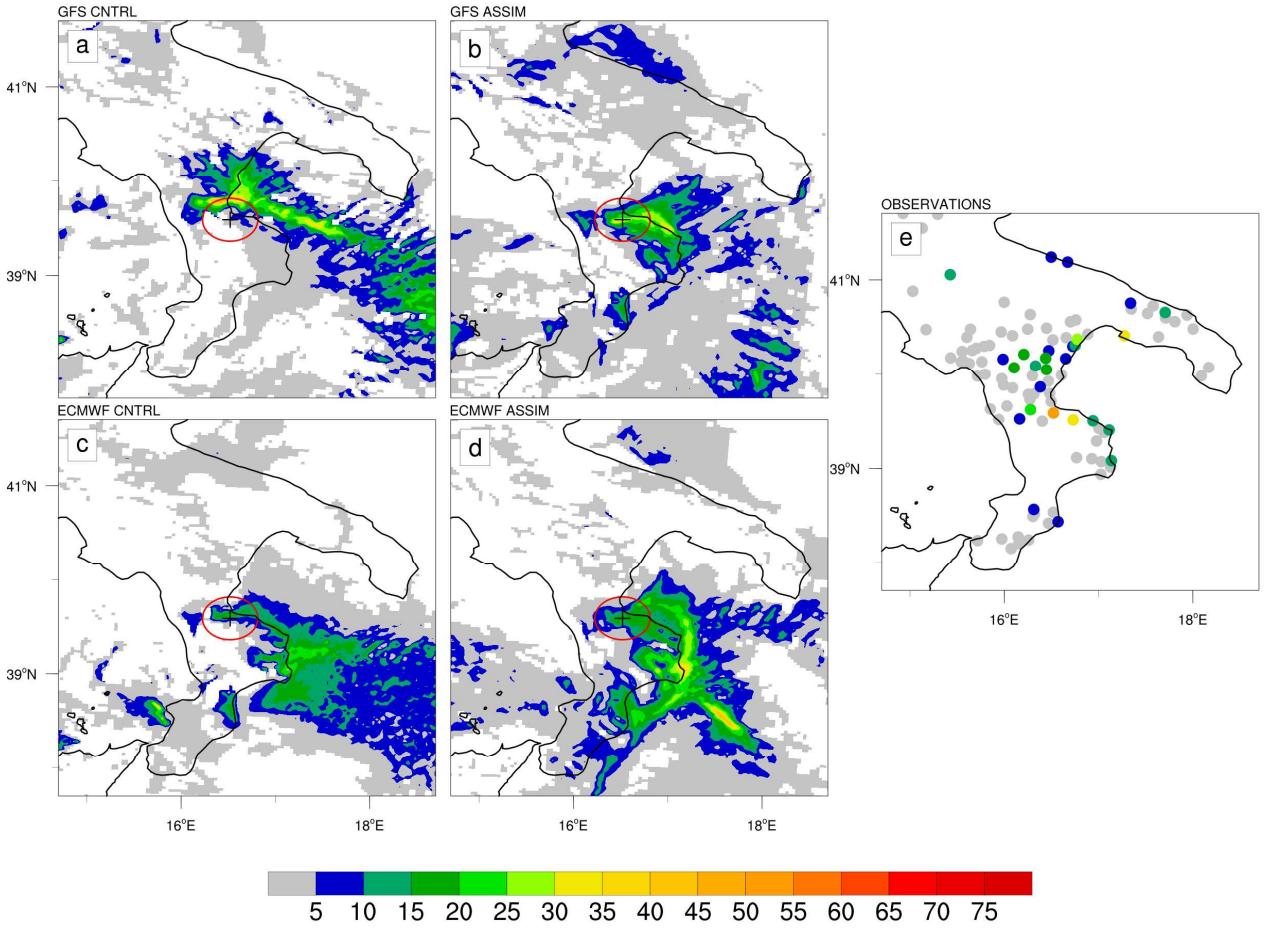


Figure 6. Simulated vertical sections across southern Sardinia at 10UTC 10 October 2018 of water vapour (g/kg, filled contours), relative humidity (%, solid black contours at 70, 90 and 99%), and vertical velocity (m/s, positive contours in solid blue, negative in dashed green, values at ±0.25, ±1, ±3 and ±5 m/s) (Vertical units in hPa).



