Received: 23 January 2019 | Revised: 17 September 2019 | Accepted: 3 October 2019 | Published on: 27 November 2019

DOI: 10.1002/qj.3679

RESEARCH ARTICLE

Quarterly Journal of the Royal Meteorological Society RMetS

Reflectivity and velocity radar data assimilation for two flash flood events in central Italy: A comparison between 3D and 4D variational methods

V. Mazzarella[1,2] | I. Maiello[3] | R. Ferretti[1] | V. Capozzi[2] | E. Picciotti[1] | P. P. Alberoni[4] | F. S. Marzano[1,5] | G. Budillon[2]

[1]Center of Excellence Telesensing of Environment and Model Prediction of Severe events (CETEMPS), Department of Physical and Chemical Sciences, University of L'Aquila, L'Aquila, Italy

[2]Department of Science and Technology, University of Naples "Parthenope", Naples, Italy

[3]Abruzzo Region, Civil Protection, L'Aquila, Italy

[4]Regional Agency for Prevention, Environment and Energy of Emilia-Romagna, Italy Climate, Weather and hydrological Service, Bologna, Italy

[5]Department of Information Engineering, Sapienza University of Rome, Rome, Italy

Correspondence
Vincenzo Mazzarella, Center of Excellence Telesensing of Environment and Model Prediction of Severe events (CETEMPS), Department of Physical and Chemical Sciences, University of L'Aquila, Via Vetoio, 67100 L'Aquila, Italy.
Email: vincenzo.mazzarella@aquila.infn.it

Abstract

The aim of this study is to provide an evaluation of the impact of two largely used data assimilation techniques, namely three- and four-dimensional variational data assimilation systems (3D-Var and 4D-Var), on the forecasting of heavy precipitation events using the Weather Research and Forecasting (WRF) model. For this purpose, two flash flood events in central Italy are analysed. The first occurred on September 14, 2012 during an Intensive Observation Period of the Hydrological cycle in the Mediterranean experiment (HyMeX) campaign, while the other occurred on May 3, 2018. Radial velocity and reflectivity acquired by C-band weather radars at Mt. Midia (central Italy) and San Pietro Capofiume (northern Italy), as well as conventional observations (SYNOP and TEMP), are assimilated into the WRF model to simulate these damaging flash flood events. In order to evaluate the impact of the 3D-Var and 4D-Var assimilation systems on the estimation of short-term quantitative precipitation forecasts, several experiments are carried out using conventional observations with and without radar data. Rainfall evaluation is performed by means of point-by-point and filtering methodologies. The results point to a positive impact of the 4D-Var technique compared to results without assimilation and with 3D-Var experiments. More specifically, the 4D-Var system produces an increase of up to 22% in terms of the Fractions Skill Score compared to 3D-Var for the first flash flood event, while an increase of about 5% is achieved for the second event. The use of a warm start initialization results in a considerable reduction in the spin-up time and a significant improvement in the rainfall forecast, suggesting that the initial precipitation spin-up problem still occurs when using 4D-Var.

KEYWORDS

3D-Var, 4D-Var, flash floods, radar data assimilation, Weather Research and Forecasting (WRF) model

1 | INTRODUCTION

In recent years the Mediterranean region, and the Italian territory in particular, has been affected by several flash flood events (Hally *et al.*, 2015; Cassola *et al.*, 2016; Silvestro *et al.*, 2016; Maiello *et al.*, 2017). Most of these events have occurred during the autumn when the Mediterranean Basin is frequently crossed by Atlantic troughs, which advect warm and moist air mass from the sea to the western regions of Italy, increasing the water vapour content in the air and

destabilizing the atmosphere. In autumn the Mediterranean Sea is still relatively warm, providing additional heat and moisture to the lower atmosphere, which triggers the development of convective phenomena (Duffourg and Ducrocq, 2011, 2013). The complex orography of the Italian Peninsula, including the Alps and Apennines mountain ranges, facilitates the lifting and consequent condensation of air, which is a key factor in the production of heavy precipitation.

The aforementioned elements, in combination with urbanized watersheds with steep slopes that characterize most of the Italian regions, increase the potential risk of flash floods. Another important factor that contributes to these events is the slow evolution of convective systems that produce heavy precipitation affecting the same area for several hours or even days, resulting in fatalities and economic damages in the range of millions of euros. It is worth also recalling the related socio-economic impacts of these events on the Italian municipalities, who spend a large amount of financial resources every year on repairing the damages caused by heavy rainfall (Lastoria *et al.*, 2006; Salvati *et al.*, 2010). Accurate forecasting of the locations and timings of precipitating cells is therefore necessary to ensure the safeguarding of human life, and to adopt a series of measures aimed at preventing and reducing the possible damages during a heavy rainfall event. This is still a challenging problem in both numerical studies and operational forecasts (Mass *et al.*, 2002; Bryan and Rotunno, 2005; Yano *et al.*, 2018).

Improvements in the initial conditions may help to solve this problem (Rabier *et al.*, 1996; Ehrendorfer, 1997; Stensrud *et al.*, 2000; Simmons and Hollingsworth, 2002; Gallus and Bresch, 2006) and, in this context, assimilation of local data can be used. Variational data assimilation allows different observation types to be combined with the background field (first guess) in order to provide the best possible estimate of the atmospheric state at the initial time. In recent years, many efforts have been made in the development of variational data assimilation systems, including three- and four-dimensional variational data assimilation systems (3D-Var and 4D-Var), to include weather radar measurements. Weather radar systems can provide a 3D structure of a storm in terms of radial velocity (by using Doppler radars) and reflectivity factor measurements with wide spatial coverage (up to 100 km) and at very high spatial resolutions (hundreds of metres). Thus, the radar data volume (i.e., the resolution) can be much larger than for conventional observations, and it can include the possibility of assimilating data within the 3D atmospheric structure, which is useful for understanding and forecasting the convective motions and structures. Xiao *et al.* (2005) developed an observation operator for radial velocity within the 3D-Var system of the Pennsylvania State University/National Center for Atmospheric Research (PSU/NCAR) fifth-generation Mesoscale Model (MM5). The results showed the positive impact of radial velocity data on short-term rainfall prediction. A radar reflectivity data assimilation scheme for the same 3D-Var system was also developed by Xiao *et al.* (2007). The assimilation of radar reflectivity measurements acquired by the Jindo radar in South Korea was assessed for rainfall estimates of Typhoon *Rusa*, highlighting the positive impact of the assimilation procedure on short-term quantitative precipitation forecasting (SQPF).

Several studies based on the 3D-Var method have been accomplished by assimilating measurements from one or more weather radars. As an example, reflectivity and radial velocity data provided by 12 radars were assimilated into the Weather Research and Forecasting (WRF) model for the simulation of a squall line over the Great Plains of the United States (Xiao and Sun, 2007). The results showed an improvement in precipitation skill and a positive impact on the localization of convective cells. Lee *et al.* (2010) evaluated the impact of multiple radar data for the simulation of a mesoscale convective system over the Korean Peninsula. The use of radial velocity and reflectivity data, acquired from two radars, improved the heavy rainfall and wind field forecasts. Over Europe, both radial velocity and reflectivity, measured by a single Doppler radar, have been assimilated for the simulation of a heavy rainfall event in central Italy (Maiello *et al.*, 2014). The authors showed that the use of weather radar data, combined with the outer-loop strategy that allows the nonlinearity of observation operators to be factored in and increases the number of assimilated observations, can have a positive impact on rainfall prediction. Wattrelot *et al.* (2014) assimilated the radar reflectivity from the French radar network using the Application of Research to Operations at Mesoscale (AROME) model. They demonstrated that the 3D-Var system improved the analysis field and the performance in terms of SQPF. Stanešić and Brewster (2016) assimilated SYNOP data and radial velocity and reflectivity data from the Bilogora radar in Croatia using the 3D-Var method assimilated into the Advanced Regional Prediction System (ARPS) high-resolution model for the simulation of a storm in northwest Croatia. They showed that the assimilation of radar data and conventional observations improved the localization and estimation of precipitation of the thunderstorm. Finally, Maiello *et al.* (2017) evaluated the impact of assimilating radar reflectivity from three radars with the 3D-Var method over low- and high-resolution domains for the simulation of a flash flood event in central Italy. They showed the positive impact of assimilating multiple radar reflectivity fields when performed over both domains.

The 4D-Var method, developed by Huang *et al.* (2009), was used by Wang *et al.* (2013a, 2013bb,) to assimilate weather radar data into the WRF model. They improved the cost function with the addition of a Kessler warm rain microphysics scheme and its tangent linear (TL) and adjoint (AD)

models. The assimilation of radial velocity and reflectivity data in a short-duration assimilation window worked well, producing a dynamically consistent analysis and improving the estimation of precipitation as compared to the 3D-Var method. Choi *et al.* (2013) used the 4D-Var method for the simulation of a heavy rainfall event in South Korea. They assimilated the radial velocity from 13 radars, adopting a strategy with more outer loops. The results showed that 4D-Var improves the rainfall prediction and the quality of the analysis field, ensuring a better forecast.

Although both 3D-Var and 4D-Var demonstrate good performance in rainfall forecasting, the required computational resources differ significantly. The 4D-Var method is very computationally demanding because, unlike 3D-Var, it uses the numerical weather prediction (NWP) model as a dynamical constraint to ensure that the local structure is retained. On the plus side, 4D-Var allows for observations distributed within the assimilation window to be assimilated, improves the convective-scale initial condition and ensures a consistent analysis field. Nevertheless, the incremental approach proposed by Courtier *et al.* (1994) and used in WRF 4D-Var relies on the TL model, which replaces the full nonlinear model as a forward model. This approach reduces the computational resources in the cost function minimization and provides greater stability of the numerical problem. Chu *et al.* (2013) carried out a preliminary comparison between the WRF 3D-Var and 4D-Var data assimilation methods, assimilating conventional observations for the prediction of two cyclonic systems in the Antarctic region. The study demonstrated the positive impact of 4D-Var after the first 24 hr of simulation and showed a better physical balance of the analysis field. Sun and Wang (2013) also compared 4D-Var and 3D-Var methods by assimilating radar reflectivity and radial velocity from the Weather Surveillance Radar-1988 Doppler (WSR-88D) network. The results showed the positive impact of 4D-Var on the precipitation forecast during the passage of a squall line and demonstrated how the assimilation of radial velocity produces a greater impact than reflectivity on rainfall estimation as well as a reduction in the precipitation spin-up time.

The aim of this study is to (a) evaluate the impact of the 4D-Var system in comparison with the 3D-Var system for the WRF model, (b) evaluate the impact of multiple weather radar data with respect to the assimilation of conventional observations, and (c) quantify the potential contributions of the 3D-Var and 4D-Var methods in terms of SQPF. Two flash flood events, occurring in central Italy, are used for this purpose. The first occurred on September 14, 2012 during the Hydrological cycle in the Mediterranean experiment (HyMeX) campaign (Ducrocq *et al.*, 2014; Ferretti *et al.*, 2014; Davolio *et al.*, 2015), while the other occurred in the spring, on May 3, 2018. Conventional observations are assimilated with or without radar reflectivity and radial velocity acquired by two C-band Doppler radars and the performances of these two approaches are evaluated in terms of SQPF. In order to evaluate the impact on the precipitation forecast, two different methods are used: a point-by-point basis (conventional) technique and a filtering (neighborhood) technique. The key novelties of this work lie (a) in the assimilation of multiple radar data using the 4D-Var technique in an orographically complex region in the Mediterranean Basin, (b) in the comparison between the 3D-Var and 4D-Var assimilation methods implemented in the WRF model, assimilating the same number of observations, and (c) in the exploitation of different methods for the statistical evaluation. Moreover, this work is the first experiment conducted with WRF 4D-Var and radar data in Italy and it represents the first effort to compare the two variational methods for two severe weather events occurring in the Mediterranean basin.

This paper is structured as follows. Section 2 provides a brief description of the WRF 3D-Var and WRF 4D-Var data assimilation systems. An overview of the case studies is presented in section 3. The WRF model configuration and the experiments carried out are presented in section 4. The description of the two statistical methods used for precipitation assessment is provided in section 5. The results are discussed in section 6. A summary and conclusions are given in section 7.

