

Parameter Sensitivity of the WRF–LETKF System for Assimilation of Radar Observations: Imperfect-Model Observing System Simulation Experiments

PAULA MALDONADO AND JUAN RUIZ

Centro de Investigaciones del Mar y la Atmósfera, Universidad de Buenos Aires–Consejo Nacional de Investigaciones Científicas y Técnicas, and Departamento de Ciencias de la Atmósfera y los Océanos, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, and Institut Franco-Argentin d'Estudes sur le Climat et ses Impacts, Unité Mixte Internationale, Buenos Aires, Argentina

CELESTE SAULO

Consejo Nacional de Investigaciones Científicas y Técnicas, and Servicio Meteorológico Nacional, and Departamento de Ciencias de la Atmósfera y los Océanos, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina

(Manuscript received 1 August 2019, in final form 17 March 2020)

ABSTRACT

Specification of suitable initial conditions to accurately forecast high-impact weather events associated with intense thunderstorms still poses a significant challenge for convective-scale forecasting. Radar data assimilation has been showing encouraging results to produce an accurate estimate of the state of the atmosphere at the mesoscale, as it combines high-spatiotemporal-resolution observations with convection-permitting numerical weather prediction models. However, many open questions remain regarding the configuration of state-of-the-art data assimilation systems at the mesoscale and their potential impact upon short-range weather forecasts. In this work, several observing system simulation experiments of a mesoscale convective system were performed to assess the sensitivity of the local ensemble transform Kalman filter to both relaxation-to-prior spread (RTPS) inflation and horizontal localization of the error covariance matrix. Realistic large-scale forcing and model errors have been taken into account in the simulation of reflectivity and Doppler velocity observations. Overall, the most accurate analyses in terms of RMSE were produced with a relatively small horizontal localization cutoff radius (~ 3.6 – 7.3 km) and large RTPS inflation parameter (~ 0.9 – 0.95). Additionally, the impact of horizontal localization on short-range ensemble forecast was larger compared to inflation, almost doubling the lead times up to which the effect of using a more accurate state to initialize the forecast persisted.

1. Introduction

Severe weather events associated with deep moist convection have been drawing vast attention because of their enormous societal and economic impacts. Subtropical South America hosts some of the deepest convective storms and most intense mesoscale convective systems (MCSs) in the world as shown by satellite observations (Zipser et al. 2006), and it frequently exhibits synoptic-scale environments favorable for the development of severe weather as shown by reanalysis datasets (Brooks et al. 2003). In particular, MCSs account for 90% of the precipitation over La Plata Basin

and 50% over the Amazon region and produce hazards such as strong winds, large hail, tornadoes, lightning, and flooding (Matsudo and Salio 2011; Mezher et al. 2012; Rasmussen et al. 2014). Thus, to enable accurate short-range forecasts of these high-impact convective weather events is a primary concern in our region. One possible approach to improve these forecasts is by coupling high-resolution numerical weather prediction models with data assimilation systems (Yano et al. 2018).

Over the past two decades, data assimilation of Doppler radar observations has proven to be a feasible technique for producing storm-scale analyses to initialize dynamical forecasts that can predict the evolution of convective systems in the following 0–6 h (Gustafsson et al. 2018; Sun et al. 2014, and references therein). In this direction, variational data assimilation techniques

Corresponding author: Paula Maldonado, paula.maldonado@cima.fcen.uba.ar

(i.e., 3D-Var and 4D-Var) were initially used with promising results (e.g., Sun and Crook 1997, 1998; Xue et al. 2003; Caya et al. 2005; Schmitz and Wulfmeyer 2014, among others). In recent years, the use of ensemble-based methods, like the ensemble Kalman filter (EnKF; Evensen 1994) and its variants, have also shown great potential for convective-scale data assimilation as they allow the description of the flow-dependent evolution of the background error covariances. Since the nature of these covariances is not as well known as on the synoptic scale, where geostrophic and hydrostatic balances play an important role, being able to estimate these covariances is one of the main advantages of ensemble methods for mesoscale and convective-scale applications. Moreover, as these methods are based on an ensemble of model states, they also provide valuable information about the analysis and forecast uncertainty, which can ultimately be used to initialize short-range ensemble forecasts.

The first application of the EnKF in convective-scale radar data assimilation was carried out by Snyder and Zhang (2003) for a perfect-model experiment of a splitting supercell storm, assimilating simulated observations of radial velocity. The authors revealed the ability of the EnKF to produce accurate analyses, even for unobserved model variables such as vertical velocity and temperature. Similarly, Tong and Xue (2005) used a fully compressible cloud model and additionally assimilated simulated reflectivity observations. The authors demonstrated the potential of the EnKF to retrieve all water and ice species associated with a multiclass ice microphysics scheme as well as the wind and thermodynamic variables. The best results were obtained when both radial velocity and reflectivity data, including reflectivity information outside of the precipitation regions, were used. Encouraging results were also shown when applying the EnKF to assimilate real radar observations. Dowell et al. (2004) were the first to report on this application of the EnKF and obtained good-quality analyses of the main updraft and mesocyclone of a supercell storm. Aksoy et al. (2009) were among the first to apply the EnKF to assimilate real radar data for cases whose behaviors span supercellular, linear, and multicellular organization. The authors reported robust results across all cases when assimilating radial velocity and reflectivity, including “no precipitation” observations (i.e., reflectivity observations with values small enough to indicate the absence of precipitation). Further, EnKF radar data assimilation has also become a key component of operational high-resolution numerical weather prediction systems (e.g., Schraff et al. 2016; Gustafsson et al. 2018).

As the EnKF makes use of a finite ensemble size sampling errors arise, leading to suboptimal performance of the method and degrading the estimation of

the background error covariances. Thus, to account for unrepresented error sources (i.e., sampling as well as model errors), localization and inflation methods are employed in ensemble Kalman filters. Localization consists of limiting the impact of observations to nearby model grid points since correlations between an observation and distant grid points have significant errors and should be disregarded. Sobash and Stensrud (2013, hereafter SS13) explored the sensitivity of the EnKF to the localization cutoff radius in the context of convective-scale data assimilation and perfect-model experiments. The authors showed that increasing the horizontal localization and decreasing the vertical localization produced analyses with the smallest error for most of the state variables and discussed the role of model error in the choice of the optimal localization radius. Lange and Craig (2014) also performed perfect-model experiments and investigated the limit of predictability in precipitation forecasts by comparing analysis schemes that resolve different length scales. Results showed that high-resolution analysis led to better forecasts for the first hour, but after 3 h, the forecast quality of the experiments became indistinguishable due to rapid error growth of the small scales. Inflation consists of increasing the ensemble variance in each assimilation cycle to prevent the forecast from deviating from the truth since less weight is given to observations if the ensemble spread becomes too small. Potentially, this could lead to a total denial of new observations and ultimately cause the problem of filter divergence (Anderson and Anderson 1999). Harnisch and Keil (2015) tested various inflation methods to account for unrepresented error sources in an EnKF using the Consortium for Small-Scale Modeling (COSMO) limited-area model. The authors showed that using either multiplicative covariance inflation or relaxation inflation methods enhances the analysis spread and provides initial conditions that produce more consistent ensemble forecasts.

The first goal of this study is to assess the performance of a convective-scale radar data assimilation system by carrying out imperfect-model observing system simulation experiments (OSSEs; Lord et al. 1997) of an intense MCS. Synthetic radial velocity and reflectivity observations are extracted from a reference simulation and assimilated using the local ensemble transformed Kalman filter (LETKF; Hunt et al. 2007). This ensemble-based method is an efficient implementation of the EnKF that has also been applied to radar data assimilation at convective scale with promising results (e.g., Yang et al. 2012, 2013; Tsai et al. 2014; Lien et al. 2017). We also examine the impact of radar data assimilation upon short-range ensemble forecasts initialized from the generated analyses and the predictability of this MCS. The second objective is to evaluate the sensitivity of the convective-scale data

assimilation system to the specification of some adjustable parameters of the LETKF. In particular, we explore the sensitivity of the analysis and short-range ensemble forecasts to the horizontal localization cutoff radius and the magnitude of adaptive multiplicative inflation that is introduced based on a relaxation-to-prior spread approach (RTPS; Whitaker and Hamill 2012). In this work, we complement previous OSSE studies by taking model errors into account and using a realistic large-scale forcing.

The rest of the paper is outlined as follows: Section 2 describes the design of our OSSEs, including the reference simulation and model setup, the synthetic radar observations, and the assimilation procedure. In section 3, the impact of assimilating radar observations on both analysis ensemble mean and short-range forecast ensemble mean is examined. Section 4 and section 5 present the findings of the sensitivity experiments to horizontal localization and adaptive multiplicative inflation, respectively. A concluding section is given at the end of the paper.

2. Experimental design

The experiments presented in this work are based on an OSSE approach. In most convective-scale OSSEs, an idealized setting is considered regarding model error (i.e., twin experiments) and observations (i.e., simulated on model grid points). In our experiments, model error and a realistic large-scale environment are taken into account, and radar observations are sampled on radar coordinates considering a realistic observing strategy.

a. Reference simulation

The Weather Research and Forecasting (WRF) Model, version 3.6 (Skamarock et al. 2008), is used to create a reference simulation (also referred to as nature run or truth) of a well-organized MCS. Initial and boundary conditions for the nature run are taken from the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) operational analysis data with 0.5° horizontal resolution every 6 h, performing a dynamical downscaling using three nested domains with 10-km, 2-km, and 500-m horizontal grid spacing, respectively (Fig. 1). In particular, the highest resolution domain (D03 in Fig. 1) covers the central region of Argentina, spanning 500 km in both horizontal directions, and 20 km in the vertical direction using 60 vertical levels. The physical parameterizations include the WRF single-moment 6-class microphysics scheme (WSM6; Hong and Lim 2006), the Rapid Radiative Transfer Model (RRTM; Mlawer et al. 1997) for longwave radiation, the Dudhia (1989) shortwave radiation scheme, the Unified Noah Land Surface Model (Tewari et al. 2004), and the Yonsei University planetary boundary layer scheme (YSU; Hong et al. 2006).

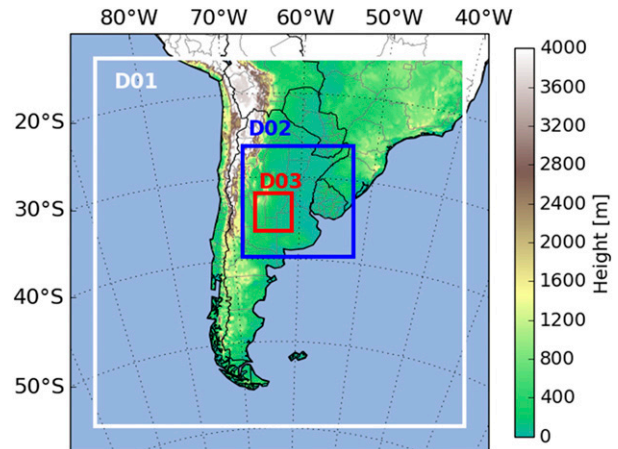


FIG. 1. WRF Model terrain height (shaded). Domain D03 (red line) is used to run the reference simulation and all OSSEs, with 500-m and 2-km horizontal grid spacing, respectively. Domains D01 (white line) and D02 (blue line) are used to dynamically downscale the NCEP GFS initial and boundary conditions to initialize the reference simulation.

The Kain–Fritsch scheme (Kain 2004) for cumulus parameterization is only used in domain D01 as convection is explicitly resolved in domains D02 and D03.

