A Heuristic Approach for Precipitation Data Assimilation: Characterization Using OSSEs

ANDRÉS A. PÉREZ HORTAL, ISZTAR ZAWADZKI, AND M. K. YAU

Department of Atmospheric and Oceanic Sciences, McGill University, Montreal, Quebec, Canada

(Manuscript received 6 February 2019, in final form 18 June 2019)

ABSTRACT

We introduce a new technique for the assimilation of precipitation observations, the localized ensemble mosaic assimilation (LEMA). The method constructs an analysis by selecting, for each vertical column in the model, the ensemble member with precipitation at the ground that is locally closest to the observed values. The proximity between the modeled and observed precipitation is determined by the mean absolute difference of precipitation intensity, converted to reflectivity and measured over a spatiotemporal window centered at each grid point of the model. The underlying hypothesis of the approach is that the ensemble members that are locally closer to the observed precipitation are more probable to be closer to the “truth” in the state variables than the other members. The initial conditions for the new forecast are obtained by nudging the background states toward the mosaic of the closest ensemble members (analysis) over a 30 min time interval, reducing the impacts of the imbalances at the boundaries between the different selected members. The potential of the method is studied using observing system simulation experiments (OSSEs) employing a small ensemble of 20 members. The ensemble is produced by the WRF Model, run at a horizontal grid spacing of 20 km. The experiments lend support to the validity of the hypothesis and allow the determination of the optimal parameters for the approach. In the context of OSSE, this new data assimilation technique is able to produce forecasts with considerable and long-lived error reductions in the fields of precipitation, temperature, humidity, and wind.

1. Introduction

Despite all the advances in numerical weather prediction (NWP) and the techniques used to estimate the initial state of the atmosphere, accurate quantitative precipitation forecasts (QPFs) still remains a challenging task. One important limiting factor in the QPF quality is the inaccurate specification of the initial atmospheric state. At present, the most accurate initial conditions (ICs) are estimated by “optimally” combining the available observations with an initial estimate of the actual atmospheric state (background) by a process called data assimilation (DA). In simple terms, the objective of DA is to *obtain the best estimate of the probability density for the actual atmospheric state given the current (and past) observations*.

Since the estimate of the actual atmospheric state is mainly constrained by the observations of the true state, densely spaced observations networks, such as satellites or radars, are a valuable source of information. This has motivated many studies to assimilate precipitation observations over large domains to improve NWP models using variational DA approaches (Koizumi et al. 2005; Lopez 2011; Lopez and Bauer 2007; Kumar et al. 2014) and ensemble Kalman filters (EnKF; Lien et al. 2016; Kotsuki et al. 2017). These methods combine the background information and the observations by minimizing a cost function that is a sum of at least two terms: one term that penalizes the distance to the background mean and another term that penalizes the distance to the observations. The penalties terms are derived assuming Gaussian error statistics for both the background and the observations errors (Lorenc 1986; Hamill 2006).

Although precipitation observations can significantly improve the forecast quality, the assimilation of these observations is a challenging task (Bauer et al. 2011; Errico et al. 2007). One reason is that precipitation is a result of nonlinear moist physical processes, limiting the

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

Corresponding author: Andrés A. Pérez Hortal, andres.perezhortal@mail.mcgill.ca

effectiveness of variational and EnKF methods that relies on linearizations of the observation operators. Another difficulty in the assimilation of precipitation is the non-Gaussian characteristics of the background and observations errors that violate the underlying assumption of normal error statistics. Recently, to alleviate the non-Gaussianity problem, [Lien et al. \(2013, 2016\)](#) and [Kotsuki et al. \(2017\)](#) applied a Gaussian transformation to the precipitation observations with encouraging results.

Besides the variational and Kalman filter approaches, simpler and more economical diabatic initialization (nudging) methods are also used for precipitation DA. These methods modify the model's buoyancy to force the model precipitation toward the observed values by adjusting the humidity or temperature profiles (e.g., [Falkovich et al. 2000](#); [Davolio and Buzzi 2004](#); [Davolio et al. 2017](#); [Jones and Macpherson 1997](#); [Macpherson 2001](#); [Stephan et al. 2008](#); [Bick et al. 2016](#); [Jacques et al. 2018](#)). One of the most popular of these methods is latent heat nudging (LHN), which adjusts the model latent heat release according to the ratio of observed-to-model surface precipitation rate to match the observed precipitation. This method was used successfully in operational setups ([Macpherson 2001](#); [Stephan et al. 2008](#); [Jacques et al. 2018](#)). Nonetheless, the positive impacts in precipitation obtained by the DA typically last for a few hours.

An alternative is another class of ensemble-based DA methods that do not rely on Gaussianity or linearity assumptions: the particle filters (PFs; [van Leeuwen 2009](#)). Their basic idea is to describe the model probability density function (PDF) by a discrete set of model states (ensemble members or particles), instead of approximating the PDF by a Gaussian function as in EnKF or variational methods. In this manner, the evolution of the model PDF is obtained by integrating each member of the ensemble forward in time. When observations become available, each member is weighted according to its proximity to the observations, where the members that are closer to the observations receive a higher weight. Afterward, members with low weight are discarded while multiples copies of high-weight members are kept to describe the posterior PDF.

Although PF methods do not assume any particular error distributions, the standard particle filter that uses simple forms of resampling, requires an ensemble size that increases exponentially with the dimension of the system ([Snyder et al. 2008](#)). If the ensemble size is not large enough, one member receives all the weight after a few analysis steps, resulting in a meaningless posterior PDF. To avoid the collapse of the filter, [Poterjoy \(2016\)](#) introduced a localized implementation of the PF that

operates more efficiently in high-dimensional systems. In a subsequent study, [Poterjoy et al. \(2017\)](#) used this method to assimilate radar data in a cloud-permitting numerical model for an idealized squall line. The authors reported the first successful application of a PF in the context of a weather prediction model using 100 members, yielding more accurate forecast than the ensemble Kalman filter. Nevertheless, despite these promising results, more research is still needed to affirm that this method represents a possible alternative to the other available DA methods.

All the abovementioned precipitation DA techniques have shown different degrees of success. Generally, they force the precipitation toward the observed values during the assimilation windows but these improvements are quickly forgotten (e.g., [Falkovich et al. 2000](#); [Davolio and Buzzi 2004](#); [Jacques et al. 2018](#)). To achieve long-lived improvements, the ideal DA technique should modify the trajectory of a numerical model using the new and partial information on the present state of the atmosphere. This new trajectory should be closer to observed reality than it would be without the assimilation and the DA improvements should be as persistent as the intrinsic atmospheric predictability allows.

Here we will propose a new DA approach that modifies the model state trajectory given information on precipitation at the ground and is suitable for assimilating precipitation fields derived from radar composites. The proposed method is free of any restrictive a priori assumptions that cannot be easily verified using model data and it does not rely on Gaussianity or linearity assumptions. In this new DA method, the analysis state is constructed by selecting the ensemble member that is locally most consistent with the precipitation observations. As described in detail in [section 2](#), this leads to a mosaic of localized ensemble members and hence the name to the proposed method: localized ensemble mosaic assimilation (LEMA). Since the analysis is consisted of a mosaic of different members, each locally closest to the observed precipitation, the discontinuities in the mosaic may introduce imbalances into the model. To reduce the impact of these imbalances the new ensemble forecast is initialized by gradually forcing (nudging) each background member toward the analysis.

The method is studied in the context of observing system simulation experiments (OSSEs) employing a small ensemble of 20 members and four precipitation events, using the WRF Model running at a horizontal grid spacing of 20 km. By means of these experiments we validate the technique and we show that in terms of forecast quality, the LEMA method produces

FIG. 1. Frankenstate DA method summary. (a) Illustration of the local member selection procedure where the black squares denotes different observations windows. (b) Map of closest members to the precipitation observations used in the construction of the Frankenstate. (c) Analysis construction procedure. The analysis is constructed by assembling different columns from the selected members.

considerable and persistent improvement in the forecast of precipitation and the model state variables (potential temperature, vapor mixing ratio, u wind, and v wind).

The article is organized as follows. The new assimilation method is described in section 2. In section 3 we describe the OSSE experiments. This section includes a description of the model, the experimental setup, and the precipitation events used in this study. The results of the experiments are presented in section 4. Finally, in section 5 we present the discussions and conclusions.

2. Localized ensemble mosaic assimilation

The LEMA creates an analysis using only the information in the background ensemble in a direct manner, based on the local proximity of each background member to the precipitation observations. The algorithm constructs a separate analysis for each model's vertical column by assigning the vertical profile of the state variables from the ensemble member with the model precipitation locally closest to the observations.

