Predictive Capability of a High-Resolution Hydrometeorological Forecasting Framework Coupling WRF Cycling 3DVAR and Continuum

MARTINA LAGASIO, FRANCESCO SILVESTRO, LORENZO CAMPO, AND ANTONIO PARODI

CIMA Research Foundation, Savona, Italy

(Manuscript received 18 October 2018, in final form 7 May 2019)

ABSTRACT

The typical complex orography of the Mediterranean coastal areas support the formation of the so-called back-building mesoscale convective systems (MCS) producing torrential rainfall often resulting in flash floods. As these events are usually very small-scaled and localized, they are hardly predictable from a hydrometeorological standpoint, frequently causing a significant amount of fatalities and socioeconomic damage. Liguria, a northwestern Italian region, is characterized by small catchments with very short hydrological response time and is thus extremely prone to the impacts of back-building MCSs. Indeed, Liguria has been hit by three intense back-building MCSs between 2011 and 2014, causing a total death toll of 20 people and several hundred millions of euros of damages. Consequently, it is necessary to use hydrometeorological forecasting frameworks coupling the finescale numerical weather prediction (NWP) outputs with rainfall–runoff models to provide timely and accurate streamflow forecasts. Concerning the aforementioned back-building MCS episodes that recently occurred in Liguria, this work assesses the predictive capability of a hydrometeorological forecasting framework composed by a kilometer-scale cloud-resolving NWP model (WRF), including a 6-h cycling 3DVAR assimilation of radar reflectivity and conventional weather stations data, a rainfall downscaling model [Rainfall Filtered Autoregressive Model (RainFARM)], and a fully distributed hydrological model (Continuum). A rich portfolio of WRF 3DVAR direct and indirect reflectivity operators has been explored to drive the meteorological component of the proposed forecasting framework. The results confirm the importance of rapidly refreshing and data intensive 3DVAR for improving the quantitative precipitation forecast, and, subsequently, the flash flood prediction in cases of back-building MCS events.

1. Introduction

The Mediterranean region is frequently struck by severe floods and flash floods causing a significant death toll and several millions of euros of damage. The western Mediterranean area is characterized by a complex orography (Alps, Apennines, Massif Central, Pyrenees), often sitting close to the coastline, that is potentially able to enhance or even to trigger the deep convective processes originating over the warm sea in the fall season (Rebora et al. 2013; Ducrocq et al. 2014; Fiori et al. 2017). The most severe events in this area are due to a particular type of mesoscale configuration featuring a continuous redevelopment of storm cells persisting for

hours over the same area, the so-called back-building mesoscale convective systems (MCSs; Rebora et al. 2013; Ducrocq et al. 2014; Cassola et al. 2015; Fiori et al. 2017; Lagasio et al. 2017).

Having a very steep coastal orography mostly drained by very small-sized catchments (1–10 km²), the Liguria region (northwestern Italy) is particularly prone to flash floods induced by back-building MCSs: in the period between October 2010 and October 2014 alone, four events (Varazze, 4 October 2010; Cinqueterre, 25 October 2011; Genoa, 4 November 2011; and Genoa, 9 October 2014) accounted for 30 casualties and hundreds of millions of euros of damage. Consequently, the use of high-resolution hydrometeorological forecasting frameworks combining numerical weather prediction (NWP) models and rainfall–runoff models is recognized to be essential to provide timely and accurate streamflow forecasts (Silvestro et al. 2015b). A considerable effort has been made in the last few years to develop cloud-resolving NWP systems, possibly in combination with ensemble

Denotes content that is immediately available upon publication as open access.

Corresponding author: Martina Lagasio, martina.lagasio@cimafoundation.org

DOI: 10.1175/JHM-D-18-0219.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](https://www.ametsoc.org/PUBSReuseLicenses) (www.ametsoc.org/PUBSReuseLicenses).

and multiphysics approaches, to improve the short-term quantitative precipitation forecast (QPF) of convective extreme events (Ducrocq et al. 2014; Hally et al. 2015; Clark et al. 2016; Davolio et al. 2017; Fiori et al. 2017; Lagasio et al. 2017). However, a reliable forecast of these events in terms of rainfall amount, location, and timing is still an open issue (Ducrocq et al. 2014) that cannot be tackled only through the increase of the NWP models' space–time resolution.

NWP is a mathematical problem determined by its initial and boundary (in the case of limited area modeling) conditions. QPF challenges often derive from the uncertainty related to the initial state of the atmosphere at small spatiotemporal scales (Bauer et al. 2015). The inevitable model spinup often results in an inaccurate simulation of the timing, the location, and the severity of convective systems (Sugimoto et al. 2009). This challenge becomes even more relevant when the model grid spacing is approaching the kilometeric scale, mainly as a consequence of the lack of high spatiotemporal resolution observations. Several studies demonstrated that the assimilation of high spatiotemporal resolution observations such as radar reflectivity data may reduce the model spinup (Sugimoto et al. 2009). In the last few years, significant advances in forecasting heavy rainfall events have been achieved thanks to the combination of high-resolution meteorological models with the data assimilation of both in situ and radar observations (Davolio et al. 2017; Maiello et al. 2017; Mazzarella et al. 2017). More specifically, some studies investigated the influence of reflectivity data assimilation combined with conventional surface observations for heavy rainfall events in southwest England, the Korean Peninsula, and Bangladesh (Lee et al. 2010; Liu et al. 2013; Ha et al. 2011; Das et al. 2015) as well as over central Italy (Maiello et al. 2014, 2017; Mazzarella et al. 2017). In the case of the Liguria region, the effect of the nudging of radar-derived rainfall data on hydrometeorological predictive capability has been evaluated through the coupling of the meteorological forecast with the Continuum hydrological model (Davolio et al. 2015, 2017) for some events of the autumn of 2014. Their main result is that the contribution of the nudging of radar rainfall data observations is large during the assimilation period and still relevant in the following 3 h of the free forecasts, but rapidly decreases after 6 h.

This study aims to gain further insights into the hydrometeorological prediction of back-building MCSs through the combination of a high-resolution WRF Model instance including a 3DVAR data assimilation cycle—with the fully distributed Continuum hydrological model, via the Rainfall Filtered Autoregressive Model (RainFARM) stochastic downscaling procedure

(Rebora et al. 2006a,b). The specific novelty of this research resides in driving the flash flood forecasting framework with a rich portfolio of direct and indirect radar reflectivity WRF data assimilation (WRFDA)-3DVAR operators as well as in situ weather station data fed into a cloud-resolving NWP model in a cycling mode.

To the best of the authors' knowledge, this is the first work aiming to assess the improvement of precipitation and flash flood predictions, induced by highly impactful back-building MCSs in the northwestern Mediterranean area, by feeding a hydrometeorological forecasting framework with such a computationally intensive high-resolution modeling and observational data framework.

The paper is organized as it follows. In section 2 the V-shaped back-building MCSs over the Liguria region are described (section 2a) together with the available observational datasets (section 2b). Section 3 describes the model's setup and the methodological approach: more specifically, section 3a presents the hydrometeorological framework; section 3b describes the data assimilation method [section 3b(1)], WRF Model setup reporting the experiment configuration [section 3b(2)], and the validation used in this work [section 3b(3)]; and section 3c presents the stochastic downscaling model (RainFARM) that provides the probabilistic scenarios for the hydrological model (Continuum). Section 4 summarizes the results and the validation for both the meteorological forecasts (section 4a) and the hydrological peak discharge forecasts (section 4b). Conclusions are drawn in section 5.

2. Test cases: V-shaped back-building MCSs over the Liguria region

a. V-shaped back-building MCSs description

This study will focus on three extreme meteorological events that hit the Liguria region (located in the northwestern part of Italy) in 2011 and 2014. The first event occurred on 25 October 2011 when a very intense back-building MCS (470 mm of rain in 6 h) produced widespread flash flood phenomena in Cinque Terre (the red dot in Fig. 1), causing the death of 13 people and several millions of euros in damages. Ten days after, on 4 November, another back-building MCS of the same intensity (450 mm of rain in 5 h) affected Genoa's city center (black dot in Fig. 1), resulting again in a large amount of damage and the death of six people. Three years later (9 October 2014), a third flash flood again struck the very same part of Genoa. This time the meteorological event was characterized by two phases: the first one happened in the morning (between 0800 and 1200 UTC) recording rainfall amounts between 50 and 130 mm over the Bisagno catchment, while the second

FIG. 1. Data available for the assimilation. The red circle represents the area covered by the Settepani radar with the red small square indicating the radar location. Cyan dots are all the available surface observations stations recording wind speed and direction, temperature, and humidity. The gray shadow isolates the area covered by the Italian Radar Network (white circles mosaic) inside the WRF domains. The dotted and solid black lines represent WRF nested domains with spatial resolutions of 5 km (d_1) and 1 km (d_2) adopted for simulations, while the black dot indicates the Genoa city location (hit by two of the three extreme events simulated) and the red dot locates the Cinque Terre area (hit by one extreme event considered) in the Liguria region.

one occurred after a few hours and, although it had rather similar meteorological dynamics, it was even more intense, pounding again the same catchment with another 150 and 260 mm in 2 h (2000–2200 UTC). Locally, the daily maximum cumulated rainfall reached 400 mm with an average of 200 mm over the entire basin area (90 km²).

The choice of these case studies is motivated by their rather similar thermodynamic nature, featuring the occurrence of a very fine-scaled V-shaped back-building MCS, a very challenging phenomenon from the hydrometeorological predictive capability standpoint (Parodi et al. 2012; Rebora et al. 2013; Cassola et al. 2015; Davolio et al. 2015; Fiori et al. 2017; Lagasio et al. 2017). All these events were characterized by an area of very intense precipitation resembling a V-shaped pattern generated by a wind convergence line in the lower planetary boundary layer (PBL) stable enough to develop and feed the back-building process repeatedly over the same area for many hours (Parodi et al. 2012; Rebora et al. 2013). At the mesoscale level this scenario is caused by a cold and dry jet outflowing from the Po valley acting as an obstacle to the warm and moist low-level

PBL jet coming from southeast. The latter is then lifted up over the cold stable layer as it would be when in the presence of coastal topography (Fiori et al. 2017). If the cold outflow and the warm southeasterly flow have similar intensity, the phenomenon is then quasi-stationary and may persist over the same area for quite a long time (6–8 h).

The three selected test cases will allow for evaluating the impact of a cycling 3DVAR and different assimilation operators for in situ and weather radar data on the predictive capability of such systems.

b. Observations available for data assimilation and validation

The observational data to be assimilated via 3DVAR in WRF are reflectivity from weather radars and temperature, wind speed and direction, and relative humidity from surface observations.

For the radar reflectivity, the observational data are provided by the meteorological radar national mosaic operated by the Italian Civil Protection Agency (Vulpiani et al. 2008; CAPPI data on three levels: 2000, 3000, and 5000 m MSL) covering the whole Italian territory (Fig. 1, gray shadow), and by the eight levels (CAPPI data at 1500, 2000, 3000, 4000, 5000, 6000, 8000, 10 000 m MSL) of the C-band polarimetric radar located on Mount Settepani (Fig. 1, red square) covering the Liguria region (Silvestro et al. 2009). The Settepani radar is already integrated in the national mosaic, however, the idea of using its data separately when available (for instance, during the Genoa 2014 flood) is justified by their higher vertical resolution with respect to the national mosaic. The ground sensor data are provided by the Italian Civil Protection hydrometeorological network. This operational network, employed for the hydrometeorological monitoring of the Italian territory, is composed by thermometers, rain gauges, hygrometers, and anemometers and is particularly dense (about one station every 10 km²) on the Ligurian coast (Fig. 1, cyan dots).

Concerning the quantitative precipitation estimate (QPE), the 2011 case studies rely purely on rain gauge data, while the rainfall retrieval from the Settepani radar (Silvestro et al. 2009) was only possible for the Genoa 2014 flash flood. The Settepani radar therefore allowed for obtaining accurate data over the sea, a key element for the prediction of the onset of these kinds of events (Lagasio et al. 2017).