The high-resolution nature run is initialized at 1500 UTC 22 January 2014 and consists of a 7-h simulation (Fig. 2). During the first two hours, a few isolated thunderstorms rapidly grow upscale giving rise to a long-lived squall line at the western part of domain D03, with a dominant north–south orientation and eastward motion. After 1800 UTC (Figs. 3a–e), the initial squall line has reached a more mature stage as shown by the low-level cold pool structure and a strong gust front at its leading edge. Intense surface winds associated with the cold pool propagation as well as high-precipitation rates were produced. Intense convective systems like the one simulated in this work are similar to those frequently observed in southern South America during the warm season and therefore represent a proper scenario to evaluate our assimilation system.

b. Simulated radar observations

Synthetic reflectivity and radial velocity observations are extracted from the high-resolution, deterministic nature run from 1730 to 2005 UTC with a 5-min frequency. Observations are simulated in radar coordinates, and the simulated radar is located at the center of domain D03 (34.37°S, 63.98°W), with a scanning strategy consisting of 14 antenna elevations (ranging from 0.1° to 16.35°), a maximum radar range of 240 km, and a 500-m bin resolution. Gaussian distributed, random observational errors are assumed with zero mean and standard deviation of 2.5 dBZ and 1 m s⁻¹ for reflectivity and

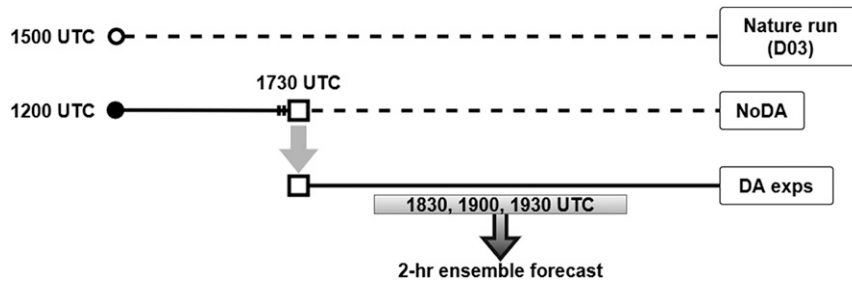


FIG. 2. Schematic design of the conducted experiments. Dashed and solid lines represent deterministic and 60-member ensemble simulations, respectively. Nonfilled and filled dots indicate that initial and boundary conditions come from the NCEP GFS operational analysis with and without downscaling, respectively. Nonfilled squares indicate that at 1730 UTC, CTRL's and NoDA's initial condition is taken from the 60-member ensemble forecast initialized at 1200 UTC.

radial velocity, respectively. The magnitude of these errors is on the order of those used in previous studies [e.g., 2 dBZ and 2 m s^{-1} in SS13; 5 dBZ and 1 m s^{-1} in Lange and Craig (2014)]. For simplicity, we also assume that observational errors are spatially uncorrelated, even though a correlation between adjacent radar pixels is known to exist in real radar observations. Moreover, attenuation and clear-air echoes are not considered, a lower limit of -20 dBZ is assigned to reflectivity observations, and radial velocity observations are assumed to be available only at locations in which model reflectivity

is over -20 dBZ . Although, this value is relatively low compared to a 0-dBZ threshold used by Tong and Xue (2005) as the lower limit for reflectivity, the -20-dBZ threshold implies barely having $\sim 8\%$ more observations than if a 0-dBZ threshold was used.

The radar observation operator described in Tong and Xue (2005) is employed to compute radial velocity and reflectivity observations, and it is assumed to be perfect (i.e., the same one is used to simulate and assimilate radar observations). Radial velocity observations are simulated using the model-simulated velocities u, v , and

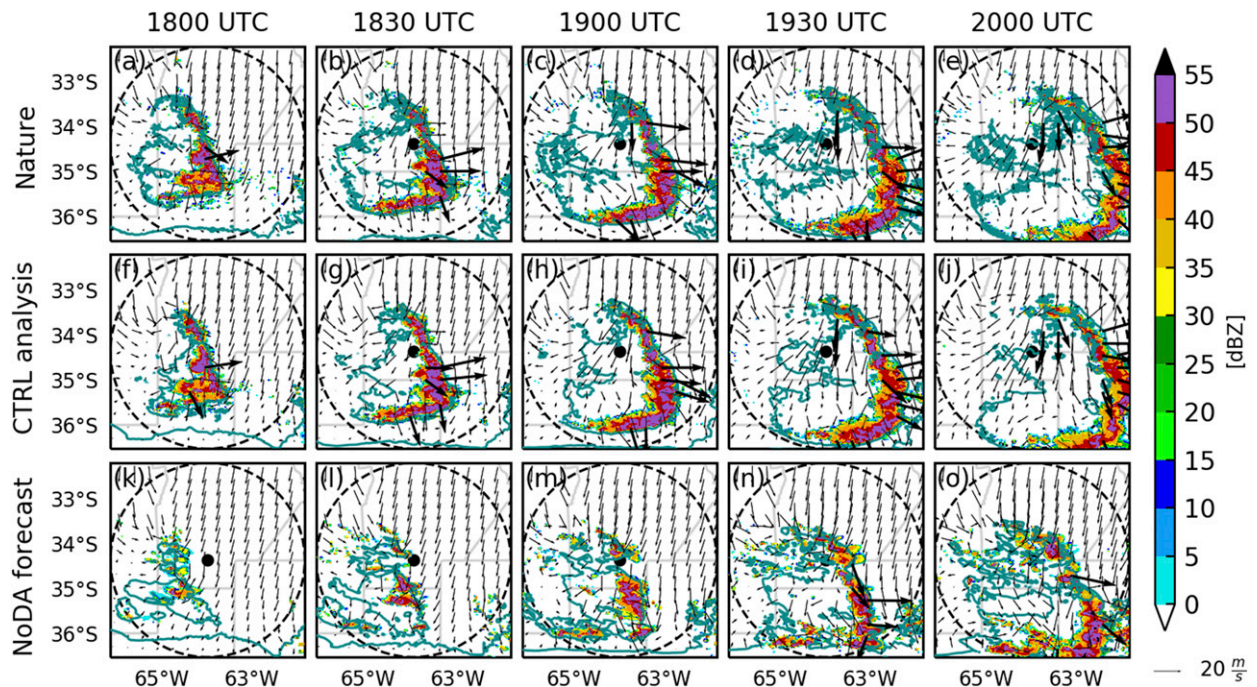


FIG. 3. Time evolution of reflectivity field (shaded, dBZ), horizontal wind field (vectors; m s^{-1}), and -2-K contour of potential temperature perturbation (cyan line) at $z = 1 \text{ km}$, from 1800 to 2000 UTC, every 30 min. Thick arrows indicate wind magnitude larger than 20 m s^{-1} . Radar maximum range (dashed black line) and location (black dot) are shown for reference. (a)–(e) Nature run, (f)–(j) CTRL analysis ensemble mean, and (k)–(o) NoDA forecast.

w , and the hydrometeor's averaged terminal velocity is taken into account following Lin et al. (1983). When more than one microphysical species is present, a reflectivity weighted average is used to estimate this terminal velocity. For reflectivity observations, the equivalent reflectivity factor consists of contributions from rainwater, snow, and graupel, which are functions of the model-simulated mixing ratios q_r , q_s , and q_g , respectively. The equivalent reflectivity factor–mixing ratio relations correspond to a S-band radar, and a Marshall–Palmer exponential raindrop size distribution is assumed for the intercept parameter.

c. Data assimilation and OSSEs

Assimilation of simulated observations is carried out by coupling the WRF Model with the LETKF assimilation method. In particular, the WRF–LETKF system developed by Miyoshi and Kunii (2012) has been extended to include the radar observation operator described in the previous subsection. Data assimilation experiments are run in domain D03 (cf. Fig. 1). To simulate the effect of model errors upon the data assimilation system, the model configuration of all assimilation experiments differ from the one used in the high-resolution nature run in three aspects: (i) horizontal grid spacing (i.e., 2 km instead of 500 m) as in Maejima et al. (2019), (ii) initial and boundary conditions, and (iii) microphysics scheme (i.e., Lin et al. 1983 single-moment scheme instead of WSM6).

When radar data are assimilated into convection-allowing models, the spatial observation number is usually reduced either by applying a thinning technique (e.g., Montmerle and Faccani 2009) or a superobbing approach (e.g., Weng and Zhang 2012). According to Tsai et al. (2014), the superobbing approach smooths out small-scale details reducing the observation representativeness error, while both techniques diminish the impact of correlated observation errors. In our OSSEs, observation errors are assumed to be uncorrelated. Still representativeness error is an issue since the nature run's horizontal grid spacing is smaller than in the data assimilation experiments. Therefore, a superobbing technique is chosen, and observations sampled in radar coordinates are smoothed using a 2-km box average to match the model horizontal grid spacing. Furthermore, reflectivity observations with values small enough to indicate the absence of precipitation, that is, below 0 dBZ in our case, are deemed no-precipitation observations. Since these observations are beneficial in suppressing spurious convection within the assimilation experiments domain (Tong and Xue 2005; Aksoy et al. 2009), they are assimilated as a 0-dBZ reflectivity value with the same observation error as for positive reflectivity observations (described in section 2b).

The assimilation of simulated superobservations is done every 5 min over a 150-min period, from 1735 to 2005 UTC, using a 60-member ensemble. The first background ensemble is initialized at 1730 UTC when the squall line has already developed in the nature run. The first ensemble is taken from a cold-started, 60-member ensemble forecast initialized from the NCEP GFS operational analysis at 1200 UTC (cf. Fig. 2). In this way, small-scale perturbations are allowed to develop throughout the domain during this 5.5-h forecast, and the spinup time of moist processes is taken into account so that convection could form inside the domain by the time the assimilation experiments begin. Ensemble boundary conditions are generated by adding two types of perturbations to the NCEP GFS operational analysis. On the one hand, balanced perturbations are generated by subtracting two random atmospheric states given by the Climate Forecast System Reanalysis (CFSR) data with 0.5° horizontal grid spacing and using an amplitude scale factor equal to 5% of the climatological value. These perturbations are applied to all model variables and are used to represent the uncertainty of the large-scale flow (i.e., synoptic scale; Dowell et al. 2004) and maintain the ensemble spread outside convective areas. These perturbations are also used to initialize the cold-started ensemble forecast at 12 UTC. In the absence of a coarser-resolution ensemble, using these types of perturbations has proven to be suitable to account for uncertainty in the boundary conditions in limited-area models (Nutter et al. 2004; Torn et al. 2006). On the other hand, spatially smoothed random perturbations are sampled from a Gaussian distribution with zero mean and a standard deviation of 2 ms⁻¹ and 2 K for velocity and temperature fields, respectively, and are applied only at the boundaries. Finally, 2-h ensemble forecasts are initialized at 1830, 1900, and 1930 UTC from the generated 60-member analysis ensemble (cf. Fig. 2).