The local proximity between the observed and the modeled precipitation is measured over a rectangular window centered at the column and over a time period preceding the analysis time (Fig. 1a). In this manner, we construct a mosaic of column states, named “Frankenstate,”¹ where for each vertical column the ensemble member

that is locally closest to the observations provides all the state variables for DA at that location (Figs. 1b,c). As mentioned above, the only underlying assumption in the Frankenstate construction is: *the member that is locally closest to the observed precipitation is more probable to be closer to the “truth” in the state variables*. Precipitation is the final result of the state variables; hence, it is quite intuitive that closeness to the precipitation observations (the Truth) should lead to a greater likelihood of the state variables being closer to the truth state. This is the only hypothesis used in LEMA and it will be verified within the ensemble forecast for the cases where we applied this DA method.

To measure the “local” proximity of a member to the observations we use the Mean Absolute Difference (MAD) of the member's surface precipitation and the observed values, transformed to reflectivity, and computed over a square region of $\Delta x = 820$ km width (41×41 grid points) and over a $\Delta t = 30$ min period preceding the analysis time (see Fig. 1a). The temporal window is used to select a member that is the closest to the observations over a time period, not only at a single time. Although precipitation intensity observations in mm h^{-1} can be used in the distance measure, preliminary experiments showed that transforming the precipitation observations to reflectivity (in dBZ units) yields better results. Therefore, to compute MAD, the precipitation values are converted to reflectivity values using the Marshall–Palmer Z ($\text{mm}^6 \text{m}^{-3}$) = $200R^{1.6}$ relation, with R expressed in mm h^{-1} . This relation expresses precipitation in a logarithmic scale, thus avoiding the high penalty of the extreme precipitation values and produces a better analysis quality than using precipitation values directly.

¹The name is inspired by Mary Shelley's novel where Victor Frankenstein assembles his monster (here analysis) from parts of corpses (here ensemble members) collected from charnel houses and morgues (here from an ensemble forecast).

Therefore, for a given member “ m ,” the distance to the observations around the “ i, j ” horizontal grid point is defined as

$$\text{MAD}_m(i, j) = \frac{1}{N_t N_x N_y} \sum_{t=1}^{N_t} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} |Z_{t,x,y}^{\text{obs}} - Z_{t,x,y}^m|, \quad (1)$$

where Z^{obs} indicates the reflectivity (converted from precipitation observations) and Z^m the m th member reflectivity, both in dBZ units. The subindex x and y denotes the x index and y index of the horizontal grid point inside the observation windows, while the subindex t indicates the observation time. The summation limits N_x and N_y denote the total number of horizontal grid points in the square window over the x and y directions, respectively, while N_t indicates the number of observation times in the temporal window where precipitation observations are available. Since in our idealized experiments the observations are available every 5 min, over the 30 min observation window used the number of observations times is $N_t = 7$ (including the extremes of the interval).

For each horizontal grid point, the ensemble member with the lowest MAD value is considered the member that is “locally closest” to the observations. During the selection process only members with a minimum precipitation coverage of $n_{\text{min}} = 35$ grid points over the spatiotemporal localization window are used. The minimum coverage ensures that only members with MAD values strictly greater than zero are used in the closest member selection. Members that do not meet this criterion are not considered as candidates in the selection process. If over the observations window no background member or the observations exceed the minimum coverage n_{min} , no member is selected as the “closest” and in consequence no information is assigned to that analysis column.

Since the Frankenstate is constructed by a mosaic of information from different members (patches in Fig. 1b), at the patch boundaries the analysis can be incompatible with the model dynamics, producing imbalances in the model. To reduce the impact of these imbalances, instead of initializing the forecast directly from an analysis (direct initialization), the new ensemble forecast is initialized by gradually forcing (nudging) each member of the background toward the Frankenstate. We will refer to these forecasts as “Frankencasts.”

The relaxation toward the Frankenstate is done by adding artificial terms to the model’s prognostic equations:

$$\left[\frac{\partial \phi(t)}{\partial t} \right]_{\text{new}} = \left[\frac{\partial \phi(t)}{\partial t} \right]_{\text{model}} + G[\phi^F - \phi(t)], \quad (2)$$

where $\phi(t)$ indicates a model variable at time t , ϕ^F the Frankenstate, and G is the nudging factor controlling the relative magnitude of the nudging term with respect to other model processes. The first term on the right side of the equation is the original model forcing (advection, Coriolis, diabatic heating, etc.) while the second term denotes the artificial forcing term, proportional to the difference between the model and the Frankenstate. The relaxation toward the analysis is applied over a time period τ preceding the analysis (preforecast nudging). If τ is too small, it forces the model state too strongly toward the Frankenstate, thus reducing the ability of the model to dampen possible imbalances introduced during the initialization. On the other end, if τ is too large, the artificial forcing terms would have a minimal effect on the evolution of the model state. For this study, we use a relaxation period of $\tau = 30$ min which was the optimal value determined through sensitivity experiments. For all the grid points where a closest member is found, we use a nudging factor $G = 1/\tau$, otherwise, $G = 0$ (no nudging). Therefore, the artificial forcing is only applied where the analysis was constructed (not empty), leaving the rest of the domain to evolve without any artificial forcing.

Although the Frankenstate can be constructed with all the model prognostic variables, forcing all of them toward a state containing imbalances may limit the ability of the model to adjust to the introduced instabilities. Previous published studies showed that potential temperature (θ), vapor mixing ratio (q_v), and horizontal winds (U and V) are the most useful variables to initialize the models (Anthes 1974; Stauffer and Seaman 1990). We confirmed that those conclusions hold for our experimental setup by running simple experiments that address the effectiveness of different combinations of state variables in the forecast initialization (see appendix A). Consequently, the Frankenstate is constructed only using these four variables.

3. Data assimilation experiments

a. The model

The numerical model used in this study is the Weather Research and Forecasting (WRF) Model with the Advanced Research WRF (ARW) dynamic solver (WRF-ARW), version 3.7.1 (Skamarock and Klemp 2008). All the simulations were performed using a coarse horizontal grid spacing of 20 km employing 300×180 grid points and 41 vertical levels, covering the contiguous United States and southern Canada. The lateral boundary conditions (LBCs) and initial conditions (ICs) are constructed by downscaling the 1° resolution Global

FIG. 2. Hourly precipitation for the four study cases. The black contours represent the sea level pressure (SLP).

Ensemble Forecast System (GEFS)² data to the WRF grid (20 km grid spacing).

The main physics options used in the experiments are the WRF single-moment 3-class microphysics scheme (WSM3; Hong and Lim 2006), the Yonsei University (YSU) boundary layer scheme (Hong et al. 2006), the Kain–Fritsch (KF) cumulus parameterization (Kain 2004), the Dudhia (1989) shortwave, and Rapid Radiative Transfer Model (RRTM) longwave radiation (Mlawer et al. 1997) schemes. Finally, the computational dynamic time step is 1 min. As for the other WRF parameters, we use the WRF default values.

b. Observing system simulation experiments

The OSSEs are designed to mimic the process of data assimilation. In these experiments, one model simulation is considered the “true” atmosphere and a different set of runs is considered as the background ensemble. Then, a complete set of surface precipitation observations is simulated from the Truth run and they are available every 5 min. To characterize the DA method two experiments

are carried out. One experiment is considered the control where a forecast is produced from the background ensemble without assimilating any observations. For the second experiment, synthetic precipitation observations are assimilated in the background ensemble to produce the forecast. To evaluate the impacts of DA, the forecast errors for precipitation and for the state variables in the second experiment are compared with the control one.

The new DA method is tested on four different precipitation events that took place in 2013. One of the four cases (case A) was selected for an in-depth characterization of the method while the other three cases were to test the robustness of the new DA method to different meteorological situations.

Case A and Case B are widespread precipitation events with precipitation driven by cyclonic systems. In Case A, an extensive squall line over the central United States took place from 1800 UTC 10 April 2013 to 1200 UTC the next day. This event was associated with a midlatitude cyclone over the eastern United States with the eastern line of precipitation caused by a cold front extending in the south–north direction from eastern Texas to central Missouri, and in the west–east direction from Missouri to the south of the New York state (Fig. 2a). In Case B, at 1800 UTC 4 April 2013, three widespread precipitation systems developed around three cyclonic systems over the United States, located in the northwest, northeast, and southeast regions of the domain, respectively (Fig. 2b).

For the two remaining cases, precipitation was produced by several mesoscale convective systems (MCSs) scattered over the United States. Case C occurred at

²The Global Ensemble Forecast System (GEFS), previously known as the GFS Global Ensemble (GENS), is a weather forecast model made up of 21 separate forecasts, or ensemble members. It is produced by the National Centers for Environmental Prediction (NCEP). The data are accessible from the National Oceanic and Atmospheric Administration (NOAA) Operational Model Archive and Distribution System (NOMADS). More information available at <https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-ensemble-forecast-system-gefs>.