3. Models, setup, and methodology

a. Hydrometeorological framework

To assess the impact of the atmospheric 3DVAR data assimilation on streamflow prediction, a hydrometeorological forecasting framework was employed. The framework is

composed by the cascade of the Advanced Research version of the Weather Research and Forecasting (WRF-ARW) Model with WRFDA for cyclic 3DVAR data assimilation, a stochastic downscaling model (RainFARM) and a hydrological model (Continuum).

b. WRF-ARW setup and WRFDA assimilation method

The atmospheric model used in this work is the WRF-ARW, a fully compressible and nonhydrostatic regional atmospheric model, with terrain-following hydrostatic pressure vertical coordinate (Skamarock et al. 2008), here deployed in its version 3.9. The WRF Model includes a data assimilation system (WRFDA version 3.9.1; Barker et al. 2012) that is used in this study to perform a cycling 3DVAR data assimilation.

1) VARIATIONAL DATA ASSIMILATION AND OBSERVATIONS OPERATORS

The data assimilation is a mathematical technique that combines a NWP output (first guess or background forecast) with observations and their respective error statistics handing out a more reliable state of the atmosphere (analysis). The variational data assimilation achieves this result through the minimization of a cost function given two sources of data: a background forecast (first guess) and observations. In WRFDA both 3DVAR and 4DVAR are available since the WRF 3.0 release (Barker et al. 2004; Skamarock et al. 2008). In this study the 3DVAR method is employed with a cycling update technique. The basic goal of this technique is to provide an optimal estimate of the true state of the atmosphere through the minimization of the cost function reported in Eq. (1) (Ide et al. 1997):

$$J(\mathbf{x}) = J^b + J^0 = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}(\mathbf{y} - \mathbf{y}^0)^T (\mathbf{E} + \mathbf{F})^{-1}(\mathbf{y} - \mathbf{y}^0). \quad (1)$$

In summary, the 3DVAR problem solution is given by the analysis state $\mathbf{x}$ that minimizes the cost function $J(\mathbf{x})$, which represents the a posteriori maximum likelihood estimate of the true state of the atmosphere, combining the two sources of data: observations $\mathbf{y}^0$ and background $\mathbf{x}^b$. Here, $\mathbf{B}$, $\mathbf{E}$, and $\mathbf{F}$ are respectively the background, observation (instrumental), and representativity error covariance matrices used as weights for the two sources of data. The representativity error is introduced from the use of an observation operator H to transform the gridded analysis to the observation space: $\mathbf{y} = H(\mathbf{x})$.

The correct estimation of the error covariance matrices is very important to obtain a good quality result. The errors linked to the observations are summarized

in a $\mathbf{R} = \mathbf{E} + \mathbf{F}$ matrix that usually is diagonal assuming the correlations between different instruments and between different observations made by the same instruments equal to zero. Furthermore, the background error covariance statistics $\mathbf{B}$ are necessary in WRFDA cost function minimization to weight errors in features of the background forecast field. It can be estimated through statistical approaches because a straightforward estimation of $\mathbf{B}$ is not possible as the variable's correlation is unknown. The WRFDA's `gen_be` utility estimates domain-specific climatological background error covariance matrix based on input training data that could be time series of forecast differences [the so-called National Meteorological Center (NMC) method; Parrish and Derber 1992] or perturbations from an ensemble prediction system (Skamarock et al. 2008, chapter 9). For the 3DVAR application the first method is commonly used; it estimates climatological background error covariances using a process that assumes background errors to be well approximated by averaged forecast difference statistics:

$$\mathbf{B} = \overline{(\mathbf{x}'^b - \mathbf{x}'^t)(\mathbf{x}'^b - \mathbf{x}'^t)^T} = \overline{\mathbf{\varepsilon}_b \mathbf{\varepsilon}_b^T} = \overline{(\mathbf{x}'^{t+24} - \mathbf{x}'^{t+12})(\mathbf{x}'^{t+24} - \mathbf{x}'^{t+12})^T}, \quad (2)$$

where $\mathbf{x}'^t$ is the true state of the atmosphere and $\mathbf{\varepsilon}_b$ is the background error. The overbar means an average over time and space. Forecast differences are valid at the same time, but one of them starts later than the other (e.g., compare the 12-h forecast that is in common of a 24-h forecast initialized at 0000 UTC with a 12-h forecast initialized at 1200 UTC).

In this work the 3DVAR system is used to assimilate reflectivity and conventional observations (surface observations from ground sensors), performing a sensitivity analysis using all the different reflectivity operator options available in WRFDA, as summarized hereafter.

The first operator used is the direct technique (Xiao et al. 2007) that assimilates reflectivity by converting the model rainwater mixing ratio into reflectivity using the total mixing ratio as control variable (Sun and Crook 1997). The relation for the observation operator is

$$\tilde{Z}(q_r) = 43.1 + 17.5 \log_{10}(\rho q_r), \quad (3)$$

where $\tilde{Z}$ is the reflectivity (dBZ), ρ is the atmospheric density (kg m^{-3}), and q_r is the rainwater mixing ratio (g kg^{-1}) [refer to Xiao et al. (2007) for further information about direct reflectivity data assimilation methods using the 3DVAR technique].

The second method is the indirect assimilation (Wang et al. 2013; Gao and Stensrud 2012), which assimilates hydrometeor mixing ratios estimated from radar

reflectivity. The forward reflectivity operator is obtained adjusting the formulation of Lin et al. (1983), Gilmore

et al. (2004), and Dowell et al. (2011), and it is represented in Eq. (4):

$$Z_e = \begin{cases} Z(q_r), & T_b > 5^\circ\text{C}, \\ Z(q_s) + Z(q_h), & T_b < -5^\circ\text{C}, \\ \alpha Z(q_r) + (1 - \alpha)[Z(q_s) + Z(q_h)], & -5^\circ\text{C} < T_b < 5^\circ\text{C}, \end{cases} \quad (4)$$

where Z_e is the equivalent reflectivity, α varies linearly between 0 at $T_b = -5^\circ\text{C}$ and 1 at $T_b = 5^\circ\text{C}$, T_b is the background temperature from an NWP model,

$$Z(q_r) = 3.63 \times 10^9 (\rho q_r)^{1.75} \quad (5)$$

is the rain component of reflectivity (Smith et al. 1975), and

$$Z(q_s) = 9.80 \times 10^8 (\rho q_s)^{1.75}, \quad T_b < 0^\circ\text{C}, \quad (6)$$

$$Z(q_s) = 4.26 \times 10^{11} (\rho q_s)^{1.75}, \quad T_b > 0^\circ\text{C} \quad (7)$$

are the snow component operators. If the temperature is lower than 0°C the dry snow operator is used [Eq. (6)] otherwise the wet snow operator [Eq. (7)] is applied. Finally, Eq. (8) represents the hail component of reflectivity (Lin et al. 1983; Gilmore et al. 2004):

$$Z(q_h) = 4.33 \times 10^{10} (\rho q_h)^{1.75}. \quad (8)$$

It is worth noticing that Eq. (8) mentions the hail component q_h , which is not necessarily predicted by all microphysics parameterizations available in the WRF modeling suite; however, WRFDA code understands and uses the q_h variable as a graupel species q_g . The last step needed is the conversion of the equivalent reflectivity Z_e (dBZ) to $\tilde{Z}$:

$$\tilde{Z} = 10 \log_{10}(Z_e). \quad (9)$$

A third experiment has been performed using the indirect assimilation combined with an option that also allows the assimilation of the in-cloud humidity estimated from reflectivity (Wang et al. 2013). In this case the observation operator is defined by Eq. (10):

$$q_v = \text{rh} \times q_s, \quad (10)$$

where q_v is the specific humidity, rh is the relative humidity, and q_s is the saturated specific humidity of water vapor. Thus, this experiment also includes the assimilation of the in-cloud humidity in addition to the hydrometeors species retrieved with the indirect method alone.

All the reflectivity operators presented above are available in the standard WRFDA-3DVAR package.

In this work an additional experiment (the fourth method for this work) has been performed by modifying the direct assimilation operator. In fact, the direct reflectivity operator [Eq. (3)] uses a warm rain scheme only and does not take into account all hydrometeors (snow, hail/graupel) like the default indirect operator [Eq. (4)] does. Thus, taking inspiration from Eqs. (5)–(8), this fourth method (hereafter named the modified direct operator) applies the same equations to compute the modeled reflectivity instead of using only Eq. (3). Finally, the reflectivity components calculated from Eqs. (5)–(8) are summed up depending on temperature ranges, similarly as Eq. (4) does for the indirect method, but in a simplified way, just discriminating between temperature values below and above zero, as it is reported in Eq. (11):

$$Z_e = \begin{cases} Z(q_r) + Z(q_{s_dry}) + Z(q_h), & T_b < 0^\circ\text{C} \\ Z(q_r) + Z(q_{s_wet}) + Z(q_h), & T_b > 0^\circ\text{C}, \end{cases} \quad (11)$$

where $Z(q_r)$ is computed as in Eq. (5), while $Z(q_{s_dry})$ is computed as in Eq. (6), $Z(q_{s_wet})$ is computed as in Eq. (7), and finally $Z(q_h)$ is computed as in Eq. (8). Finally, the last step needed is the conversion of the equivalent reflectivity Z_e (dBZ) [$\tilde{Z}$, Eq. (9)]. In this way the model reflectivity is computed using the different microphysics species and subsequently compared to the observed one for the innovation vector calculation. Thus, the moisture and hydrometeors partitioning is done similar to the indirect method, but the operation is performed on the models' variables and not on the observed reflectivity [the indirect method uses an inverse form of Eqs. (5)–(8) to obtain hydrometeors species from observed reflectivity]. Therefore, the modified operator allows obtaining a direct data assimilation of reflectivity that takes into account all the hydrometeors (snow, hail/graupel) and not only the rainwater like the default direct operator does.

2) MODEL SETUP, EXPERIMENTS DESIGN, AND VERIFICATION

The WRF Model setup is based on the previous results for the V-shaped back-building MCSs that occurred in Liguria (Fiori et al. 2014, 2017; Lagasio et al. 2017).

Two nested domains (Fig. 1) with respectively 5-km (179×200 grid points) and 1-km (475×475 grid points) grid spacing, covering the upper and lower limits of the cloud-permitting range (Arakawa 2004), have been used for all the experiments.

The number of vertical levels is set to 50, with a higher density in the first 1000-m layer of the atmosphere. Both grid spacings (5 and 1 km) allow solving explicitly many convective processes (Kain et al. 2006, 2008), so an explicit treatment of convection is chosen. Given that the observed presence of solid hydrometeors in the atmosphere is due to the strong convection that characterizes all these events (Fiori et al. 2017; Lagasio et al. 2017), the microphysics parameterization corresponding to the WRF single-moment six-class scheme (WSM6; Hong and Lim 2006) is applied. The Yonsei University (YSU) scheme is used for the PBL because it accurately simulates deeper vertical mixing in buoyancy-driven PBLs with shallower mixing in strong-wind regimes with respect to the older MRF scheme (Hong et al. 2006). Short-wave and longwave parameterization are taken into account through the Rapid Radiative Transfer Model for GCMs (RRTMG; Iacono et al. 2008). Furthermore, the land surface is parameterized by the Rapid Update Cycle (RUC) land surface model (Benjamin et al. 2004), which is a multilevel soil model with higher resolution in the top part of soil domain (0, 5, 20, 40, 160, and 300 cm as in the default configuration).

Regarding the data assimilation, a sensitivity analysis has been performed for each test case using all the available reflectivity operators plus the modified direct operator [refer to section 3b(1)], both stand-alone and coupled with surface observations data. Table 1 describes the nine sensitivity experiments.