As mentioned in section 1, localization and inflation of the error covariance matrix are needed to account for sampling errors due to limitations in ensemble size. In our experiments, R localization (Grebush et al. 2011) is used. Each element of the observation error covariance matrix \mathbf{R} is multiplied by the inverse of a Gaussian localization function as follows:

$$R_{ij}^{\text{loc}} = R_{ij} \exp\left(-\frac{d_{h,ij}^2}{2L_h^2} - \frac{d_{v,ij}^2}{2L_v^2}\right),$$

where d_h and d_v are the horizontal and vertical distance between the i th observation and the j th model grid point, respectively; L_h and L_v are the horizontal and vertical localization length, respectively; and R_{ij} and R_{ij}^{loc} are the (i, j) element of the original and localized \mathbf{R} matrix, respectively. This results in an increasing

function of observation error with distance from the model grid point being updated. In particular, observations farther than a certain cutoff radius r , defined as ~ 3.65 times L (Miyoshi et al. 2007), will not be assimilated. The horizontal localization cutoff radius r_h will vary among experiments. Vertical localization cutoff value is fixed at 7.3 km, similar to the one used in previous studies (e.g., 8 km in Tong and Xue 2005; 7.3 km in Caya et al. 2005). Additionally, inflation of the background error covariance matrix is achieved by using a multiplicative scheme. We consider a constant covariance inflation factor (Anderson and Anderson 1999) of 1.1 (10%). However, we found in previous assimilation experiments that using this scheme alone produced a rapid collapse of the ensemble spread. Therefore, an RTPS scheme (Whitaker and Hamill 2012) is also used in all experiments presented in this paper. On each element of the state vector, this scheme relaxes the analysis ensemble standard deviation back to the background ensemble spread and modifies the analysis ensemble perturbations according to

$$x_i^{t^a} \leftarrow x_i^{t^a} \left(\alpha \frac{\sigma^b - \sigma^a}{\sigma^a} + 1 \right),$$

where $x_i^{t^a}$ is the departure of the i th ensemble member from the analysis ensemble mean; σ^a and σ^b are the analysis and background ensemble standard deviation, respectively; and α is the RTPS inflation parameter, whose value will vary among experiments. For a given value of α , the multiplicative inflation factor is proportional to the amount of reduction of the ensemble spread due to the assimilation of observations, normalized by the analysis ensemble spread. Hence, it allows increasing the ensemble variance more in regions where observations produce a larger reduction of the analysis uncertainty.

The performance of the WRF-LETKF radar data assimilation system is first assessed by comparing a control experiment (CTRL) versus an experiment with no assimilation of observations (NoDA; cf. Fig. 2). In CTRL, radar observations are assimilated using an RTPS inflation parameter $\alpha = 0.9$ and a horizontal localization cutoff radius $r_h = 7.3$ km. NoDA is a deterministic forecast initialized from the mean of the 60-member ensemble forecast at 1730 UTC, thus avoiding smoothing out small-scale features as would happen if working with the ensemble mean, and the large-scale forcing comes from a randomly chosen member of the ensemble of boundary conditions used in CTRL. Additional assimilation experiments are performed to study the sensitivity of the data assimilation system to the horizontal localization cutoff radius and the magnitude of the RTPS multiplicative

TABLE 1. List of assimilation experiments. The asterisk indicates the same setting as CTRL.

Experiment name	RTPS inflation parameter (α)	Horizontal localization cutoff radius (r_h ; km)
CTRL	0.9	7.3
LOC3.6	*	3.6
LOC14.6	*	14.6
RTPS0.7	0.7	*
RTPS0.95	0.95	*

inflation parameter (Table 1). Experiments LOC3.6, CTRL, and LOC14.6, with $\alpha = 0.9$ and $r_h = 3.6$ km, $r_h = 7.3$ km, and $r_h = 14.6$ km, respectively, are used to explore the former, while experiments RTPS0.7, CTRL, and RTPS0.95, with $r_h = 7.3$ km and $\alpha = 0.7$, $\alpha = 0.9$, and $\alpha = 0.95$, respectively, are used to explore the latter.

d. Verification metrics

The performance of the WRF-LETKF system is validated against the inner-most domain nature run after applying a box-averaging interpolation technique to match the assimilation experiment's horizontal grid spacing of 2 km (the interpolated nature run is hereafter noted as x^t).

The quality of the analyses produced by the WRF-LETKF system, and the short-range ensemble forecast initialized from those estimates, is assessed by using gridpoint-based statistical metrics. Space-averaged root-mean-square error (RMSE), bias (BIAS), and ensemble spread (SPREAD) are defined as

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{j=1}^N (\bar{x}_j - x_j^t)^2}, \quad \text{BIAS} = \frac{1}{N} \sum_{j=1}^N (\bar{x}_j - x_j^t), \quad \text{and}$$

$$\text{SPREAD} = \sqrt{\frac{1}{N} \sum_{i=j}^N \left[\frac{1}{M-1} \sum_{i=1}^M (x_{j,i} - \bar{x}_j)^2 \right]},$$

where the subscript i is an index over all M ensemble members, and the subscript j is an index over all N grid points where convection is present (i.e., defined by reflectivity values greater than 0 dBZ) either in the analysis/forecast ensemble mean \bar{x} or in x^t . Defining the space-averaged computation this way provides a more thorough measure of the analysis/forecast quality because it penalizes both misses and false alarms.

The impact of radar data assimilation on different spatial scales and its effect on forecast predictability are studied by applying a spectral decomposition technique to the analysis/forecast mean errors, defined as the difference with x^t . To compute the spectra, a linear trend removal method following that in Errico (1985) is

applied to the longitudinal error series, since the model-simulated fields are not periodic given that a limited-area model is being used. A discrete Fourier transform is then applied to determine the spectral power associated with each horizontal scale. The resulting spectra obtained at different latitudes are averaged to obtain a one-dimensional error spectrum. A time average of the analysis error spectra is computed between 1830 and 2000 UTC, not including the spinup period of the filter. Short-range forecast error spectra are averaged over the three forecast initializations available (i.e., 1830, 1900, and 1930 UTC). In the following sections, the ratio of these averaged analysis/forecast error spectrum to the spectral power of x' spatial variability will be presented and referred to as “error relative spectral power” (ERSP).

3. Assimilation impact

In this section, we present the results regarding the performance of the WRF-LETKF on the chosen MCS case study by assessing the impact of radar data assimilation upon analysis and short-range ensemble forecasts, comparing CTRL and NoDA experiments.

a. Analysis performance

Figure 3 shows the time evolution of the reflectivity field, the -2-K contour of potential temperature perturbation, computed as the difference with respect to the mean value of the potential temperature field at 1 km, to broadly outline the low-level cold pool, and the wind field at an altitude of 1 km for CTRL analysis mean (Figs. 3f–j) and NoDA forecast (Figs. 3k–o).

In general, CTRL correctly retrieves the main characteristics of the MCS, such as the convective mode, the squall line eastward propagation, and the low-level circulation, while in NoDA the motion of convective cells roughly agrees with the nature run and the reflectivity can reach a comparable level of intensity. However, in terms of wind speed at the leading edge of the cold pool (i.e., the gust front), the system’s intensity is comparatively weaker in NoDA than in CTRL. By 1800 UTC, a line-structured convective pattern is present in the analysis mean after six assimilation cycles (i.e., 30 min), while isolated cells are depicted in NoDA forecast. It takes approximately two hours for convective cells to organize into a squall line in NoDA (i.e., 1930 UTC), suggesting that assimilation of radar observations helped to represent convection organization in CTRL. Moreover, the analysis ensemble mean locates the squall line closer to the reference simulation than NoDA as a systematic lagging of the convective cells is observed in the latter with no aid from observations to correct it in successive times. Furthermore, the gust front propagation

velocity is smaller in NoDA than in the nature run between 1930 and 2000 UTC due to a possible underestimation of the cold pool intensity during that period, contributing to the misplacement of the system.

It is worth noticing that from 1830 UTC onward, CTRL continuously improves the cold pool extension and captures much of the small-scale structures in the nature run. As no temperature observations are being assimilated, this variable could be improved in two ways: (i) by updating the temperature values based on cross correlations with observed variables (i.e., reflectivity), especially close to the convective system during the assimilation step, and (ii) by a good representation of the mesoscale dynamics of the convective system during the forecast step. While these two mechanisms are intrinsically related to each other, and their relative contribution is difficult to infer from Fig. 3, a good representation of the cold pool (and of any variable that is not directly related to observed variables) is an indication that the system is converging to a solution consistent with the actual dynamics of the MCS. Moreover, in regions far from the convective system where radar data are sparse (except for no-precipitation observations), the presence of a cold pool is partly a result of a good representation of the MCS structure driven by the model dynamical evolution as it adequately propagates information from convective clouds where observations are available.

The impact of radar data assimilation on different spatial scales during the assimilation period is studied by comparing the error spectral power of CTRL and NoDA (Fig. 4). We also explore the impact on variables that are directly and indirectly related to those being assimilated by showing reflectivity and the zonal component of the velocity field (i.e., directly related to Doppler velocity as the MCS has a dominant north–south orientation) for the former and temperature for the latter. All power spectra present a maximum at the longest wavelength (i.e., 250 km) with decreasing values toward smaller scales. NoDA error spectrum amplitudes are larger than the energy spectrum of the nature run at nearly all wavelengths and in particular for wavelengths between 10 and 100 km, meaning that there is hardly any improvement regarding the spatial scales that can be correctly represented by NoDA compared to the ones present in the truth simulation. On the other hand, errors in CTRL experiment show lower energy than in nature run for wavelengths larger than roughly 6 km for zonal wind, 15 km for temperature, and 8 km for reflectivity, suggesting that mesoscale phenomena are better represented by the analysis mean. Further examination showed that the maximum wavelength at which the spectral power of the analysis/forecast error saturates or becomes similar to the energy spectrum of the nature run

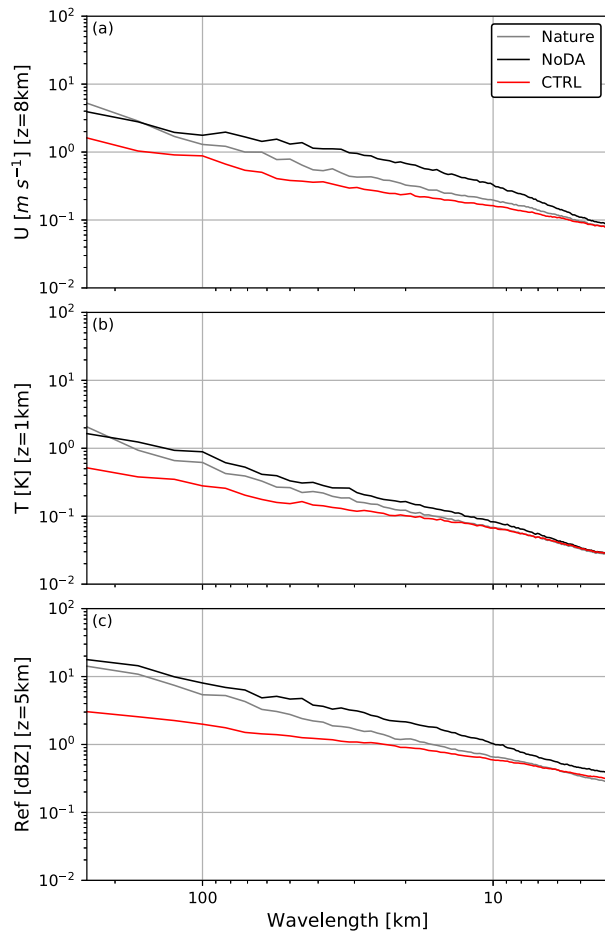


FIG. 4. Error spectral power as a function of wavelength, for CTRL analysis mean (red line) and NoDA forecast (black line). Nature's run spectral power is shown for reference (gray line). (a) Zonal wind at $z = 8$ km, (b) temperature at $z = 1$ km, and (c) reflectivity at $z = 5$ km.