2 | DATA ASSIMILATION IN THE WRF MODEL

Radar measurements and conventional observations are assimilated using the WRF model data assimilation system (WRFDA; Barker *et al.*, 2012). This system allows different observation types to be ingested and a domain-dependent forecast error covariance matrix (B matrix) to be estimated. Moreover, the system is designed to be used with the two variational approach methodologies, 3D-Var and 4D-Var, as described in the following subsections.

2.1 | General description of WRF 3D-Var and 4D-Var

The variational data assimilation methods 3D-Var and 4D-Var are widely employed in ocean and atmospheric sciences, mostly in NWP. The WRF 3D-Var data assimilation system (Barker *et al.*, 2004; Skamarock *et al.*, 2008) is based on the minimization of a penalty or cost function that reduces the misfit between the background forecast (or the first guess) and observations. With respect to the atmospheric state vector $\mathbf{x}$, the cost function J is defined as follows:

$$J(\mathbf{x}) = \frac{1}{2}\{(\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + [\mathbf{y}_0 - \mathbf{H}(x)]^T\mathbf{R}^{-1}[\mathbf{y}_0 - \mathbf{H}(\mathbf{x})]\}, \quad (1)$$

where $\mathbf{y}_0$ is the observation vector, $\mathbf{x}_b$ is the background vector field and $\mathbf{H}$ is the observation operator. $\mathbf{B}$ and $\mathbf{R}$ are the background and observation error covariance matrices, respectively. Moreover, in WRF 3D-Var, all observations falling within the assimilation window are assimilated at its centre.

In order to reduce the computational cost, WRF 3D-Var adopts the incremental approach (Courtier *et al.*, 1994) in which the cost function minimization is performed in a preconditioned variable space. The preconditioned control variables used for this purpose are unbalanced temperature, stream function, pseudo-relative humidity, unbalanced potential velocity and unbalanced surface pressure.

The WRF 4D-Var data assimilation system uses a numerical model as a constraint with the aim of simulating the trajectory (Huang *et al.*, 2009; Zhang *et al.*, 2013). This choice allows observations to be assimilated into the entire assimilation window. The state analysis is estimated by minimizing the following cost function:

$$\begin{aligned} J(\mathbf{x}_0) = {} & \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) \\ & + \frac{1}{2}\sum_{k=0}^{n} \{\boldsymbol{y}_k - \mathbf{H}_k[\mathbf{M}_k(\mathbf{x}_0)]\}^T \\ & \times \mathbf{R}^{-1}\{\boldsymbol{y}_k - \mathbf{H}_k[\mathbf{M}_\mathbf{k}(\mathbf{x}_0)]\}, \end{aligned} \tag{2}$$

where the term $\mathbf{x}_0$ is the initial state vector at time t_0 and $\mathbf{x}_0^b$ is the background state vector provided by a previous forecast. The assimilation window is divided in n sub-windows, so that $\boldsymbol{y}_k$ is the observation in each of them, and $\mathbf{H}_k$ and $\mathbf{M}_k$ are the observation and forward nonlinear model operators, respectively. The WRF 4D-Var system, as already mentioned, adopts the incremental approach proposed by Courtier *et al.* (1994), so that the forward nonlinear forecast model is replaced with the TL and AD models. This approach decreases the computational resources necessary for minimizing the cost function and enhances the efficiency in the estimation of the model trajectory, as well as in the computation of the cost function. The use of the TL and AD models handles the nonlinear processes within the assimilation window, improving the conditioning of the cost function. The nonlinear trajectory, which is ignored in the TL model, is updated at the end of the minimization processes before the next outer loop is run. Therefore, the two variational schemes have their advantages and disadvantages: WRF 4D-Var has the advantage of treating the observations at the correct time and the increments are calculated at the time of each observation. The simulation of the model trajectory, a specific feature of 4D-Var, ensures better consistency between the dynamical and microphysical fields. On the other hand, WRF 3D-Var has the advantage of being fast and computationally cheap.

2.2 | Reflectivity and radial velocity observation operators

The radar reflectivity and radial velocity are assimilated using the 3D-Var and 4D-Var methods by means of two different observation operators implemented in WRFDA. The radial velocity operator V_r takes into consideration the three components of the wind field (u,v,w), the distance D between the radar antenna and the target and, finally, the hydrometeor terminal fall speed V_t:

$$V_r = \frac{1}{D}[(x_d - x_r)u + (y_d - y_r)v + (z_d - z_r)(w - V_t)], \tag{3}$$

where

$$D = [(x_d - x_r)^2 + (y_d - y_r)^2 + (z_d - z_r)^2]^{1/2}, \tag{4}$$

with (x_d,y_d,z_d) and (x_r,y_r,z_r) being the coordinates of each grid point and the radar location, respectively. V_t is computed using the relationship between the rainwater mixing ratio and a correction factor (Sun and Crook, 1997).

The reflectivity Z (in dBZ) is assimilated using the following observation operator:

$$\mathrm{Z} = a_1 + a_2\, log_{10}(\rho q_r), \tag{5}$$

where ρ is the air density (kg·m^{-3}) and q_r is the rainwater mixing ratio (g·kg^{-1}). The coefficient a_1 (equal to 43.1) and a_2 (equal to 17.5) are determined according to the hydrometeor distributions defined by Marshall and Palmer (1948). Equation 5 was introduced by Sun and Crook (1997) and was implemented for the first time in WRF 3D-Var and 4D-Var by Xiao *et al.* (2007).

2.3 | Estimation of the background matrix B

The estimation of the B matrix is a critical point for both the 3D-Var and 4D-Var methods. The assimilation procedure combines observations with a short-range previous forecast affected by errors. Therefore, the evaluation of the error in the forecast field is a key action to ensure that the assimilated data produce dynamically consistent increments at each grid point.

The National Meteorological Center (NMC) method (Parrish and Derber, 1992) is widely used in the literature (e.g., Berre, 2000; Gustafsson *et al.*, 2001; Barker *et al.*, 2012) and is adopted in this work. This technique is based on an ensemble of forecasts over a long time period (weeks or months). More specifically, the method evaluates the differences between different sets of 24 and 12 hr forecasts being verified at the same time. For the case study occurring in September 2012, the B matrix has been computed to consider the whole period of the HyMeX campaign (from September

to November 2012). Conversely, for the second case a period of 2 weeks was taken into account.

Moreover, the tuning of the background error factors, named LEN_SCALING and VAR_SCALING, has a significant influence in 3D/4D-Var techniques. Both these factors are meaningful to the analysis (Jianfeng *et al.*, 2005) and can strengthen the impact of radar measurements. LEN_SCALING controls the spatial decorrelation scale and VAR_SCALING tunes the background error covariance for the following five variables: unbalanced velocity potential, unbalanced temperature, pseudo-relative humidity, unbalanced surface pressure and stream function. The covariance tuning has been performed through a sensitivity study by varying the LEN_SCALING parameter in 3D-Var and 4D-Var experiments with radar data. The water vapour mixing ratio increments (i.e., the difference between analysis and first guess) have been evaluated with three different LEN_SCALING factors: 1, 0.75 and 0.5, respectively. The use of a LEN_SCALING factor of 0.5 reduces the variable perturbation length scale by 50%, ensuring that the water vapour increment is comparable with the weather radar range; this is why this value is chosen for the simulations. For simplicity, the same LEN_SCALING factor is used for the following variables: unbalanced velocity potential, unbalanced temperature, pseudo-relative humidity, unbalanced surface pressure and stream function. In contrast, the default values are set for the VAR_SCALING factors.

3 | DESCRIPTION OF THE CASE STUDIES

Two flash flood events are analysed, the first of which occurred on September 14, 2012 (case study 1), in autumn during the HyMex campaign, while the second occurred on May 3, 2018 (case study 2) in the spring. Synoptic descriptions of the two case studies are presented in the following sections.

3.0.1 | Case study 1: September 14, 2012

The HyMeX program was aimed at improving our knowledge of hydrological cycles in the Mediterranean Basin and, in particular, the predictability and modelling of extreme events and interdecadal or multidecadal variabilities in the atmospheric circulation. The First Special Observation Period (SOP1) field campaign took place during the period between September and November 2012, and focused on heavy rainfall and flash flooding events occurring in the western Mediterranean area and their impacts in terms of casualties and damage caused. During the campaign, nine Intensive Observation Periods (IOPs) were identified over the Italian territory. In this study, we take into consideration the fourth Intensive Observation Period (IOP4), which occurred in central Italy on September 14, 2012.

An anticyclonic ridge characterized the synoptic scenario of IOP4 over western Europe, which separated two different cyclonic areas, one located over the Azores Islands and the other over the central Mediterranean Basin (Figure 1a). On September 14, 2012 at 0000 UTC, the Mediterranean low-pressure system, resulting from a strong meridional oscillation of the polar front, evolved into a cut-off. The latter was associated with two different relative minima on the surface pressure field, the first of which was located over the southern part of the Tyrrhenian Sea (off the coast of Campania), while the second occurred over the central Adriatic Sea (Figure 2a). Strong and organized convective activity was observed around the two relative surface pressure minima. In the central Adriatic region the convection was further enhanced by a low-level wind convergence line, generated by the interaction between cool northeastern winds (the bora) and moist and warm southeastern flow (the sirocco) formed along the coast of Abruzzo (Figure 2b). The convergence line increased the instability of the air column, triggering convective cells. Furthermore, in the inland areas of the Abruzzo and Marche regions, a relevant orographic lifting occurred along the eastern slopes of the Apennine mountains.

The local meteorological features, embedded with the synoptic and mesoscale structures, produced heavy and persistent precipitation during the morning of September 14, 2012 over the central Adriatic region. A flash flood event occurred along the coastline of the Abruzzo and Marche regions, causing damage to houses, railways, streets, hospitals and highways, with estimated economic damages of over one million euros: in several municipalities of the Pescara urban area a state of emergency was declared. Because of the complex orography of the area in question, peak rainfall was recorded at the Majella and Gran Sasso mountains and along the coastal areas of the Abruzzo and Marche regions (Figure 3). Rain-gauges located at the coastal cities of Chieti, Pescara Colli and Ortona measured 83, 80 and 64 $mm \cdot 6\,h^{-1}$, respectively. Relevant accumulations were also reported in the Gran Sasso and Majella mountain regions, namely at Roccacasale ($80\,mm \cdot 6\,h^{-1}$), Passo Lanciano ($78\,mm \cdot 6\,h^{-1}$) and Campo Imperatore ($70\,mm \cdot 6\,h^{-1}$).

3.0.2 | Case study 2: May 3, 2018

The second flash flood occurred in central Italy on May 3, 2018. At 0000 UTC, the large-scale flow showed a stretched trough (with its axis oriented from the northeast to the southwest), extending from northwestern Europe to the inland sectors of Algeria (Figure 1b). A strong meridional oscillation of the polar jet stream occurring within the previous 48 hr, together with an upper-level divergence and a strong temperature gradient at low levels, resulted in the deepening of a

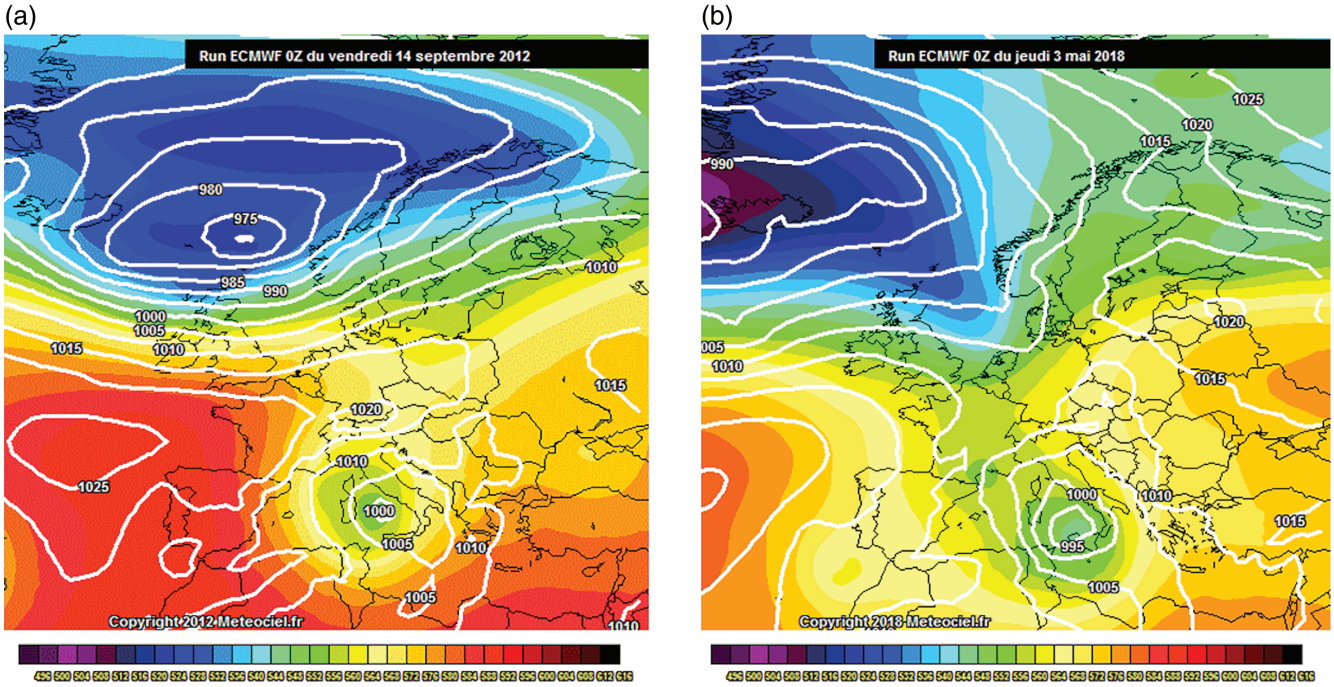

FIGURE 1 ECMWF analyses: geopotential height at 500 hPa (coloured areas) and sea level pressure fields (white solid lines) on (a) September 14, 2012 at 0000 UTC and (b) May 3, 2018 at 0000 UTC. Images courtesy of http://www.meteociel.fr

cyclone (992 hPa) on the western side of Sicily (Figure 2c). Over the next 12 hr, the trough evolved into a cut-off system.