0600 UTC 18 May 2013 where the precipitation was located in the northern and southeastern United States (Fig. 2c). For Case D, at 0000 UTC 27 May 2013, several MCSs developed over the central and northwestern United States, along with a cyclonic precipitation system located in the northeastern part of the domain (Fig. 2d).

For each study case, we produce an ensemble forecast of 21 members (0–20) using ICs/LBCs downscaled from the GEFS forecast data. The GEFS forecast used was initialized 24 h prior to assimilating the observations. The WRF runs are initialized 12 h after the GEFS initialization to allow the spread in the GEFS members to grow. In this way, the initial WRF ensemble dispersion is inherited from the GEFS data during the downscaling process. Afterward, an ensemble forecast is created by running the model for 24 h. The member 0 of this ensemble is considered as the “Truth” while the other 20 members are considered the background. A description of the WRF initialization setup is shown in Fig. 3.

For the DA experiment, we use the precipitation observations to construct the Frankenstate 12 h after the WRF Model was initialized (end of the spinup period). Afterward, each member of the background is relaxed toward this analysis to initialize a 12-h ensemble forecast (Frankencasts). Figure 3 shows a summary of the general OSSE setup used in all the study cases while a detailed description of the setup for each case is given in Table 1.

Although the OSSEs represent an efficient manner to characterize the DA methods, the results may not hold in real observation assimilation experiments. In our OSSE setup we use perfect observations (no error) that are available over the entire domain and a “perfect model” scenario (no model errors). Therefore, the model reproduces the true structure and characteristics of the observations and this compatibility may not hold for real DA experiments. The background runs are constructed by downscaling the forecast data from the GEFS members 1 to 20. The GEFS forecasts are produced from ICs that contain different orthogonal perturbations around the member 0 ICs (Wei et al. 2008). The sum of all the perturbations is equal to zero to ensure that the resulting global ensemble ICs are centered around the member 0. As will be shown later, using a background centered around the truth results in an overdispersive ensemble,³ meaning that the ensemble spread overestimates the actual forecast uncertainty.

³ The correct ensemble dispersion is typically considered when a set of observations are indistinguishable from the members in the ensemble forecast, averaged over a sufficiently large number of members and observations made at different times (Fortin et al. 2014).

FIG. 3. Observing system simulation experiments. The horizontal lines indicate the GEFS forecast data initialized from the different analyses (circles) while the numbers inside the circles represent the GEFS member. The simulation period where the model ICs/LBCs are downscaled from the GEFS forecast data are denoted in light blue.

However, real ensemble prediction systems tend to be underdispersive and the ensemble spread does not correctly represent the actual forecast uncertainty (Fortin et al. 2014), especially for storm-scale ensembles (Vié et al. 2011; Clark et al. 2011; Johnson et al. 2014). In addition, in real DA experiments, errors in the precipitation observations will affect the calculation of the local distance of each ensemble members to the observations [MAD; Eq. (1)] and, in consequence, it may affect the selection of the closest member to the observations. In appendix B we show that LEMA is robust to the observation errors if good-quality precipitation estimates are used.

Despite the limitations of our OSSE setup, this controlled environment allows us to find the optimal configuration for the new DA method and to understand how and why it works.

4. Results

We first show in section 4a the characterization of the LEMA method using Case A. This characterization includes the hypothesis testing of the method [section 4a(1)], the determination of the optimal observation window size [section 4a(2)], and the impacts of the DA in the forecast quality [section 4a(3)]. In addition, we present the advantages and limitations of initializing the forecast by relaxing (nudging) the background toward the analysis [section 4a(4)]. In section 4b we show the impacts of the DA for Cases B, C, and D.

a. In-depth study—Case A

1) HYPOTHESIS TESTING

We begin the method’s characterization by testing our hypothesis that the ensemble members that are locally

TABLE 1. Precipitation events.

Case	Type	Simulation start date	Frankenstate date	Simulation end date
A	Synoptic front	1800 UTC 10 Apr	0000 UTC 11 Apr	1200 UTC 11 Apr
B	Cyclonic systems	0600 UTC 4 Apr	1800 UTC 4 Apr	0600 UTC 5 Apr
C	MCSs	1800 UTC 17 May	0600 UTC 18 May	1800 UTC 18 May
D	MCSs	1200 UTC 26 May	0000 UTC 27 May	1200 UTC 27 May

closer to the observed surface precipitation (expressed as reflectivity) are probably closer to the “truth” in the other state variables. To validate this hypothesis, we use the background ensemble to estimate the joint probability of having a decrease in the state variables errors along with the corresponding decrease in precipitation MAD with respect to the locally “closest member” using different observation window sizes. The error for each state variables is measured using the RMSE, an error metric commonly used in the literature, while for the precipitation error we use the same measure of the proximity to the observations used in the Frankenstate construction (MAD over the observation window). The hypothesis is first tested for each individual variable using the smallest observation window possible ($\Delta x = 20$ km and $n_{\min} = 1$). Then, the study is extended to larger window sizes (up to $\Delta x = 1020$ km and $n_{\min} = 35$) using a state error measure that represents the distance of the four variables (U, V, θ, q_v) to the truth by a single parameter.

The RMSE (ε) for a given variable (ϕ) and for the Frankenstate (“ F ”) is computed over each vertical column as

$$\varepsilon_{\phi}^m = \sqrt{\frac{\sum_k^{N_z} [\phi_m(k) - \phi_t(k)]^2}{N_z}},$$

$$\varepsilon_{\phi}^F = \sqrt{\frac{\sum_k^{N_z} [\phi_F(k) - \phi_t(k)]^2}{N_z}}, \quad (3)$$

where $\phi_m(k)$ is the value of a given member “ m ” at the level “ k ” and N_z the number of vertical levels in the troposphere. The Frankenstate and the truth values are denoted by $\phi_F(k)$ and $\phi_t(k)$, respectively. Using the above definitions, the decrease in the column’s RMSE for an ensemble member “ m ,” with respect to the Frankenstate RMSE for the same column, is computed as

$$\text{Decrease in RMSE} = \Delta\varepsilon_{\phi}^m = \varepsilon_{\phi}^m - \varepsilon_{\phi}^F, \quad (4)$$

where ε_{ϕ}^m is the RMSE for a state variable ϕ computed over the column and ε_{ϕ}^F is the Frankenstate RMSE.

Similarly, the decrease in precipitation MAD for the member “ m ,” computed over a square window centered at a given column is

$$\text{Decrease in MAD} = \Delta\text{MAD}_m = \text{MAD}_m - \text{MAD}_F, \quad (5)$$

where MAD_F denotes the error of the closest member to the observations (the Frankenstate). Defined in this way, positive values are associated with a decrease in the errors (positive gains) while negative ones are associated with an error increase (negative gains).

Therefore, the joint probability $p(\Delta\varepsilon_{\phi}, \Delta\text{MAD})$ of decreasing the state variable error by $\Delta\varepsilon_{\phi}$ and the precipitation error by ΔMAD is estimated by the bidimensional histogram of the $\Delta\varepsilon_{\phi}, \Delta\text{MAD}$ pairs (columns), sampled over all the background members. For each member, we exclude areas where the Frankenstate selected that member simply because they provide no information (0% error reduction by definition).

Figure 4 shows the joint probability $p(\Delta\varepsilon_{\phi}, \Delta\text{MAD})$ when a 20 km observation window is used. The joint probability indicates that selecting the closest member to the observed precipitation (positive MAD gains) results, on average, in a decrease in the state variable RMSE (black line in Fig. 4), which validates the basic hypothesis of LEMA. The average RMSE gain is approximately zero for small MAD gains, and increases as the reduction in MAD increases. Nevertheless, there are grid points in the domain where the errors of these variables increase even when the MAD decreases. There are three reasons that can explain this behavior. First, although the state variables in the Kain–Fritsch parameterization determine the surface precipitation in a deterministic way, the inverse is not true. The relationship of observations to the state variables is stochastic (different states can lead to the same precipitation). Second, the limited size of the ensemble could be a source of noise in the probability estimation. Finally, since precipitation is determined by several state variables, precipitation values closer to the truth can be a result of some variables being closer to the truth while the other variables being farther from it. As will be discussed later, we associate these negative gains with the “noise” of the “closest member” selection method.