(hereafter referred to as cutoff wavelength) not only depend on model variable but also present a dependence with height. For instance, the cutoff wavelength decreases with height for zonal wind, increases with height for temperature, and has small variations with height for reflectivity. These cutoff wavelengths will be determined more precisely in the following sections. It is important to note that the error spectral power can be regarded as a scale decomposition of the root-mean-square error and is quite sensitive to both intensity and position errors. This partially explains why NoDA error spectrum is higher than the nature run spectrum, indicating almost no skill even when some aspects of the convective system evolution are well captured by NoDA (cf. Fig. 3).

b. Short-range ensemble forecast performance

The ensemble forecast initialized at 1830 UTC from CTRL analyses ensemble, valid at 1900, 1930, and 2000

UTC is shown in Fig. 5. As expected, CTRL forecast mean losses skill with increasing forecast lead time (cf. Figs. 3c–e) but is still distinctly better than NoDA (cf. Figs. 3m–o). During the first forecast hour, the location and intensity of the squall line are well represented. In particular, strong near-surface winds associated with the leading edge of the cold pool are well captured up to 90 min in advance. However, for lead times longer than one hour, a considerable difference between the truth and the forecasted evolution of the system is found in the southern sector of the squall line. This region presents a significant underestimation of the MCS reflectivity and surface winds and becomes more significant as the system approaches the southeastern boundary. Partly, it could be explained because the CAPE ahead of the convective line in the reference simulation is higher than in the analysis used to initialize this forecast (not shown). Another important aspect related to the weakening of the system is that the forecast boundary conditions come from the NCEP GFS analysis available every 6 h, while the nature run employs a two-way nesting between domains D02 and D03 (cf. Fig. 1), allowing a continuous interaction between the two in each model integration.

Quantification of predictability loss throughout the forecast period for different spatial scales is assessed through the forecast ensemble mean ERSP of reflectivity at an altitude of 5 km (Fig. 6). Values greater or equal than unity imply a complete loss of predictability for CTRL's forecast ensemble mean. In the case of NoDA's deterministic forecast, the loss of predictability is reached for values greater than $\sqrt{2}$ (Kalnay 2003). The initial analysis mean from CTRL is able to skillfully represent the different spatial scales of the spectrum except for those smaller than the cutoff wavelength (i.e., 8 km for reflectivity), while a forecast initialized from a state with no assimilation of radar observations shows no skill for almost all resolved wavelengths. In CTRL forecast, the most predictable scales are the ones corresponding to the β -mesoscale part of the spectrum (i.e., wavelengths larger than 50 km), while in NoDA forecast, predictability is limited to the larger spatial scales, that is, wavelengths higher than 90 km, that mainly respond to the forcing from the boundaries. Many studies have shown that errors at convective scale are eventually responsible for the loss of predictability on much larger scales as forecast range increases (e.g., Zhang et al. 2003; Selz and Craig 2015). For instance, in the present case study, a 1-h forecast cannot accurately describe phenomena with a characteristic length smaller than 40 km, and for a 2-h forecast, this threshold increases to 80 km, suggesting a limit of predictability that roughly reduces the solvable spatial scales at a rate of 20 km per half hour of forecast time. A similar error

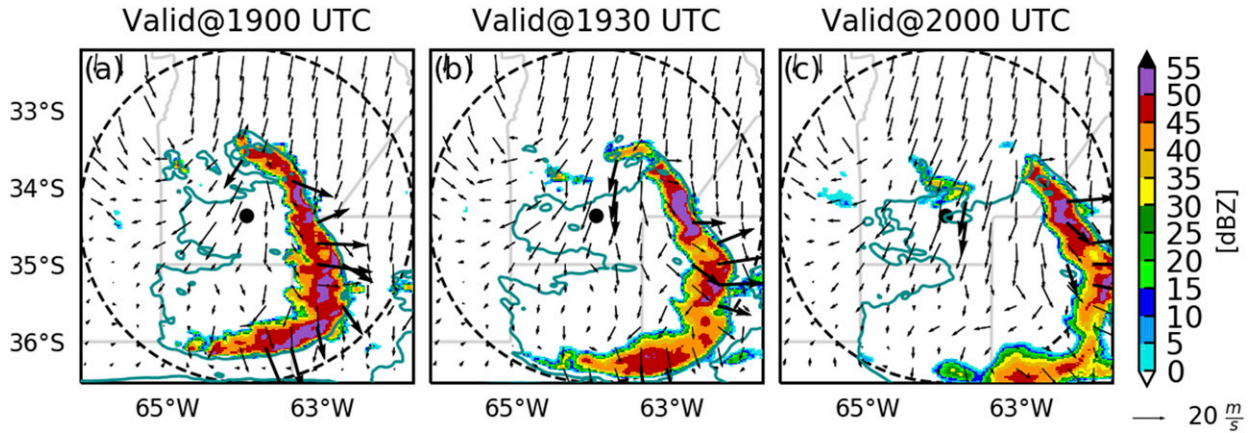


FIG. 5. As in Fig. 3, but for the ensemble forecast initialized at 1830 UTC from CTRL analysis ensemble. Forecast ensemble mean valid at (a) 1900 UTC, (b) 1930 UTC, and (c) 2000 UTC. Comparison should be made to Figs. 3c–e and 3m–o.

growth rate was found by Zhang et al. (2003), as they showed that errors first grew as small-scale differences associated with moist convection and then spread up-scale as their growth began to slow down.

4. Sensitivity to horizontal localization

We explore the sensitivity to horizontal localization by comparing CTRL (i.e., 7.3-km cutoff radius) with experiments LOC3.6 and LOC14.6, all of which use a vertical localization cutoff radius of 7.3 km.

a. State-space statistics

Space-averaged statistics of zonal wind, temperature, and the sum of rain, snow, and graupel mixing ratios

(defined as Ref-condensates) for NoDA, CTRL, LOC3.6, and LOC14.6 are shown in Fig. 7. Figures 7a, 7c, and 7e show the analysis and background scores in a single line giving the classical “sawtooth” appearance representing the error growth during the forecast step and the corrections introduced in the analysis step. In particular, zonal wind and Ref-condensates exhibit larger corrections than temperature as those variables are directly related to observed variables. Figures 7b, 7d, and 7f show the statistics for short-range forecast, which are averaged over three forecasts initialized at 1830, 1900, and 1930 UTC.

During the first 60 min of assimilation, all experiments show a gradual decrease in RMSE and more steady behavior is reached afterward. Thus, it takes the analysis almost one hour to converge to a state closer to the truth,

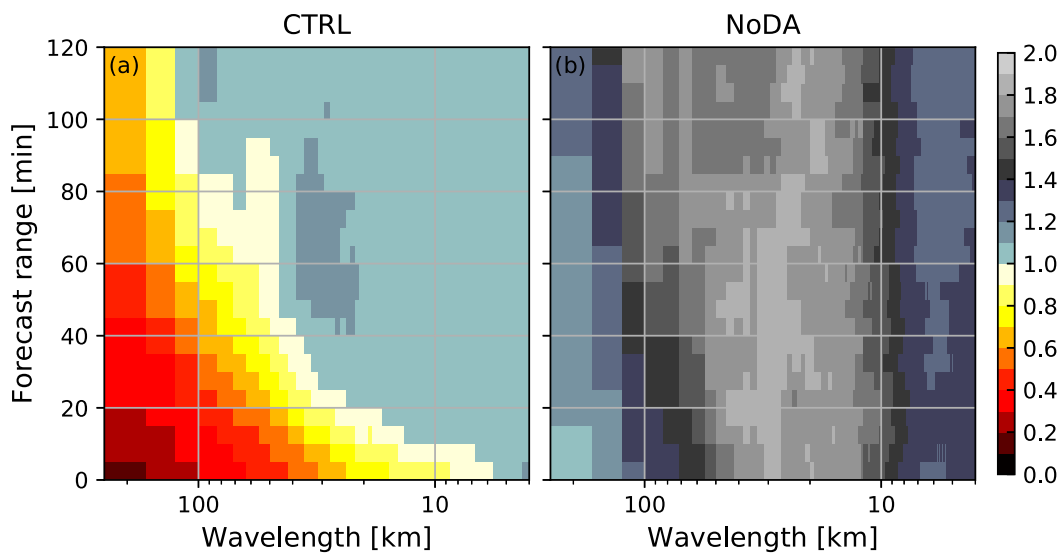


FIG. 6. Time evolution of the forecast ERSP of reflectivity at $z = 5$ km. (a) Forecast ensemble mean initialized from the CTRL analysis ensemble and (b) NoDA forecast.

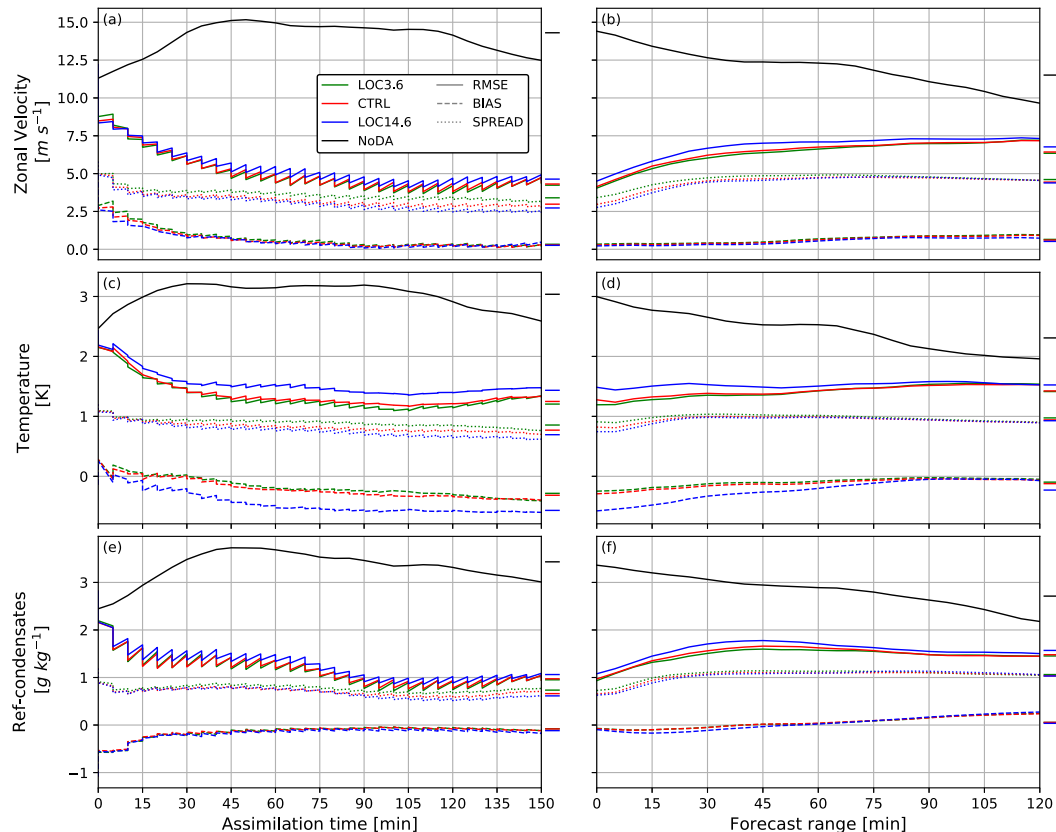


FIG. 7. Space-averaged RMSE (solid line), BIAS (dashed line), and SPREAD (dotted line) for (top) zonal wind, (middle) temperature, and (bottom) Ref-condensates for LOC3.6 (green line), CTRL (red line), and LOC14.6 (blue line). NoDA RMSE (solid black line) is shown for reference. The average includes grid points where reflectivity is greater than 0 dBZ either in the assimilation/forecast experiment or in the nature run. (a),(c),(e) Analysis ensemble mean and (b),(d),(f) forecast ensemble mean. Forecast mean statistics are averaged over three initializations (i.e., 1830, 1900, and 1930 UTC).

which is related to a slow convergence of the error covariance matrix but also possibly linked to a small initial ensemble spread compared to the magnitude of the initial RMSE. All assimilation experiments improve the RMSE compared with NoDA, and at the end of the assimilation period Ref-condensates exhibit higher error reduction than temperature and lower than zonal wind. LOC3.6 and CTRL present the smallest errors for all variables, indicating that assimilation of a smaller number of observations helps to reduce the analysis error. Sensitivity to horizontal localization is more evident in temperature than in zonal wind and Ref-condensates, suggesting that distant convective-scale covariances are poorly estimated, especially for nonobserved variables. A slight improvement of LOC3.6 respect to CTRL is shown toward the end of the assimilation period on zonal wind and Ref-condensates and, more notably, on temperature RMSE.