Because of this synoptic scenario, the central Adriatic region was affected by a southeasterly flow at upper- and midtropospheric levels (300, 500 and 700 hPa) and by easterly winds at lower-pressure heights (950 and 1,000 hPa), which supplied a large amount of warm and moist air. Moreover, as highlighted by Figure 2d, strong low northeasterly winds forced upward by the Apennine mountain ridge in the central Adriatic region further increased the instability of the air column. As illustrated in Figure 3b, relevant peak rainfall was recorded in northern Abruzzo: rain-gauges in Mozzano and Force measured, respectively, 55 and 58 mm·6 h^{-1}, while along the coast rain-gauges measured 63 mm·6 h^{-1} at Pineto and 48 mm·6 h^{-1} at Pescara Colli. In the inland areas, the orographic precipitation contributed to an increase in the accumulated rainfall, especially on the eastern side of the Sibillini mountains (indicated by a red rectangle in Figure 3b).

4 | DATA AVAILABILITY AND NUMERICAL SIMULATIONS

The available data used for assimilation and numerical experiments are described in the following subsections. The dataset consists of reflectivity and radial velocity data, collected by two C-band Doppler radars, as well as data from conventional observations, such as SYNOP (surface synoptic observations) and TEMP (upper-level temperature, humidity and wind observations), provided by the Global Telecommunication System of the World Meteorological Organization. In addition, the WRF model configuration and the set-up of the numerical experiments are also presented.

4.1 | Description of the assimilated dataset

C-band Doppler radars at Mt. Midia (MM) and San Pietro Capofiume (SPC) (Figure 4a), which form part of the Italian radar network, acquire reflectivity and radial velocity data that are assimilated into the 3D-Var and 4D-Var methods. The MM radar, which is managed by the Abruzzo region, is located in central Italy at the top of Mt. Midia (42.057°N, 13.177°E, 1,710 m above sea level). The instrument operates with a frequency of 5.5 GHz and a resolution in the range of 500 m. Four volume scans are performed every 15 min at different elevation angles (0°, 1°, 2° and 3°), with a maximum range of 120 km. The SPC radar (44.654°N, 11.623°E, 31 m above sea level), managed by Arpae Emilia Romagna, is installed in San Pietro Capofiume in northern Italy. The instrument covers a large part of the Po Valley and the coastal areas of the northern Adriatic Sea, with a maximum range of 125 km. The radar scanning strategy provides six elevation angles: 0.5°, 1.4°, 2.3°, 3.2°, 4.2° and 5°, and operates with a frequency of 5.6 GHz and a resolution in the range of 250 m.

The measurements obtained by the two weather radars are processed with a quality control algorithm in order to remove the main sources of errors that may affect radar measurements (e.g., ground clutter, signal attenuation, beam blocking and radio interference). The MM and SPC radars have different

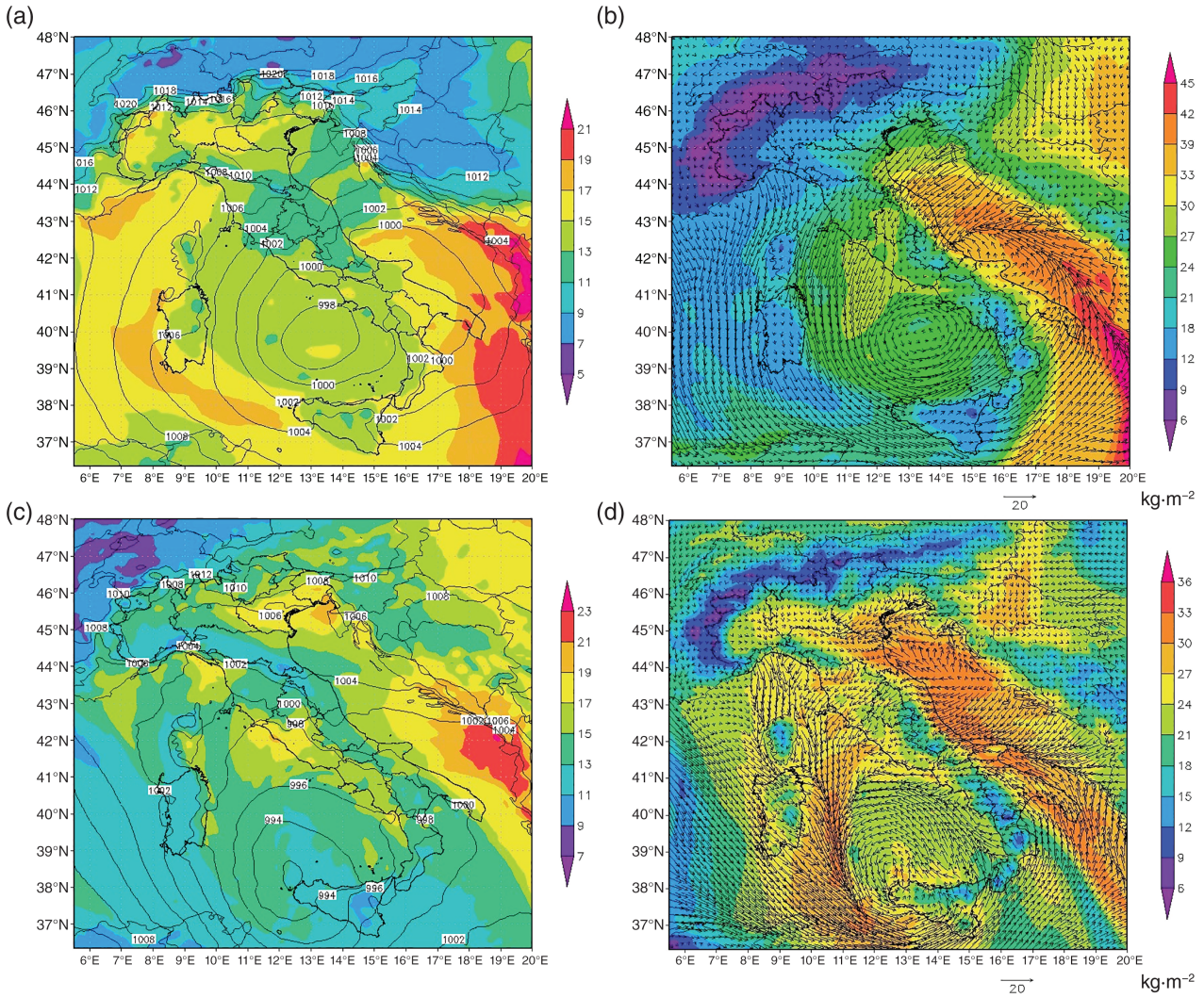

FIGURE 2 Analyses of 950 hPa temperature field in °C (coloured areas) and sea level pressure in hPa (black solid lines) from ECMWF at (a) 0600 UTC on September 14, 2012 and (c) 1200 UTC on May 3, 2018. Wind fields ($m \cdot s^{-1}$) at 10 m and total column and water vapour ($kg \cdot m^{-2}$) analyses from ECMWF at (b) 0600 UTC on September 14, 2012 and (d) 1200 UTC on May 3, 2018

spatial resolutions and their data are collected as polar coordinates. For these reasons, the reflectivity and radial velocity data are converted to Cartesian coordinates and horizontally thinned by interpolating them over the domain grid with a spatial resolution of 12 km (described in the next section). The thinning of the radial velocity and reflectivity data is needed to reduce the high density of the radar measurements and to ensure numerical stability of the assimilation techniques. SYNOP and TEMP data are also quality-controlled, through the observation preprocessing module implemented in WRF, and then assimilated.

4.2 | WRF model configuration

The Advanced Research WRF (ARW) model (version 3.7.1), which was developed at the National Center for Atmospheric Research (NCAR), is used for the numerical simulations. The model is characterized by a fully compressible non-hydrostatic set of equations, terrain-following hydrostatic-pressure vertical coordinates, Arakawa C-grid staggering and a third-order Runge–Kutta time-integration scheme. Additional technical information can be found in Skamarock *et al.* (2008). A two-way nesting configuration with two domains is used for this study: the first domain (D01) covers central western Europe and the Mediterranean Basin, with a spatial resolution of 12 km; the nested one (D02) covers central Italy with a horizontal resolution of 3 km (Figure 4b). The initial and boundary conditions for D01 are provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) with a horizontal resolution of 0.125°, while the D02 uses the initial and boundary conditions generated from the mother domain. For each domain, 37 unequally spaced vertical levels are used, from ground level up to 100 hPa.

Several physics options are eligible for the WRF model. In this work, we have used the same parametrization schemes

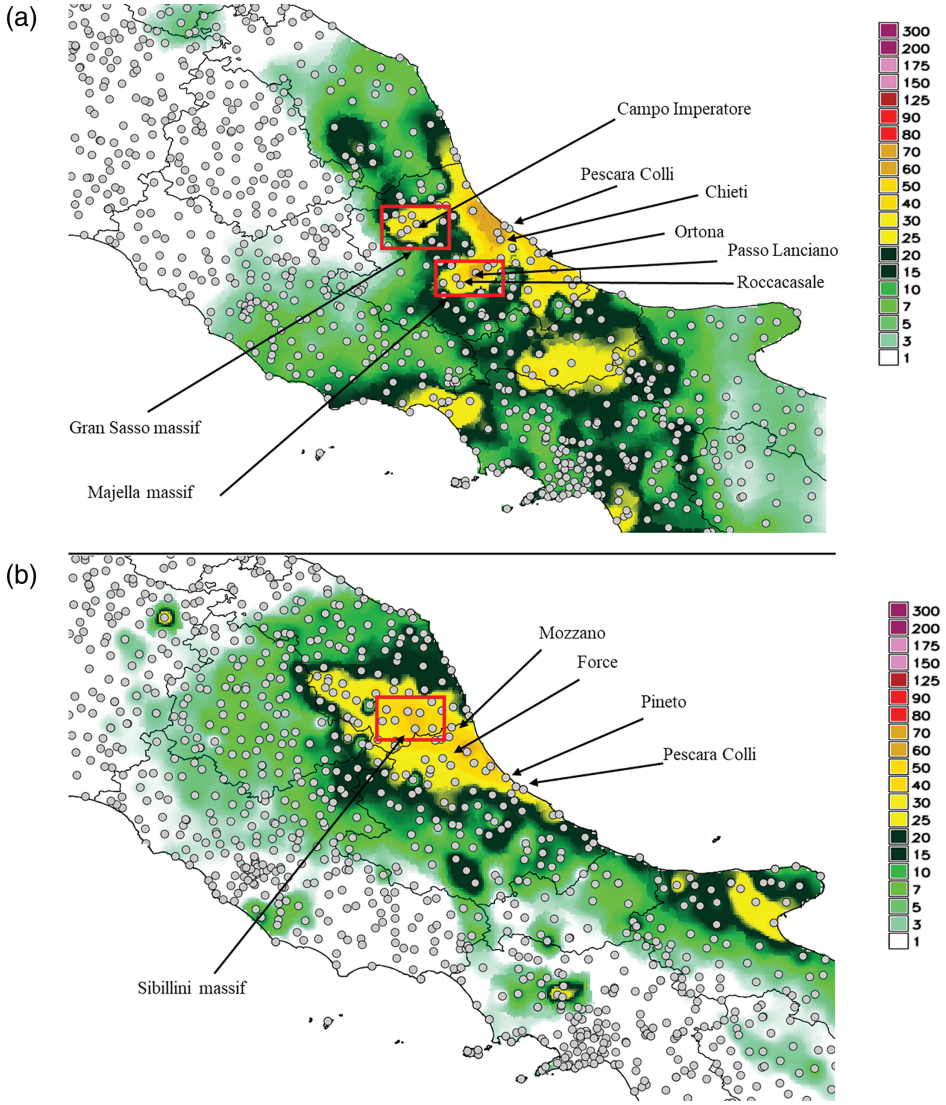

FIGURE 3 Precipitation amount (mm) (a) from 0000 to 0600 UTC on September 14, 2012 and (b) from 0600 to 1200 UTC on May 3, 2018. Grey points represent the rain-gauges of the Italian Civil Protection Department [Colour figure can be viewed at wileyonlinelibrary.com].