The run with the modified direct reflectivity operator is implemented only for the assimilation of reflectivity alone (Table 1) due to the fact that, on one side, the main aim is to compare its behavior with the other operators and, on the other one, because the best results for the other operators have been achieved with the assimilation of reflectivity alone. The 3DVAR is applied every 6 h in a cycling mode, and for each cycle the forecast lead time is the end of the day of interest (Fig. 2). Referring to section 3b(1), the **B** matrix plays a fundamental role for the good quality of data assimilation results. In this work the Control Variable option 5 (CV5) of the WRFDA package is used in this work (for more details refer to WRF Users Group 2016) for the **B** matrix calculation using the NMC method (Wang et al. 2014) over the entire month of October 2013 with a 24-h lead time for the forecasts starting at 0000 UTC and a 12-h lead time for the ones initialized at 1200 UTC of the same day. The differences between the two forecasts ($t + 24$ and $t + 12$) valid

TABLE 1. List of simulations that compose the sensitivity for each test case and the corresponding abbreviation that will be used in the text.

Run abbreviation	Run description
Open Loop	Run without data assimilation
ALL-direct	Assimilation of reflectivity and surface observations with direct method
ALL-indirect	Assimilation of reflectivity and surface observations with indirect method
ALL-indirect-rqv	Assimilation of reflectivity and surface observations with indirect method adding the in-cloud humidity estimation
Radar-direct	Assimilation of reflectivity only with direct method
Radar-direct-modif	Assimilation of reflectivity only with direct method using the modified reflectivity operator
Radar-indirect	Assimilation of reflectivity only with indirect method
Radar-indirect-rqv	Assimilation of reflectivity only with indirect method adding the in-cloud humidity estimation
Stations-only	Assimilation of surface observations stations only

for the same reference time are used to calculate the domain's specific error statistics.

When coming to the test cases, the initialization time depends on the timing of each event (see section 2); for the two episodes of 2011 the runs are initialized at 1200 UTC of the day before (24 October and 3 November) while the 2014 case is initialized at 0000 UTC of the same day (9 October). Initial and boundary conditions for all the simulations are provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS) with a spatial resolution of $0.125^\circ \times 0.125^\circ$, and the boundary conditions are updated every 3 h. In terms of operational framework, IFS analyses are available 6 and 5 h after the initialization time in the cases that happened during October and November, respectively. The main advantage of maintaining the same IFS analysis over the entire forecasting period, while updating the model with a 6-h cycling 3DVAR of observations, is that the corresponding forecast is available 4 or 5 h in advance with respect to the forecast run every time, with the most recent IFS analysis during the day. Consequently, the hydrometeorological chain can provide updated forecasts during the entire event in a nowcasting framework.

3) METEOROLOGICAL OUTPUT VALIDATION

To validate all the modeling experiments and identify the most convenient WRF-3DVAR setup for each case study, the Method for Object-Based Evaluation (MODE; Davis et al. 2006a,b) is applied by comparing the quantitative

FIG. 2. Six-hour cycling 3DVAR assimilation scheme for the selected test cases: (a) Genoa 2014; (b) Cinqueterre 2011 and Genoa 2011.

precipitation forecast (QPF) of WRF with the QPE offered by rain gauges. MODE identifies precipitation structures in both forecast and observed fields and performs a spatial evaluation of the model capability of reproducing the identified observed objects. The evaluation of MODE is summarized as output indices such as centroid distance, angle difference, area ratio, symmetric difference, and percentile intensity (in this work above the 90th percentile threshold); for a complete description of the indices used for this part of the validation, refer to [Table 2](#).

MODE also provides some classical statistical scores retrieved from contingency tables. In this work each meteorological simulation has been validated using the following (further description with formulation of each index in the [appendix](#)):

- frequency bias (FBIAS): measures the ratio of the frequency of forecast events to the frequency of observed events, indicates whether the forecast system has a tendency to underforecast ($\text{FBIAS} < 1$) or overforecast ($\text{FBIAS} > 1$) events. FBIAS does not measure how well the forecast corresponds to the observations, only measures relative frequencies;
- probability of detection yes (PODY): the fraction of events that were correctly forecasted to occur (range: 0–1, perfect value = 1);
- false alarm ratio (FAR): the proportion of forecasts of the event occurring for which the event did not occur (range: 0–1, perfect value = 0);
- critical success index (CSI): the ratio of the number of times the event was correctly forecasted to occur to

TABLE 2. MODE indices description (in part courtesy of [Lagasio et al. 2017](#)).

Index	Description
CEN DIST	Centroid distance: provides a quantitative sense of spatial displacement of forecast (best score 0)
ANG DIFF	For noncircular object gives measure of orientation errors (best score 0)
AREA RATIO	Provides an objective measure of whether there is an over or underprediction of areal extent of forecast (best score 1)
INT AREA	Area of intersection between corresponding objects (best value equal to observed area)
UNION AREA	Total area of two corresponding objects summed together (best value equal to observed area)
SYMM DIFF	Provides a good summary statistic for how well forecast and observed objects match (best value small)
P90 INT	Provides objective measures of near-peak (90th percentile) intensities found in objects (best score 1)

TABLE 3. Characteristics of the considered basins and of the spatial and time model implementation. The values of two skill scores calculated in the validation period are also shown.

Basin	Station	Area (km ²)	Spatial resolution Dx (m)	Time resolution (min)	Period of validation	NS	REHF
Bisagno	Passerella Firpo	92	480	60	1 Jan 2013–31 Dec 2014	0.26	0.16
Vara	Nasceto	202	480	60	1 Jan 2013–31 Dec 2014	0.83	0.1
Magra	Calamazza	960	480	60	1 Jan 2013–31 Dec 2014	0.81	0.14

the number of times it was either forecasted or occurred (range: 0–1, perfect value = 1);

- Hanssen and Kuipers discriminant (HK): measures the ability of the forecast to discriminate between (or correctly classify) events and nonevents (range: from -1 to 1, perfect value = 1);
- Heidke skill score (HSS): a skill score based on accuracy, where the accuracy is corrected by the number of correct forecasts that would be expected by chance (range: from $-\infty$ to 1, perfect value = 1).

The main goal of this meteorological validation is to select the best meteorological forecast out of the whole set of the sensitivity experiments to be downscaled by means of RainFARM and eventually fed to the Continuum hydrological model. The most reliable meteorological forecast was selected as in [Lagasio et al.](#)

(2017): all the indices and statistical scores described above are calculated for each sensitivity experiment, then the times in which a simulation has been the best for each score is counted. Finally, the run ranking as the best for the higher number of times will be further used to feed the rest of the chain in comparison to the Open Loop forecast.

c. The hydrological framework: RainFARM and Continuum

The hydrological framework is constituted by a rainfall downscaling model and a hydrological model both widely described in previous publications ([Silvestro et al. 2011](#); [Laiolo et al. 2014](#); [Silvestro et al. 2016](#)). Continuum is a continuous and distributed hydrological model, developed by [Silvestro et al. \(2013, 2015\)](#), while

FIG. 3. Comparison among the 25 Oct 2011 24-h QPE from (a) rain gauges interpolation, (b) the Open Loop QPF, and the QPF of each member of the sensitivity experiments: (c) ALL-direct, (d) ALL-indirect, (e) ALL-indirect-rqv, (f) Radar-direct, (g) Radar-direct-modif, (h) Radar-indirect, (i) Radar-indirect-rqv, and (j) Stations-only.

FIG. 4. Representation of the objects obtained through the MODE application for the (left) 72- and (right) 96-mm threshold for the 25 Oct 2011, comparing in each panel the object obtained from the QPE (solid red) with the QPFs (blue contour) for each simulation: (a),(b) Open Loop, (c),(d) ALL-direct, (e),(f) ALL-indirect, (g),(h) ALL-indirect-rqv, (i),(j) Radar-direct, (k),(l) Radar-direct-modif, (m),(n) Radar-indirect, (o),(p) Radar-indirect-rqv, and (q),(r) Stations-only.

TABLE 5. Summary of the sensitivity performances; the times in which each forecast has the best result for each score is counted for each threshold and summarized in a total count that is used to find the best simulation (shown in bold font) for the Cinque Terre 2011 event.

72 mm	96 mm	Total	
1	4	5	ALL-direct
0	1	1	ALL-indirect
1	1	2	ALL-indirect-rqv
0	0	0	Radar-direct
8	6	14	Radar-direct-modif
2	0	2	Radar-indirect
0	0	0	Radar-indirect-rqv
0	0	0	Stations-only
2	1	3	Open Loop

the configuration adopted in this work is described in Davolio et al. (2017) together with its calibration, particularly focused on floods and flow peak events. Table 3 reports the main characteristics of the implementation for three basins affected by the considered events and where streamflow observations were available. The table also reports the values of skill scores for the validation period (Davolio et al. 2017):

- Nash–Sutcliffe (NS) coefficient (Nash and Sutcliffe 1970):

$$NS = 1 - \frac{\sum_{t=1}^{T_{\max}} [Q_m(t) - Q_0(t)]^2}{\sum_{t=1}^{T_{\max}} [Q_0(t) - \bar{Q}_0]^2}, \quad (12)$$

where $Q_m(t)$ and $Q_o(t)$ are the modeled and observed streamflows at time t , $\bar{Q}_0$ is the mean observed streamflow, and $T_{\max}$ is the number of time steps of the entire simulation.

- Relative error of high flows (REHF):

$$REHF = \frac{1}{N \text{ values}} \left[\sum_{i=1}^{N \text{ values}} \frac{|Q_m(t) - Q_0(t)|}{Q_0(t)} \right]_{Q > Q_{th}}, \quad (13)$$

where Q_{th} is chosen as the 99th percentile of the observed hydrograph along the considered period and N values is the number of time steps where $Q > Q_{th}$.

The state variables of the hydrological model at the beginning of each of the considered events were evaluated doing a seamless run from 1 January 2011 until 31 December 2014 feeding the model with gauges (rainfall, air temperature, solar radiation, air relative

FIG. 5. Comparison among the 4 Nov 2011 24-h OPE from (a) rain gauges interpolation, (b) the Open Loop QPF, and the QPF of each member of the sensitivity experiments: (c) ALL-direct, (d) ALL-indirect, (e) ALL-indirect-rqv, (f) Radar-direct, (g) Radar-direct-modif, (h) Radar-indirect, (i) Radar-indirect-rqv, and (j) Stations-only. The black bold contour highlights the Bisagno catchment subjected to the flood.

FIG. 6. Representation of the objects obtained through the MODE application for the (left) 72- and (right) 96-mm threshold for the 25 Oct 2011, comparing in each panel the object obtained from the QPE (solid red) with the QPFs (blue contour) for each simulation: (a),(b) Open Loop, (c),(d) ALL-direct, (e),(f) ALL-indirect, (g),(h) ALL-indirect-rqv, (i),(j) Radar-direct, (k),(l) Radar-direct-modif, (m),(n) Radar-indirect, (o),(p) Radar-indirect-rqv, and (q),(r) Stations-only.

TABLE 6. Spatial and statistical indices calculated through MODE to evaluate the sensitivity forecasts with respect to the Open Loop run for the Genoa extreme event of 4 Nov 2011. The best performance for each score is highlighted in bold.