The expected sensitivity to localization in SPREAD is observed in all variables. The SPREAD value increases

as the localization cutoff radius decreases since fewer observations impact the model state during the update at each grid point. Experiments with greater SPREAD (i.e., LOC3.6 and CTRL) reach smaller RMSE values, suggesting that changes in the ensemble spread can partly explain the sensitivity to the horizontal localization cutoff radius. Moreover, comparison of the SPREAD with the RMSE indicates an underdispersive ensemble, especially during the first 45 min. The ratio between ensemble variance and ensemble mean square error is below the optimal value [defined as $N/(N + 1)$ by Murphy 1988] of 0.98 in the case of our 60-member ensemble. The multiple sources of model error considered in these experiments can partly explain the underestimation of ensemble spread. Figure 7 shows that the SPREAD growth rate is lower than the forecast error growth rate since the latter is probably enlarged by model errors that are not explicitly being accounted for in the evolution of the ensemble perturbations during model integration.

During the whole assimilation period, all experiments produce a dry and cold BIAS and a systematic overestimation of the zonal velocity. Sensitivity to localization is evident in temperature, with the smallest BIAS for LOC3.6, while zonal wind and Ref-condensates present hardly any sensitivity during the entire assimilation period. While zonal velocity and Ref-condensates BIAS are significantly reduced with assimilation time, reaching values close to zero once the analysis has converged to the nature run, an intriguing feature is that radar data assimilation slightly increases the temperature cold BIAS particularly during the spinup of the filter. This behavior is unexpected, and its cause is still unclear. Furthermore, the vertical distribution of the temperature bias presents positive values in the first 4 km and negative values between 6 and 12 km (not shown), suggesting weaker low-level evaporative cooling and mid-high-level release of latent heat, respectively. Additionally, a negative bias in vertical velocity is present during the entire assimilation period and the sensitivity to localization is also observed, with the largest BIAS for LOC14.6 (not shown). Therefore, experiments with larger systematic errors (that also exhibit smaller ensemble spread) lead to a less intense convective system, with weaker updrafts and cold pools.

In short-range forecasts (Figs. 7b,d,f), NoDA also shows worse results than a forecast whose initial state has information on the mesoscale. In particular, the decrease of NoDA RMSE with time is related to the squall line leaving the computational domain toward the end of the forecasted period. Zonal wind and Ref-condensates RMSE exhibit rapid error growth during the first 30 min for all experiments. Afterward, the error growth rate reduces, and the error stabilizes around a value of 7 m s^{-1} and 1.5 g kg^{-1} , respectively. This behavior is observed first in LOC14.6, and later in CTRL and LOC3.6, as the error still increases between 30 and 65 min for these two experiments. As mentioned for NoDA, the halt in error growth might be related to the MCS leaving the domain. Regarding the BIAS in zonal velocity and Ref-condensates, model states are comparable among experiments during the entire forecast period. All experiments exhibit a negative temperature BIAS decrease during the forecast range as implied through the assimilation period when the assimilation of observations seemed to increase the analysis BIAS. Sensitivity among experiments is observed in all variables during the entire forecast time for RMSE, and during the first 75 min for temperature BIAS.

b. Ensemble-mean error spectral decomposition

We examine the influence of the localization cutoff radius on different spatial scales by comparing the

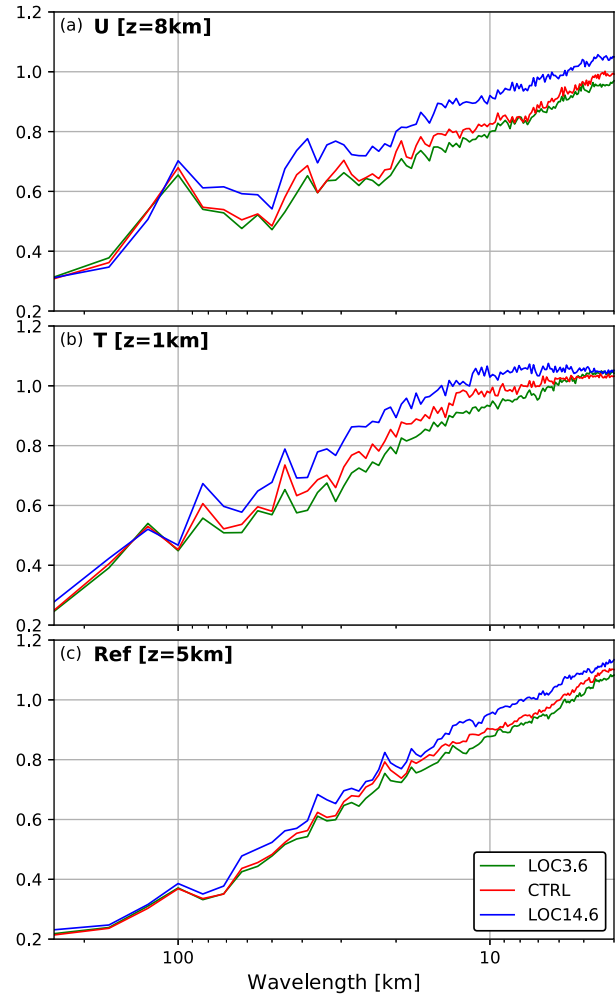


FIG. 8. Analysis mean ERSF for LOC3.6 (green line), CTRL (red line), and LOC14.6 (blue line). (a) Zonal wind at $z = 8 \text{ km}$, (b) temperature at $z = 1 \text{ km}$, and (c) reflectivity at $z = 5 \text{ km}$.

analysis ensemble mean ERSF of LOC3.6, CTRL, and LOC14.6 (Fig. 8). The smallest wavelength from which the analysis mean cannot accurately represent small-scale phenomena (i.e., values higher than unity) differs for variables types and localization scales. For upper-level zonal wind, this cutoff wavelength is approximately 6 km for LOC14.6, and 4 km for CTRL and LOC3.6. For low-level temperature, it is roughly 12–16 km for LOC14.6, 8 km for CTRL, and 6 km for LOC3.6. For midlevel reflectivity, it is nearly 7–8 km for LOC14.6, 6 km for CTRL, and 5–6 km for LOC3.6. As bigger localization radius is used, the analysis cannot accurately depict the smaller scales as well as LOC3.6, and therefore LOC14.6 inadequately represents the small-scale structures of this MCS (e.g., wind gusts, intensity of cold pool and updrafts), which could also explain LOC14.6 producing a less intense convective system. Thus, small-scale

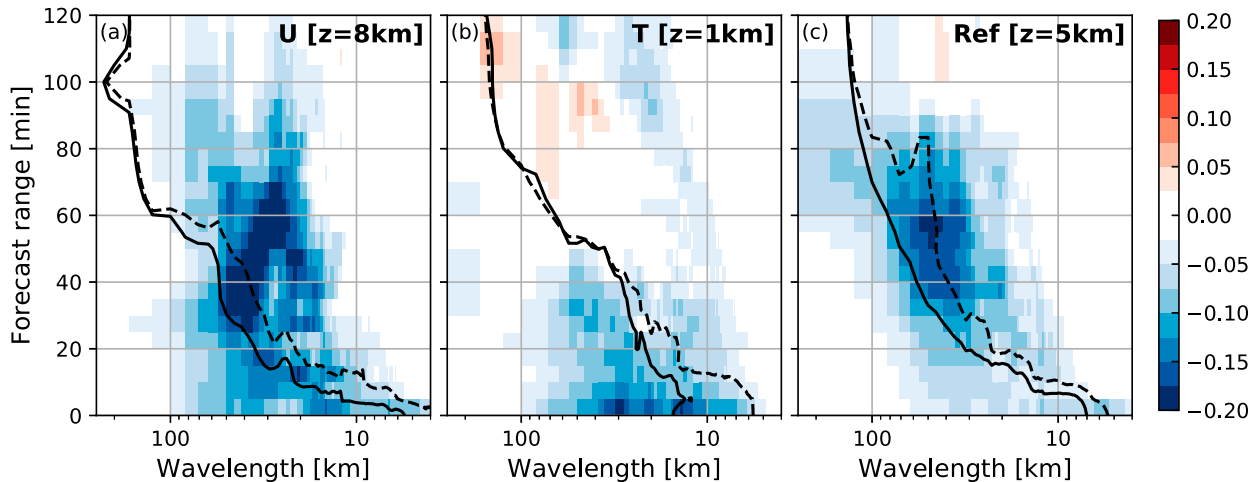


FIG. 9. Time evolution of the difference of forecast ensemble mean ERSR between LOC3.6 and LOC14.6. (a) Zonal velocity at $z = 8$ km, (b) temperature at $z = 1$ km, and (c) reflectivity at $z = 5$ km. The black line shows the forecast error saturation for LOC3.6 (dashed) and LOC14.6 (solid).

phenomena tend to be not properly represented in the analysis ensemble mean when distant observations are used, and this effect is even more evident when we consider model variables that are not directly observed.

The time evolution of the difference of forecast mean error relative spectral power between LOC3.6 and LOC14.6 is shown for zonal wind, temperature, and reflectivity (Fig. 9). Using a smaller localization scale produces more skillful forecasts (as shown by negative values) and extends the limit of predictability to almost 60 min for zonal velocity, 50 min for temperature, and 100 min for reflectivity (as shown by black lines). Moreover, model variables exhibit a considerable contrast on the spatial scales the forecast can accurately represent. For instance, the impact of localization on zonal velocity and temperature is bounded for wavelengths between their respective cutoff wavelengths and 80 km, and up to 40-min forecast lead time. Whereas the impact on reflectivity is recognized for wavelengths larger than 30 km and lead times longer than 40 min.

In particular, the most significant difference on reflectivity error spectral power between LOC3.6 and LOC14.6 is reached at larger scales (i.e., 50–100 km) and around 1-h lead times, possibly associated to a worse representation of the southern region of the squall line in LOC14.6 than in LOC3.6. Additionally, the impact of localization on larger scales is reached at longer lead times, which might be related to the upscale error growth given the nonlinearity in moist processes (Zhang et al. 2003). Therefore, it seems that improving the representation of the smaller scales in the analysis by using a smaller localization radius has a positive effect on the larger scales,

even though the analysis sensitivity to localization on those scales is low.

As previously mentioned, LOC14.6 forecasts a less intense convective system than LOC3.6. Figure 10 shows the ensemble forecast probability of 60-min accumulated precipitation and 10-m wind from the forecast initialized at 1830 UTC and valid at 1930 UTC for LOC3.6 and LOC 14.6. Both experiments can accurately forecast high 60-min precipitation rates and strong wind gusts. However, the southern region of the squall line (i.e., 36°S) exhibits the lowest probabilities in both accumulated rain and surface wind for LOC14.6, suggesting a worse representation of the system intensity and/or position between LOC14.6 ensemble members than LOC3.6.

5. Sensitivity to adaptive multiplicative inflation

We explore the sensitivity to adaptive multiplicative inflation, in particular to RTPS inflation parameter α , by comparing CTRL (i.e., $\alpha = 0.9$) with RTPS0.7 and RTPS0.95. All experiments use a 7.3-km horizontal localization cutoff radius. Figure 11 shows the space-averaged statistics for these experiments. Similar features as in localization experiments are found. However, modifying the RTPS inflation parameter does not seem to have an impact as significant as tuning the horizontal localization radius does, in both analysis and short-range forecasts errors. Almost no difference is observed between CTRL and RTPS0.95 in terms of analysis/forecast errors, but there is a distinct improvement compared to RTPS0.7.