as set for the WRF operational chain of the Center of Excellence in Telesensing of Environment and Model Prediction of Severe Events (CETEMPS) (Ferretti *et al.*, 2014). For long wave radiation, the Rapid Radiative Transfer Model (RRTM) scheme (Mlawer *et al.*, 1997) is applied, while the Dudhia scheme (Dudhia, 1989) is selected for short wave solar radiation. The Grell 3D ensemble (G3D) scheme (Grell and Devenyi, 2002) is also adopted, although no cumulus parametrization is used on the higher resolution domain. The new Thompson scheme (Thompson *et al.*, 2008) with six different types of hydrometers is used for the microphysics. The Yonsei University (YSU) scheme (Hong *et al.*, 2006) is chosen to parametrize the vertical turbulent fluxes within the planetary boundary layer (PBL). Finally, the Noah land surface model (Noah LSM; Chen and Dudhia, 2001) is chosen for the parametrization of land surface processes. The parametrizations used in this study are summarized in Table 1.

4.3 | Design of experiments

For case study 1, the numerical experiments are initialized at 0000 UTC on September 14, 2012; case study 2 starts at 0600 UTC on May 3, 2018. All simulations are initialized using ECMWF analyses and forecasts. The 4D-Var method is applied to the low-resolution domain (D01, 12 km) only, because of a too-high computational cost for the D02 domain.

In order to factor in the nonlinearity of the observation operators and increase the number of assimilated observations, the outer-loop procedure (Rizvi *et al.*, 2008) is applied in this work. This method reduces the gap between first guess and observations; therefore, some measurements previously rejected by the WRFDA quality control (Andersson and Järvinen, 1999) can be assimilated in the next try. The outer-loop technique may also improve the quality of the analysis field (Guo *et al.*, 2006; Rosmond and Xu, 2006; Massart *et al.*,

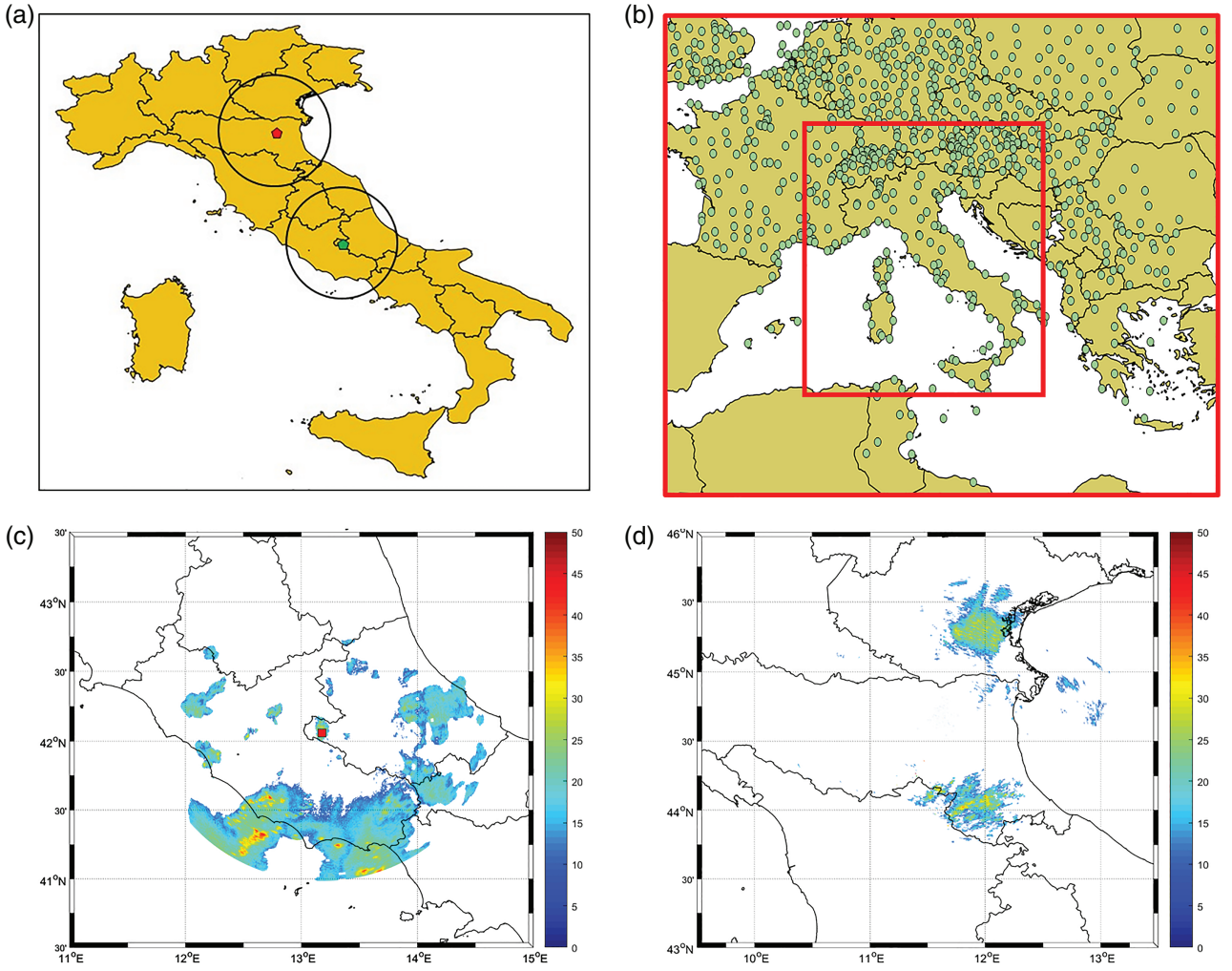

FIGURE 4 (a) Locations and spatial coverage of radars at Mt. Midia (MM; green point) and San Pietro Capofiume (SPC; red point). (b) The two domains used for the simulations (red rectangles) and the spatial distribution of assimilated SYNOP data (green points). (c, d) Two examples of assimilated data at 0000 UTC on September 14, 2012. Radar reflectivity is measured from (c) MM and (d) SPC radars at 1° and 1.4° elevation angles, respectively

TABLE 1 Parametrization adopted for subgrid-scale processes

Physics options	D01	D02
Microphysics	Thompson	Thompson
Cumulus parametrization	Grell 3D Ensemble Scheme	Resolved
Long wave radiation	RRTM	RRTM
Short wave radiation	Dudhia	Dudhia
PBL	YSU	YSU
Land surface physics	Noah LSM	Noah LSM

2010) by retaining local features. Two outer loops are applied for the assimilation experiments.

For each case study, seven numerical experiments are performed to assess the impact of 3D/4D-Var methods on the rainfall forecast. The same assimilation window of 1 hr is used for both variational techniques, assimilating radar data (Figure 4c,d) and conventional observations (Figure 4b) from 0000 to 0100 UTC and from 0600 to 0700 UTC in case studies 1 and 2, respectively.

For the 3D-Var experiments, conventional observations, radar reflectivity and radial velocity are assimilated at 0000 and 0100 UTC for case study 1. The same observations are

TABLE 2 Description of the seven simulations performed

Experiment	Assimilation method	Radar data	Conventional observations	Start mode
CTL	None	No	No	Cold
3DVAR-CON	3D-Var	No	Yes	Cold
3DVAR-CONMMSPC	3D-Var	Yes	Yes	Cold
4DVAR-CON	4D-Var	No	Yes	Cold
4DVAR-CONMMSPC	4D-Var	Yes	Yes	Cold
3DVAR-CONMMSPC_warm	3D-Var	Yes	Yes	Warm
4DVAR-CONMMSPC_warm	4D-Var	Yes	Yes	Warm

assimilated at 0600 and 0700 UTC for case study 2. To evaluate the impact of the radar data, two experiments are carried out:

1. 3DVAR-CON, assimilating SYNOP and TEMP data only;
2. 3DVAR-CONMMSPC, also including the radial velocity and reflectivity volumes provided by the MM and SPC radars.

Similarly to the 3D-Var experiments, the following numerical simulations are performed using 4D-Var:

1. 4DVAR-CON, assimilating SYNOP data at 0000 and 0100 UTC (0600 and 0700 UTC for case study 2) and TEMP data at 0000 UTC (not available for case study 2);
2. 4DVAR-CONMMSPC, accounting for the same conventional observations, but also including the radial velocity and reflectivity volumes acquired at 0000 and 0100 UTC (0600 and 0700 UTC for case study 2).

With the aim of using an equal number of observations for the numerical experiments, the same assimilation window is used for the 3D-Var and 4D-Var methods. A total of 1,353 SYNOP, 38 TEMP and 2,748 radar observations are assimilated for case study 1, while 1,674 SYNOP and 2,748 radar observations are assimilated for case study 2. Furthermore, a control experiment (CTL) without assimilation is carried out for both events.

Finally, in order to evaluate the performance of 4D-Var and 3D-Var with a warm start initialization, two additional experiments (4DVAR-CONMMSPC_warm and 3DVAR-CONMMSPC_warm), including radar measurements (reflectivity and radial velocity) and conventional observations, are performed for both case studies. For this simulation, the WRF forecast fields, initialized 6 hr previously, are used as the first guess for the two assimilation techniques.

In order to compare the different variational methods in terms of SQPF, all simulations are run for 6 hr. A summary of the performed experiments is reported in Table 2.

5 | VERIFICATION METHODS

The impact of 3D-Var and 4D-Var is evaluated by comparing the observed precipitation, collected by the rain-gauge network of the Italian Civil Protection Department (DPC), with the rainfall forecasted by the WRF simulations. A quality control is applied to the rain-gauge data before using them for the statistical analysis. Possible errors due to instrument failures, connection problems and false measures are removed. The verification is performed over a sub-area of the D02 domain, referred to hereafter as the Study Area (SA). The SA extends from the coastal region of the central Adriatic Sea to the coastal region of the central Tyrrhenian Sea (Figure 5), mainly covering central Italy.

The analysis of the 6 hr precipitation field, obtained through a spatial interpolation of rain-gauge data, shows rainfall only in the SA of the D02 domain (Figure 5). The statistical analysis is therefore performed over the SA. In addition, two different sub-regions (Figure 5) are identified within the SA: Lazio-Abruzzo (LA) and Marche-Abruzzo (MA). The Abruzzo region, which is characterized by the highest amount of precipitation, is included in both sub-regions and is associated with both Lazio and Marche because the latter regions are either totally or partially covered by the Mt. Midia radar and affected by precipitation in this case study. The use of the two sub-regions, on the one hand, reduces the number of no-rain data that can affect the statistical analysis and, on the other hand, allows the impact of data assimilation on the leeward and upwind sides of the Gran Sasso and Majella massifs (central Apennines) to be evaluated.