FIG. 4. Joint probability of the decrease in the RMSE ($\Delta\epsilon$) for selected state variables when the error in precipitation is decreased (ΔMAD) by assigning at each grid point the ensemble member with the smallest MAD respect to the “true” precipitation (20 km observation window). The joint probability is shown for (a) potential temperature θ , (b) vapor mixing ratio q_v , (c) u -wind U , and (d) v -wind V . The black curve indicates the mean value of the decrease in the RMSE, for a variable ϕ , as a function of ΔMAD : $\langle\Delta\epsilon_\phi\rangle = \sum_{\Delta\epsilon_\phi} \Delta\epsilon_\phi p(\Delta\epsilon_\phi, \Delta\text{MAD})$. The probability is computed using 100 bins in each axis direction. The values $\langle\Delta\epsilon_\phi\rangle$ are amplified 100 times to fit the y scale of the joint probability.

To test the validity of the hypothesis for different observations window sizes we introduce a single parameter measure that represents the distance of each model column to the truth. The distance between two column states is measured with a metric similar to the Euclidean distance but with the error at each grid point

normalized by its variance. We define our column state as $\Psi = (\mathbf{U}, \mathbf{V}, \theta, \mathbf{q}_v)$, where each vector inside the parenthesis denotes the values along the column for each variable. Therefore, the column-state distance between a state Ψ and a reference state Ψ_t is computed as

$$\|\Psi - \Psi_t\| = \epsilon_\Psi = \sqrt{\frac{1}{N_z} \sum_{z=1}^{N_z} \frac{(\theta - \theta_t)^2}{\sigma_\theta^2} + \frac{(q_v - q_{v_t})^2}{\sigma_{q_v}^2} + \frac{(U - U_t)^2}{\sigma_U^2} + \frac{(V - V_t)^2}{\sigma_V^2}}, \quad (6)$$

where N_z is the number of vertical levels and σ_ϕ^2 the variance of ϕ at level z , computed using the background. This measure is a simplified version of the Mahalanobis metric (De Maesschalck et al. 2000) ignoring the cross

covariances between different variables and different grid points.

Similar to the definition for a single variable, the decrease for the state error with respect to the

FIG. 5. (a) Joint probability of the decrease in the state error ($\Delta\epsilon_\psi$) and when the error in precipitation is decreased (ΔMAD) by selecting the ensemble member with the smallest difference to “true” precipitation using a 20 km wide observation window. The probability is computed using 100 bins in each axis direction. (b) As in (a), but for an 820 km wide observation window. The black curve indicates the mean value of the decrease in the state error: $\langle\Delta\epsilon_\psi\rangle = \sum_{\Delta\epsilon_\psi} \Delta\epsilon_\psi p(\Delta\epsilon_\psi, \Delta\text{MAD})$. The values $\langle\Delta\epsilon_\psi\rangle$ are amplified 100 times to fit the y scale of the joint probability.

Frankenstate for each member “ m ” and each column is computed as

Decrease in State Error

$$= \Delta\epsilon_\psi = \|\Psi_m - \Psi_t\| - \|\Psi_F - \Psi_t\|, \quad (7)$$

where Ψ_F denotes the Frankenstate column.

Figure 5a shows the joint probability $p(\Delta\epsilon_\psi, \Delta\text{MAD})$ of decreasing the state error by $\Delta\epsilon_\psi$ and the distance to the observations by ΔMAD for a 20 km wide observation window. Similarly, Fig. 5b shows the same joint probability but for the 820 km observation window, providing evidence that the hypothesis hold for the analyzed case. For both observation windows, there is a higher probability of having a decrease in the state error when the error in precipitation is reduced. The average decrease in the state error increases when the MAD reduction becomes more important (black line in Fig. 5).

As in Figs. 4 and 5 also shows that there are grid points of the domains where errors in state variables increase even though the selected member is closer to the observations. For very small ΔMAD values, the probability of having positive and negative state gains are approximately symmetric around the $y = 0$ axis, resulting in an approximately zero average reduction in the state error. We consider this zero-net reduction to be a result of the noise of the best member selection method. Therefore, we will assume that the symmetry for the noise extends to all the ΔMAD values, this implies that the negative state gains due to the noise of the method have an equal but positive counterpart.

2) OPTIMAL LOCALIZATION WINDOW SIZE

We now turn our attention to the impact of increasing the observation window size on the analysis quality. Preliminary experiments showed that the temporal extent of the window played a minor role when compared to the spatial extent. To facilitate the interpretation of the results only the width Δx of the observation window is varied, keeping the temporal interval equal to 30 min.

The change in the observation window is done by spatially smoothing (moving-window average) the absolute error of the reflectivity values computed at the grid resolution (MAD). Since the small scales of MAD are filtered out, the closest members to the observations are sought using only the large-scale component of the MAD. Comparing the two panels of Fig. 5 we see that at the 820 km scale, the joint probability distribution is more compact and more strongly peaked than the one at 20 km resolution. The average state gain increased with decreasing MAD resolution (increasing window size). Furthermore, as illustrated in Fig. 6, increasing the observation window from 20 to 820 km also extends the area over which the Frankenstate can be constructed outside the precipitation area (see reduction of gray area in Fig. 6a with respect to Fig. 6b).

Figure 7 show the Frankenstate total area as a function of the observation windows width (black line). For the 20 km observation window (single gridpoint MAD), the Frankenstate is constructed over 17% of the domain, which is approximately the precipitation area coverage. Due to the spatial smoothing, increasing the observation windows extends the area where the Frankenstate can

FIG. 6. Precipitation distance (MAD) for Member 1 measured using a (a) 20 km and (b) 820 km observation window. The areas where the Frankenstate was left empty are indicated in gray.

be constructed with an increase from 17% of the domain for the 20 km window, to 88% for the 820 km window. Extending the Frankenstate area increases the total area with positive gains in the state error (red shaded areas in Fig. 7). However, it also increases the area with negative impacts (blue shaded region in Fig. 7). As mentioned previously, the negative gain (error increase) results from the “noise” of the method and we assumed that the same noise producing negative gains has also an equal counterpart that results in positive gains. The difference between the positive and the negative gain areas represents the “net gain” area. This net gain area increases with the window size, reaching a maximum value of 30% of the domain for an 820 km window (green line in Fig. 7). Since a larger window size does not extend the net gain area, for our DA experiments, we considered 820 km as the optimal window size.

Note that reducing the errors at larger scales (using a large observation window) has a double benefit. On the one hand, it extends the area of the Frankenstate, by which a larger portion of the domain is directly benefitted from LEMA (Fig. 7). Also, it increases the magnitude of the benefit as seen by comparing Figs. 5a and 5b.

Figure 8 shows the maps of decrease in the state error with respect to the Frankenstate for selected members

when the optimal window size is used. In general terms, regions with positive gain cover a higher portion of the domain than negative gain regions, in agreement with Fig. 7. Moreover, the regions with positive and negative gains are different for each ensemble member as a result of the stochastic component of the closest member selection method (noise). For each member, the decrease in the state error (positive or negative) can be interpreted as different perturbations that are applied to each member during the nudging initialization (relaxation toward the Frankenstate). In the next section, we will show that the combined effect of these perturbations and the nudging initialization introduces differences among the Frankencast members that grow with time increasing the spread in the ensemble.

3) IMPACT OF DA IN FORECAST QUALITY

In the preceding section, we showed that the Frankenstate constructed using the optimal window size produces a decrease in the state error in approximately 60% of the domain. This section extends the previous study by analyzing the improvements on the forecast quality when the new ensemble forecast (Frankencasts) is initialized by gradually relaxing (nudging) each background member toward the Frankenstate. To evaluate the impact in the precipitation forecast quality we compare

FIG. 7. Frankensate total area as a function of the observation window width (black line). The gray shadow area denotes the areas where the Frankensate was left empty. The blue and the red shadow areas indicate the area with negative or positive gains, respectively. The green line indicates the net gain area (positive minus negative gains areas). The positive and negative gains area represents the average value over all the background members. For each member, areas where the Frankensate selected that member as the “closest” are excluded because they do not provide any information (0% error reduction by definition).

the reflectivity observations against the model values using the RMSE and three contingency scores.

In terms of the precipitation forecasts, the LEMA method produces improvements in the RMSE of $\sim 15\%$ that persist for our entire 12 h forecast (Figs. 9a,b). Consistent improvements are also obtained in the classical contingency scores (Wilks 1995) such as equitable threat score (ETS; Figs. 9c and 9d), probability of detection (POD; Figs. 9e and 9f), and false alarm ratio (FAR; Figs. 9g and 9h). The detection threshold used in the contingency score is 0.3 mm h^{-1} ($\sim 14 \text{ dBZ}$).

For potential temperature, vapor mixing ratio and u winds, on average, a $\sim 25\%$ reduction in the RMSE is achieved by the end of the nudging period (Fig. 10). Note that most of the error reduction obtained by the DA is still present after 12 h of forecast. Similar results are found for the v wind (not shown).