Zonal wind and Ref-condensates RMSE and BIAS present low sensitivity to RTPS inflation parameter α

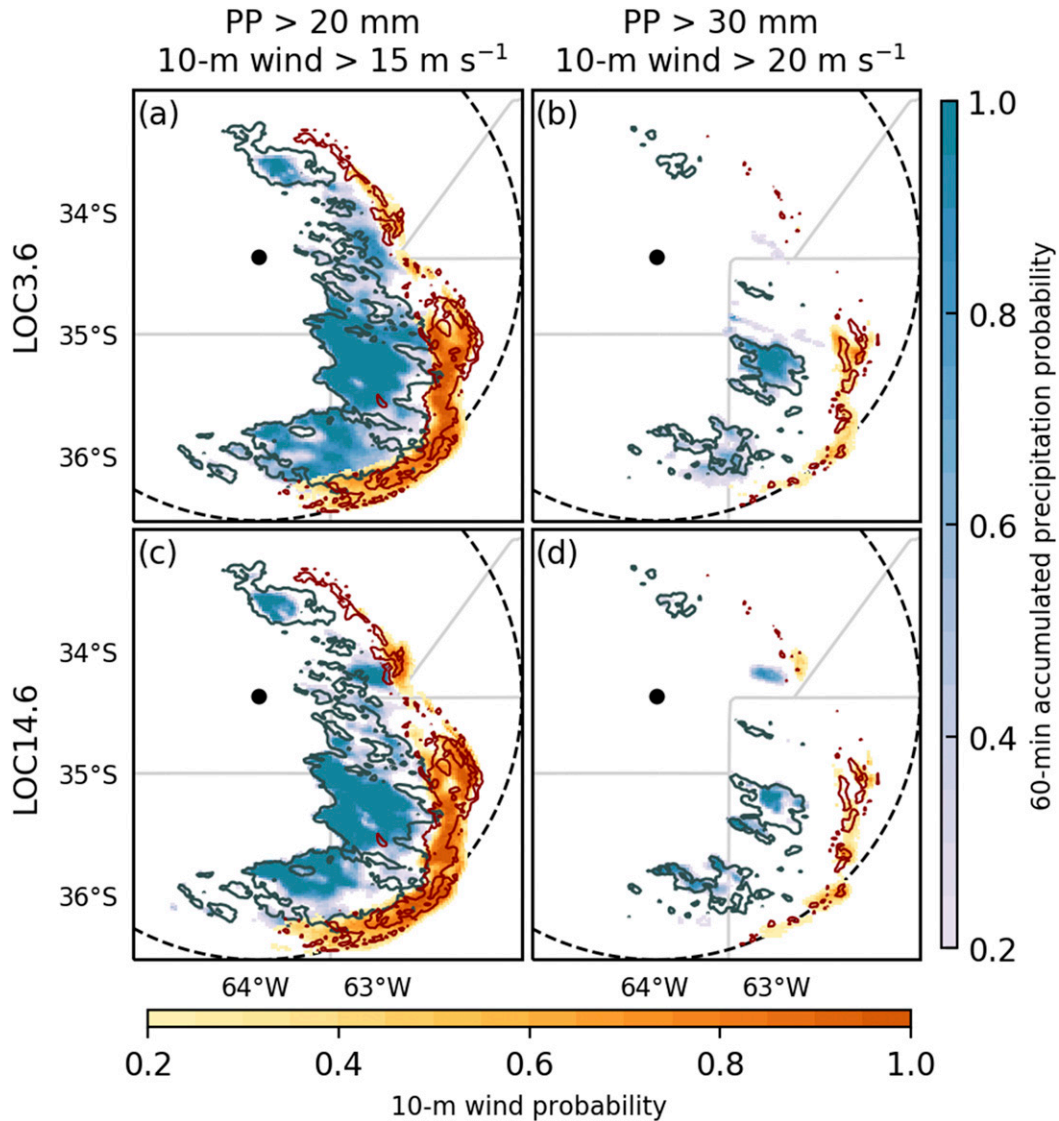


FIG. 10. Probabilistic forecast of 60-min accumulated precipitation (cold colors shaded) and 10-m wind probability (warm colors shaded) for (top) LOC3.6 and (bottom) LOC14.6, for the ensemble forecast initialized at 1830 UTC, valid at 1930 UTC. Contours represent nature's run corresponding threshold value for 60-min accumulated precipitation (blue) and 10-m wind (red). (a),(c) 20-mm precipitation and 15 m s^{-1} wind thresholds; (b),(d) 30-mm precipitation and 20 m s^{-1} wind thresholds.

and smaller errors for RTPS0.95 and CTRL than for RTPS0.7 during the first 75 min from the start of the assimilation. However, the sensitivity is higher for temperature, showing larger improvement for RTPS0.95 and CTRL than for RTPS0.7 during the entire assimilation period. In terms of the SPREAD of the analysis, the expected behavior is observed as SPREAD increases with increasing inflation parameter. Furthermore, unlike fix multiplicative inflation methods in which the ensemble spread is increased in the same amount for all model variables, RTPS inflation method produces a more significant response for zonal wind than for temperature or

Ref-condensates. Both using a larger amount of inflation and reducing the number of observations that impact the update by using a smaller localization radius allow the ensemble spread to become larger. For instance, LOC3.6 (i.e., $\alpha = 0.9$, $r_h = 3.6 \text{ km}$) produces slightly more SPREAD than RTPS0.95 (i.e., $\alpha = 0.95$, $r_h = 7.3 \text{ km}$). Experiments with smaller analyses error only produce better forecasts during the first 30–45 min, and roughly no distinction is observed among experiments afterward. This indicates that the impact of RTPS inflation upon the forecast is shorter than the impact of localization (that lasts at least up to 75 min, cf. Figs. 7b,d,f), suggesting that

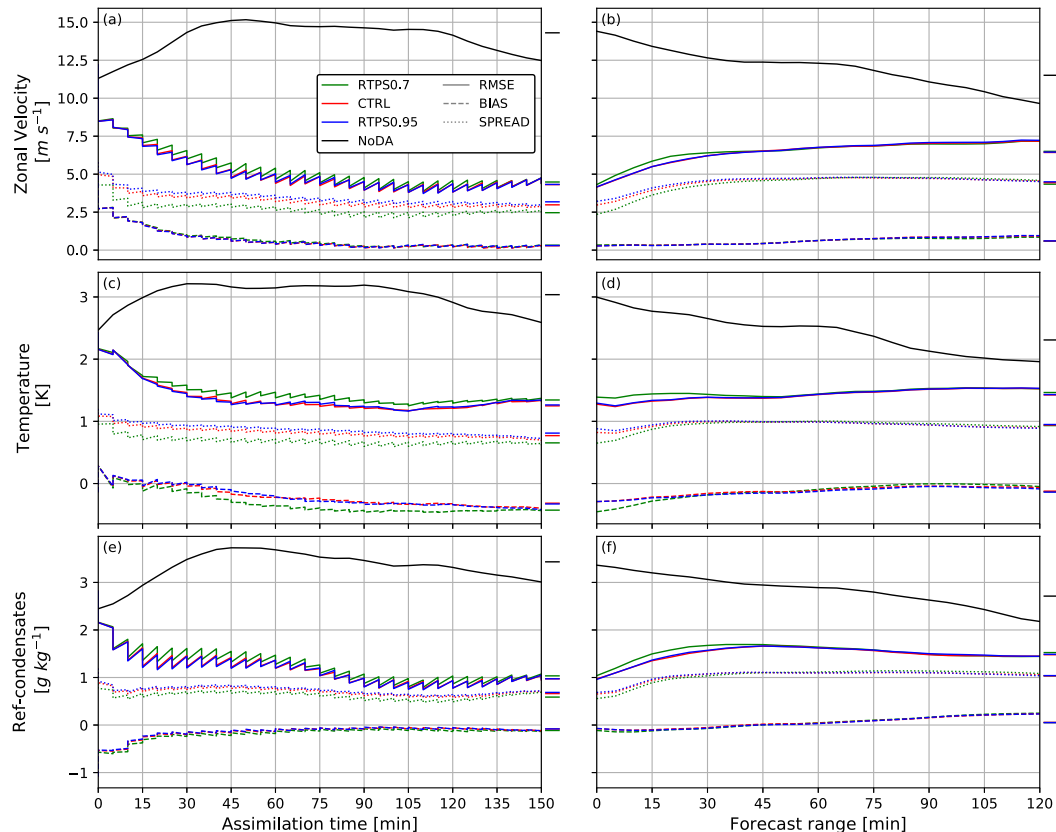


FIG. 11. As in Fig. 7, but for RTPS0.7 (green line), CTRL (red line), and RTPS0.95 (blue line).

for this MCS case, tuning the latter might be essential to improve short-range forecasts errors.

The analysis mean error relative spectral power of each experiment is shown in Fig. 12. Changing the magnitude in which the analysis ensemble is relaxed back to the prior's impacts the spectrum at wavelengths roughly between 10 and 100 km for both zonal wind and temperature and between 8 and 20 km for reflectivity. However, in contrast to localization experiments, almost no sensitivity to α is found at shorter wavelengths as inflation modifies the ensemble spread magnitude but not the spatial scales present in the update. In agreement with the state-space statistics, the main difference among experiments is observed between RTPS0.7 and the rest of the experiments.

Regarding the difference of the forecast mean error relative spectral power between RTPS0.7 and RTPS0.95 (Fig. 13), a lower value of α performs worse as it decreases the forecast skill through most spatial scales (i.e., wavelengths shorter than 50 km for zonal wind and temperature, and between 25 and 100 km for reflectivity) during the first 30 min for zonal velocity and 60 min for temperature and reflectivity. Moreover, the wavelength range in which a change in α impacts the forecast is narrower than in localization experiments, particularly for zonal wind and

temperature. As discussed in section 4b, the impact of adaptive inflation seems to follow the upscale error growth dynamics, that is, a better representation of the smaller scales on the analysis improves the larger scales at longer lead times increasing their predictability.

Joint sensitivity with horizontal localization

As both inflation and localization impact the ensemble spread, additional experiments were performed to explore the joint impact of localizing and inflating the covariance matrix by covering all possible combinations between the horizontal localization cutoff radius and the RTPS inflation parameters used throughout this work.