627
628 Figure 7. Time evolution of mean columnar contents of water vapor (a), rain water (b) and ice (c),
629 and maximum vertical windspeed (d) during the 10 October 2018 for the area of southern Sardinia
630 as indicated inside panel (a).
631



632 Figure 8. Simulated precipitation (mm) for 06-09UTC 12 August 2015 for the experiments with
633 GFS (top) and ECMWF (bottom) input data, and for the control simulations (left) and assimilation
634 runs (middle column). The red circle represents the area of 25-km radius around the rain gauge
635 (black cross) with maximum measured rainfall for this period. The panel on the right shows the
636 observed precipitation for the same time period.
637
638

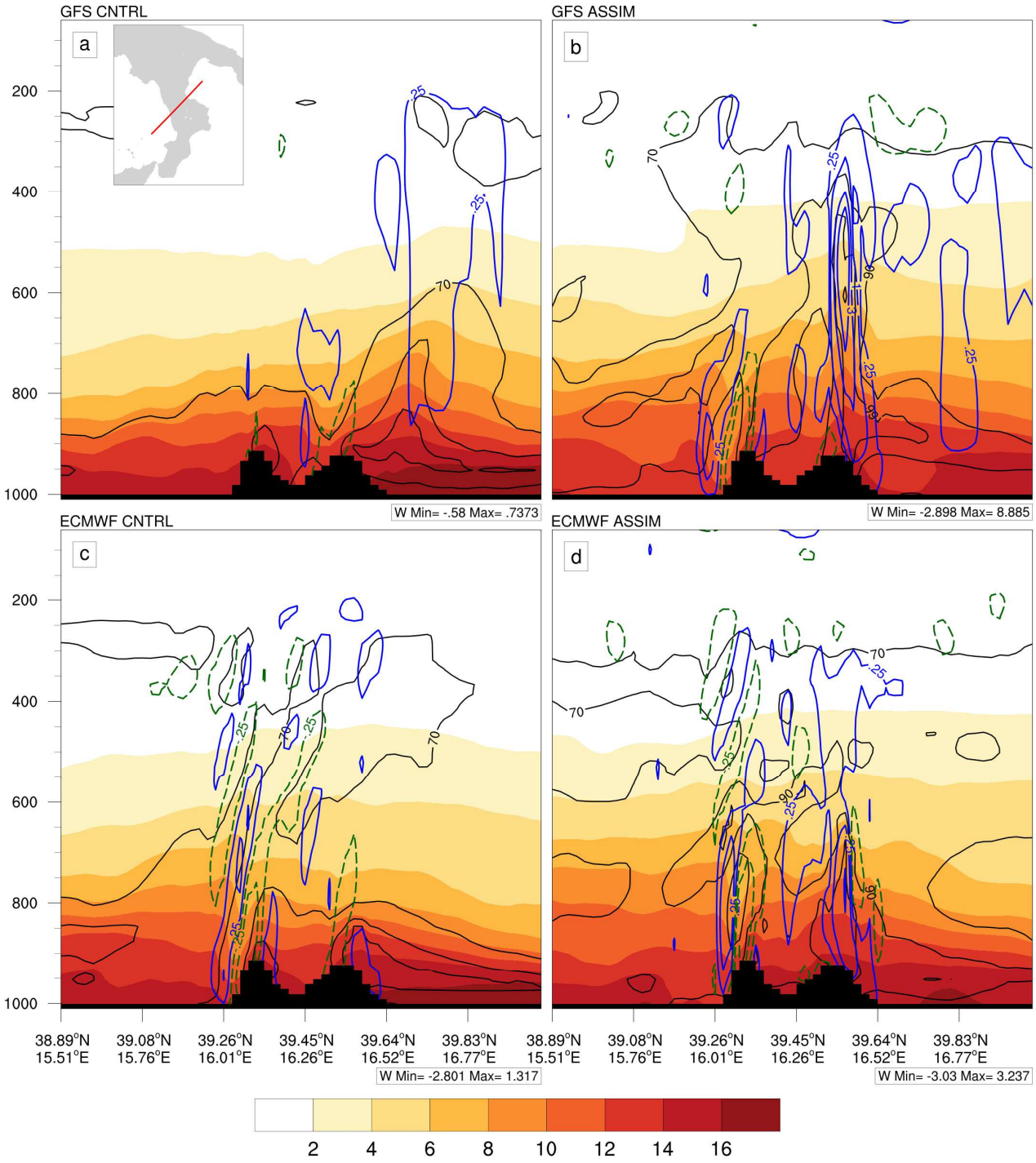


Figure 9. Simulated vertical sections across the Citrea Creek in Calabria at 09UTC 12 August 2015 of water vapour (g/kg, filled contours), relative humidity (%, solid black contours at 70, 90 and 99%), and vertical velocity (m/s, positive contours in solid blue, negative in dashed green, values at ±0.25, ±1, ±3 and ±5 m/s) (Vertical units in hPa).