In addition to the reduction in the forecast error, in an optimal ensemble prediction system (EPS) the ensemble spread should correctly quantify the forecast uncertainties due to errors in the ICs. Therefore, to maintain the correct ensemble dispersivity during the forecast, the differences between the ensemble members should increase with time to capture the growth of the forecast errors. We will show that the presented DA method creates spread in the ensemble due to the combined effect of the nudging initialization and the different corrections applied to every member (Fig. 8).

Figures 11a and 11b show the background and the Frankencast ensemble spread for potential temperature, vapor mixing ratio, u wind, and reflectivity (transformed from surface precipitation). Although nudging each background member toward the analysis severely reduces the ensemble spread, some differences among the members persist. Part of this dispersion can be attributed to the areas where the analysis is not constructed (the Frankensate’s empty regions). Over those regions, each background member evolves without any artificial forcing, maintaining the spread values similar to the background ones. Another source of spread arises from the use of dynamic initialization (nudging) gradually forcing the model toward the analysis, but due to the original model tendencies, the Frankensate with a residual is achieved at the end of the nudging period. In addition, only 4 state variables are nudged, while the other variables can contribute to the spread.

To evaluate if the ensemble dispersion correctly represents the uncertainty in the forecast, we compute the ratio of the RMSE of the ensemble mean and the average ensemble spread. The correct dispersivity in the ensemble results when the RMSE of the ensemble mean matches the ensemble spread, with the spread computed as (Fortin et al. 2014):

$$\text{spr}_{\phi} = \sqrt{N^{-1} \sum_i \left\{ (M-1)^{-1} \sum_m [\phi_m(i) - \overline{\phi(i)}]^2 \right\}},$$

where ϕ denotes a variable, $\phi_m(i)$ its value at grid point “ i ” for the m th member, and the overbar indicates the ensemble average. The total number of grid points and ensemble members are denoted by N and M , respectively. For potential temperature and horizontal winds, the background ensemble is overdispersive (dispersivity > 1) during the entire forecast period (Fig. 11c). After the DA, the dispersivity of the Frankencast ensemble is severely reduced, becoming underdispersive (dispersivity < 1). Nevertheless, due to the fast growth of the perturbations, the ensemble becomes overdispersive after 4 h. For vapor mixing ratio and reflectivity (Fig. 11d) the ensemble also becomes underdispersive after the DA but, contrary to potential temperature and u wind, 12 h is required to regain the correct dispersivity. This slow dispersivity development for vapor and reflectivity can be explained by the slow spread growth for those variables (Fig. 11b), compared with the spread increase for u wind and potential temperature (Fig. 11a).

4) NUDGING VERSUS DIRECT INSERTION FORECAST INITIALIZATION

In this section, we show the benefits and the limitations of using nudging to initialize the forecast by

FIG. 8. Maps showing the decrease in the state error with respect to the Frankenstate for members 1, 5, 10, and 15. The Frankenstate is constructed using the optimal window size ($820 \text{ km} \times 820 \text{ km} \times 30 \text{ min}$).

comparing this initialization method with the direct insertion of the Frankenstate in the background (no nudging). To this end, the Case A experiment is repeated twice. In one of the experiments, the forecasts are initialized directly from an ensemble of analyses created by replacing in each background member the values of potential temperature, humidity, and horizontal winds from the Frankenstate. In areas where the Frankenstate is not constructed the background members are left unchanged. In the other experiment, the ensemble forecast is produced by relaxing each background member toward the Truth state by forcing the same 4 variables used in the Frankenstate over the entire domain, representing the best possible DA.

Since the Frankenstate is constructed using the information from the ensemble members in a direct

manner, the regions where a member is the closest to the precipitation are “locally” in balance with the model dynamics (patches in Fig. 1b). Nonetheless, at the boundaries of those regions where a transition between different selected members occurs (patch boundaries in Fig. 1b), imbalances may be produced in the model.

A commonly used parameter to measure imbalances is the magnitude of the time derivative of surface pressure, averaged over the model domain (e.g., Stauffer and Seaman 1990; Bick et al. 2016):

$$S_t = \left\langle \left| \frac{dP_s}{dt} \right| \right\rangle, \quad (8)$$

where P_s is the surface pressure and $\langle \rangle$ denotes the domain average. These pressure perturbations can be

FIG. 9. DA impacts on precipitation forecasts for Case A. (left) The background (colors) and the Frankencasts (black) errors, measured by (a) reflectivity RMSE, (c) ETS, (e) POD, and (g) FAR. (right) The error improvement achieved by each Frankencast member, with respect to the corresponding background member. The improvements corresponds to: (b) reflectivity relative decrease in RMSE, (d) increase in ETS ($ETS_F^m - ETS_B^m$), (f) increase in POD ($POD_F^m - POD_B^m$), and (h) decrease in FAR ($FAR_B^m - FAR_F^m$). In the above equations, the subindexes "F" and "B" indicate the error for the "Frankencast" and "Background" and "m" the member number. A threshold of 0.3 mm h^{-1} is used for the ETS, POD, and FAR computations. Gray shaded areas indicate the nudging period.

FIG. 10. Case A: DA impacts on forecast quality for (a),(b) potential temperature, (c),(d) vapor mixing ratio, and (e),(f) u wind. (a),(c),(e) The RMSE for each variable for each background (colors) and Frankencasts (black) member. (b),(d),(f) The relative decrease in RMSE for each Frankencast member, with respect to the corresponding background member used in the forecast initialization. Gray shaded area indicates the nudging period.

related to mass adjustments that restore the model balance through the continuity equation, with higher values indicating the presence of imbalances.

The pressure perturbations introduced by nudging the background members toward the Frankenstate are an order of magnitude smaller than the ones produced by the direct insertion initialization (Fig. 12). Moreover, in the dynamic initialization, the pressure perturbations have the same order of magnitude as the background values (no DA), indicating that the nudging reduces model shocks to the minimum.

In addition, the nudging initialization produces a higher decrease in the state RMSE than the direct insertion initialization (Fig. 13). A possible reason for this is that during the forcing period the model has time to adjust to the new state (avoiding model shocks and information rejection) and is able to propagate the information to other variables and to areas outside the Frankenstate where no nudging was applied.

It is important to remark that even in the best possible scenario (nudging toward the Truth values in the four state variables used in the Frankenstate) the maximum

FIG. 11. (a),(b) Ensemble spread and (c),(d) dispersivity for Case A. (left) The results for potential temperature (black) and u wind (red), and (right) the results for reflectivity (black) and vapor mixing ratio (red). In all the plots, the solid lines indicate the background spread or dispersivity while the dashed ones denote the Frankencast ones.

attainable reduction in the RMSE is 55%–70% for the state variables and 15% for reflectivity (see dot–dashed lines in Fig. 13). This is an expected result since only four variables are nudged while the others adjust freely to the changes in the state, and also since the artificial forcing needs to compete with the actual model tendencies.

b. Other test cases

In the preceding section, we performed an in-depth characterization and tuning of the method using a single study case (Case A). To evaluate the performance of the DA method under different meteorological situations, we extended the study of the impacts of DA in the forecast quality for three more precipitation events (Case B, C, and D in Table 1).

Figure 14 shows the reduction in the forecast errors for precipitation, potential temperature, and u wind, for Cases B, C, and D. The results are similar for vapor mixing ratio and v wind (not shown). In general terms, the results for Case A hold for the other analyzed cases: the assimilation of precipitation observations produces considerable and persistent improvement in precipitation and in the state variables.

For Case B, an approximately constant reflectivity improvement of 7% (averaged over the ensemble) is

present during the entire 12 h forecast (Fig. 14a) which can be associated with the long-lived reduction in the state error (Figs. 14d,g). Similar results were obtained for Case D (Figs. 14c,f,i).

For Case C, improvements in reflectivity remain between 5% and 20% over the entire forecast period (Fig. 14b). The error reduction in the state variables is similar to the ones obtained for Case B (Figs. 14e,h), with the exception that the assimilation increased the error potential temperature RMSE for one of the ensemble members (member 16, dashed green line in Fig. 14e). Nevertheless, since the error in the other state variables is reduced (Fig. 14h for u wind, similar results for v wind and vapor mixing ratio), persistent improvements are obtained for that member in precipitation.

5. Discussion and conclusions

We present a simple data assimilation technique named localized ensemble mosaic assimilation (LEMA) that in the context of OSSE gives considerable improvements of ~15% in precipitation forecast and better for the state variables that persists up to 12 forecast hours. These improvements in precipitation can be associated with the long-lived reduction in the state

FIG. 12. Ensemble mean domain-averaged magnitude of the time derivative of the surface pressure for the background (black line), the forecast dynamically initialized by nudging the background members toward the Frankenstate (Frankencast, blue), and the forecast by direct initialization using the Frankenstate (red).

variable errors (potential temperature, vapor mixing ratio, u wind, and v wind).