To evaluate the short-term (first 6 hr of simulation) response to the 3D-Var and 4D-Var assimilation techniques in the SA, MA and LA sub-regions, SQPF is employed. For case study 1, hourly precipitation measurements from 0000 to 0600 UTC on September 14, 2012 are used, while for case study 2, those from 0600 to 1200 UTC on May 3, 2018 are used. In this respect, the conventional point-by-point approach and the filtering method are applied.

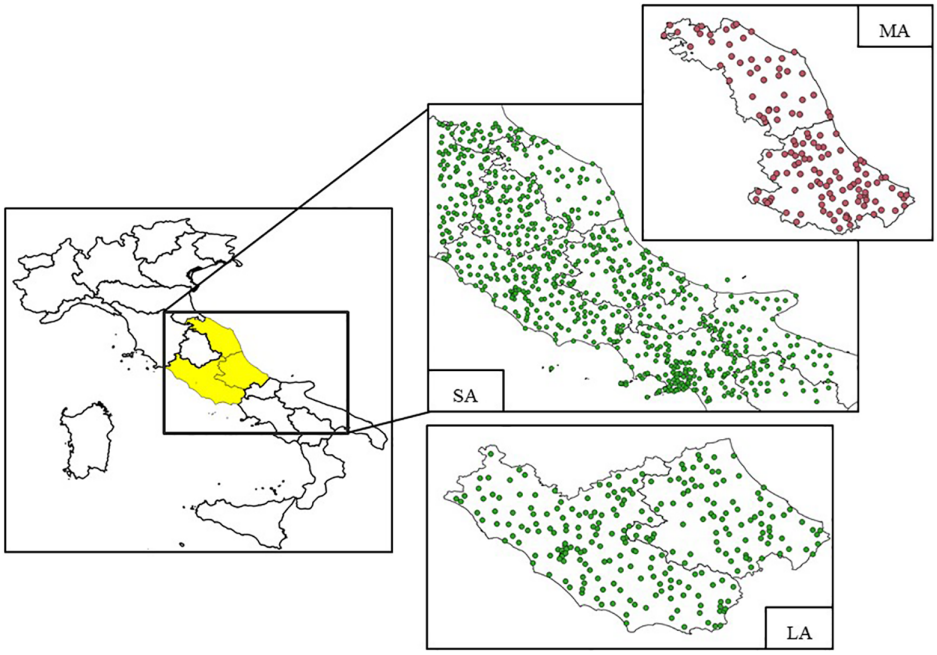

FIGURE 5 Description of the three sub-regions: Study Area (SA), Marche-Abruzzo (MA) and Lazio-Abruzzo (LA). The points represent the locations of rain-gauges forming part of the Italian Civil Protection network [Colour figure can be viewed at wileyonlinelibrary.com].

5.1 | Point-by-point approach

The point-by-point (conventional) approach provides a quantitative assessment of the precipitation estimated by the WRF simulations. Such an approach consists in comparing the predicted and observed rainfall at the same location through point-by-point matching. A contingency table is introduced to account for possible combinations between the rainfall amounts and one or more threshold values. The elements of the contingency table are described in the following:

- Hits: the observed and forecasted values are greater than the threshold;
- False alarms: only the forecasted value is greater than the threshold;
- Misses: only the observed value is greater than the threshold;
- Correct negatives: the observed and forecasted values are lower than the threshold.

Different statistical scores may be calculated using this table (Wilks, 2006, p. 627). However, the Equitable Threat Score (ETS; Schaefer, 1990), which is widely used in the literature (Sousounis *et al.*, 2004; Jankov *et al.*, 2007; Hong *et al.*, 2010; Pennelly *et al.*, 2014), is adopted for this study. The ETS represents the ratio between the hits and the sum of all hits, false alarms and misses, corrected by a random term that reduces the probability of random hits in wet or dry climates. The score ranges from −1/3 to 1, with 1 being the best score.

5.2 | Filtering method

The point-by-point method shows significant limitations in the assessment of precipitation fields at high spatial and temporal resolutions (Roberts, 2003). Therefore, another methodology – the filtering method – is used to assess the WRF precipitation forecast. The rainfall measurements observed by the rain-gauge network and those estimated by the WRF simulations are spatially interpolated on a regular grid with a horizontal resolution of 3 km (the same as the D02 domain) and matched using a neighbourhood approach, comparing the forecasts and observations considering 3×3 grid cells around each grid point. The Fractions Skill Score (FSS; Roberts and Lean, 2008), which is frequently used in the literature for the assessment of precipitation, is computed here (Baldauf *et al.*, 2011; Rennie *et al.*, 2011; Romine *et al.*, 2013; Wang *et al.*, 2013a, 2013b). The FSS index ranges from 0 to 1, with 1 being the best score.

6 | RESULTS

Evaluating the impact of the data assimilation methods involves the two statistical analyses previously described in sections 5.1 and 5.2. The FSS and ETS are calculated using a threshold value of 1 $mm \cdot h^{-1}$ to assess the model's ability to reproduce the onset of precipitation. A comparison between the FSS and the ETS in the SA region is performed with the aim of assessing the response of the two statistical indexes in estimating precipitation with the two assimilation methods. Moreover, to evaluate the impact of data assimilation

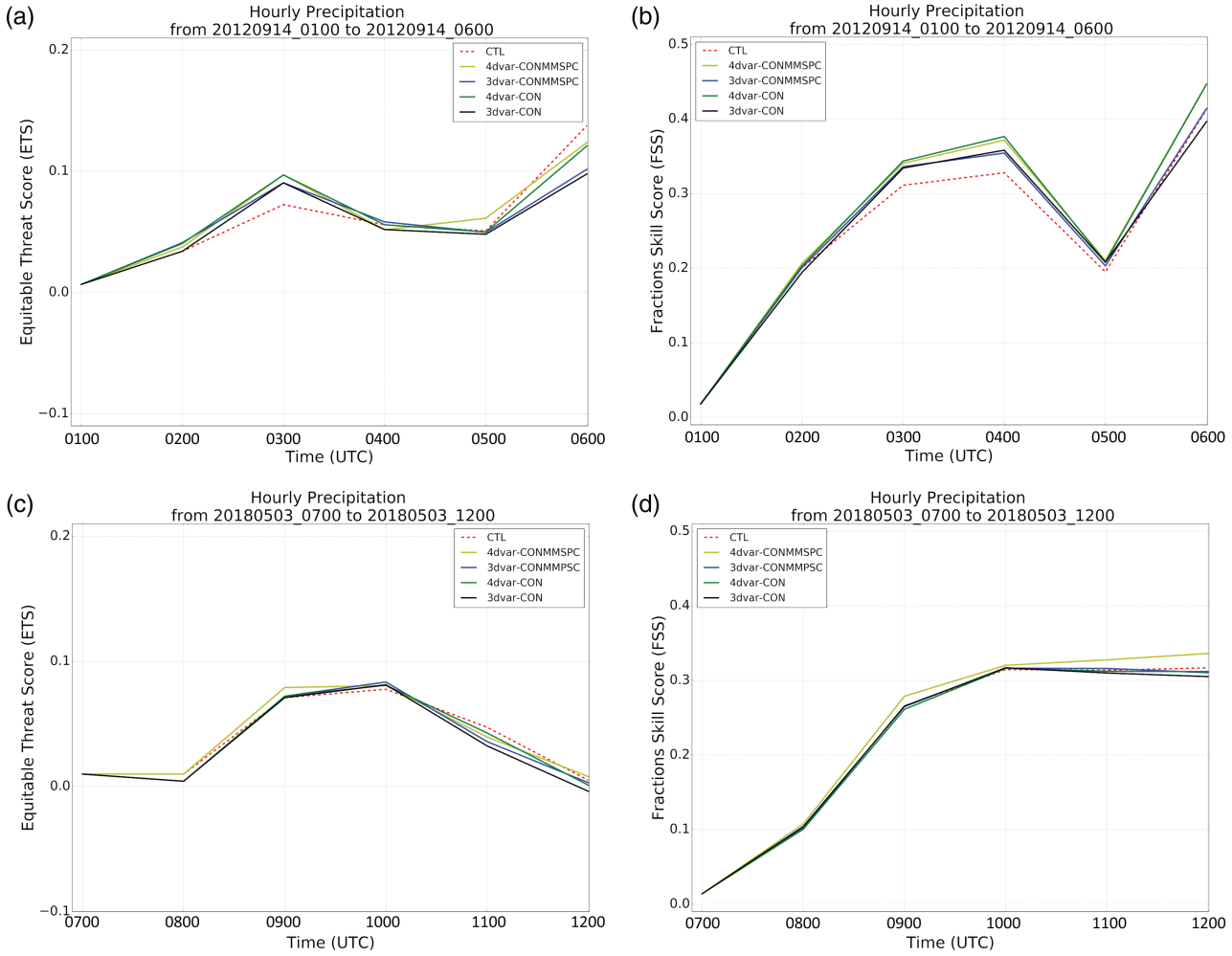

FIGURE 6 Upper panels: evolution of (a) ETS and (b) FSS for the threshold >1 mm·h^{-1} in the SA region for the case study occurring on September 14, 2012. Lower panels: evolution of (c) ETS and (d) FSS for the same threshold in the SA region for the case study occurring on May 3, 2018. Dashed red lines represent the CTL, yellow lines 4DVAR-CONMMSPC, green lines 4DVAR-CON, black lines 3DVARCON and blue lines 3DVAR-CONMMSPC

with 3D/4D-Var, the FSS are also calculated, using the same threshold value, for the LA and MA regions, where peak rainfall is observed. Finally, a sensitivity test to assess the behaviour of the assimilation in a warm start mode is carried out. The results are described in the following three subsections.

6.1 | Statistical analysis: ETS versus FSS

6.1.1 | Case study 1

The ETS is calculated for the 1 mm·h^{-1} threshold in the SA region (Figure 6a). The 3D-Var and 4D-Var simulations show higher values of the ETS than the CTL (dashed line) in the interval from 0000 to 0400 UTC. However, the impact of assimilation is reduced in the last 2 hr of simulations. On the other hand, the FSS calculated in the same region using the 1 mm·h^{-1} threshold (Figure 6b) from 0000 to 0500 UTC displays a greater improvement for the 3D/4D-Var experiments than for the CTL, and a positive impact of the 4D-Var simulations over the whole simulation period. This is in contrast with the ETS after 3 hr of simulation. This behaviour highlights the different responses of the statistical methods in the assessment of the precipitation given the different approaches. Indeed, the point-by-point method is performed considering a non-homogeneous spatial distribution and a limited number of rain-gauges, whereas the spatial verification uses a regular grid, ensuring the estimation of precipitation over all regions, as well as a neighbourhood approach that smooths the precipitation field.

6.1.2 | Case study 2

The ETS, computed in the SA region for the 1 mm·h^{-1} threshold, is provided in Figure 6c. The 4DVAR-CONMMSPC simulation (yellow line) shows higher values of the ETS than the CTL from 0700 to 1000 UTC. This behaviour highlights a slight improvement of 4D-Var with radar data and

conventional observations in terms of SQPF, even though the impact is limited to the first 4 hr of simulation. Later, the impact of assimilation is reduced and the CTL performs better. As for case study 1, the FSS computed for the same threshold (Figure 6d) confirms the good performance for 4DVAR-CONMMSPC. However, similarly to what is obtained for the case study 1, the use of the FSS extends the positive impact for the whole simulation period.

These results suggest high sensitivity of the statistical methods in the assessment of precipitation; hence, greater attention needs to be paid to the choice of methods and in the conclusions that can be drawn from them (Baldwin and Kain, 2006).

6.2 | Statistical analysis: 4D-Var versus 3D-Var

6.2.1 | Case study 1

For the comparison of 4D-Var and 3D-Var, two restricted areas with significant rainfall are used: the LA and MA regions. To this purpose, the FSS for the 1 mm·h^{-1} threshold is calculated. Figure 7a shows the good performance of the 4D-Var simulations in the LA region. This result suggests that there is a clear benefit in using 4D-Var compared to 3D-Var, although the 3D-Var experiments do produce a slight improvement in the FSS compared to the CTL in the interval from 0300 to 0600 UTC. In addition, the FSS for 3DVAR-CONMMSPC (blue line) is greater than that for 3DVAR-CON (red line), which is the simulation with only conventional observations. This behaviour highlights the benefits of also including radar data in the 3D-Var experiment.