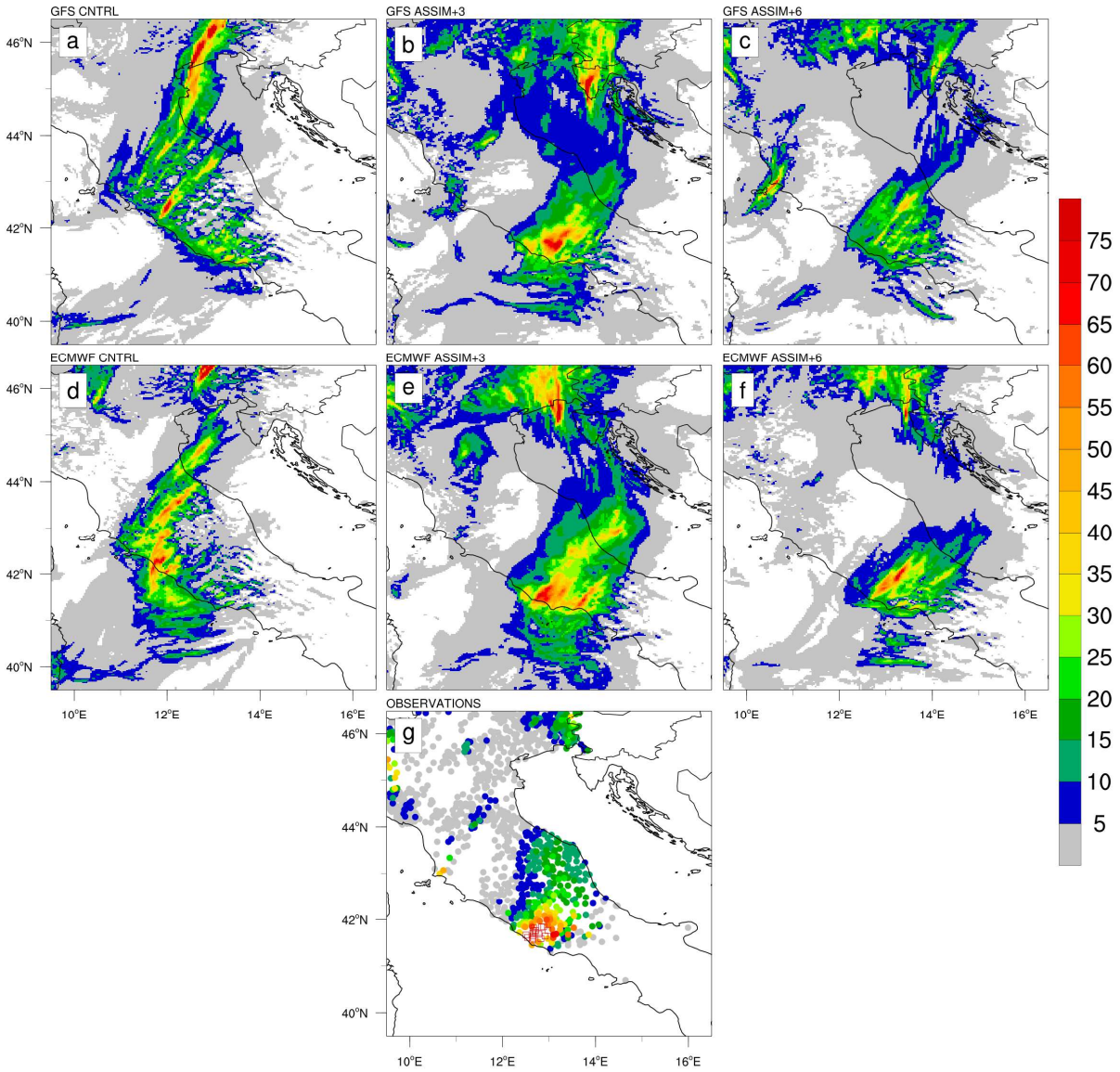


Figure 10. Simulated precipitation (mm) for 09-12UTC 10 September 2017 for experiments with GFS (top) and ECMWF (middle row) input data, and for the control (left), assimilation +3 (middle) and assimilation +6 (right) simulations. The panel on the bottom row shows the observed precipitation for the same time period (red square marks indicate values ≥ 75 mm).