Run	CENTROID		ANGLE	AREA	SYMMETRIC	INTERSECTION	UNION	P90	FBIAS	PODY	FAR	CSI	HK	HSS
	DIST (best small)	DIFF (best small)	RATIO (best = 1)	RATIO (best = 1)	DIFF (best small)	AREA (best big)	AREA (best small)	RATIO (best = 1)	(best = 1)	(best = 1)	(best = 0)	(best = 1)	(best = 1)	(best = 1)
72 mm														
ALL-direct	2.94	20.99	0.65		740	628	1368	0.72	0.82	0.50	0.39	0.38	0.47	0.50
ALL-indirect	13.17	3.58	0.87		1141	655	1796	0.74	0.94	0.49	0.48	0.34	0.44	0.45
ALL-indirect-rqv	13.84	9.42	0.97		1320	628	1948	0.73	0.96	0.47	0.51	0.32	0.42	0.42
Radar-direct	4.31	16.44	0.76		617	757	1374	0.70	0.94	0.60	0.36	0.45	0.56	0.58
Radar-direct-modif	7.52	45.95	0.97		660	858	1518	0.67	1.06	0.68	0.36	0.49	0.63	0.62
Radar-indirect	12.84	9.20	0.96		1222	670	1892	0.73	0.94	0.50	0.46	0.35	0.45	0.47
Radar-indirect-rqv	13.95	10.36	1.00		1316	650	1966	0.74	0.99	0.49	0.51	0.33	0.43	0.43
Stations-only	11.77	44.03	0.86		1084	580	1664	0.76	0.94	0.49	0.48	0.33	0.43	0.44
Open Loop	12.64	3.87	0.92		1162	675	1837	0.72	0.98	0.51	0.48	0.35	0.45	0.46
96 mm														
ALL-direct	2.98	11.12	0.53		529	355	884	0.71	0.61	0.44	0.28	0.37	0.43	0.52
ALL-indirect	9.13	9.31	0.67		701	327	1028	0.82	0.82	0.40	0.51	0.28	0.38	0.41
ALL-indirect-rqv	2.16	9.40	0.46		577	298	875	0.94	0.83	0.38	0.54	0.26	0.35	0.38
Radar-direct	9.21	9.15	0.68		430	464	894	0.69	0.86	0.58	0.32	0.46	0.56	0.60
Radar-direct-modif	2.16	13.11	0.93		384	587	971	0.66	1.02	0.72	0.29	0.56	0.70	0.70
Radar-indirect	8.98	10.11	0.69		679	340	1019	0.83	0.81	0.42	0.48	0.30	0.39	0.43
Radar-indirect-rqv	10.49	10.77	0.69		740	313	1053	0.86	0.86	0.39	0.55	0.26	0.36	0.38
Stations-only	9.37	10.66	0.67		680	335	1015	0.84	0.81	0.41	0.49	0.30	0.39	0.42
Open Loop	9.32	8.07	0.70		679	347	1026	0.80	0.83	0.43	0.48	0.31	0.40	0.44

TABLE 7. Summary of the sensitivity performances: the times in which each forecast has the best result for each score is counted for each threshold and summarized in a total count that is used to find the best simulation (shown in bold font) for the Genoa 2011 event.

72 mm	96 mm	Total	
2	1	3	ALL-direct
1	0	1	ALL-indirect
0	2	2	ALL-indirect-rqv
1	0	1	Radar-direct
6	9	15	Radar-direct-modif
0	0	0	Radar-indirect
2	0	2	Radar-indirect-rqv
1	0	1	Stations-only
0	1	1	Open Loop

humidity, wind velocity) interpolated with a simple kriging method.

The rainfall downscaling model (RainFARM; Rebora et al. 2006a,b) has been used in many applications (Davolio et al. 2015; Silvestro and Rebora 2014). Its

workflow follows, in brief, the following steps: (i) the rainfall field predicted by the NWP model is aggregated at spatial and time scales (hereafter L_r and t_r), which are considered averagely reliable; (ii) the aggregated rainfall field is then downscaled to spatial and time scales, which are generally equal or finer than those of NWP; and (iii) a stochastic component allows us to produce an ensemble of equiprobable rainfall scenarios. Using these equiprobable rainfall scenarios as input to the hydrological model, an ensemble of equiprobable streamflow scenarios can be obtained.

RainFARM has two parameters estimated directly from the power spectrum of the predicted rainfall field, so that they can vary for each event. The L_r and t_r values are assumed as in Davolio et al. (2015, 2017): $L_r = 15$ km and $t_r = 6$ h. Each rainfall scenario has a fine spatio-temporal coherent structure, which maintains the characteristics of the NWP rainfall field in terms of (i) volume of precipitation and (ii) spatial and time structure at the scales L_r and t_r .

FIG. 7. Comparison among the 9 Oct 2014 24-h QPE from (a) Settepani radar, (b) the Open Loop QPF, and the QPF of each member of the sensitivity experiments: (c) ALL-direct, (d) ALL-indirect, (e) ALL-indirect-rqv, (f) Radar-direct, (g) Radar-direct-modif, (h) Radar-indirect, (i) Radar-indirect-rqv, and (j) Stations-only. The black bold contour highlights the Bisagno catchment hit subjected to the flood.

Since in this study a high-resolution NWP is dealt with, RainFARM has the main role to manage the uncertainty in spatial and time structure of the original QPF, consequently the final spatial and time resolution of the rainfall field is the same of the NWP, $Dx = 1$ km, $Dt = 1$ h.

4. Results and validation

The first part of this section compares the results of the 3DVAR operators for reflectivity/in situ observations and assesses their performances with respect to the Open Loop simulation. The overarching goal is to identify the 3DVAR operator that allows obtaining the greater improvement in terms of forecasts skills with respect to the Open Loop run. The second part focuses on the evaluation of the impact on the hydrological forecast accuracy of the best 3DVAR-driven meteorological simulation for each case study.

a. Meteorological evaluation of the 3DVAR sensitivity

For each event, the 24-h QPFs provided by the Open Loop run and the 3DVAR operators in rapid refresh mode are compared with the available 24-h accumulation QPE of rain gauges and Settepani radar (only for the 2014 case). However, it is important to mention that the Genoa 2014 flash flood had a peculiar spatiotemporal evolution as it was characterized by two distinct phases: the first in the morning and the second in the evening. It is noteworthy to remember that the second phase of the event was completely missed by the operational models and also by several WRF hindcast simulations in Open Loop mode (Fiori et al. 2017; Lagasio et al. 2017). Thus, for this case study, in addition to the 24-h QPFs, the 12-h QPFs of the morning (0000–1200 UTC) and the afternoon (1200–2400 UTC) have been evaluated also. Chronologically, the first event to happen is the Cinque Terre flash flood on 25 October 2011, when up to 470 mm of rainfall were observed in 24 h (Fig. 3a). Figure 3 shows a general good agreement of all the simulations (as well as the Open Loop, in Fig. 3b) with respect to the observed rain map, despite the forecasts misplaced and underestimated the precipitation peak recorded.

To offer a quantitative estimation of the sensitivity experiments performances, a spatial and statistical evaluation is performed through the use of MODE [see section 3b(3)]: two rainfall accumulations thresholds of 72 and 96 mm, respectively, over 24 h are adopted to compare the QPE and the QPFs. The use of the rainfall thresholds allows us to isolate the intense part of the precipitation pattern and to obtain a set of comparable

FIG. 8. Representation of the objects obtained through the MODE application for the (left) 72- and (right) 96-mm threshold for the 0000–1200 UTC cumulated rainfall of the 9 Oct 2014 event, comparing in each panel the object obtained from the QPE (solid red) with the QPFs (blue contour) for each simulation: (a),(b) Open Loop, (c),(d) ALL-direct, (e),(f) ALL-indirect, (g),(h) ALL-indirect-rqv, (i),(j) Radar-direct, (k),(l) Radar-direct-modif, (m),(n) Radar-indirect, (o),(p) Radar-indirect-rqv, and (q),(r) Stations-only.

TABLE 8. Spatial and statistical indices calculated through MODE to evaluate the sensitivity forecasts with respect to the Open Loop run for the first phase (0000–1200 UTC) of the Genoa extreme event of 9 Oct 2014. The best performance for each score is highlighted in bold. It is worth noting that some scores can appear equal due to the approximation made to simplify reading the table.

Run	CENTROID		ANGLE	AREA	SYMMETRIC	INTERSECTION		UNION	P90	FBIAS	PODY	FAR	CSI	HK	HSS
	DIST (best small)	DIFF (best small)	RATIO (best = 1)	(best = 1)	(best small)	AREA (best big)	AREA (best small)	RATIO (best = 1)	(best = 1)	(best = 1)	(best = 1)	(best = 0)	(best = 1)	(best = 1)	(best = 1)
72 mm															
ALL-direct	15.80		7.44	0.81	1090	452	1542	1.01	1.18	0.48	0.59	0.29	0.48	0.44	0.44
ALL-indirect	8.96		4.00	0.53	723	350	1073	1.14	0.53	0.37	0.29	0.32	0.37	0.37	0.49
ALL-indirect-rqv	11.05		7.24	0.93	931	478	1409	0.94	1.05	0.51	0.51	0.33	0.51	0.51	0.50
Radar-direct	7.31		2.17	0.86	808	442	1250	0.84	0.84	0.47	0.44	0.35	0.47	0.51	
Radar-direct-modif	10.60		4.22	0.52	769	292	1061	1.08	0.49	0.31	0.37	0.26	0.31	0.42	
Radar-indirect	6.73		1.97	0.60	748	351	1099	1.07	0.58	0.38	0.35	0.31	0.37	0.47	
Radar-indirect-rqv	7.90		5.48	0.82	1002	505	1507	0.72	1.26	0.54	0.57	0.32	0.54	0.48	
Stations-only	7.56		1.92	0.48	695	315	1010	1.13	0.50	0.34	0.32	0.29	0.34	0.45	
Open Loop	7.70		0.36	0.52	781	298	1079	0.92	0.50	0.32	0.37	0.27	0.32	0.42	
96 mm															
ALL-direct	14.12		23.71	0.70	770	229	999	0.95	1.40	0.44	0.68	0.23	0.44	0.37	
ALL-indirect	9.35		19.46	0.62	474	172	646	1.08	0.61	0.33	0.45	0.26	0.33	0.41	
ALL-indirect-rqv	11.77		25.05	0.78	745	206	951	0.89	1.26	0.40	0.68	0.21	0.40	0.35	
Radar-direct	7.96		21.57	0.83	552	186	738	0.92	0.81	0.36	0.56	0.25	0.36	0.40	
Radar-direct-modif	11.10		21.38	0.55	516	135	651	1.01	0.54	0.26	0.52	0.20	0.26	0.34	
Radar-indirect	7.95		15.74	0.66	496	172	668	1.02	0.65	0.33	0.49	0.25	0.33	0.40	
Radar-indirect-rqv	5.60		22.49	0.99	579	219	798	0.68	1.02	0.43	0.58	0.27	0.42	0.42	
Stations-only	8.22		18.73	0.51	474	145	619	1.13	0.52	0.28	0.46	0.23	0.28	0.37	
Open Loop	6.08		7.36	0.40	532	89	621	0.98	0.45	0.17	0.62	0.13	0.17	0.24

FIG. 9. Comparison among the 9 Oct 2014 12-h (0000–1200 UTC) QPE from (a) Settepani radar, (b) the Open Loop QPF, and the QPF of each member of the sensitivity experiments: (c) ALL-direct, (d) ALL-indirect, (e) ALL-indirect-rqv, (f) Radar-direct, (g) Radar-direct-modif, (h) Radar-indirect, (i) Radar-indirect-rqv, and (j) Stations-only. The black bold contour highlights the Bisagno catchment hit subjected to the flood.

objects for each threshold, one in the observation field and the second in each forecast field (Fig. 4).

The objects' overlapping (Fig. 4) confirms the general agreement between the QPFs and the QPE but also highlights the misplacing of the most intense precipitation core, located too close to the coastline by some experiments (and more prominently for the 96-mm threshold, second column in Fig. 4), such as the Open Loop in Fig. 4b, the run with indirect methods in Figs. 4d, 4f, 4h, 4n, and 4p, and with stations-only data assimilation in Fig. 4r. These first considerations are confirmed by the statistical and spatial scores calculated by MODE and reported in Table 4.

The MODE results confirm the good quality of the Open Loop run that has a PODY of about 75% for the 72-mm threshold and reveal that in general the 3DVAR best performance is obtained with the modified operator for the direct data assimilation (Radar-direct-modif)

achieving the best result in 14 out of 26 scores (summing the scores for the 72- and 96-mm thresholds, Table 5) followed by the ALL-direct simulation with best values in six scores. In this case the change in the direct reflectivity operator allowed obtaining a rainfall pattern with better shape-related parameters (CENT DIST, ANGLE DIFF, SYMM DIFF), a better agreement between forecasted rainfall points and the observed ones (CSI, HK, HSS) and a good FAR.