Figure 14 shows the analysis ensemble mean, time, and space-averaged root-mean-square error for all possible combinations of horizontal localization radius (i.e., LOC3.6, LOC7.3, and LOC14.6) and RTPS inflation parameters (i.e., RTPS0.7, RTPS0.9, and RTPS0.95) for zonal velocity, temperature, and Ref-condensates. For all variables, the largest error is reached for LOC14.6-RTPS0.7, while the smallest error is reached for LOC3.6-RTPS0.9/RTPS0.95, suggesting that experiments with the largest ensemble spread perform better. The optimal localization radius is the smallest one (i.e., LOC3.6) for all

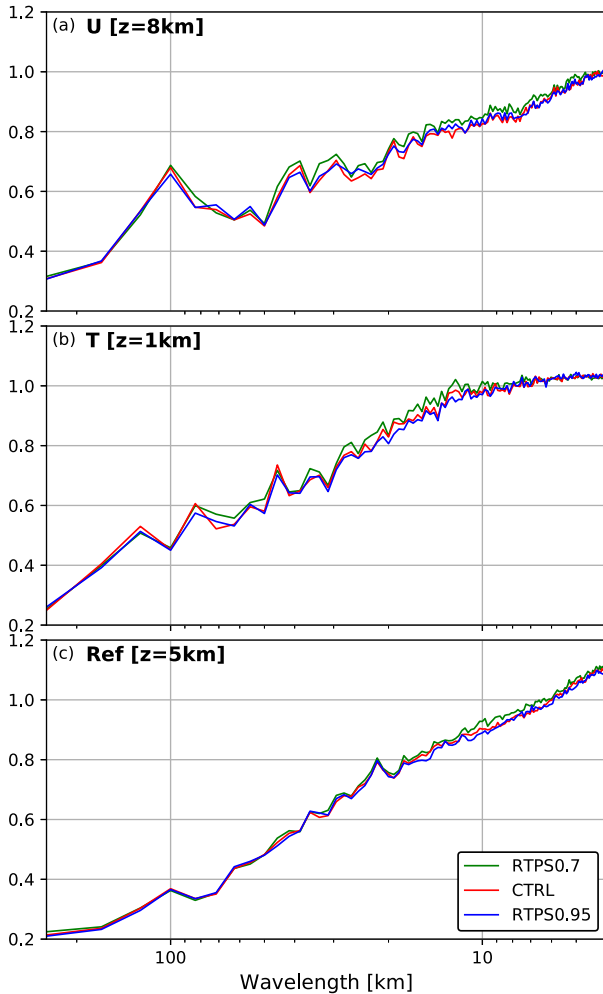


FIG. 12. As in Fig. 8, but for RTPS0.7 (green line), CTRL (red line), and RTPS0.95 (blue line).

inflation parameters, while the optimal RTPS inflation parameter is RTPS0.9 or RTPS0.95 with none/small difference, except for LOC14.6. Moreover, stronger sensitivity to the RTPS parameter is found when LOC14.6 is used instead of LOC7.3 and to the horizontal localization radius when RTPS0.7 is used instead of RTPS0.9.

The results confirm the sensitivity observed in sections 4 and 5, suggesting that the optimal value of one parameter does not vary depending on the value of the other. However, this does not imply that the impact of these two parameters is independent of each other, but instead it can be shown that this effect cannot be properly described with the resolution used to sample the parameter space in this study.

6. Summary and conclusions

In this work, we performed radar data assimilation experiments with simulated observations using the

WRF–LETKF system for a mesoscale convective system case in the central plains of Argentina. Model error was taken into account in these OSSEs as the nature run and the assimilation experiments differ in horizontal resolution, initial and boundary conditions, and the representation of microphysical processes. Based on this model setting, we explored the impact of localization and inflation of the error covariance matrix on both analyses and very short-range ensemble forecasts and quantified these sensitivities through state-space statistics and a spectral decomposition of the analyses/forecast errors.

In agreement with what several studies have been showing in the last two decades (cf. section 1), radar data assimilation produces accurate estimates in the mesoscale and surely adds value to very short-range forecasts compared to high-resolution forecasts initialized from analysis with wrong information on the storm scale. In particular, the impact of radar data assimilation on the analysis/forecast, varies significantly on both variables types and height, likely related to the spatial scales present in the covariance matrix derived from the ensemble that is also limited by the spatial scales the model can actually resolve.

A key aspect of convective-scale forecasting is verification. Given our definition of error metrics and predictability, a double-counting penalty is being applied. Therefore, a feature displaced in space scores worse than either a complete miss or false alarm. Using this type of verification method could lead to disregard a forecast (in terms of its quality), that could have had some value if an object-oriented verification technique would have been applied instead.

Overall the experiments herein suggest that a relatively small horizontal localization cutoff radius (~ 3.6 – 7.3 km) and large inflation RTPS parameter (~ 0.9 – 0.95) produce the most accurate analysis and the 2-h ensemble forecast benefits from this initial state. A stronger initial-condition sensitivity of the 2-h mesoscale ensemble forecast is observed when the localization cutoff radius is modified rather than when the RTPS inflation parameter is, almost doubling the lead times up to which the effect of using a more accurate state to initialize the forecast persists. Another interesting aspect regarding sensitivity to both localization and inflation is that their impact is mainly observed during the first 40-min forecast in spatial scales between 2 and 50 km. For reflectivity, this impact propagates to larger scales with increasing forecast lead time, resembling the initial error upscale growth. Therefore, improvement of smaller scales could generate a positive impact on larger scales even after the smaller-scale error has reached saturation.

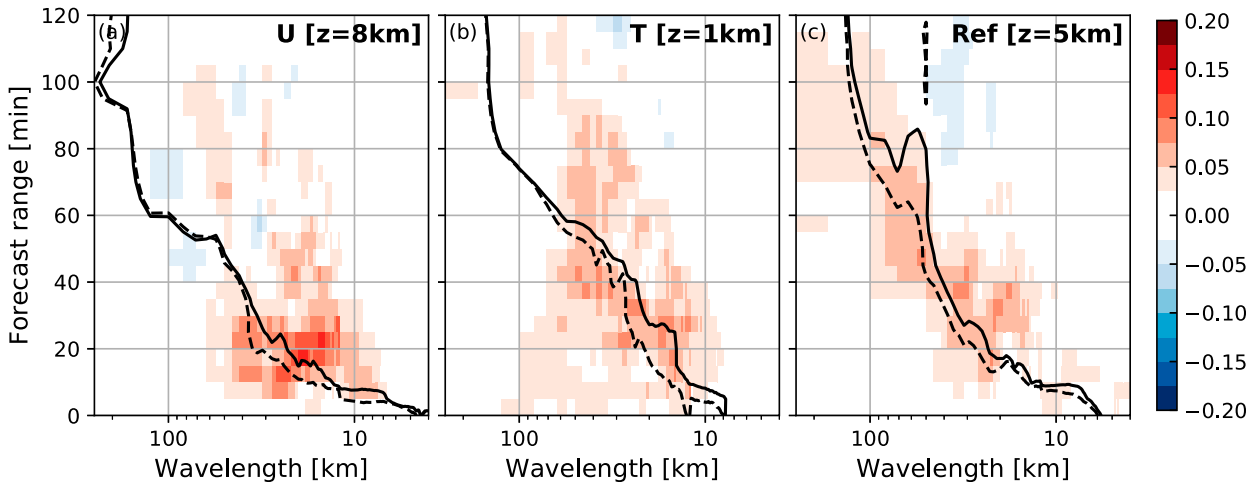


FIG. 13. As in Fig. 9, but for the difference between RTPS0.7 and RTPS0.95.

In terms of the horizontal localization cutoff radius that produce the best results in these experiments, some differences were found with the results obtained by SS13. Even though SS13 simulated a convective organization mode similar to ours, they found smaller RMSE values when using a bigger horizontal localization radius (i.e., 18 km, Fig. 3 in SS13). There are at least two possible explanations to account for the difference in the optimal localization radius between our experiments and the ones presented in SS13. On the one hand, the optimal localization scale depends on the localization technique implemented. As Greybush et al. (2011)

showed, when B localization is used (as in SS13) the optimal localization scale can be 1.5 times larger than in the case of R localization. On the other hand, in the presence of model error, the optimal localization scale is reached at lower values as showed by Metref et al. (2019) in terms of RMSE when using a toy model. Model error reduces forecast skill, thus requiring a greater amount of ensemble spread to describe forecast uncertainty correctly. Therefore, localization scales associated with larger ensemble spread may perform better, especially as the ensemble used herein is underdispersive compared to SS13 ensemble. Another possible issue

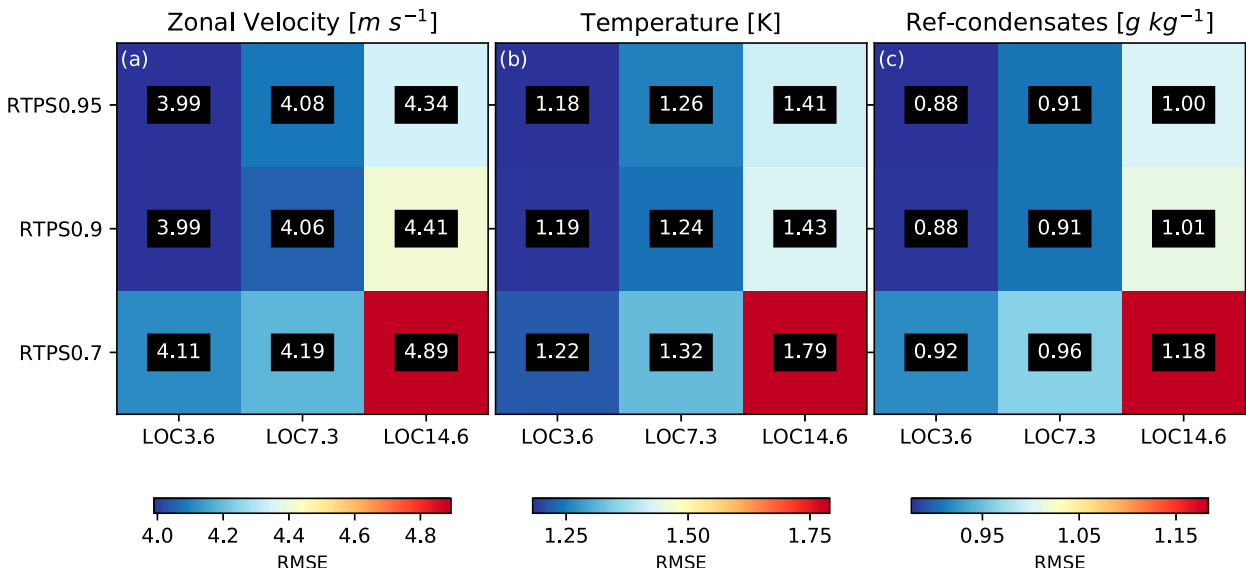


FIG. 14. Analysis mean, time, and space-averaged root-mean-square error for all possible combinations of horizontal localization radius (i.e., LOC3.6, LOC7.3, and LOC14.6) and RTPS inflation parameters (i.e., RTPS0.7, RTPS0.9, and RTPS0.95). Time average is computed between 1830 and 2000 UTC. Space average includes grid points where reflectivity is greater than 0 dBZ either in the assimilation experiment or in the nature run. (a) Zonal wind, (b) temperature, and (c) Ref-condensates.

related to model error has to do with the structure of the background error covariance matrix, and to which extent not including an explicit representation of these errors may lead to an overestimation of those covariances. For instance, if model errors in temperature and reflectivity are not correlated, then an ensemble that does not take model errors into account will tend to overestimate the strength of the covariance between these two variables. Another aspect that should be considered when choosing a localization scale is the impact of sampling errors. Using a larger ensemble size reduces the issue of sampling error and therefore a larger localization scale would be beneficial, as showed by Miyoshi et al. (2014).

The results herein suggest that efforts should be made to optimally tune localization as its importance in accurately representing mesoscale phenomena seems to be considerable. Moreover, vertical localization is also a key parameter in ensemble data assimilation, and exploring the impact of model error upon the optimal value of this parameter should be considered. Furthermore, work is needed to extend these results to other convective modes (e.g., isolated convective cells) and types of synoptic-scale forcing. These aspects are being considered in real data experiments to study the impact of ensemble-based radar data assimilation on short-range predictability of convective systems over central Argentina.

Acknowledgments. The authors thank three anonymous reviewers whose valuable comments/suggestions significantly helped improve and clarify this manuscript. This research was supported by Agencia Nacional de Promoción Científica y Tecnológica (Argentina) Grants PICT 2014-1000 and PICT 2017-2233, Consejo Nacional de Investigaciones Científicas y Técnicas (Argentina) and Universidad de Buenos Aires (Argentina).