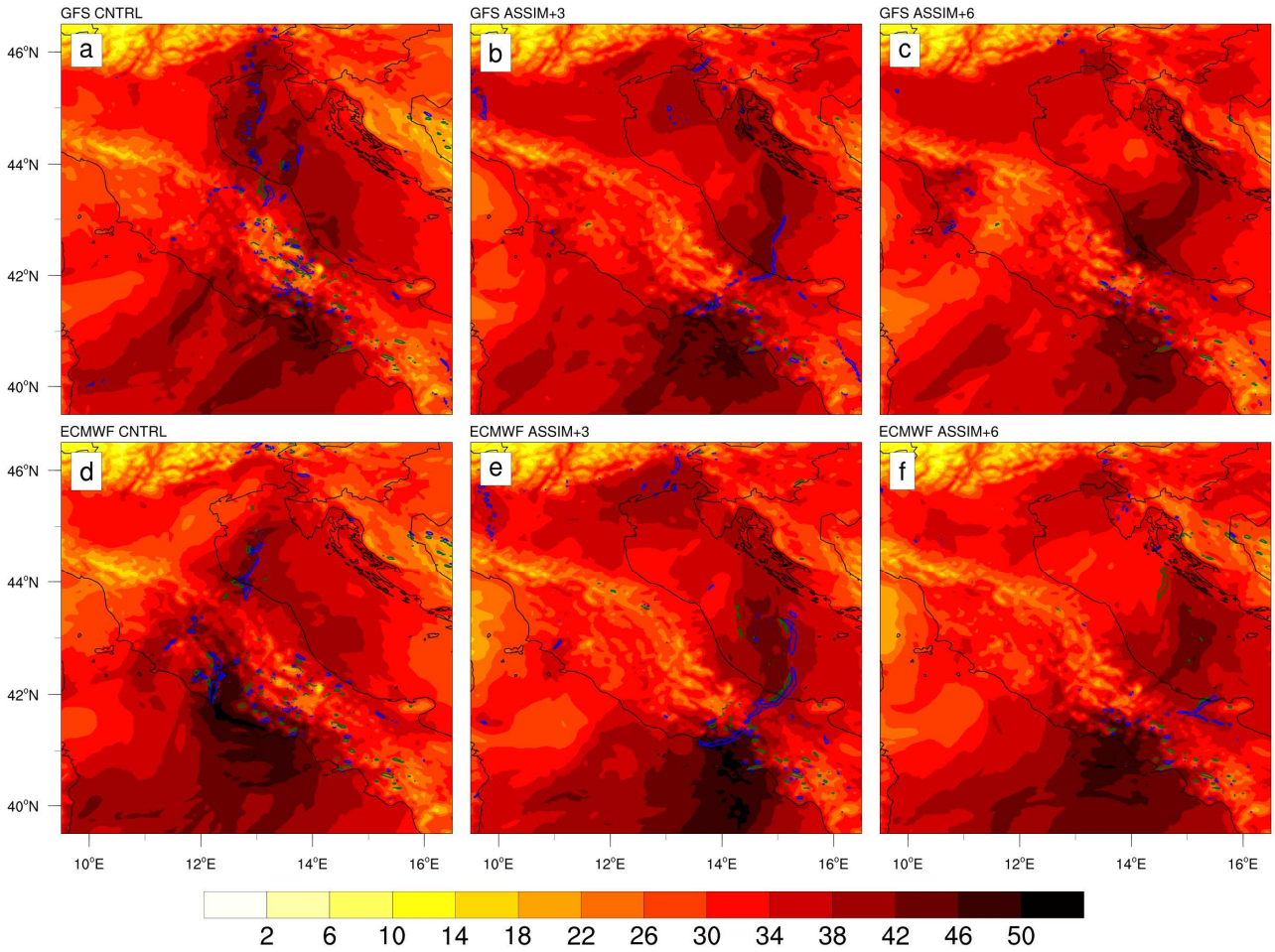


Figure 11. Precipitable water (mm, filled contours) and vertical velocity at 850 hPa (positive contours in blue, negative in green, at ±1.5 m/s) at 12 UTC 10 September 2017.