The second event here considered is the Genoa flash flood of 4 November 2011 that recorded about 450 mm of precipitation in the central hours (0900–1500 UTC) of the day (Fig. 5a) on the Bisagno catchment (black bold contour in all panels of Fig. 5).

In this case the Open Loop run (Fig. 5b) shows a good agreement in terms of rainfall peaks but the most intense core of the precipitation pattern is again misplaced: the higher QPF value within the Bisagno catchment is less

FIG. 10. Comparison among the 9 Oct 2014 12-h (1200–2400 UTC) QPE from (a) Settepani radar, (b) the Open Loop QPF, and the QPF of each member of the sensitivity experiments: (c) ALL-direct, (d) ALL-indirect, (e) ALL-indirect-rqv, (f) Radar-direct, (g) Radar-direct-modif, (h) Radar-indirect, (i) Radar-indirect-rqv, and (j) Stations-only. The black bold contour highlights the Bisagno catchment hit subjected to the flood.

than 100 mm (dark yellow color). This behavior does not change with the stations-only sensitivity experiment (Fig. 5j) or using the indirect reflectivity operator (Figs. 5d,h) also in combination with the in-cloud humidity operator (Figs. 5e,i). Conversely, the use of the direct assimilation of radar reflectivity alone (Figs. 5f,g) improves the rainfall pattern as the areal averaged QPF over the Bisagno catchment increases significantly of about 150–200 mm. These considerations are confirmed by the QPE and QPFs comparison using MODE with both 72- and 96-mm thresholds (Fig. 6): indeed, the use of the direct data assimilation with radar alone allows a better localization (Figs. 6i–l) of the precipitation maxima to be obtained.

Furthermore, MODE statistical scores show that the direct assimilation with the modified reflectivity operator significantly enhanced the PODY and HSS, still ranking amongst the best performing experiments in

terms of FAR (Table 6). The improvement is visible also in spatial scores such as CENT DIST (especially for the 96-mm threshold), AREA RATIO, and INT AREA (Table 6). The summary in Table 7 upholds the first qualitative evaluation as the Radar-direct-modif achieved the best results on 15 out of 26 scores, while the others simulations equally shared the remainders.

The third test case regards the Genoa 2014 flash flood, when again more than 400 mm of precipitation was recorded in a day. The QPE from Settepani radar is available for this event, and it is then used to gain a deeper understanding of the rainfall patterns over the sea (Fiori et al. 2017). From the rainfall daily accumulation it is possible to infer the significant underestimation of the Open Loop run (Fig. 7b).

All the tested 3DVAR operators improve the precipitation volumes, despite that in some cases the results are affected by an inland QPF overestimation downshear of

the Apennines (see upper-right corner of Figs. 7c,e,f,i). When the two distinct phases of the event are addressed separately, both the object comparison (Fig. 8) and MODE scores showed (Table 8) a general good agreement in terms of QPF patterns and volumes for the 0000–1200 UTC period (Fig. 9). Yet, a clearly better run does not stand out: for example, the Radar-indirect-rqv has a good POD but the ALL-indirect has the best FAR. Furthermore, Fig. 8 reveals a slight underestimation in terms of spatial extent and orientation (major axis of the structure) by the majority of 3DVAR sensitivity experiments. This behavior is improved by the Radar-indirect-rqv simulation (Figs. 8o,p), which has the best CENT DIST for both thresholds and the best AREA RATIO for the 96-mm threshold (Table 8).

In the second phase of the event (1200–2400 UTC), the different simulations show rather erratic performances (Fig. 10). The Open Loop run (Fig. 10b) does not reproduce the intensity or the location of the event, as inside the Bisagno catchment the rainfall peak is about 70 mm with respect to the 250 mm observed. Instead, in 4 of 8 data assimilation options, there is a significant improvement in terms of QPF performances (Figs. 10c,e,f,i) with even a good localization of the most intense part of event (Figs. 10c,i) and peaks of more than 200 mm within the Bisagno catchment. Yet, an overestimation downshear of the Apennines is more persistent in the ALL-direct and Radar-direct runs than in the ALL-indirect-rqv and Radar-indirect-rqv (Figs. 10c,e,f,i).

The object comparison (Fig. 11) reveals that, despite a significant improvement in QPF forecast for most of the 3DVAR sensitivity experiments of the inland portion of the rainfall, a reliable reproduction of the event over the sea area is still missed out upon, as the pattern shifted more landwards (Figs. 11c,d,g,h,i,j,o,p). Nevertheless, 3DVAR outperforms the Open Loop run that in turn does not exceed or even reach the 96-mm threshold in any point of the domain (Fig. 11b). The reflectivity assimilation allowed us to maintain a significant amount of precipitation on the coastline when also considering the 96-mm threshold, while the runs using the in-cloud humidity estimation (ALL-indirect-rqv in Figs. 11g,h and Radar-indirect-rqv in Figs. 11o,p) largely decrease the overestimation downshear of the Apennines.

The qualitative considerations provided by the object maps comparison in Fig. 11 are quantitatively confirmed by the scores computed by MODE (Table 9), where the Radar-indirect-rqv outperforms the other simulations, especially in terms of statistical scores (POD, FAR, CSI, HK, HSS) for both thresholds (Table 9) with a good

FIG. 11. Representation of the objects obtained through the MODE application for the (left) 72- and (right) 96-mm threshold for the 1200–2400 UTC cumulated rainfall of the 9 Oct 2014 event, comparing in each panel the object obtained from the QPE (solid red) with the QPFs (blue contour) for each simulation: (a),(b) Open Loop, (c),(d) ALL-direct, (e),(f) ALL-indirect, (g),(h) ALL-indirect-rqv, (i),(j) Radar-direct, (k),(l) Radar-direct-modif, (m),(n) Radar-indirect, (o),(p) Radar-indirect-rqv, and (q),(r) Stations-only.

TABLE 9. Spatial and statistical indices calculated through MODE to evaluate the sensitivity forecasts with respect to the Open Loop run for the second phase (1200–2400 UTC) of the Genoa extreme event of 9 Oct 2014. The best performance for each score is highlighted in bold. It is worth noting that some scores can appear equal due to the approximation made to simplify reading the table.

Run	CENTROID		ANGLE	AREA	SYMMETRIC	INTERSECTION	UNION	P90	FBIAS	PODY	FAR	CSI	HK	HSS
	DIST (best small)		DIFF (best small)	RATIO (best = 1)	DIFF (best small)	AREA (best big)	AREA (best small)	RATIO (best = 1)	(best = 1)	(best = 1)	(best = 0)	(best = 1)	(best = 1)	(best = 1)
72 mm														
ALL-direct	44.98		11.21	0.58	2981	528	3509	1.17	1.75	0.36	0.80	0.15	0.34	0.25
ALL-indirect	19.77		8.17	0.29	1734	91	1825	0.64	0.29	0.06		0.05	0.06	0.09
ALL-indirect-rqv	32.99		19.65	0.80	2380	482	2862	0.88	1.26	0.33	0.74	0.17	0.32	0.28
Radar-direct	49.34		7.41	0.49	3438	523	3961	1.16	2.07	0.35	0.83	0.13	0.34	0.22
Radar-direct-modif	21.82		2.06	0.45	1792	180	1972	0.71	0.45	0.12	0.73	0.09	0.12	0.16
Radar-indirect	22.32		3.83	0.32	1796	84	1880	0.58	0.32	0.06	0.83	0.04	0.05	0.08
Radar-indirect-rqv	32.84		14.82	0.90	2036	546	2582	1.07	1.12	0.37	0.67	0.21	0.36	0.34
Stations-only	20.91		4.11	0.33	1734	117	1851	0.55	0.33	0.08	0.76	0.06	0.08	0.11
Open Loop	16.65		8.47	0.02	1456	30	1486	0.39	0.18	0.02	0.89	0.02	0.02	0.03
96mm														
ALL-direct	43.05		20.39	0.62	2150	393	2543	1.21	1.59	0.35	0.78	0.15	0.34	0.26
ALL-indirect	18.89		2.43	0.20	1297	31	1328	0.67	0.20	0.03	0.87	0.02	0.03	0.04
ALL-indirect-rqv	28.28		22.63	0.96	1589	355	1944	0.89	1.03	0.31	0.70	0.18	0.31	0.30
Radar-direct	45.46		17.77	0.54	2477	372	2849	1.26	1.87	0.33	0.83	0.13	0.32	0.22
Radar-direct-modif	23.21		6.61	0.36	1390	72	1462	0.71	0.36	0.06	0.82	0.05	0.06	0.09
Radar-indirect	21.17		7.29	0.20	1330	14	1344	0.61	0.20	0.01	0.94	0.01	0.01	0.02
Radar-indirect-rqv	24.02		10.49	0.84	1285	394	1679	1.17	0.85	0.35	0.59	0.23	0.34	0.37
Stations-only	19.98		15.49	0.21	1282	39	1321	0.57	0.20	0.03	0.83	0.03	0.03	0.05
Open Loop	—		—	—	—	—	—	—	0.08	0.00	1.00	0.00	0.00	0.00

TABLE 10. Spatial and statistical indices calculated through MODE to evaluate the sensitivity forecasts with respect to the Open Loop run for the daily accumulation (0000–2400 UTC) of the Genoa extreme event of 9 Oct 2014. The best performance for each score is highlighted in bold. It is worth noting that some scores can appear equal due to the approximation made to simplify reading the table.

Run	CENTROID DIST (best small)	ANGLE DIFF (best small)	AREA RATIO (best = 1)	SYMMETRIC DIFF (best small)	INTERSECTION AREA (best big)	UNION AREA (best small)	P90 RATIO (best = 1)	FBIAS (best = 1)	PODY (best = 1)	FAR (best = 0)	CSI (best = 1)	HK (best = 1)	HSS (best = 1)
						72 mm							
ALL-direct	37.85	4.32	0.52	3699	1252	4951	0.86	1.94	0.59	0.70	0.25	0.57	0.39
ALL-indirect	14.21	15.40	0.51	1888	665	2553	0.84	0.51	0.31	0.39	0.26	0.31	0.41
ALL-indirect-rqv	28.72	7.51	0.70	2667	1250	3917	0.86	1.47	0.59	0.60	0.31	0.57	0.47
Radar-direct	41.01	3.00	0.44	4176	1407	5583	0.75	2.34	0.66	0.72	0.25	0.64	0.38
Radar-direct-modif	14.92	9.22	0.66	2159	684	2843	0.68	0.66	0.32	0.51	0.24	0.32	0.38
Radar-indirect	14.13	20.58	0.60	1838	781	2619	0.76	0.60	0.37	0.39	0.30	0.36	0.45
Radar-indirect-rqv	26.73	3.52	0.69	2512	1364	3876	0.86	1.49	0.64	0.57	0.34	0.63	0.50
Stations-only	15.42	14.84	0.57	2004	674	2678	0.72	0.57	0.32	0.45	0.25	0.31	0.40
Open Loop	15.09	20.18	0.64	1952	768	2720	0.53	0.64	0.36	0.43	0.28	0.36	0.43
						96mm							
ALL-direct	34.25	14.26	0.58	2802	944	3746	0.94	1.72	0.54	0.68	0.25	0.53	0.39
ALL-indirect	14.88	7.67	0.44	1720	385	2105	0.87	0.44	0.22	0.50	0.18	0.22	0.30
ALL-indirect-rqv	24.76	13.00	0.80	2023	936	2959	0.92	1.26	0.54	0.57	0.31	0.53	0.47
Radar-direct	34.88	12.86	0.55	2930	965	3895	0.84	1.84	0.56	0.70	0.24	0.54	0.38
Radar-direct-modif	14.80	2.08	0.61	1889	443	2332	0.69	0.60	0.25	0.58	0.19	0.25	0.31
Radar-indirect	14.56	16.90	0.51	1691	470	2161	0.78	0.51	0.27	0.47	0.22	0.27	0.35
Radar-indirect-rqv	22.21	6.62	0.84	1832	972	2804	0.95	1.20	0.56	0.53	0.34	0.55	0.50
Stations-only	15.13	10.31	0.48	1709	425	2134	0.73	0.49	0.24	0.50	0.20	0.24	0.32
Open Loop	15.13	16.57	0.48	1715	424	2139	0.53	0.48	0.24	0.49	0.20	0.24	0.32

FIG. 12. Representation of the objects obtained through the MODE application for the (left) 72- and (right) 96-mm threshold for the 0000–2400 UTC cumulated rainfall of the 9 Oct 2014 event, comparing in each panel the object obtained from the QPE (solid red) with the QPFs (blue contour) for each simulation: (a), (b) Open Loop, (c), (d) ALL-direct, (e), (f) ALL-indirect, (g), (h) ALL-indirect-rqv, (i), (j) Radar-direct, (k), (l) Radar-direct-modif, (m), (n) Radar-indirect, (o), (p) Radar-indirect-rqv, and (q), (r) Stations-only.