The heuristic approach presented has three fundamental components. The first element is the construction of an analysis, named “Frankenstate,” by assigning to each grid point the information from the ensemble member that is locally closest to the precipitation observations. The second important aspect is how the “local proximity” is measured: using the large-scale component

of the mean absolute difference (MAD) between the modeled and the observed precipitation, by computing MAD over a rectangular window centered at the column (820 km window). The last component of LEMA is the initialization of the new ensemble forecast by gradually forcing (nudging) each member of the background toward the Frankenstate (“Frankencasts”), thus reducing the impact of any imbalances present in the Frankenstate.

LEMA was introduced here in its more simple terms—as it actually originated—based on the construction of a mosaic of ensemble members each chosen so that at every pixel the contributing member is the closest one to observations. In our experience this simple description always elicits the comment “why nobody said it before”. This is indicative of the very intuitive concept behind LEMA. However, the idea that proximity to precipitation observations must lead to better proximity in state variables is not necessarily obvious nor simple. In more rigorous terms LEMA is defined by the joint probabilities $p(\Delta\epsilon_\phi, \Delta\text{MAD})$, where ϕ is a state variable, obtained from model data, such as shown in Fig. 4. These joint probabilities play in LEMA the same role as the background covariance matrices and the observation operator in classical assimilation methods: the transmission of the information from observations to the state variables. However, in LEMA no assumptions of Gaussianity nor linearity are necessary to propagate the information. It is all rolled into a single container of the joint probability distribution.

As discussed in section 4, the joint pdf has a “noise” component that generates increases in state errors.

FIG. 13. Value of the reduction in the forecast RMSE by different forecast initialization methods, averaged over all the ensemble members, for (a) potential temperature θ (red) and u -wind U (blue), and (b) vapor mixing ratio q_v (red) and reflectivity dBZ (blue). The solid lines indicate the forecasts initialized by nudging each background member toward the Frankenstate while the dashed lines denote the initialization by direct insertion of the Frankenstate in the background. Finally, the dot-dashed line denotes the forecast initialized by nudging the background toward the truth state.

FIG. 14. DA impacts on forecasts quality for (left) Case B, (middle) Case C, and (right) Case D. The impacts are expressed as the reduction in RMSE for (a)–(c) reflectivity (dBZ), (d)–(f) potential temperature (θ), and (g)–(i) u wind (U) for each ensemble member. For all plots, the gray shaded area indicates the nudging period.

But LEMA's decrease in state variables' errors is due to the net effect of $p(\Delta\epsilon_\phi, \Delta\text{MAD})$ that is not readily seen in 2D joint pdf. To better understand the mechanism of the transfer of information in LEMA the expected value of ϵ_ϕ is shown in Fig. 15. Here, the graph for potential

temperature is shown but similar figures hold for the other three variables. Figure 15 shows that the smoothing of MAD from 20 to 820 km decreases the range of ΔMAD by about one order of magnitude but in both cases around 2.5 decades of ΔMAD values are present.

FIG. 15. Average transmission of the information from ΔMAD to ε_{θ} for the 20 km and the 820 km observations windows implicit in the joint probability $p(\Delta \varepsilon_{\theta}, \Delta \text{MAD})$. The black curve indicates the mean value of the decrease in the RMSE $\Delta \varepsilon$ as a function of ΔMAD : $\langle \Delta \varepsilon_{\theta} \rangle = \sum_{\Delta \varepsilon_{\theta}} \Delta \varepsilon_{\theta} p(\Delta \varepsilon_{\theta}, \Delta \text{MAD}) = \sum_{\Delta \varepsilon_{\theta}} \Delta \varepsilon_{\theta} p(\Delta \varepsilon_{\theta} | \Delta \text{MAD}) p(\Delta \text{MAD})$. The green line shows the mean RMSE reduction normalized by the ΔMAD probability: $\langle \Delta \varepsilon_{\theta} \rangle / p(\Delta \text{MAD})$. The values $\langle \Delta \varepsilon_{\theta} \rangle$ are amplified 100 times to fit the same y scale as the normalized expectation values (green line).

From the joint pdf alone (as in Figs. 4 and 5) it is not apparent that a greater ΔMAD leads to greater decrease in the state variable errors. However, the computation of the expected value of $\Delta \varepsilon_{\theta}$, $\langle \Delta \varepsilon_{\theta} \rangle = \sum_{\Delta \varepsilon_{\theta}} \Delta \varepsilon_{\theta} p(\Delta \varepsilon_{\theta}, \Delta \text{MAD})$ clearly shows that on the average this is the case for both resolutions, although at 820 km resolution the relationship extends over a larger range of ΔMAD . The gain in state variables could originate in an average positive relationship between ε_{θ} and ΔMAD as well as a greater number of grid points with larger ΔMAD . The black curve measures the average effect of both. In the $\langle \Delta \varepsilon_{\theta} \rangle / p(\Delta \text{MAD})$ (the green curve) the second effect is eliminated by computing the expectation value of $\Delta \varepsilon$ using the conditional probability $p(\Delta \varepsilon_{\theta} | \Delta \text{MAD})$ instead of the joint probability. Hence, the green line shows the average relationship $\Delta \varepsilon_{\theta} = f(\Delta \text{MAD})$ per grid point of the Frankenstate. We now see clearly the better performance of the transmission of information at the 820 km resolution: even for very low gain in MAD there is a gain in state variables and for the last decade of ΔMAD the power-law relationship is quite steeper at the lower resolution.

It is interesting to note that by matching model and observation at the large scales leads to a more effective transmission of information from precipitation to the state variables than at the 20 km scale. On the average, the $\langle \Delta \varepsilon_{\theta} \rangle$ curve indicates that the transmission of information follows a power-law growth between decrease in MAD

and decrease in state variables error for most of the range of MAD. By construction we directly decrease the error in precipitation but the better performance of LEMA when the 820 km window is used arise entirely from the mosaic of the ensemble members generated by the model at the original 20 km grid spacing scales. This result appears to be consistent with the idea that decreasing errors at the large scales has an enhanced benefit at smaller scales (Durrán et al. 2013).

The results summarized in Fig. 15 hold under less ideal conditions of the experiments. When truth is selected at the edge of the ensemble instead of its center (as in a forecast error) relationships similar to Fig. 15 are obtained but with lower gains in state variables.

In this work, only one variable is assimilated. Nevertheless, LEMA can be extended to more observed variables in which case the multidimensional joint probability distributions must take into account the relative reliability of each assimilated variable. But even with only radar-derived precipitation as assimilated measurement it may be useful to consider the uncertainty of the transformation of dBZ to precipitation rate. Lee et al. (2007) have shown that this uncertainty can be used to generate an ensemble of R fields from a single Z field and in this manner can generate various Frankencasts and hence increases the number of ensemble members in the model forecast. This will be explored in a future work.

Moreover, since only precipitation is used, in cases of weather with little or no precipitation in the domain, it is possible that none of the ensemble members will have a precipitation coverage over the observation windows that exceed the minimum threshold needed to construct the Frankenstate (e.g., more than 35 grid points with precipitation for the 820 km observation windows). In this case there is not enough information to construct the Frankenstate and therefore each member of the ensemble evolves without any modification.

The other question that may arise is whether weights should be assigned to the ensemble members contributing to the mosaic of Frankenstate. In fact, the ensemble members that are closer to the truth have automatically a greater weight by the fact that they cover a greater number of pixels (more and/or larger tiles of the mosaic). Nevertheless, it could be interesting to explore the idea of constructing an ensemble of Frankenstates using different combinations of the locally closest “ n ” members instead of just using the closest one.

Finally, LEMA belongs to the class of DA methods based purely on the information in the ensemble forecasts with no additional physical constraints. This is a fundamental weakness: the method relies on the full coverage of reality by the ensemble. In real situations with considerable model errors and underdispersive

backgrounds we know that this is not to be the case. One possible way to overcome this limitation is to augment the ensemble used to construct the analysis in LEMA with members from ensemble forecasts initialized at different times (lagged forecasts) as well as states at different times close to the analysis time. Nevertheless, if reality is outside the set of analogs of the NWP ensemble, if the forecast does not predict precipitation or the predicted location of precipitation is totally off, the ensemble LEMA will not be able to correct it.

Acknowledgments. The research reported here has been supported by the NSERC/Hydro-Quebec Industrial Research Chair program. We acknowledge the comments and suggestions of Dr. Dominik Jacques and three anonymous reviewers for their help in improving the paper.