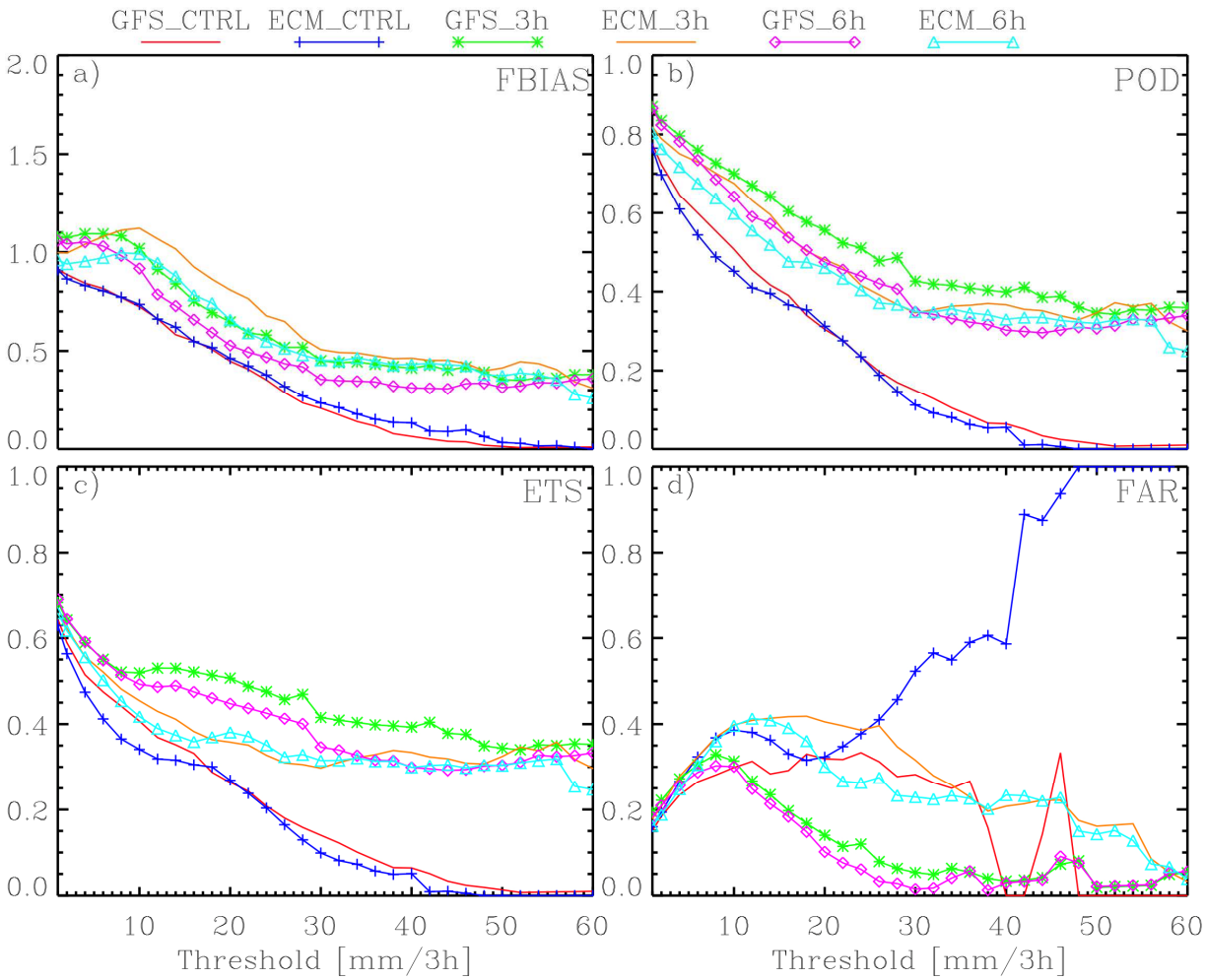
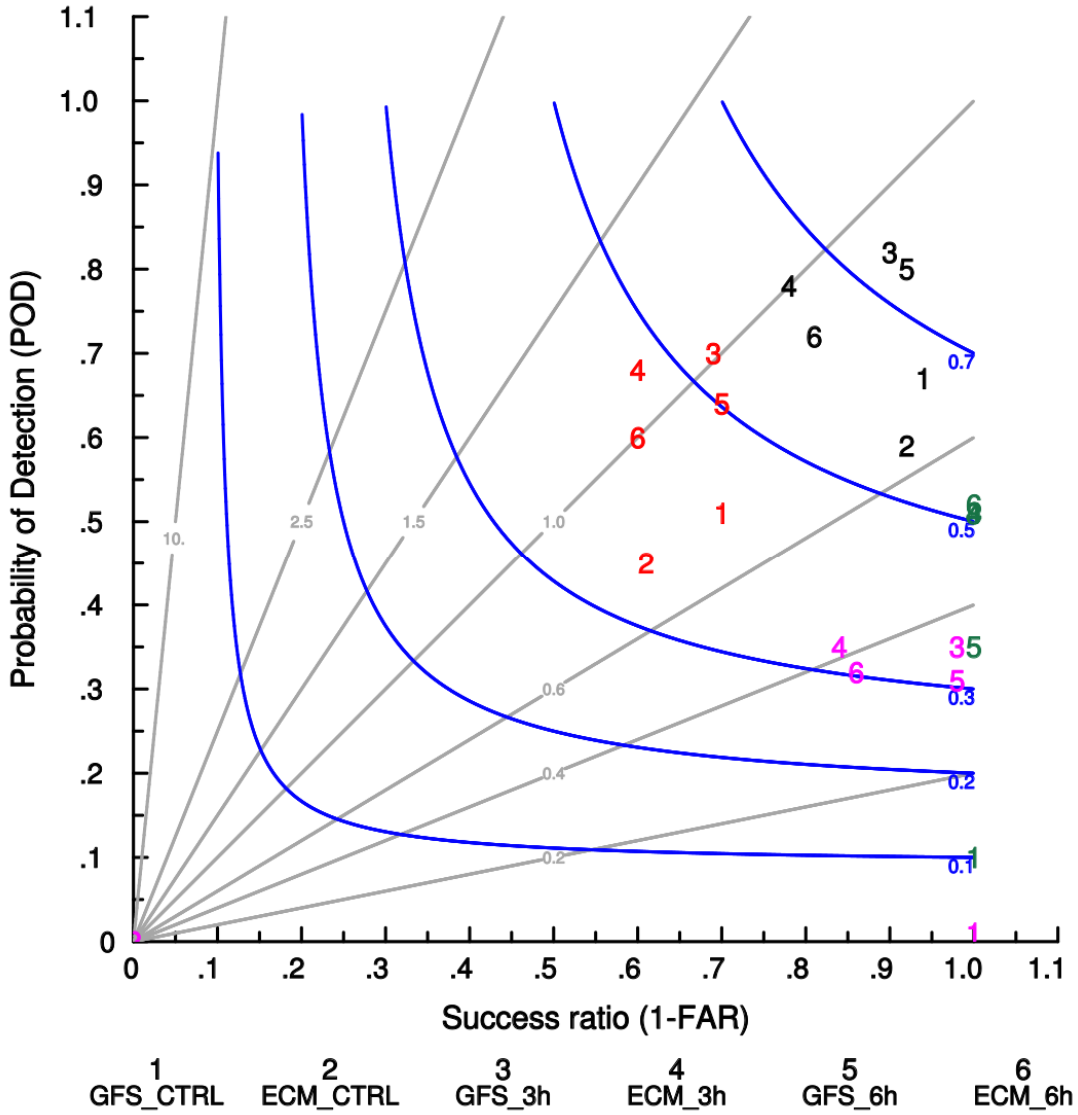