TABLE 11. Summary of the sensitivity performances; the times in which each forecast has the best result for each score is counted for each threshold and summarized in a total count (summing Tables 8–10) that is used to find the best simulation (shown in bold font) for the Genoa 2014 event.

72 mm	96 mm	Total	
0	2	2	ALL-direct
3	5	8	ALL-indirect
4	3	7	ALL-indirect-rqv
6	0	6	Radar-direct
2	2	4	Radar-direct-modif
5	3	8	Radar-indirect
14	19	33	Radar-indirect-rqv
1	4	5	Stations-only
4	1	5	Open Loop

overlap of the objects area (INTER AREA, Table 9). Furthermore, the reduced overestimation downshear of the Apennines and the best localization of the precipitation pattern are confirmed by a lower FAR (Table 9).

Overall, considering the 24-h cumulated rainfall (Fig. 7), the Radar-indirect-rqv is the best-performing experiment, particularly for the highest threshold (Table 10). Indeed, the objects comparison highlights that the Radar-indirect-rqv simulation better reproduces the event, both in terms of cumulated rainfall (Fig. 7) and precipitation pattern orientation (Figs. 12o,p) with respect to the Open Loop run (Figs. 12a,b).

The Radar-indirect-rqv experiment has the best POD for the 96-mm threshold, providing the best localization for the most intense precipitation core. Furthermore, despite the overestimation on the Apennines, the Radar-indirect-rqv experiment has acceptable FAR values with respect to the run using the direct assimilation with the standard reflectivity operator. Conversely, the Radar-direct-modified experiment is the only one removing the large downshear overestimation that affects the other runs adopting the direct method, that is, it shows one of the lowest FARs for the second phase and a better FBIAS (Table 9). This consideration holds when considering the whole 24-h cumulated rainfall (Table 10), even if the misplacing of the pattern in Fig. 12 is to be accounted for the poor results of PODY and HSS with respect to the run using the indirect method (Tables 9, 10).

From the summary of the scores calculated for all the considered cumulated rainfall (Table 11), the Radar-indirect-rqv stands out as the best for the Genoa 2014 case. If a time frame during the second phase of the event (2000 UTC) is considered, a deeper insight of the impact of the Radar-indirect-rqv data assimilation simulation with respect to the Open Loop run (Fig. 13) is gained.

FIG. 13. Comparison between the Open Loop–simulated structure with respect to the Radar-direct-rqv simulated structure at 2000 UTC. (left) The 3D simulated structure composed by rainwater (cyan), graupel (yellow), and snow (gray) microphysics species respectively for (a) Open Loop and (c) Radar-indirect-rqv simulations with the horizontal 10-m wind intensity for the Open Loop in (a) and the Radar-indirect-rqv in (c). The black line in (a) and (c) indicates the location of the vertical section of the two structures to investigate the reflectivity values in the mean of the convective structure in (b) Open Loop and (d) Radar-indirect-rqv.

The use of data assimilation actually provides an enhancement of the wind intensity (Fig. 13c), supporting a more evident convergence line and in turn a more intense convection (Fig. 13c). This results in the production of a widespread and more intense area of snow and graupel that was nearly absent in the Open Loop (Fig. 13a). Open Loop (Fig. 13b) has a shallower and more disorganized convective structure run while, through the use of the reflectivity data assimilation (Fig. 13d), a simulated deep moist and convective storm is in very good agreement with the observed one.

In conclusion, the best result for the two 2011 flash floods is achieved with the Radar-direct-modif taking into account also the ice species, which are known to be crucial in thermodynamics of several back-building MCS (Fiori et al. 2017; Lagasio et al. 2017). Interestingly, for the 2014 flood the Radar-direct-modif approach also exhibits better performances than standard direct assimilation. However, it is worth highlighting that for the 2014 flash flood the in-cloud humidity assimilation associated to the indirect method achieves the best

performance both in terms of cumulated rainfall and pattern location.

To better understand this difference in the results, the columnar contents of the different hydrometeors are computed for a temporal snapshot corresponding to the main phases of these events (0900 UTC for the 2011 cases and 2000 UTC for the 2014 flood). Only radar reflectivity data assimilation operators are hereafter considered. The two use cases of 2011 are characterized by quite low-lying 0°C isotherms (around 2000–2500 m). Thus, not surprisingly, for both the 25 October event (Fig. 14) and the 4 November event (Fig. 15), the Radar-direct-modif (Figs. 14c,h,m,r,w and Figs. 15c,h,m,r,w) simulations produce significant (around 8–10 mm) amounts of graupel columnar content (upshear) on the Tyrrhenian side of the Ligurian Apennines largely coinciding with large (around 4–5 mm) columnar rainwater content over the same areas. Then the corresponding mixed-phase clouds experience a seeder–feeder mechanism in which the graupel and its falling crystals act as condensation nuclei for cloud water generation via

FIG. 14. Columnar contents analysis for 0900 UTC 25 Oct 2011 of (first row) graupel (QG), (second row) ice (QI), (third row) snow (QS), (fourth row) rain (QR), and (fifth row) cloud water (QC). Comparison between the (a),(f),(k),(p),(u) Open Loop simulation and the results achieved with the different reflectivity operators: (b),(g),(l),(q),(v) Radar-direct, (c),(h),(m),(r),(w) Radar-direct-modif, (d),(i),(n),(s),(x) Radar-indirect, and (e),(j),(o),(t),(y) Radar-indirect-rqv. The magenta contour line indicates the position of the observed convective structure using the 3000-m reflectivity CAPPI above 20 dBZ at 0900 UTC. The black dot highlights the city most hit by the flash flood (Monterosso in this case).

heterogeneous nucleation, thus overall increasing precipitating efficiency. The direct modified takes better account of this phenomenon with respect to the Radar_direct simulation, which refers only to warm rain processes (Figs. 14b,g,l,q,v and Figs. 15b,g,l,q,v). The Radar-indirect (Figs. 14d,i,n,s,x and Figs. 15d,i,n,s,x) and Radar-indirect-rqv (Figs. 14e,j,o,t,y and Figs. 15e,j,o,t,y) are able to capture to some extent the same mechanism, but they locate the structure in the wrong position (more evident for the 25 October event than in the 4 November case).

Concerning the 2014 event (Fig. 16), the 0°C isotherm is at about 4000 m, thus resulting into a more “warm rain” case, at least in lower to middle troposphere, while the ice species are conversely located in the upper level of the convective anvil located downshear the Liguria Apennines. Consequently, the simulations using operators weighting more the liquid part of the structures (Radar_direct in Figs. 16b,g,l,q,v and Radar_indirect_rqv in Figs. 16e,j,o,t,y) are in a better position to capture the predominantly observed “warm rain” mechanisms. However, the Radar_direct simulation is still penalized by the aforementioned overestimation downshear that is mitigated by the use of the Radar_direct_modif (Figs. 16c,h,m,r,w).

In the next section these best runs will be evaluated in terms of hydrological impact improvement with respect to their associated Open Loop simulations.

b. Hydrological impact evaluation of data assimilation

To assess how the best performing 3DVAR configuration affects the hydrological prediction (RainFARM+Continuum), the results are presented by means of a box plot of the peak flows. For the 2011 event in Cinqueterre, the most affected basin was Magra (basin of 1686 km² crossing Toscana and Liguria regions): in this case, 3DVAR Radar-direct-modif operator does not enhance very much the streamflow prediction which is quite good also in the Open Loop configuration (Fig. 17b). On the contrary, in the Vara basin (Fig. 17d) the observed peak flow was not particularly severe, yet Open Loop configuration overestimated it. The data assimilation experiment helps in reducing the overestimation, especially in 0000 and 0600 UTC assimilation cycles. In Fig. 17 red dots represent the observed peak, while blue crosses display the peak obtained with the hydrological model fed with observations.

FIG. 15. Columnar contents analysis for 0900 UTC 4 Nov 2011 of (first row) graupel (QG), (second row) ice (QI), (third row) snow (QS), (fourth row) rain (QR), and (fifth row) cloud water (QC). Comparison between the (a),(f),(k),(p),(u) Open Loop simulation and the results achieved with the different reflectivity operators: (b),(g),(l),(q),(v) Radar-direct, (c),(h),(m),(r),(w) Radar-direct-modif, (d),(i),(n),(s),(x) Radar-indirect, and (e),(j),(o),(t),(y) Radar-indirect-rqv. The magenta contour line indicates the position of the observed convective structure using the 3000-m reflectivity CAPPI above 20 dBZ at 0900 UTC. The black dot highlights the city most hit by the flash flood (Genoa in this case).

Streamflow predictions for the Genoa 2011 event also benefit from 3DVAR application (viz., Radar-direct-modif; Fig. 17a), with the 75% interquartile increasing from about 180 to 220 m³ s⁻¹; similarly, the upper boundary of the distribution (whiskers) increases from 830 to more than 1000 m³ s⁻¹. In this case 3DVAR Radar-direct-modif operator cannot localize the intense rainfall core with high accuracy on the catchment, but the downscaling observed peaks are nonetheless included in the tail of the predicted peaks distribution through the application of the downscaling algorithm, a quite common circumstance when the basin targeted by the prediction is so small-sized (Siccardi et al. 2005).

Figure 17c shows the results for the Genoa 2014 event again on Bisagno basin: the streamflow forecast obtained with WRF in Open Loop configuration is compared with the ones obtained with the 3DVAR Radar-indirect-rqv experiment performed every 6 h. Black dots represent the observed peak while blue diamonds display the peak obtained with the hydrological model fed with observations. The 3DVAR Radar-indirect-rqv experiment effect is negligible at 0000 UTC while improving the prediction at 0600 UTC and then at 1200 UTC. These latter two seem to be particularly good results, also from an operational standpoint, since both observed and simulated

peak flow are inside the interquartile. The 1800 UTC DA improves results, but in this case we are really close to the observed peak, which occurred in the evening at 2200 UTC.

The Genoa 9 October 2014 event was very challenging in terms of predictability of the rainfall intensity, the second phase of the event especially. Thus, for this event the whole hydrographs are analyzed after each assimilation cycle in order to evaluate the progressive improvement of the discharge forecast not only in terms of discharge peaks, as in the boxplots, but also addressing their time evolution (Fig. 18). Results must be read accounting for the performance achieved in terms of precipitation (see Figs. 8 and 9).

At 0000 and 0600 UTC DA is not impacting significantly on peak discharge timing, when compared with the Open Loop. In these cycles the forecast framework overestimates the rainfall between 0000 and 1200 UTC and therefore the discharge peak. Looking at Fig. 18, it is in fact evident that the time window where the gray bands reach the higher values of streamflow (about 0800–1400 UTC) is similar in Figs. 18a–c.