The positive impact obtained using the 3D-Var and 4D-Var methods decreases in the final hours, confirming that data assimilation has a significant impact for only a limited time interval. For the MA region, the FSS shows good performance of the 4D-Var experiments over the whole simulation interval (Figure 7b). On the other hand, the FSS for the 3D-Var simulations shows a worsening when compared to the CTL. Similarly to the results for the LA sub-region, 3DVAR-CONMMSPC performs a bit better than 3DVAR-CON, while 4DVAR-CONMMSPC shows a slight improvement only until 0300 UTC.

6.2.2 | Case study 2

Similarly to case study 1, FSS is analysed for the LA region, with a threshold of 1 mm·h^{-1}, for case study 2 (Figure 7c). 4DVAR-CONMMSPC clearly shows good performance, with the highest FSS values for the whole period. On the other hand, the FSS for the 3D-Var and 4DVAR-CON experiments does not differ greatly from the CTL. These results suggest that the use of the 4D-Var method in combination with the radar data benefits the SQPF for this case. Finally, the FSS for the same threshold value is presented in Figure 7d for the MA region. According to the result in the SA, the 4DVAR-CONMMSPC experiment confirms the positive impact of assimilating radar data. Higher values of the FSS as compared to the CTL are found for 4DVAR-CONMMSPC during the simulation period, suggesting that the radar data have a positive impact on the estimation of precipitation for this case study.

In conclusion, the use of the FSS allows us to identify the key role of 4D-Var in the estimation of precipitation by showing a clear signal in the results, which is also found for the point-by-point score. Very low values of the FSS are found in the first few hours of simulation, suggesting that the model is unable to produce realistic precipitation because of the spin-up, which also occurs in simulations using 3D-Var or 4D-Var assimilation techniques.

6.2.3 | A brief comparison of 4D-Var with respect to the ECMWF initial conditions for D01

In order to evaluate the impact of 4D-Var on the initial condition (IC) with respect to the ECMWF IC, we compare the relative humidity over D01, predicted by CTL (no assimilation, i.e. ECMWF IC) and 4DVAR-CONMMSPC, with the relative humidity measured by automatic weather stations provided by the DPC. A point-by-point verification approach is adopted for this purpose. More specifically, the comparison is performed using the root mean square error (RMSE) only in the MA region, where a significant amount of precipitation is observed. The 4DVAR-CONMMSPC simulation shows a reduction in the RMSE of about 1% compared to the CTL (not shown) for both flash flood events. This result points out the positive impact of 4D-Var with respect to the initial and lateral boundary conditions provided by ECMWF for D01.

6.3 | Sensitivity to the warm start initialization

The previous statistical analysis suggests that the 4D-Var technique, in combination with the radar data and conventional observations, can improve precipitation estimation in the assimilation window. Two additional simulations for each flash flood event, named 4DVAR-CONMMSPC_warm and 3DVAR-CONMMSPC_warm, are carried out to evaluate the impacts of 3D-Var and 4D-Var methods in the warm start mode. To this purpose, the initial conditions are provided by a WRF simulation initialized 6 hr earlier. The same conventional observations and radar data as used in the previous simulations are assimilated in 4DVAR-CONMMSPC_warm and 3DVAR-CONMMSPC_warm. The FSS is used to compare the new simulations with the CTL, 4DVAR-CONMMSPC and 3DVAR-CONMMSPC (both in cold start mode) for the 1 mm·h^{-1} threshold value and the MA and LA sub-regions previously defined.

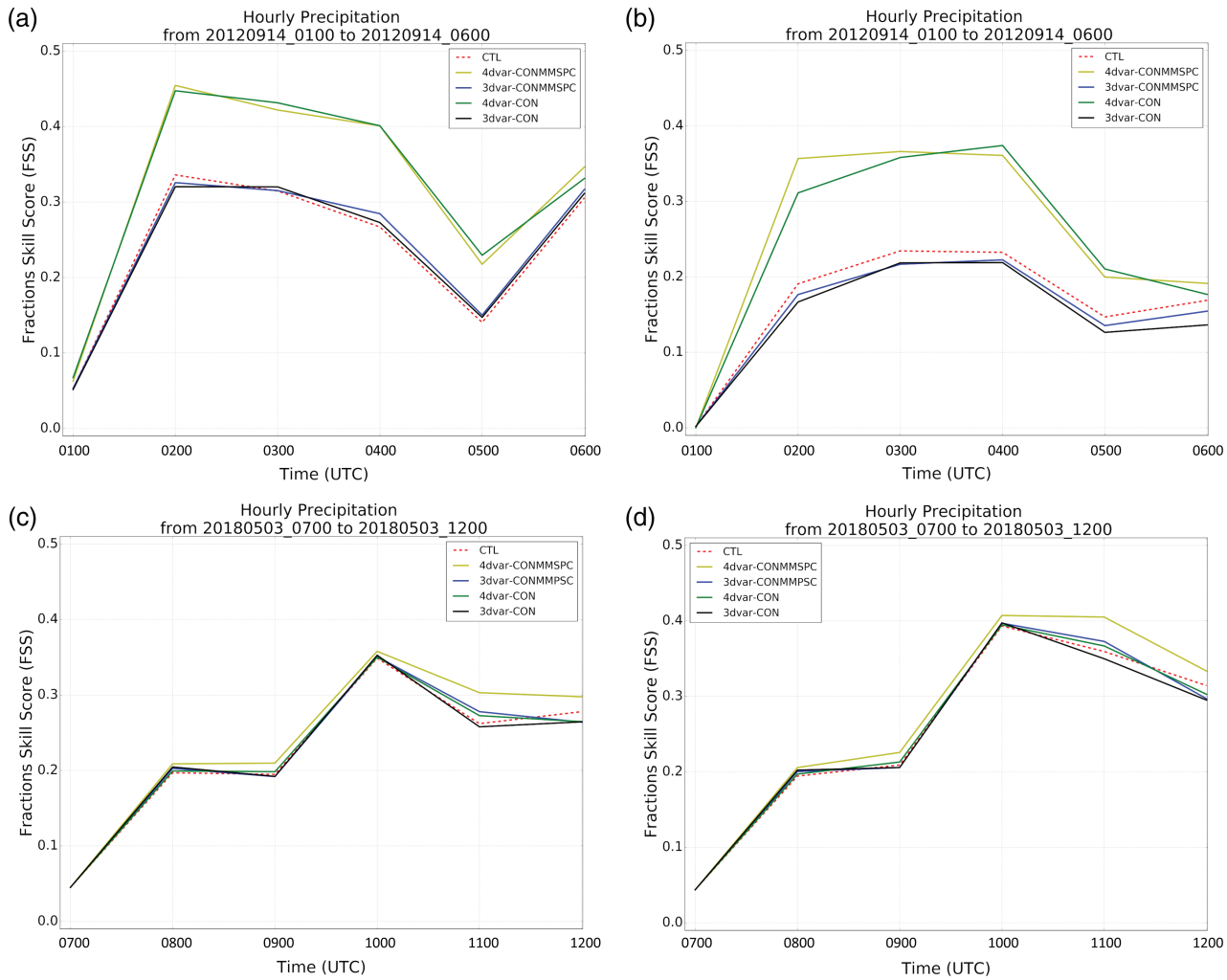

FIGURE 7 Upper panels: evolution of the FSS for the threshold >1 mm·h^{-1} in (a) the LA region and (b) the MA region for the case study occurring on September 14, 2012. Lower panels: evolution of the FSS for the same threshold in (c) the LA region and (d) the MA region for the case study occurring on May 3, 2018. Dashed red lines represent the CTL, yellow lines 4DVAR-CONMMSPC, green lines 4DVAR-CON, black lines 3DVAR-CON and blue lines 3DVAR-CONMMSPC

6.3.1 | Case study 1

The FSS, with a threshold of 1 mm·h^{-1} (Figure 8a), shows the highest value for 4DVAR-CONMMSPC_warm over the whole simulated period for the LA region. Moreover, a reduction in the spin-up time is found in the first few hours of simulation. The FSS for 3DVAR-CONMMSPC_warm is greater than that for 3DVAR-CONMMSPC. This result, in combination with the positive impact of 4DVAR-CONMMSPC_warm, suggests that a spin-up period is required for both assimilation techniques. However, the simulations with a warm start initialization converge to the CTL in the final hours of simulation, highlighting that the influence of 4D-Var and 3D-Var is restricted to a short time range. The FSS for the same threshold confirms the reduction in the spin-up time for the 3D-Var and 4D-Var experiments and the positive impact of 4DVAR-CONMMSPC_warm in the estimation of the rainfall field in the MA region (Figure 8b). The differences between the warm start and cold start experiments indicate that a spin-up time is still required for the MA region.

6.3.2 | Case study 2

The FSS is also used for the second flash flood event.The FSS is provided for the threshold value of 1 mm for the LA region (Figure 8c). According to the results for case study 1, experiments with a warm start show a reduction in the spin-up time except at 0800 UTC, where the FSS values are very low. This behaviour is due to the low accumulated rainfall from 0700 to 0800 UTC in the Abruzzo and Marche regions, where increased precipitation occurs later. The warm start initialization improves the precipitation estimation for both 3DVAR-CONMMSPC_warm and 4DVAR-CONMMSPC_warm when compared to the CTL and the simulations in cold start mode. On the other hand, no difference between 4DVAR-CONMMSPC_warm

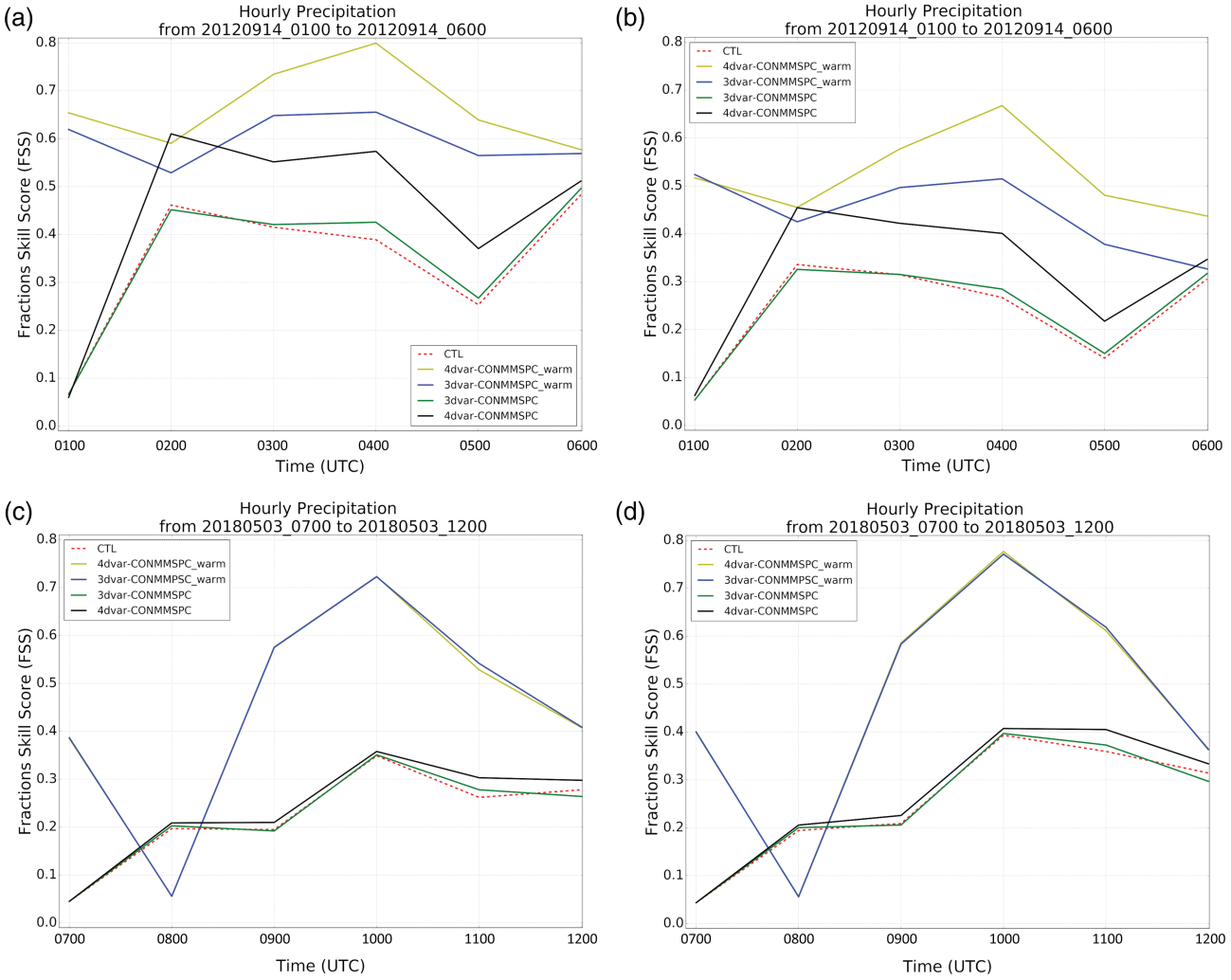

FIGURE 8 Upper panels: evolution of the FSS for the threshold >1 mm·h^{-1} in (a) the LA region and (b) the MA region for the case study occurring on September 14, 2012. Lower panels: evolution of the FSS for the same threshold in (c) the LA region and (d) the MA region for the case study occurring on May 3, 2018. Dashed red lines represent the CTL, yellow lines 4DVAR-CONMMSPC, green lines 4DVAR-CON, black lines 3DVAR-CON and blue lines 3DVAR-CONMMSPC

and 3DVAR-CONMMSPC_warm is found, in contrast to the results for case study 1. Figure 8d shows that the highest FSS values for the MA region are associated with 4DVAR-CONMMSPC_warm and 3DVAR-CONMMSPC_warm and a significant gap with respect to the other simulations is found. Moreover, the impact of a warm start initialization is reduced in the final hours of the simulation period, where the simulations converge to CTL.