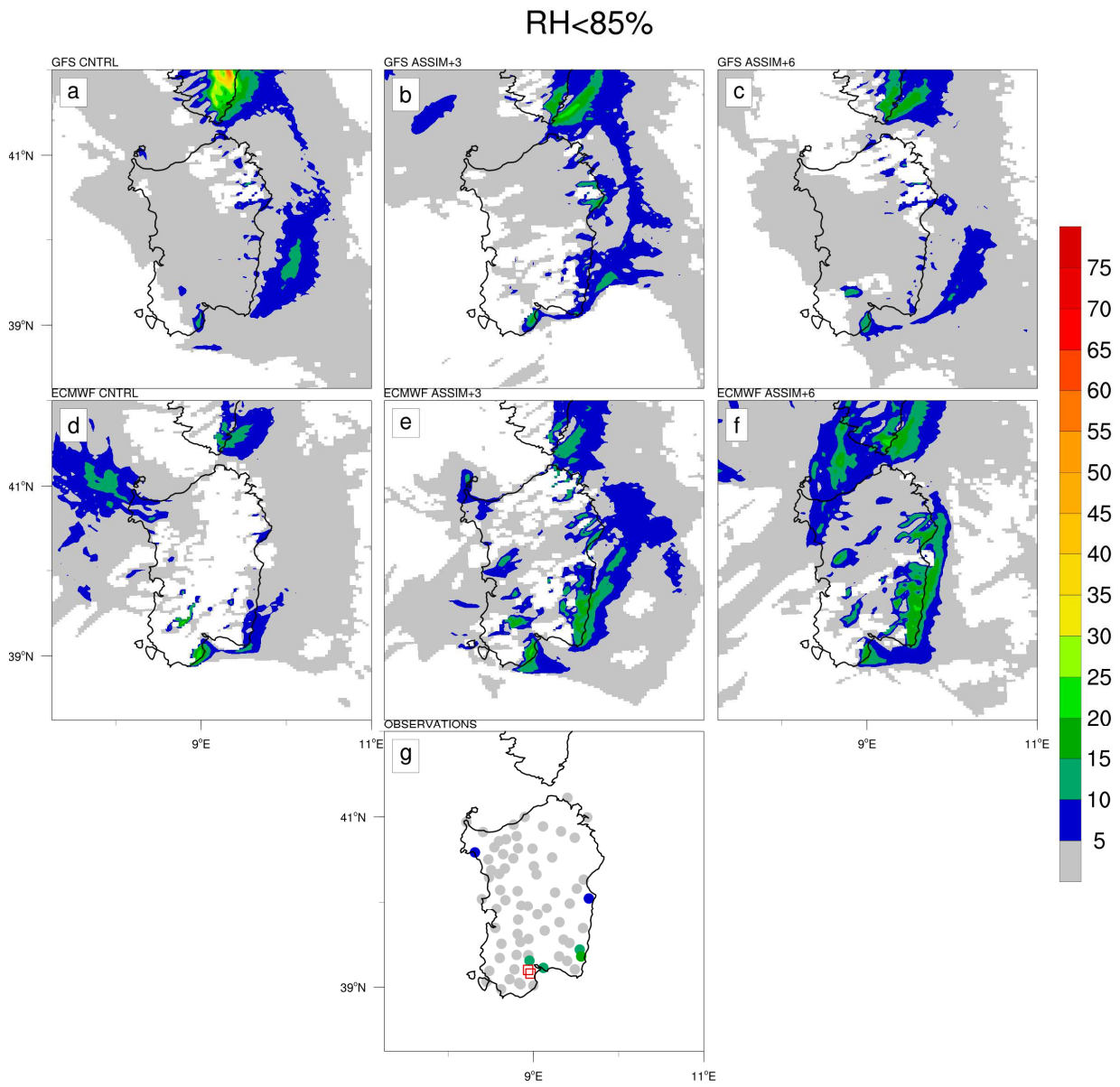