DA performed at 1200 UTC, which would have been available from an operational point of view around 1500 UTC, namely, 5–6 h earlier than the run forced with 1200 UTC analysis, improves significantly the rainfall prediction between 1200 and 2400 UTC, thus

FIG. 16. Columnar contents analysis for 2000 UTC 9 Oct 2014 (first row) graupel (QG), (second row) ice (QI), (third row) snow (QS), (fourth row) rain (QR), and (fifth row) cloud water (QC). Comparison between the (a),(f),(k),(p),(u) Open Loop simulation and the results achieved with the different reflectivity operators: (b),(g),(l),(q),(v) Radar-direct, (c),(h),(m),(r),(w) Radar-direct-modif, (d),(i),(n),(s),(x) Radar-indirect, and (e),(j),(o),(t),(y) Radar-indirect-rqv. The magenta contour line indicates the position of the observed convective structure using the 3000-m reflectivity CAPPI above 20 dBZ at 2000 UTC. The black dot highlights the city most hit by the flash flood (Genoa in this case).

leading to an improvement of the discharge forecast accuracy. The 95th percentile is around $1200 \text{ m}^3 \text{ s}^{-1}$ and the average peak timing is around 1800 UTC, much closer to the observed one, significantly improving also the finding of [Parodi et al. \(2017\)](#).

Also DA at 1800 UTC, available from an operational point of view around 1900 UTC, would have been very important from a physically based short-range nowcasting perspective allowing us to understand the evolution in the next few hours during the most intense phases of the rainfall and discharge phenomena.

Generally, the application of the 3DVAR in cycling mode has, at least for this case study, a relevant impact on the hydrometeorological results for the next 8–9 h, lasting longer than in [Davolio et al. \(2017\)](#).

5. Conclusions

The back-building MCSs frequently affecting the Mediterranean coastal regions are very challenging from a

predictive ability point of view. For this reason, this work addressed three back-building MCSs that occurred in Liguria between 2011 and 2014, causing 20 casualties and several hundreds of millions of euros of damage. The impact of a 6-h cycling 3DVAR data assimilation scheme on the high resolution (1 km) WRF simulations feeding the Continuum hydrological model via the RainFARM stochastic downscaling has been evaluated.

The innovation of this work is represented by the use of different 3DVAR operators for the direct and indirect radar reflectivity data assimilation together with surface observations aiming to identify the best-performing setup for MCSs prediction in terms of both QPF patterns and amounts. Subsequently, the best-performing QPF 3DVAR sensitivity experiments, evaluated through MODE, are fed into the RainFARM and the Continuum hydrological model to forecast peak discharge. The simulated discharge is used to validate the NWP performance at each assimilation step so as to highlight the added value of the use of a 6-h cycling 3DVAR.

FIG. 17. Results of hydrological verification in terms of peak flows. The x axis reports the time of assimilation or the Open Loop NWP runs and y axes report peak flows. DA1 stands for data assimilation, OLP stands for Open Loop. The boxplot represents the predicted peaks distribution, the red dot the observed peak, and the blue cross the simulated peak obtained using observations as input to hydrological model. Each panel refers to a basin and to one of the considered events.

From a meteorological point of view, the 3DVAR assimilation of radar reflectivity has a greater impact on the forecasts in comparison to the use of surface observations data: radar data in fact provide information at many elevations within the troposphere, while the ground sensor data account for surface observations only. An additional advantage of radar observation is its geographical location: reflectivity observations cover the sea, where the convective cells develop, while ground sensors provide observational data only above land once the convective cells are developed. Furthermore, the modified direct operator allows achieving the best performance for the two study cases of 2011, improving the forecast made with standard direct operator. This positive impact is probably due to the fact that, the 0°C isotherm was quite low lying (850–900 hPa), thus supporting a relevant production of solid-phase hydrometeors.

For the Genoa 2014 case study, the main challenge was the reproduction of the second phase of the event, completely missed by the operational Open Loop simulation. The use of the indirect reflectivity

operator, coupled with the in-cloud humidity retrieval, achieved the best performance providing an enhancement of the kinematics, that is, the prominent convergence line that triggered an even more intense deep convection in the second phase of the event.

The best meteorological simulations for each case study (Radar-direct-modif for both 2011 events and Radar-indirect-rqv for the 2014 event) was then fed into the Continuum hydrological model after the application of the RainFARM stochastic downscaling: peak discharge improves significantly even when the Open Loop already provided a good forecast (like the Cinqueterre 2011 use case).

It is possible to conclude that the use of the hydro-meteorological framework coupling a high-resolution WRF simulation including a 6-h cycling 3DVAR of radar reflectivity, possibly using an ensemble of reflectivity operators, with the Continuum hydrological model can help to obtain more timely and accurate streamflow forecasts for back-building MCSs. Whenever there is not the possibility to use the full portfolio of 3DVAR radar reflectivity operators, the

FIG. 18. Hydrograph related to the Genoa 2014 event: (a) Open Loop, (b) DA at 0000 UTC, (c) DA at 0600 UTC, (d) DA at 1200 UTC, and (e) DA at 1800 UTC. Dark gray represents the ensemble area between 0% and 100%, light gray represents the ensemble area between 5% and 95%, the red line refers to the ensemble mean, the blue line is the observation at Passerella Firpo station, and the light blue line represents the streamflow computed using observed meteorological variables as input to the hydrological model.

Radar-direct-modif setup turns out to be the best compromises solution.

Acknowledgments. This work was supported by the Italian Civil Protection Department and by the Ligurian Environmental Agency. We acknowledge the Italian Civil Protection Department for providing us with the Italian Radar and Weather Stations Network. Thanks are due to LRZ Supercomputing Centre, Garching, Germany, where the numerical simulations were performed on the SuperMUC Petascale System, Project-ID: pr62ve.

APPENDIX

Statistical Scores Description

In the manuscript different statistical scores are used; these parameters were derived from a contingency table that shows the frequency of “yes” and “no” rain forecasts and occurrences. The four combinations of forecasts (yes or no) and observations (yes or no) generate four different output of the table:

- 1) hit: rain forecasted and occurred;
- 2) miss: rain not forecasted and occurred;
- 3) false alarm: rain forecasted and not occurred;
- 4) correct negative: rain not forecasted and not occurred.

From this output it is possible to compute the statistical scores with the following formulations:

$$\text{FBIAS} = \frac{\text{hits} + \text{false alarms}}{\text{hits} + \text{misses}},$$

$$\text{PODY} = \frac{\text{hits}}{\text{hits} + \text{misses}},$$

$$\text{FAR} = \frac{\text{false alarms}}{\text{hits} + \text{false alarms}},$$

$$\text{CSI} = \frac{\text{hits}}{\text{hits} + \text{misses} + \text{false alarms}},$$

$$\text{HK} = \frac{\text{hits}}{\text{hits} + \text{misses}} - \frac{\text{false alarms}}{\text{false alarms} + \text{correct negative}},$$

and

$$\text{HSS} = \frac{(\text{hits} + \text{correct negatives}) - (\text{expected correct})_{\text{random}}}{\text{total} - (\text{expected correct})_{\text{random}}},$$

where

$$(\text{expected correct})_{\text{random}} = \frac{1}{\text{total}} \left[\frac{(\text{hits} + \text{misses})(\text{hits} + \text{false alarms}) + (\text{correct negatives} + \text{misses})(\text{correct negatives} + \text{false alarms})}{2} \right].$$

REFERENCES

- Arakawa, A., 2004: The cumulus parameterization problem: Past, present, and future. *J. Climate*, **17**, 2493–2525, [https://doi.org/10.1175/1520-0442\(2004\)017<2493:RATCPP>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<2493:RATCPP>2.0.CO;2).
- Barker, D. M., W. Huang, Y. R. Guo, and Q. N. Xiao, 2004: A three-dimensional (3DVAR) data assimilation system for use with MM5: Implementation and initial results. *Mon. Wea. Rev.*, **132**, 897–914, [https://doi.org/10.1175/1520-0493\(2004\)132<0897:ATVDAS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<0897:ATVDAS>2.0.CO;2).
- , and Coauthors, 2012: The Weather Research and Forecasting Model's Community Variational/Ensemble Data Assimilation System: WRFDA. *Bull. Amer. Meteor. Soc.*, **93**, 831–843, <https://doi.org/10.1175/BAMS-D-11-00167.1>.
- Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525**, 47–55, <https://doi.org/10.1038/nature14956>.
- Benjamin, S. G., G. A. Grell, J. M. Brown, and T. G. Smirnova, 2004: Mesoscale weather prediction with the RUC hybrid isentropic-terrain-following coordinate model. *Mon. Wea. Rev.*, **132**, 473–494, [https://doi.org/10.1175/1520-0493\(2004\)132<0473:MWPWTR>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<0473:MWPWTR>2.0.CO;2).
- Cassola, F., F. Ferrari, and A. Mazzino, 2015: Numerical simulations of Mediterranean heavy precipitation events with the WRF model: A verification exercise using different approaches. *Atmos. Res.*, **164–165**, 210–225, <https://doi.org/10.1016/j.atmosres.2015.05.010>.
- Clark, P., N. Roberts, H. Lean, S. P. Ballard, and C. Charlton-Perez, 2016: Convection-permitting models: A step-change in rainfall forecasting. *Meteor. Appl.*, **23**, 165–181, <https://doi.org/10.1002/met.1538>.
- Das, M. K., M. A. M. Chowdhury, S. Das, S. K. Debsarma, and S. Karmakar, 2015: Assimilation of Doppler weather radar data and their impacts on the simulation of squall events during premonsoon season. *Nat. Hazards*, **77**, 901–931, <https://doi.org/10.1007/s11069-015-1634-9>.
- Davis, A. C., B. Brown, and R. Bullock, 2006a: Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, **134**, 1772–1784, <https://doi.org/10.1175/MWR3145.1>.
- , —, and —, 2006b: Object-based verification of precipitation forecasts. Part II: Application to convective rain system. *Mon. Wea. Rev.*, **134**, 1785–1795, <https://doi.org/10.1175/MWR3146.1>.
- Davolio, S., F. Silvestro, and P. Malguzzi, 2015: Effects of increasing horizontal resolution in a convection permitting model on flood forecasting: The 2011 dramatic events in Liguria (Italy). *J. Hydrometeorol.*, **16**, 1843–1856, <https://doi.org/10.1175/JHM-D-14-0094.1>.
- , —, and T. Gastaldo, 2017: Impact of rainfall assimilation on high-resolution hydrometeorological forecasts over Liguria, Italy. *J. Hydrometeorol.*, **18**, 2659–2680, <https://doi.org/10.1175/JHM-D-17-0073.1>.
- Dowell, D. C., L. J. Wicker, and C. Snyder, 2011: Ensemble Kalman filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale analyses. *Mon. Wea. Rev.*, **139**, 272–294, <https://doi.org/10.1175/2010MWR3438.1>.
- Ducrocq, V., and Coauthors, 2014: HyMeX-SOP1: The field campaign dedicated to heavy precipitation and flash flooding in the northwestern Mediterranean. *Bull. Amer. Meteor. Soc.*, **95**, 1083–1100, <https://doi.org/10.1175/BAMS-D-12-00244.1>.
- Fiori, E., A. Comellas, L. Molini, N. Rebora, F. Siccardi, D. Gochis, S. Tanelli, and A. Parodi, 2014: Analysis and hindcast simulations of an extreme rainfall event in the Mediterranean area: The Genoa 2011 case. *Atmos. Res.*, **138**, 13–29, <https://doi.org/10.1016/j.atmosres.2013.10.007>.
- , L. Ferraris, L. Molini, F. Siccardi, D. Kranzlmüller, and A. Parodi, 2017: Triggering and evolution of a deep convective system in the Mediterranean Sea: Modelling and observations at a very fine scale. *Quart. J. Roy. Meteor. Soc.*, **143**, 927–941, <https://doi.org/10.1002/qj.2977>.
- Gao, J., and D. J. Stensrud, 2012: Assimilation of reflectivity data in a convective-scale, cycled 3DVAR framework with hydrometeor classification. *J. Atmos. Sci.*, **69**, 1054–1065, <https://doi.org/10.1175/JAS-D-11-0162.1>.
- Gilmore, M. S., J. M. Straka, and E. N. Rasmussen, 2004: Precipitation and evolution sensitivity in simulated deep convective storms: Comparisons between liquid-only and simple ice and liquid phase microphysics. *Mon. Wea. Rev.*, **132**, 1897–1916, [https://doi.org/10.1175/1520-0493\(2004\)132<1897:PAESIS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<1897:PAESIS>2.0.CO;2).
- Ha, J.-H., H.-W. Kim, and D.-K. Lee, 2011: Observation and numerical simulations with radar and surface data assimilation for heavy rainfall over central Korea. *Adv. Atmos. Sci.*, **28**, 573–590, <https://doi.org/10.1007/s00376-010-0035-y>.
- Hally, A., O. Caumont, L. Garrote, E. Richard, A. Weerts, F. Delogu, and M. Ivković, 2015: Hydrometeorological multi-model ensemble simulations of the 4 November 2011 flash flood event in Genoa, Italy, in the framework of the DRIHM project. *Nat. Hazards Earth Syst. Sci.*, **15**, 537–555, <https://doi.org/10.5194/nhess-15-537-2015>.
- Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42**, 129–151.
- , Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, <https://doi.org/10.1175/MWR3199.1>.
- Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.*, **113**, D13103, <https://doi.org/10.1029/2008JD009944>.