REFERENCES

- Aksoy, A., D. C. Dowell, and C. Snyder, 2009: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Storm-scale analyses. *Mon. Wea. Rev.*, **137**, 1805–1824, <https://doi.org/10.1175/2008MWR2691.1>.
- Anderson, J. L., and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741–2758, [https://doi.org/10.1175/1520-0493\(1999\)127<2741:AMCIOT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2).
- Brooks, H. E., J. W. Lee, and J. P. Craven, 2003: The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmos. Res.*, **67–68**, 73–94, [https://doi.org/10.1016/S0169-8095\(03\)00045-0](https://doi.org/10.1016/S0169-8095(03)00045-0).
- Caya, A., J. Sun, and C. Snyder, 2005: A comparison between the 4DVAR and the ensemble Kalman filter techniques for radar data assimilation. *Mon. Wea. Rev.*, **133**, 3081–3094, <https://doi.org/10.1175/MWR3021.1>.
- Dowell, D. C., F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma, supercell: Ensemble Kalman filter experiments. *Mon. Wea. Rev.*, **132**, 1982–2005, [https://doi.org/10.1175/1520-0493\(2004\)132<1982:WATRIT>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<1982:WATRIT>2.0.CO;2).
- Dudhia, J., 1989: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107, [https://doi.org/10.1175/1520-0469\(1989\)046<3077:NSOCOD>2.0.CO;2](https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2).
- Errico, R. M., 1985: Spectra computed from a limited area grid. *Mon. Wea. Rev.*, **113**, 1554–1562, [https://doi.org/10.1175/1520-0493\(1985\)113<1554:SCFALA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1985)113<1554:SCFALA>2.0.CO;2).
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162, <https://doi.org/10.1029/94JC00572>.
- Greybush, S. J., E. Kalnay, T. Miyoshi, K. Ide, and B. R. Hunt, 2011: Balance and ensemble Kalman filter localization techniques. *Mon. Wea. Rev.*, **139**, 511–522, <https://doi.org/10.1175/2010MWR3328.1>.
- Gustafsson, N., and Coauthors, 2018: Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Quart. J. Roy. Meteor. Soc.*, **144**, 1218–1256, <https://doi.org/10.1002/qj.3179>.
- Harnisch, F., and C. Keil, 2015: Initial conditions for convective-scale ensemble forecasting provided by ensemble data assimilation. *Mon. Wea. Rev.*, **143**, 1583–1600, <https://doi.org/10.1175/MWR-D-14-00209.1>.
- 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>.
- Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126, <https://doi.org/10.1016/j.physd.2006.11.008>.
- Kain, J. S., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.*, **43**, 170–181, [https://doi.org/10.1175/1520-0450\(2004\)043<0170:TKCPAU>2.0.CO;2](https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2).
- Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 341 pp.
- Lange, H., and G. C. Craig, 2014: The impact of data assimilation length scales on analysis and prediction of convective storms. *Mon. Wea. Rev.*, **142**, 3781–3808, <https://doi.org/10.1175/MWR-D-13-00304.1>.
- Lien, G.-Y., T. Miyoshi, S. Nishizawa, R. Yoshida, H. Yashiro, S. A. Adachi, T. Yamamura, and H. Tomita, 2017: The near-real-time SCALE-LETKF system: A case of the September 2015 Kanto-Tohoku heavy rainfall. *SOLA*, **13**, 1–6, <https://doi.org/10.2151/SOLA.2017-001>.
- 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).
- Lord, S. J., E. Kalnay, R. Daley, G. D. Emmitt, and R. Atlas, 1997: Using OSSEs in the design of future generation integrated observing systems. Preprints, *First Symp. on Integrated Observing Systems*, Long Beach, CA, Amer. Meteor. Soc., 45–47.
- Maejima, Y., T. Miyoshi, M. Kunii, H. Seko, and K. Sato, 2019: Impact of dense and frequent surface observations on 1-minute-update severe rainstorm prediction: A simulation

- study. *J. Meteor. Soc. Japan*, **97**, 253–273, <https://doi.org/10.2151/JMSJ.2019-014>.
- Matsudo, C., and P. Salio, 2011: Severe weather reports and proximity to deep convection over northern Argentina. *Atmos. Res.*, **100**, 523–537, <https://doi.org/10.1016/j.atmosres.2010.11.004>.
- Metref, S., A. Hannart, J. Ruiz, M. Bocquet, A. Carrassi, and M. Ghil, 2019: Estimating model evidence using ensemble-based data assimilation with localization—The model selection problem. *Quart. J. Roy. Meteor. Soc.*, **145**, 1571–1588, <https://doi.org/10.1002/qj.3513>.
- Mezher, R. N., M. Doyle, and V. Barros, 2012: Climatology of hail in Argentina. *Atmos. Res.*, **114–115**, 70–82, <https://doi.org/10.1016/j.atmosres.2012.05.020>.
- Miyoshi, T., and M. Kunii, 2012: The local ensemble transform Kalman filter with the weather research and forecasting model: Experiments with real observations. *Pure Appl. Geophys.*, **169**, 321–333, <https://doi.org/10.1007/s00024-011-0373-4>.
- , S. Yamane, and T. Enomoto, 2007: Localizing the error covariance by physical distances within a local ensemble transform Kalman filter (LETKF). *SOLA*, **3**, 89–92, <https://doi.org/10.2151/SOLA.2007-023>.
- , K. Kondo, and T. Imamura, 2014: The 10,240-member ensemble Kalman filtering with an intermediate AGCM. *Geophys. Res. Lett.*, **41**, 5264–5271, <https://doi.org/10.1002/2014GL060863>.
- Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16 663–16 682, <https://doi.org/10.1029/97JD00237>.
- Montmerle, T., and C. Faccani, 2009: Mesoscale assimilation of radial velocities from Doppler radars in a preoperational framework. *Mon. Wea. Rev.*, **137**, 1939–1953, <https://doi.org/10.1175/2008MWR2725.1>.
- Murphy, J. M., 1988: The impact of ensemble forecasts on predictability. *Quart. J. Roy. Meteor. Soc.*, **114**, 463–493, <https://doi.org/10.1002/qj.49711448010>.
- Nutter, P., D. Stensrud, and M. Xue, 2004: Effects of coarsely resolved and temporally interpolated lateral boundary conditions on the dispersion of limited-area ensemble forecasts. *Mon. Wea. Rev.*, **132**, 2358–2377, [https://doi.org/10.1175/1520-0493\(2004\)132<2358:EOCRAT>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<2358:EOCRAT>2.0.CO;2).
- Rasmussen, K. L., M. D. Zuluaga, and R. A. Houze, 2014: Severe convection and lightning in subtropical South America. *Geophys. Res. Lett.*, **41**, 7359–7366, <https://doi.org/10.1002/2014GL061767>.
- Schraff, C., H. Reich, A. Rhodin, A. Schomburg, K. Stephan, A. Perriñez, and R. Potthast, 2016: Kilometre-scale ensemble data assimilation for the COSMO model (KENDA). *Quart. J. Roy. Meteor. Soc.*, **142**, 1453–1472, <https://doi.org/10.1002/qj.2748>.
- Schwitalla, T., and V. Wulfmeyer, 2014: Radar data assimilation experiments using the IPM WRF rapid update cycle. *Meteor. Z.*, **23**, 79–102, <https://doi.org/10.1127/0941-2948/2014/0513>.
- Selz, T., and G. C. Craig, 2015: Upscale error growth in a high-resolution simulation of a summertime weather event over Europe. *Mon. Wea. Rev.*, **143**, 813–827, <https://doi.org/10.1175/MWR-D-14-00140.1>.
- Skamarock, W., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., <http://doi.org/10.5065/D68S4MVH>.
- Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **131**, 1663–1677, <https://doi.org/10.1175/2555.1>.
- Sobash, R. A., and D. J. Stensrud, 2013: The impact of covariance localization for radar data on EnKF analyses of a developing MCS: Observing system simulation experiments. *Mon. Wea. Rev.*, **141**, 3691–3709, <https://doi.org/10.1175/MWR-D-12-00203.1>.
- 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).
- , and —, 1998: Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. part ii: Retrieval experiments of an observed Florida convective storm. *J. Atmos. Sci.*, **55**, 835–852, [https://doi.org/10.1175/1520-0469\(1998\)055<0835:DAMRFD>2.0.CO;2](https://doi.org/10.1175/1520-0469(1998)055<0835:DAMRFD>2.0.CO;2).
- , and Coauthors, 2014: Use of NWP for nowcasting convective precipitation: Recent progress and challenges. *Bull. Amer. Meteor. Soc.*, **95**, 409–426, <https://doi.org/10.1175/BAMS-D-11-00263.1>.
- Tewari, M., and Coauthors, 2004: Implementation and verification of the unified Noah land surface model in the WRF model. *20th Conf. on Weather Analysis and Forecasting/16th Conf. on Numerical Weather Prediction*, Seattle, WA, Amer. Meteor. Soc., 14.2a, https://ams.confex.com/ams/84Annual/techprogram/paper_69061.htm.
- Tong, M., and M. Xue, 2005: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. *Mon. Wea. Rev.*, **133**, 1789–1807, <https://doi.org/10.1175/MWR2898.1>.
- Torn, R. D., G. J. Hakim, and C. Snyder, 2006: Boundary conditions for limited-area ensemble Kalman filters. *Mon. Wea. Rev.*, **134**, 2490–2502, <https://doi.org/10.1175/MWR3187.1>.
- Tsai, C.-C., S.-C. Yang, and Y.-C. Liou, 2014: Improving quantitative precipitation nowcasting with a local ensemble transform Kalman filter radar data assimilation system: Observing system simulation experiments. *Tellus*, **66A**, 21804, <https://doi.org/10.3402/tellusa.v66.21804>.
- Weng, Y., and F. Zhang, 2012: Assimilating airborne Doppler radar observations with an ensemble Kalman filter for convection-permitting hurricane initialization and prediction: Katrina (2005). *Mon. Wea. Rev.*, **140**, 841–859, <https://doi.org/10.1175/2011MWR3602.1>.
- Whitaker, J. S., and T. M. Hamill, 2012: Evaluating methods to account for system errors in ensemble data assimilation. *Mon. Wea. Rev.*, **140**, 3078–3089, <https://doi.org/10.1175/MWR-D-11-00276.1>.
- Xue, M., D. Wang, J. Gao, K. Brewster, and K. K. Droegemeier, 2003: The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation. *Meteor. Atmos. Phys.*, **82**, 139–170, <https://doi.org/10.1007/s00703-001-0595-6>.
- Yang, S.-C., E. Kalnay, and T. Miyoshi, 2012: Accelerating the EnKF spinup for typhoon assimilation and prediction. *Wea. Forecasting*, **27**, 878–897, <https://doi.org/10.1175/WAF-D-11-00153.1>.
- , K.-J. Lin, T. Miyoshi, and E. Kalnay, 2013: Improving the spin-up of regional EnKF for typhoon assimilation and forecasting with typhoon Sinlaku (2008). *Tellus*, **65A**, 20804, <https://doi.org/10.3402/tellusa.v65i0.20804>.
- Yano, J.-I., and Coauthors, 2018: Scientific challenges of convective-scale numerical weather prediction. *Bull. Amer. Meteor. Soc.*, **99**, 699–710, <https://doi.org/10.1175/BAMS-D-17-0125.1>.
- Zhang, F., C. Snyder, and R. Rotunno, 2003: Effects of moist convection on mesoscale predictability. *J. Atmos. Sci.*, **60**, 1173–1185, [https://doi.org/10.1175/1520-0469\(2003\)060<1173:EOMCOM>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)060<1173:EOMCOM>2.0.CO;2).
- Zipser, E. J., D. J. Cecil, C. Liu, S. W. Nesbitt, and D. P. Yorty, 2006: Where are the most intense thunderstorms on Earth? *Bull. Amer. Meteor. Soc.*, **87**, 1057–1072, <https://doi.org/10.1175/BAMS-87-8-1057>.