Figure 12. Scores for the 3 h precipitation thresholds (mm/3h) for all simulations; Frequency bias (a), probability of detection (b), equitable threat score (c) and false alarm ratio (d).



692 Figure 13. Performance diagram for all simulations (shown with numbers 1 through 6). In red (in
693 black) for the 8.5 (20) km nearest neighbourhood at the 10 mm/3h rainfall threshold, and in
694 magenta (in green) for the 8.5 (20) km nearest neighbourhood at the 50 mm/3h threshold.
695
696



697 Figure 14. Same as Figure 5 but for the experiment of LDA activated where relative humidity was
698 lower than 85 % instead of 95 %.

699
700 Figure 15. Same as Figure 7 but for the experiment of LDA activated where relative humidity was
701 lower than 85 % instead of 95 %.
702

References

Alexander, G. D., Weinman, J. A., Karyampoudi, V. M., Olson, W. S., and Lee, A. C. L.: The effect of assimilating rain rates derived from satellites and lightning on forecasts of the 1993 superstorm, Mon. Weather Rev., 127, 1433–1457, 1999.

Benjamin, S. G., Dévényi, D., Weygandt, S. S., et al.: An Hourly Assimilation–Forecast Cycle: The RUC. Mon. Weather Rev., 132(2), 495-518, 2004.

Betz, H. D., Schmidt, K., Laroche, P., Blanchet, P., Oettinger, W. P., Defer, E., Dziewit, Z., and Konarski, J.: LINET – An international lightning detection network in Europe, Atmos. Res., 91, 564–573, 2009.

Betz, H. D., Schmidt, K., Oettinger, P., Wirz, M.: Lightning detection with 3D-discrimination of intracloud and cloud-to-ground discharges. J. Geophys. Res. Lett. 31 L11108. doi:10.1029/2004GL019821, 2004.

Bryan , G. H. , J. C. Wyngaard , and J. M. Fritsch , 2003 : On adequate resolution for the simulation of deep moist convection. Mon. Wea. Rev, 131, 2394 –2416.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. Mon. Wea. Rev., 129, 569–585.

Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. J. Atmos. Sci., 46, 3077–3107.

Federico, S., Avolio, E., Petracca, M., Panegrossi, G., Sanò, P., Casella, D., and Dietrich, S.: Simulating lightning into the RAMS model: implementation and preliminary results, Nat. Hazards Earth Syst. Sci., 14, 2933-2950, doi:10.5194/nhess-14-2933-2014, 2014

Federico, S., Petracca, M., Panegrossi, G., and Dietrich, S.: Improvement of RAMS precipitation forecast at the short-range through lightning data assimilation, Nat. Hazards Earth Syst. Sci., 17, 61–76, https://doi.org/10.5194/nhess-17-61-2017, 2017a.

Federico, S., Petracca, M., Panegrossi, G., Transerici, C., and Dietrich, S.: Impact of the assimilation of lightning data on the precipitation forecast at different forecast ranges. Adv. Sci. Res., 14, 187–194, 2017b.

Federico, S., Torcasio, R. C., Avolio, E., Caumont, O., Montopoli, M., Baldini, L., Vulpiani, G., and Dietrich, S.: The impact of lightning and radar reflectivity factor data assimilation on the very short-term rainfall forecasts of RAMS@ISAC: application to two case studies in Italy, Nat. Hazards Earth Syst. Sci., 19, 1839–1864, https://doi.org/10.5194/nhess-19-1839-2019, 2019.

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

Fierro, A. O., Gao, J., Ziegler, C. L., Mansell, E. R., Macgorman, D. R., and Dembek, S. R.: Evaluation of a cloud scale lightning data assimilation technique and a 3DVAR method for the analysis and short-term forecast of the 29 June 2012 derecho event, Mon. Weather Rev., 142, 183–202, doi:10.1175/MWR-D-13-00142.1, 2014.

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.

Gerard, L., 2007. An integrated package for subgrid convection, clouds and precipitation compatible with meso-gamma scales. Q. J. R. Meteor. Soc. 133, 711–730.