APPENDIX A

Variable Importance in Nudging Initialization

Previous studies showed that potential temperature (θ), vapor mixing ratio (q_v), and horizontal winds (U and V) are the most useful variables when the forecasts are initialized by nudging the background toward an analysis (Anthes 1974; Stauffer and Seaman 1990). In this appendix, we confirm these results for our model configuration using simple experiments.

The experiments are similar to the ones used in the OSSE setup for Case A (described in section 3b). The only difference is that only one member is used as background (member 1). To characterize the impacts, two simulations are carried out. We consider one of the simulations as the control by producing a forecast from the background ensemble without assimilating any observations. For the second simulation, a forecast is initialized by relaxing the background toward the truth. This experiment is repeated four times, nudging different combinations of state variables that have the greatest impact for the initialization. The nudging is performed at every grid point in the domain using a relaxation time $\tau = 30$ min. The following combinations of variables are used for each experiment:

- potential temperature (θ),
- vapor mixing ratio (q_v),
- horizontal winds (U and V), and
- all the variables above (θ , q_v , U , and V).

The effectiveness of the nudging initialization is measured by the decrease in the RMSE for each variable between the simulation with nudging and the background simulation [Eq. (4)].

Figure A1 shows the decrease in the RMSE for each variable and each experiment. At the end of the nudging period, nudging only θ (Fig. A1a) results in a $\sim 60\%$ improvement on that variable. But, after the nudging took place, most of the improvements are lost within 4 h. Only nudging θ induces small improvements in the horizontal winds ($\sim 10\%$). Similar results occur when we only adjust q_v (Fig. A1b). But only marginal corrections are propagated to the other variables (θ , U , and V). Nudging θ and q_v gives similar results to nudging only θ (not shown), with also similar improvements in q_v than the ones shown in Fig. A1b. Better results are obtained when we nudge the horizontal winds (Fig. A1c). The error reduction at the end of the nudging period is $\sim 60\%$ for the nudged variables. In the subsequent hours, the errors in the winds (U and V) and the mass fields (θ and q_v) balance each other during the first 4 h. After that period, a 20%–30% gain persists for the rest of the forecast. The best results are obtained when the four variables are nudged (Fig. A1d) with 55%–70% of improvements after the nudging. In contrast to the other experiments, most of these improvements persist during the forecast period. Nudging other prognostic variables like vertical velocity, surface pressure, and geopotential height has little impacts on the forecast quality (results not shown).

In conclusion, even under the ideal conditions where we nudge toward the truth on the entire domain, no single variable is sufficient to obtain considerable impacts. Potential temperature, vapor mixing ratio, and horizontal winds must be adjusted to obtain a persistent improvement in the forecast quality.

APPENDIX B

Sensitivity to Observations Errors

Precipitation estimates from ground-based radar networks are typically derived from 2D reflectivity composites constructed by combining the reflectivity (Z) data from the different radars onto a common grid. Nonetheless, the precipitation composites derived from radar data are affected by different sources of errors (e.g., Zawadzki 1984; Joss et al. 1990). The main sources of error are the uncertainty associated with the Z – R transformations used to convert reflectivity into surface rainfall and the radar range-dependent errors (Berenguer and Zawadzki 2008). Moreover, these errors have a strong dependence with range and that their structure is scale dependent (Berenguer and Zawadzki 2009). For lower heights the errors in the Z – R relation dominates, while for elevated observations the range-dependent errors are dominant.

FIG. A1. Nudging variable importance experiments. Impacts on forecast quality using the following combinations of variables during the nudging initialization: (a) potential temperature θ , (b) vapor mixing ratio q_v , (c) horizontal winds (U and V), and (d) θ , q_v , U , and V . For all plots, the gray shaded area indicates the nudging period.

In LEMA, these errors will affect the calculation of the local distance of each ensemble members to the observations [MAD; Eq. (1)] and, in consequence, it may affect the selection of the closest member to the observations. To evaluate the sensitivity of LEMA to the errors in the observations we compared expectation values of the decrease in the state error $\Delta\epsilon_\Psi$ using perfect (no errors) and “imperfect” observations (observations + Gaussian noise). The errors were simulated by adding an unbiased Gaussian noise field with $\sigma_{\text{noise}} = 3$ dBZ standard deviation to the reflectivity observations used in the MAD computation (converted from the observed precipitation). For simplicity, no spatial correlations on the error field were considered. The standard deviation for the errors was taken from Berenguer and Zawadzki (2008), and can be considered as an upper bound for low-altitude observations.^{B1}

Figure B1 shows the expectation values $\Delta\epsilon_\Psi$ computed using the conditional probability $p(\Delta\text{MAD}|\Delta\epsilon_\Psi)$

^{B1} In Berenguer and Zawadzki (2008), Eqs. (3) and (6) expressed the errors in dBR units. Those equations can be converted to the corresponding reflectivity errors using the Marshall–Palmer relationship.

for the 820 km observation windows (i.e., the expectation values are normalized by $p(\Delta\text{MAD})$). The expectation values for $\Delta\epsilon_\Psi$ are only slightly affected by the simulated errors, indicating that the LEMA hypothesis still holds: the member that is locally closest to the observed precipitation is more probable to be closer to the “truth” in the other state variables. The low sensitivity to the simulated errors can be explained by the large observation windows used, where the MAD values are mostly dominated by the actual differences between the observed (truth) and the modeled precipitation patterns, with the observations error contribution playing a minor role.

The above results show that LEMA is robust with respect to the observations errors considered here. Nevertheless, as is shown in Berenguer and Zawadzki (2008), for elevated radar observations located far away from the radar, the error in the precipitation estimates can be considerable larger than the ones considered here. Under those circumstances, one can simply exclude the regions with very low-quality observations (Lopez 2011) or assign quality index (weight between 0 and 1) to the individual observations (Jacques et al. 2018) and use them as weights on each term in the MAD computation. Another alternative is to use the observations

FIG. B1. Expectation values of the decrease in the state error $\Delta\epsilon_\psi$ computed using the conditional probability $p(\Delta\epsilon_\psi|\Delta\text{MAD})$, where ΔMAD denotes the decrease in the precipitation error. The expectation values are computed as $\langle\Delta\epsilon_\psi\rangle/p(\Delta\text{MAD}) = \sum_{\Delta\epsilon_\psi} \Delta\epsilon_\psi p(\Delta\epsilon_\psi|\Delta\text{MAD}) = \sum_{\Delta\epsilon_\psi} \Delta\epsilon_\psi p(\Delta\epsilon_\psi, \Delta\text{MAD})/p(\Delta\text{MAD})$ for the 820 km observation windows using perfect observation (black line) and observations with error (red).

errors to select different members that fall inside the error's margins. The selected members can be either averaged to construct the analysis or be used to generate an analysis ensemble by using different members combinations. However, using the observation's errors in LEMA is outside the scope of this study and is left as future work.

Nevertheless, it should be pointed out that in LEMA both the model and the observations errors affect the MAD between model and observations. Although errors in radar precipitation have been reasonably assessed (e.g., Berenguer and Zawadzki 2008, 2009), model errors are a more difficult problem that requires a great deal of consideration. In real situations, it is likely that these model errors represent a dominant factor in the MAD computation, especially due to the uncertainties in the convective parameterization.