These results highlight a possible limitation of 4D-Var when assimilating a reduced number of radars over a large area, which might reduce its capability to have a significant impact.

7 | SUMMARY AND CONCLUSIONS

In this paper, seven numerical experiments are presented to evaluate the impact of 3D-Var and 4D-Var in terms of SQPF for two flash flood events occurring in central Italy. The same number of radar measurements, reflectivity and radial velocity from Doppler radars at Mt. Midia and San Pietro Capofiume, in combination with conventional observations, are assimilated into the WRF model. Two different techniques for objectively analysing and comparing the results of the experiments are used: the point-by-point (or conventional) method with the ETS score, and the filtering method with the FSS score. Three different regions are selected for the evaluation: the Study Area (SA) and the Lazio-Abruzzo (LA) and Marche-Abruzzo (MA) sub-regions, using a threshold value of 1 mm·h^{-1}.

The conventional approach highlights the positive impact of 4D-Var in the first few hours of simulation for both case studies in the SA region. This improvement reduces in time, ending with the CTL performing better than the two assimilation methods. However, the results obtained with the conventional approach depend strongly on the number of rain-gauges

and their spatial distributions, which can affect the statistical analysis and the computation of the ETS. In this respect, the filtering method (FSS) allows the signal partially found with the conventional method to be clarified. Indeed, the 4D-Var experiments show higher values than the CTL for the FSS over the whole simulation period. This behaviour confirms that the SQPF evaluation depends on the statistical methods. Moreover, it demonstrates a positive impact of radar data assimilation in the 4D-Var experiment for case study 2. The FSS is calculated for the LA and MA regions to evaluate the performance of the two assimilation methods in the areas affected by relevant rainfall. The FSS for the 4D-Var experiments show greater values than for the 3D-Var experiments for both case studies. This result suggests a clear benefit of using 4D-Var as compared to 3D-Var, even if the same number of observations are assimilated. The FSS calculated for case study 2 shows a positive impact on the precipitation forecast when radar data are assimilated using the 4D-Var technique, although such a clear signal is not found for case study 1. Moreover, the FSS for 3DVAR-CONMMSPC is greater than that for 3DVAR-CON, suggesting a positive impact of radar data assimilation in the 3D-Var experiments. Therefore, to quantify these results, comparing the FSS for 4D-Var with radar data and conventional observations with that of 3D-Var results in an increment (expressed as a percentage) equal to 22.5 and 15.9% for the MA and LA regions, respectively, for the case study occurring on September 14, 2012. For the second case study occurring on May 3, 2018, the score increases by 5.5% in both regions. The different responses of each of the two experiments may be related to the computation of the B matrix, which covers 3 months for case study 1 but only 1 week for case study 2 due to limited computational resources.

Finally, two experiments are performed to examine the role of the warm start in relation to both 4D-Var and 3D-Var radar data assimilation. In these experiments, a further improvement in the FSS compared to the cold start and CTL experiments is found. Moreover, a tendency for 3DVAR-CONMMSPC_warm and 4DVAR-CONMMSPC_warm to converge to the CTL in the final hours indicates that the positive influence of using a warm start mode for 4D-Var and 3D-Var is restricted to a short time range. The warm start mode simulations show that 4D-Var performs better than 3D-Var in case study 1. This result suggests a greater impact when the observations are assimilated using 4D-Var. Conversely, no differences with respect to 3D-Var in the warm start mode are found for case study 2. Overall, these results allow us to identify an improvement with the warm start both in terms of reducing the spin-up time and increasing the FSS values during the simulations across all regions.

In conclusion, the experiments suggest that 4D-Var assimilation has a positive impact on the precipitation forecast for both case studies in cold start mode. If a warm start initialization is used with 4D-Var and 3D-Var, a considerable reduction in the spin-up time and a significant improvement in the rainfall forecast is obtained. This suggests that the initial precipitation spin-up problem still occurs when using 4D-Var. Moreover, the simulations in warm start mode indicate that 4D-Var performs better than 3D-Var for case study 1, while for case study 2 the difference is less clear. However, case study 2 is characterized by a lower amount of precipitation, which might cause the benefit of 4D-Var to be reduced.

The next step in this work will be to perform the assimilation with 4D-Var and 3D-Var over a high-resolution domain (3 km), increasing the frequency of the assimilated radar data (e.g., every 10 min) to further investigate the impact of both techniques on the estimation of precipitation. Moreover, the development of a procedure for a cycling assimilation could additionally improve the impact of data assimilation, allowing initial conditions to be continuously updated and thereby producing a better rainfall forecast. However, 4D-Var requires a huge amount of computational resource and computational time, meaning that its usage in operational contexts is still a challenge. In this respect, a comparison between the two assimilation methods and their benefits could be very useful for guiding future investments in the computational resources for small forecast centres. Finally, further investigations could focus on the assimilation of more than two radars and on the use of a larger number of outer loops, as in Maiello *et al.* (2017).

ACKNOWLEDGEMENTS

The authors thank Gran Sasso Laboratories for access to the computational resources and scientific computing, as well as NCAR for providing the WRF and WRFDA code.

ORCID

V. Mazzarella https://orcid.org/0000-0002-7596-4694
V. Capozzi https://orcid.org/0000-0002-8279-9922
P. P. Alberoni https://orcid.org/0000-0003-2107-0289

REFERENCES

Andersson, E. and Järvinen, H. (1999) Variational quality control. *Quarterly Journal of the Royal Meteorological Society*, 125, 697–722.

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M. and Reinhardt, T. (2011) Operational convective-scale numerical weather prediction with the COSMO model: description and sensitivities. *Monthly Weather Review*, 139(12), 3887–3905.

Baldwin, M.E. and Kain, J.S. (2006) Sensitivity of several performance measures to displacement error, bias, and event frequency. *Weather Forecasting*, 21, 636–648. https://doi.org/10.1175/WAF933.1.

Barker, D., Huang, X., Liu, Z., Auligné, T., Zhang, X., Rugg, S., Ajjaji, R., Bourgeois, A., Bray, J., Chen, Y., Demirtas, M., Guo, Y., Henderson, T., Huang, W., Lin, H., Michalakes, J., Rizvi, S. and Zhang, X. (2012) The Weather Research and Forecasting model's Community

Variational/Ensemble Data Assimilation System: WRFDA. *Bulletin of the American Meteorological Society*, 93, 831–843. https://doi.org/10.1175/BAMS-D-11-00167.1.

Barker, D.M., Huang, W., Guo, Y.-R., Bourgeois, A. and Xiao, Q. (2004) A three-dimensional variational data assimilation system for MM5: implementation and initial results. *Monthly Weather Review*, 132, 897–914.

Berre, L. (2000) Estimation of synoptic and mesoscale forecast error covariances in a limited area model. *Monthly Weather Review*, 128, 644–667.

Bryan, G.H. and Rotunno, R. (2005) Statistical convergence in simulated moist absolutely unstable layers. In: *Proceedings of the 11th Conference on Mesoscale Processes*, Albuquerque, NM: American Meteorological Society, 1M.6.

Cassola, F., Ferrari, F., Mazzino, A. and Miglietta, M.M. (2016) The role of the sea on the flash floods events over Liguria (northwestern Italy). *Geophysical Research Letters*, 43(7), 3534–3542.

Chen, F. and Dudhia, J. (2001) Coupling an advanced land-surface/hydrology model with the Penn State/NCAR MM5 modeling system. Part I: model description and implementation. *Monthly Weather Review*, 129, 569–585.

Choi, Y., Lim, G.-H. and Lee, D.-K. (2013) Radar radial wind data assimilation using the time-incremental 4D-Var method implemented to the WRFDA system. *Tellus A: Dynamic Meteorology and Oceanography*, 65(1), 19677. https://doi.org/10.3402/tellusa.v65i0.19677.

Chu, K., Xiao, Q. and Liu, C. (2013) Experiments of the WRF three-/four-dimensional variational (3/4DVAR) data assimilation in the forecasting of Antarctic cyclones. *Meteorology and Atmospheric Physics*, 120, 145–156. https://doi.org/10.1007/s00703-013-0243-y.

Courtier, P., Thépaut, J.-N. and Hollingsworth, A. (1994) A strategy for operational implementation of 4D-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, 120, 1367–1387.

Davolio, S., Ferretti, R., Baldini, L., Casaioli, M., Cimini, D., Ferrario, M.E., Gentile, S., Loglisci, N., Maiello, I., Manzato, A., Mariani, S., Marsigli, C., Marzano, F.S., Miglietta, M.M., Montani, A., Panegrossi, G., Pasi, F., Pichelli, E., Pucillo, A. and Zinzi, A. (2015) The role of the Italian scientific community in the first HyMeX SOP: an outstanding multidisciplinary experience. *Meteorologische Zeitschrift*, 24, 261–267.

Ducrocq, V., Braud, I., Davolio, S., Ferretti, R., Flamant, C., Jansa, A., Kalthoff, N., Richard, E., Taupier-Letage, I., Ayral, P., Belamari, S., Berne, A., Borga, M., Boudevillain, B., Bock, O., Boichard, J., Bouin, M., 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., 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. (2014) HyMeX-SOP1: the field campaign dedicated to heavy precipitation and flash flooding in the northwestern Mediterranean. *Bulletin of the American Meteorological Society*, 95, 1083–1100. https://doi.org/10.1175/BAMS-D-12-00244.1.

Dudhia, J. (1989) Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *Journal of the Atmospheric Sciences*, 46, 3077–3107.

Duffourg, F. and Ducrocq, V. (2011) Origin of the moisture feeding the Heavy Precipitating Systems over Southeastern France. *Natural Hazards and Earth System Sciences*, 11, 1163–1178. https://doi.org/10.5194/nhess-11-1163-2011.

Duffourg, F. and Ducrocq, V. (2013) Assessment of the water supply to Mediterranean heavy precipitation: a method based on finely designed water budgets. *Atmospheric Science Letters*, 14, 133–138. https://doi.org/10.1002/asl2.429.

Ehrendorfer, M. (1997) Predicting the uncertainty of numerical weather forecasts: a review. *Meteorologische Zeitschrift*, 6, 147–183.

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. (2014) Overview of the first HyMeX Special Observation Period over Italy: observations and model results. *Hydrology and Earth System Sciences*, 18, 1953–1977. https://doi.org/10.5194/hess-18-1953-2014.

Gallus, W.A. and Bresch, J.F. (2006) Comparison of impacts of WRF dynamic core, physics package, and initial conditions on warm season rainfall forecasts. *Monthly Weather Review*, 134, 2632–2641. https://doi.org/10.1175/MWR3198.1.