Giannaros, T. M., Kotroni, V., and Lagouvardos, K.: WRFLTNGDA: A lightning data assimilation technique implemented in the WRF model for improving precipitation forecasts, Environ. Model. Softw., 76, 54–68, doi:10.1016/j.envsoft.2015.11.017, 2016.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. Mon. Wea. Rev., 134, 2318–2341.

Hu, J., A. O. Fierro, Y. Wang, J. Gao, and E. R. Mansell, 2020: Exploring the assimilation of GLM derived water vapor mass in a cycled 3DVAR framework for the short-term forecasts of high impact convective events. Mon. Wea. Rev., 148, 1005-1028.

Huang, X. Y., Xiao, Q., Barker, D. M., Zhang, X., Michalakes, J., Huang, W., Henderson, T., Bray, J., Chen, Y., Ma, Z., Dudhia, J., Guo, Y., Zhang, X., Won, D. J., Lin, H. C., and Kuo, Y. H.: Four-Dimensional Variational Data Assimilation for WRF: Formulation and Preliminary Results, Mon. Weather Rev., 137, 299–314, doi:10.1175/2008MWR2577.1, 2009.

Johnson, A., X. Wang, J. Carely, L. Wicker, and C. Karstens, 2015: A comparison of multiscale GSI-based EnKF and 3DVar data assimilation using radar and conventional observations for mid-latitude convective-scale precipitation forecasts. Mon. Wea. Rev., 143, 3087–3108, https://doi.org/10.1175/MWR-D-14-00345.1.

Kain, J. S., 2004: The Kain-Fritsch convective parameterization: An update. J. Appl. Meteor., 43, 170–181.

Kain, S.J., Weiss, J., Levit, J., Baldwin, M.E., Bright, D.R., 2006. Examination of convection-allowing configurations of the WRF model for the prediction of severe convective weather: the SPC/NSSL Spring Program 2004. Weather Forecast. 21, 167–181.

Kain, et al., 2008. Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. Weather Forecast. 23, 931–952.

Li, X., Pu, Z., 2009. Sensitivity of numerical simulations of the early rapid intensification of Hurricane Emily to cumulus parameterization schemes in different model horizontal resolutions. J. Meteor. Soc. Japan 87, 403–421.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmosphere: RRTM, a validated correlated-k model for the longwave. J. Geophys. Res., 102D, 16 663–16 682.

Monin, A.S. and A. M. Obukhov, 1954: Basic turbulence mixing laws in the atmospheric surface layer. Tr. Inst. Teor. Geofiz, Akad, Nauk, SSSR, 24, 163-187 (English translation available in V. N. Bespalyi, Ed., 2001: Turbulence and Atmospheric Dynamics, J. L. Lumley, 164-194).

Papadopoulos, A., Chronis, T.G., Anagnostou, E.N.. Improving convective precipitation forecasting through assimilation of regional lightning measurements in a mesoscale model. Mon. Weather Rev. 133, 1961-1977, 2005.

Park, S. K., and L. Xu, Eds., 2013: Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications. Vol. II, Springer-Verlag, 730 pp., https://doi.org/10.1007/978-3-642-35088-7.

Parodi, A., Tanelli, S., 2010. Influence of turbulence parame- terizations on high-resolution numerical modeling of tropical convection observed during the TC4 field campaign. J. Geophys. Res. 115.

Pessi, A.T. and S. Businger: Relationships among Lightning, Precipitation, and Hydrometeor Characteristics over the North Pacific Ocean. J. Appl. Meteor. Climatol., 48, 833–848, https://doi.org/10.1175/2008JAMC1817.1, 2009.

Qie, X., Zhu, R., Yuan, T., Wu, X., Li, W., and Liu, D.: Application of total-lightning data assimilation in a mesoscale convective system based on the WRF model, Atmos. Res., 145–146, 255–266, 2014.

Roebber, P. J.: Visualizing multiple measures of forecast quality, Weather Forecast., 24, 601–608, 2009.

Skamarock, W. C. et al. (2019): A Description of the Advanced Research WRF Model Version 4 (No. NCAR/TN-556+STR). doi:10.5065/1dfh-6p97.

Stensrud, F. J., and Coauthors, 2009: Convective-scale Warn-On-Forecast system: A vision for 2020. Bull. Amer. Meteor. Soc., 90, 1487–1499, https://doi.org/10.1175/2009BAMS2795.1.

801 Thompson, G., Field, P.R., Rasmussen, R.M., Hall, W.D., 2008. Explicit forecasts of winter
802 precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow
803 parameterization. Mon. Weather Rev. 136, 5095–5115.
804 Wyngaard, J. C., 2004: Toward Numerical Modeling in the “Terra Incognita”. J. Atmos. Sci., 61,
805 1816–1826.