REFERENCES

- Anthes, R. A., 1974: Data assimilation and initialization of hurricane prediction models. *J. Atmos. Sci.*, **31**, 702–719, [https://doi.org/10.1175/1520-0469\(1974\)031<0702:DAAIOH>2.0.CO;2](https://doi.org/10.1175/1520-0469(1974)031<0702:DAAIOH>2.0.CO;2).
- Bauer, P., G. Ohring, C. Kummerow, and T. Auligne, 2011: Assimilating satellite observations of clouds and precipitation into NWP models. *Bull. Amer. Meteor. Soc.*, **92**, ES25–ES28, <https://doi.org/10.1175/2011BAMS3182.1>.
- Berenguer, M., and I. Zawadzki, 2008: A study of the error covariance matrix of radar rainfall estimates in stratiform rain. *Wea. Forecasting*, **23**, 1085–1101, <https://doi.org/10.1175/2008WAF2222134.1>.
- , and —, 2009: A study of the error covariance matrix of radar rainfall estimates in stratiform rain. Part II: Scale dependence. *Wea. Forecasting*, **24**, 800–811, <https://doi.org/10.1175/2008WAF2222210.1>.
- Bick, T., and Coauthors, 2016: Assimilation of 3D radar reflectivities with an ensemble Kalman filter on the convective scale. *Quart. J. Roy. Meteor. Soc.*, **142**, 1490–1504, <https://doi.org/10.1002/qj.2751>.
- Clark, A. J., and Coauthors, 2011: Probabilistic precipitation forecast skill as a function of ensemble size and spatial scale in a convection-allowing ensemble. *Mon. Wea. Rev.*, **139**, 1410–1418, <https://doi.org/10.1175/2010MWR3624.1>.
- Davolio, S., and A. Buzzi, 2004: A nudging scheme for the assimilation of precipitation data into a mesoscale model. *Wea. Forecasting*, **19**, 855–871, [https://doi.org/10.1175/1520-0434\(2004\)019<0855:ANSFTA>2.0.CO;2](https://doi.org/10.1175/1520-0434(2004)019<0855:ANSFTA>2.0.CO;2).
- , F. Silvestro, 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>.
- De Maesschalck, R., D. Jouan-Rimbaud, and D. L. Massart, 2000: The Mahalanobis distance. *Chemo. Intell. Lab. Syst.*, **50**, 1–18, [https://doi.org/10.1016/S0169-7439\(99\)00047-7](https://doi.org/10.1016/S0169-7439(99)00047-7).
- 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).
- Durrán, D. R., P. A. Reinecke, and J. D. Doyle, 2013: Large-scale errors and mesoscale predictability in Pacific Northwest snowstorms. *J. Atmos. Sci.*, **70**, 1470–1487, <https://doi.org/10.1175/JAS-D-12-0202.1>.
- Errico, R. M., P. Bauer, and J.-F. Mahfouf, 2007: Issues regarding the assimilation of cloud and precipitation data. *J. Atmos. Sci.*, **64**, 3785–3798, <https://doi.org/10.1175/2006JAS2044.1>.
- Falkovich, A., E. Kalnay, S. Lord, and M. B. Mathur, 2000: A new method of observed rainfall assimilation in forecast models. *J. Appl. Meteor.*, **39**, 1282–1298, [https://doi.org/10.1175/1520-0450\(2000\)039<1282:ANMOOR>2.0.CO;2](https://doi.org/10.1175/1520-0450(2000)039<1282:ANMOOR>2.0.CO;2).
- Fortin, V., M. Abaza, F. Anctil, and R. Turcotte, 2014: Why should ensemble spread match the RMSE of the ensemble mean? *J. Hydrometeorol.*, **15**, 1708–1713, <https://doi.org/10.1175/JHM-D-14-0008.1>.
- Hamill, T. M., 2006: Ensemble-based atmospheric data assimilation. *Predictability of Weather and Climate*, T. Palmer and R. Hagedorn, Eds., Cambridge University Press, 124–156.
- Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42** (2), 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>.
- Jacques, D., D. Michelson, J.-F. Caron, and L. Fillion, 2018: Latent heat nudging in the Canadian regional deterministic prediction system. *Mon. Wea. Rev.*, **146**, 3995–4014, <https://doi.org/10.1175/MWR-D-18-0118.1>.
- Johnson, A., and Coauthors, 2014: Multiscale characteristics and evolution of perturbations for warm season convection-allowing precipitation forecasts: Dependence on background flow and method of perturbation. *Mon. Wea. Rev.*, **142**, 1053–1073, <https://doi.org/10.1175/MWR-D-13-00204.1>.
- Jones, C., and B. Macpherson, 1997: A latent heat nudging scheme for the assimilation of precipitation data into an operational mesoscale model. *Meteor. Appl.*, **4**, 269–277, <https://doi.org/10.1017/S1350482797000522>.

- Joss, J., A. Waldvogel, and C. G. Collier, 1990: Precipitation measurement and hydrology. *Radar in Meteorology*, D. Atlas, Ed., Amer. Meteor. Soc., 577–606.
- 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).
- Koizumi, K., Y. Ishikawa, and T. Tsuyuki, 2005: Assimilation of precipitation data to the JMA mesoscale model with a four-dimensional variational method and its impact on precipitation forecasts. *SOLA*, **1**, 45–48, <https://doi.org/10.2151/sola.2005-013>.
- Kotsuki, S., T. Miyoshi, K. Terasaki, G.-Y. Lien, and E. Kalnay, 2017: Assimilating the global satellite mapping of precipitation data with the Nonhydrostatic Icosahedral Atmospheric Model (NICAM). *J. Geophys. Res. Atmos.*, **122**, 631–650, <https://doi.org/10.1002/2016JD025355>.
- Kumar, P., C. M. Kishtawal, and P. K. Pal, 2014: Impact of satellite rainfall assimilation on Weather Research and Forecasting model predictions over the Indian region. *J. Geophys. Res. Atmos.*, **119**, 2017–2031, <https://doi.org/10.1002/2013JD020005>.
- Lee, G. W., A. W. Seed, and I. Zawadzki, 2007: Modeling the variability of drop size distributions in space and time. *J. Appl. Meteor. Climatol.*, **46**, 742–756, <https://doi.org/10.1175/JAM2505.1>.
- Lien, G.-Y., E. Kalnay, and T. Miyoshi, 2013: Effective assimilation of global precipitation: Simulation experiments. *Tellus*, **65A**, 19915, <https://doi.org/10.3402/tellusa.v65i0.19915>.
- , T. Miyoshi, and E. Kalnay, 2016: Assimilation of TRMM multisatellite precipitation analysis with a low-resolution NCEP global forecast system. *Mon. Wea. Rev.*, **144**, 643–661, <https://doi.org/10.1175/MWR-D-15-0149.1>.
- Lopez, P., 2011: Direct 4D-Var assimilation of NCEP Stage IV radar and gauge precipitation data at ECMWF. *Mon. Wea. Rev.*, **139**, 2098–2116, <https://doi.org/10.1175/2010MWR3565.1>.
- , and P. Bauer, 2007: “1D+4DVAR” assimilation of NCEP Stage-IV radar and gauge hourly precipitation data at ECMWF. *Mon. Wea. Rev.*, **135**, 2506–2524, <https://doi.org/10.1175/MWR3409.1>.
- Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **112**, 1177–1194, <https://doi.org/10.1002/qj.49711247414>.
- Macpherson, B., 2001: Operational experience with assimilation of rainfall data in the Met Office mesoscale model. *Meteor. Atmos. Phys.*, **76**, 3–8, <https://doi.org/10.1007/s007030170035>.
- 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>.
- Poterjoy, J., 2016: A localized particle filter for high-dimensional nonlinear systems. *Mon. Wea. Rev.*, **144**, 59–76, <https://doi.org/10.1175/MWR-D-15-0163.1>.
- , R. A. Sobash, and J. L. Anderson, 2017: Convective-scale data assimilation for the weather research and forecasting model using the local particle filter. *Mon. Wea. Rev.*, **145**, 1897–1918, <https://doi.org/10.1175/MWR-D-16-0298.1>.
- Skamarock, W. C., and J. B. Klemp, 2008: A time-split non-hydrostatic atmospheric model for weather research and forecasting applications. *J. Comput. Phys.*, **227**, 3465–3485, <https://doi.org/10.1016/j.jcp.2007.01.037>.
- Snyder, C., T. Bengtsson, P. Bickel, and J. Anderson, 2008: Obstacles to high-dimensional particle filtering. *Mon. Wea. Rev.*, **136**, 4629–4640, <https://doi.org/10.1175/2008MWR2529.1>.
- Stauffer, D. R., and N. L. Seaman, 1990: Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I: Experiments with synoptic-scale data. *Mon. Wea. Rev.*, **118**, 1250–1277, [https://doi.org/10.1175/1520-0493\(1990\)118<1250:UOFDDA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1990)118<1250:UOFDDA>2.0.CO;2).
- Stephan, K., S. Klink, and C. Schraff, 2008: Assimilation of radar-derived rain rates into the convective-scale model COSMO-DE at DWD. *Quart. J. Roy. Meteor. Soc.*, **134**, 1315–1326, <https://doi.org/10.1002/qj.269>.
- van Leeuwen, P. J., 2009: Particle filtering in geophysical systems. *Mon. Wea. Rev.*, **137**, 4089–4114, <https://doi.org/10.1175/2009MWR2835.1>.
- Vié, B., O. Nuissier, and V. Ducrocq, 2011: Cloud-resolving ensemble simulations of mediterranean heavy precipitating events: Uncertainty on initial conditions and lateral boundary conditions. *Mon. Wea. Rev.*, **139**, 403–423, <https://doi.org/10.1175/2010MWR3487.1>.
- Wei, M., Z. Toth, R. Wobus, and Y. Zhu, 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. *Tellus*, **60A**, 62–79, <https://doi.org/10.1111/j.1600-0870.2007.00273.x>.
- Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences: An Introduction*. International Geophysics Series, Vol. 59, Elsevier, 467 pp.
- Zawadzki, I., 1984: Factors affecting the precision of radar measurement of rain. *Preprints, 22nd Int. Conf. on Radar Meteorology*, Zurich, Switzerland, Amer. Meteor. Soc., 251–256.