Grell, G.A. and Devenyi, D. (2002) A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophysical Research Letters*, 29, 38-1–38-4. https://doi.org/10.1029/2002GL015311.

Guo, Y.-R., Lin, H.-C., Ma, X.X., Huang, X.-Y., Terng, C.T. and Kuo, Y.-H. (2006) Impact of WRF-Var (3DVar) Background Error Statistics on typhoon analysis and forecast. In: *WRF Users' Workshop*, Boulder, CO: NCAR, P4.2.

Gustafsson, N., Berre, L., Hörnquist, S., Huang, X.-Y., Lindskog, M., Navascués, B., Mogensen, K.S. and Thorsteinsson, S. (2001) Three-dimensional variational data assimilation for a limited area model. Part I: general formulation and the background error constraint. *Tellus A: Dynamic Meteorlogy and Oceanography*, 53, 425–446.

Hally, A., Caumont, O., Garrote, L., Richard, E., Weerts, A., Delogu, F., Fiori, E., Rebora, N., Parodi, A., Mihalović, A., Ivković, M., Dekić, L., van Verseveld, W., Nuissier, O., Ducrocq, V., D'Agostino, D., Galizia, A., Danovaro, E. and Clematis, A. (2015) Hydrometeorological multi-model ensemble simulations of the November 4, 2011 flash flood event in Genoa, Italy, in the framework of the DRIHM project. *Natural Hazards and Earth System Sciences*, 15, 537–555. https://doi.org/10.5194/nhess-15-537-2015.

Hong, S.–.Y., Noh, Y. and Dudhia, J. (2006) A new vertical diffusion package with an explicit treatment of entrainment processes. *Monthly Weather Review*, 134, 2318–2341.

Hong, S.Y., Lim, K.S.S., Lee, Y.H., Ha, J.C., Kim, H.W., Ham, S.J. and Dudhia, J. (2010) Evaluation of the WRF Double-Moment 6-Class microphysics scheme for precipitating convection. *Advances in Meteorology*, 2010, 707253.

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. (2009) Four-dimensional variational data assimilation for WRF: formulation and preliminary results. *Monthly Weather Review*, 137, 299–314.

Jankov, I., Gallus, W.A., Segal, M. and Koch, S.E. (2007) Influence of initial conditions on the WRF–ARW model QPF response to physical parameterization changes. *Weather Forecasting*, 22, 501–519. https://doi.org/10.1175/WAF998.1.

Jianfeng, G.U., Xiao, Q., Kuo, Y.-H., Barker, D.M., Jishan, X. and Xiaoxing, M.A. (2005) Assimilation and simulation of Typhoon Rusa (2002) using the WRF system. *Advances in Atmospheric Sciences*, 22, 415–427. https://doi.org/10.1007/BF02918755.

Lastoria, B., Simonetti, M.R., Casaioli, M., Mariani, S. and Monacelli, G. (2006) Socio-economic impacts of major floods in Italy from 1951 to 2003. *Advances in Geosciences*, 7, 223–229.

Lee, J.-H., Lee, H.-H., Choi, Y., Kim, H.-W. and Lee, D.-K. (2010) Radar data assimilation for the simulation of mesoscale convective systems. *Advances in Atmospheric Sciences*, 27, 1025–1042. https://doi.org/10.1007/s00376-010-91628.

Maiello, I., Ferretti, R., Gentile, S., Montopoli, M., Picciotti, E., Marzano, F.S. and Faccani, C. (2014) Impact of radar data assimilation for the simulation of a heavy rainfall case in central Italy using WRF–3DVAR. *Atmospheric Measurement Techniques*, 7, 2919–2935. https://doi.org/10.5194/amt-7-2919-2014.

Maiello, I., Gentile, S., Ferretti, R., Baldini, L., Roberto, N., Picciotti, E., Alberoni, P.P. and Marzano, F.S. (2017) Impact of multiple radar reflectivity data assimilation on the numerical simulation of a flash flood event during the HyMeX campaign. *Hydrology and Earth System Sciences*, 21, 5459–5476. https://doi.org/10.5194/hess-21-5459-2017.

Marshall, J.S. and Palmer, W.M. (1948) The distribution of raindrops with size. *Journal of Meteorology*, 5, 165–166.

Mass, C.F., Ovens, D., Westrick, K. and Colle, B.A. (2002) Does increasing horizontal resolution produce more skillful forecasts? *Bulletin of the American Meteorological Society*, 83, 407–430. https://doi.org/10.1175/1520-0477(2002)083<0407:DIHRPM>2.3.CO;2.

Massart, S., Pajot, B., Piacentini, A. and Pannekoucke, O. (2010) On the merits of using a 3D-FGAT assimilation scheme with an outer loop for atmospheric situations governed by transport. *Monthly Weather Review*, 138, 4509–4522.

Mlawer, E.J., Taubman, S.J., Brown, P.D., Iacono, M.J. and Clough, S.A. (1997) Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research: Atmospheres*, 102, 16663–16682.

Parrish, D.F. and Derber, J.C. (1992) The National Meteorological Center's spectral statistical-interpolation analysis system. *Monthly Weather Review*, 120, 1747–1763.

Pennelly, C., Reuter, G. and Flesch, T. (2014) Verification of the WRF model for simulating heavy precipitation in Alberta. *Atmospheric Research*, 135, 172–192.

Rabier, F., Klinker, E., Courtier, P. and Hollingsworth, A. (1996) Sensitivity of forecast errors to initial conditions. *Quarterly Journal of the Royal Meteorological Society*, 122, 121–150. https://doi.org/10.1002/qj.49712252906.

Rennie, S.J., Dance, S.L., Illingworth, A.J., Ballard, S.P. and Simonin, D. (2011) 3D-Var assimilation of insect-derived Doppler radar radial winds in convective cases using a high-resolution model. *Monthly Weather Review*, 139(4), 1148–1163.

Rizvi, S.R., Guo, Y.R., Shao, H., Demirtas, M. and Huang, X.Y. (2008) Impact of outer loop for WRF data assimilation system (WRFDA). *WRF Users' Workshop*. Boulder, CO: NCAR. Available at: http://www.mmm.ucar.edu/wrf/users/workshops/WS2008/abstracts/P5-03.pdf [Accessed 14th January 2019].

Roberts, N.M. (2003) *The impact of a change to the use of the convection scheme to high-resolution simulations of convective events*. Met Office Forecasting Research Technical Report number: 407.

Roberts, N.M. and Lean, H.W. (2008) Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Monthly Weather Review*, 136, 78–97.

Romine, G.S., Schwartz, C.S., Snyder, C., Anderson, J.L. and Weisman, M.L. (2013) Model bias in a continuously cycled assimilation system and its influence on convection-permitting forecasts. *Monthly Weather Review*, 141, 1263–1284.

Rosmond, T. and Xu, L. (2006) Development of NAVDAS-AR: non-linear formulation and outer loop tests. *Tellus A: Dynamic Meteorology and Oceanography*, 58, 45–58.

Salvati, P., Bianchi, C., Rossi, M. and Guzzetti, F. (2010) Societal landslide and flood risk in Italy. *Natural Hazards and Earth System Sciences*, 10, 465–483. https://doi.org/10.5194/nhess-10-465-2010.

Schaefer, J.T. (1990) The critical success index as an indicator of warning skill. *Weather Forecasting*, 5, 570–575.

Silvestro, F., Rebora, N., Giannoni, F., Cavallo, A. and Ferraris, L. (2016) The flash flood of the Bisagno Creek on 9th October 2014: an "unfortunate" combination of spatial and temporal scales. *Journal of Hydrology*, 541, 50–62.

Simmons, A.J. and Hollingsworth, A. (2002) Some aspects of the improvement in skill of numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 128, 647–677. https://doi.org/10.1256/003590002321042135.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Duda, M.G., Huang, X.-Y., Wang, W. and Powers, J.G. (2008) *A description of the Advanced Research WRF Version 3*, NCAR Technical Note TN 475+STR. Available at: http://www2.mmm.ucar.edu/wrf/users/docs/arw_v3.pdf [Accessed 15th January 2019].

Sousounis, P.J., Hutchinson, T.A. and Marshall, S.F. (2004) A comparison of MM5, WRF, RUC, and ETA performance for Great Plains heavy precipitation events during the spring of 2003. *24, 20th Conference on Weather Analysis and Forecasting/16th Conference on Numerical Weather Prediction*, Meteor. Soc., Amer, p. 6.

Stanešić, A. and Brewster, K.A. (2016) Impact of radar data assimilation on the numerical simulation of a severe storm in Croatia. *Meteorologische Zeitschrift*, 25, 37–53.

Stensrud, D.J., Bao, J. and Warner, T.T. (2000) Using initial condition and model physics perturbations in short-range ensemble simulations of mesoscale convective systems. *Monthly Weather Review*, 128, 2077–2107. https://doi.org/10.1175/1520-0493(2000)128<2077:UICAMP>2.0.CO;2.

Sun, J. and Crook, N.A. (1997) Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part I: model development and simulated data experiments. *Journal of the Atmospheric Sciences*, 54, 1642–1661.

Sun, J. and Wang, H. (2013) Radar data assimilation with WRF 4DVar. Part II: comparison with 3D-Var for a squall line over the US Great Plains. *Monthly Weather Review*, 11, 2245–2264. https://doi.org/10.1175/MWR-D-12-00169.1.

Thompson, G., Field, P.R., Rasmussen, R.M. and Hall, W.D. (2008) Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: implementation of a new snow parameterization. *Monthly Weather Review*, 136, 5095–5115. https://doi.org/10.1175/2008MWR2387.1.

Wang, H., Sun, J., Fan, S. and Huang, X.Y. (2013a) Indirect assimilation of radar reflectivity with WRF 3D-Var and its impact on prediction of

four summertime convective events. *Journal of Applied Meteorology and Climatology*, 52, 889–902.

Wang, H., Sun, J., Zhang, X., Huang, X. and Auligne, T. (2013b) Radar data assimilation with WRF 4D-Var. Part I: system development and preliminary testing. *Monthly Weather Review*, 141, 2224–2244.

Wattrelot, E., Caumont, O. and Mahfouf, J. (2014) Operational implementation of the 1D+3D-Var assimilation method of radar reflectivity data in the AROME model. *Monthly Weather Review*, 142, 1852–1873. https://doi.org/10.1175/MWR-D-13-00230.1.

Wilks, D.S. (2006) *Statistical Methods in the Atmospheric Sciences*, 2nd edition. New York: Academic Press.

Xiao, Q., Kuo, Y.-H., Sun, J. and Lee, W.-C. (2005) Assimilation of Doppler radar observations with a regional 3DVAR system: impact of Doppler velocities on forecasts of a heavy rainfall case. *Journal of Applied Meteorology*, 44, 768–788.

Xiao, Q., Kuo, Y.-H., Sun, J., Lee, W.-C., Barker, D.M. and Lim, E. (2007) An approach of radar reflectivity data assimilation and its assessment with the inland QPF of Typhoon Rusa (2002) at landfall. *Journal of Applied Meteorology and Climatology*, 46, 14–22.

Xiao, Q. and Sun, J. (2007) Multiple-radar data assimilation and short-range quantitative precipitation forecasting of a squall line observed during IHOP_2002. *Monthly Weather Review*, 135, 3381–3404.

Yano, J., Ziemiański, M.Z., Cullen, M., Termonia, P., Onvlee, J., Bengtsson, L., Carrassi, A., Davy, R., Deluca, A., Gray, S.L., Homar, V., Köhler, M., Krichak, S., Michaelides, S., Phillips, V.T., Soares, P.M. and Wyszogrodzki, A.A. (2018) Scientific challenges of convective-scale numerical weather prediction. *Bulletin of the American Meteorological Society*, 99, 699–710. https://doi.org/10.1175/BAMS-D-17-0125.1.

Zhang, X., Huang, X.-Y. and Pan, N. (2013) Development of the upgraded tangent linear and adjoint of the Weather Research and Forecasting (WRF) model. *Journal of Atmospheric and Oceanic Technology*, 30, 1180–1188. https://doi.org/10.1175/JTECH-D-12-00213.1.

How to cite this article: Mazzarella V, Maiello I, Ferretti R, *et al*. Reflectivity and velocity radar data assimilation for two flash flood events in central Italy: A comparison between 3D and 4D variational methods. *Q J R Meteorol Soc*. 2020;146:348–366. https://doi.org/10.1002/qj.3679