- Ide, K., P. Courtier, M. Ghil, and A. C. Lorenc, 1997: Unified notation for data assimilation: Operational, sequential and variational. *J. Meteor. Soc. Japan*, **75**, 181–189, https://doi.org/10.2151/jmsj1965.75.1B_181.
- Kain, J. S., S. J. Weiss, J. J. Levit, M. E. Baldwin, and D. R. Bright, 2006: Examination of convection-allowing configurations of the WRF model for the prediction of severe convective weather: The SPC/NSSL Spring Program 2004. *Wea. Forecasting*, **21**, 167–181, <https://doi.org/10.1175/WAF906.1>.
- , and Coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection allowing NWP. *Wea. Forecasting*, **23**, 931–952, <https://doi.org/10.1175/WAF2007106.1>.
- Lagasio, M., A. Parodi, R. Procopio, F. Rachidi, and E. Fiori, 2017: Lightning Potential Index performances in multimicrophysical cloud-resolving simulations of a back-building mesoscale convective system: The Genoa 2014 event. *J. Geophys. Res. Atmos.*, **122**, 4238–4257, <https://doi.org/10.1002/2016JD026115>.
- Laiolo, P., S. Gabellani, N. Rebora, R. Rudari, L. Ferraris, S. Ratto, H. Stevenin, and M. Cauduro, 2014: Validation of the Flood-PROOFS probabilistic forecasting system. *Hydrol. Processes*, **28**, 3466–3481, <https://doi.org/10.1002/hyp.9888>.
- Lee, J.-H., H.-H. Lee, Y. Choi, H.-W. Kim, and D.-K. Lee, 2010: Radar data assimilation for the simulation of mesoscale convective systems. *Adv. Atmos. Sci.*, **27**, 1025–1042, <https://doi.org/10.1007/s00376-010-9162-8>.
- Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.*, **22**, 1065–1092, [https://doi.org/10.1175/1520-0450\(1983\)022<1065:BPOTSF>2.0.CO;2](https://doi.org/10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2).
- Liu, J., M. Bray, and D. Han, 2013: A study on WRF radar data assimilation for hydrological rainfall prediction. *Hydrol. Earth Syst. Sci.*, **17**, 3095–3110, <https://doi.org/10.5194/hess-17-3095-2013>.
- Maiello, I., R. Ferretti, S. Gentile, M. Montopoli, E. Picciotti, F. S. Marzano, and C. Faccani, 2014: Impact of radar data assimilation for the simulation of a heavy rainfall case in central Italy using WRF-3DVAR. *Atmos. Meas. Tech.*, **7**, 2919–2935, <https://doi.org/10.5194/amt-7-2919-2014>.
- , S. Gentile, R. Ferretti, L. Baldini, N. Roberto, E. Picciotti, P. P. Alberoni, and F. S. Marzano, 2017: Impact of multiple radar reflectivity data assimilation on the numerical simulation of a flash flood event during the HyMeX campaign. *Hydrol. Earth Syst. Sci.*, **21**, 5459–5476, <https://doi.org/10.5194/hess-21-5459-2017>.
- Mazzarella, V., I. Maiello, V. Capozzi, G. Budillon, and R. Ferretti, 2017: Comparison between 3D-Var and 4D-Var data assimilation methods for the simulation of a heavy rainfall case in central Italy. *Adv. Sci. Res.*, **14**, 271–278, <https://doi.org/10.5194/asr-14-271-2017>.
- Nash, J. E., and J. V. Sutcliffe, 1970: River flood forecasting through conceptual models. Part I: A discussion of principles. *J. Hydrol.*, **10**, 282–290, [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6).
- Parodi, A., G. Boni, L. Ferraris, F. Siccardi, P. Pagliara, E. Trovatore, and D. Kranzmueller, 2012: The “perfect storm”: From across the Atlantic to the hills of Genoa. *Eos, Trans. Amer. Geophys. Union*, **93**, 225–226, <https://doi.org/10.1029/2012EO240001>.
- , and Coauthors, 2017: DRIHM (2US): An e-science environment for hydrometeorological research on high-impact weather events. *Bull. Amer. Meteor. Soc.*, **98**, 2149–2166, <https://doi.org/10.1175/BAMS-D-16-0279.1>.
- Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center’s spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763, [https://doi.org/10.1175/1520-0493\(1992\)120<1747:TNMCS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1992)120<1747:TNMCS>2.0.CO;2).
- Rebora, N., L. Ferraris, J. H. Hardenberg, and A. Provenzale, 2006a: Rainfall downscaling and flood forecasting: a case study in the Mediterranean area. *Nat. Hazards Earth Syst. Sci.*, **6**, 611–619, <https://doi.org/10.5194/nhess-6-611-2006>.
- , —, —, and —, 2006b: The RainFARM: Rainfall downscaling by a filtered auto regressive model. *J. Hydrometeor.*, **7**, 724–738, <https://doi.org/10.1175/JHM517.1>.
- , and Coauthors, 2013: Extreme rainfall in the Mediterranean: What can we learn from observations? *J. Hydrometeor.*, **14**, 906–922, <https://doi.org/10.1175/JHM-D-12-083.1>.
- Siccardi, F., G. Boni, L. Ferraris, and R. Rudari, 2005: A reference framework for probabilistic flood forecast. *J. Geophys. Res.*, **110**, D05101, <https://doi.org/10.1029/2004JD005314>.
- Silvestro, F., and N. Rebora, 2014: Impact of precipitation forecast uncertainties and initial soil moisture conditions on a probabilistic flood forecasting framework. *J. Hydrol.*, **519**, 1052–1067, <https://doi.org/10.1016/j.jhydrol.2014.07.042>.
- , —, and L. Ferraris, 2009: An algorithm for real-time rainfall rate estimation using polarimetric radar: RIME. *J. Hydrometeor.*, **10**, 227–240, <https://doi.org/10.1175/2008JHM1015.1>.
- , —, and —, 2011: Quantitative flood forecasting on small and medium size basins: A probabilistic approach for operational purposes. *J. Hydrometeor.*, **12**, 1432–1446, <https://doi.org/10.1175/JHM-D-10-05022.1>.
- , S. Gabellani, F. Delogu, R. Rudari, and G. Boni, 2013: Exploiting remote sensing land surface temperature in distributed hydrological modelling: The example of the Continuum model. *Hydrol. Earth Syst. Sci.*, **17**, 39, <https://doi.org/10.5194/hess-17-39-2013>.
- , —, —, —, P. Laiolo, and G. Boni, 2015: Uncertainty reduction and parameter estimation of a distributed hydrological model with ground and remote-sensing data. *Hydrol. Earth Syst. Sci.*, **19**, 1727–1751, <https://doi.org/10.5194/hess-19-1727-2015>.
- , N. Rebora, G. Cummings, and L. Ferraris, 2015b: Experiences of dealing with flash floods using an ensemble hydrological nowcasting framework: Implications of communication, accessibility and distribution of the results. *J. Flood Risk Manage.*, **10**, 446–462, <https://doi.org/10.1111/jfr3.12161>.
- , —, F. Giannoni, A. Cavallo, and L. Ferraris, 2016: The flash flood of the Bisagno Creek on 9th October 2014: An “unfortunate” combination of spatial and temporal scales. *J. Hydrol.*, **541A**, 50–62, <https://doi.org/10.1016/j.jhydrol.2015.08.004>.
- Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., <https://doi.org/10.5065/D68S4MVH>.
- Smith, P. L., Jr., C. G. Myers, and H. D. Orville, 1975: Radar reflectivity factor calculations in numerical cloud models using bulk parameterization of precipitation processes. *J. Appl. Meteor.*, **14**, 1156–1165, [https://doi.org/10.1175/1520-0450\(1975\)014<1156:RRFCIN>2.0.CO;2](https://doi.org/10.1175/1520-0450(1975)014<1156:RRFCIN>2.0.CO;2).
- Sugimoto, S., N. A. Crook, J. Sun, Q. Xiao, and D. M. Barker, 2009: An examination of WRF 3DVAR radar data assimilation on its capability in retrieving unobserved variables and forecasting precipitation through observing system simulation experiments. *Mon. Wea. Rev.*, **137**, 4011–4029, <https://doi.org/10.1175/2009MWR2839.1>.
- Sun, J., and N. A. Crook, 1997: Dynamical and microphysical retrieval from Doppler RADAR observations using a cloud model and its adjoint. Part I: Model development and

- simulated data experiments. *J. Atmos. Sci.*, **54**, 1642–1661, [https://doi.org/10.1175/1520-0469\(1997\)054<1642:DAMRFD>2.0.CO;2](https://doi.org/10.1175/1520-0469(1997)054<1642:DAMRFD>2.0.CO;2).
- Vulpiani, G., and Coauthors, 2008: The Italian radar network within the national early-warning system for multi-risks management. *Proc. of Fifth European Conference on Radar in Meteorology and Hydrology*, Helsinki, Finland, ERAD, https://www.researchgate.net/profile/Gianfranco_Vulpiani/publication/242304008_The_Italian_radar_network_within_the_national_early-warning_system_for_multi-risks_management/links/5955f020aca272fbb37c868a/The-Italian-radar-network-within-the-national-early-warning-system-for-multi-risks-management.pdf.
- Wang, H., J. Sun, S. Fan, and X.-Y. Huang, 2013: Indirect assimilation of radar reflectivity with WRF 3D-Var and its impact on prediction of four summertime convective events. *J. Appl. Meteor. Climatol.*, **52**, 889–902, <https://doi.org/10.1175/JAMC-D-12-0120.1>.
- , X. Y. Huang, J. Sun, D. Xu, M. Zhang, S. Fan, and J. Zhong, 2014: Inhomogeneous background error modeling for WRF-Var using the NMC method. *J. Appl. Meteor. Climatol.*, **53**, 2287–2309, <https://doi.org/10.1175/JAMC-D-13-0281.1>.
- WRF Users Group, 2016: WRF data assimilation. User's Guide for the Advanced Research WRF (ARW) Modeling System Version 3.9, http://www2.mmm.ucar.edu/wrf/users/docs/user_guide_V3.8/contents.html.
- Xiao, Q., Y. Kuo, J. Sun, W. Lee, D. M. Barker, and E. Lim, 2007: An Approach of Radar Reflectivity Data Assimilation and Its Assessment with the Inland QPF of Typhoon Rusa (2002) at Landfall. *J. Appl. Meteor. Climatol.*, **46**, 14–22, <https://doi.org/10.1175/JAM2439.1>.