A heuristic data assimilation strategy for precipitation observations

Andrés A. Pérez Hortal

Department of Atmospheric & Oceanic Sciences
McGill University
Montreal, Canada

November 2020

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Doctorate of Philosophy in Atmospheric Sciences.

ABSTRACT

Although precipitation is a difficult quantity to forecast, it is also a well-observed quantity. Currently, several networks of ground-based radars provide precipitation estimates over large regions. These precipitation estimates offer a valuable source of information to improve precipitation forecasts because they contain detailed information on the spatial and temporal scales of these events. Therefore, the assimilation of precipitation observations into Numerical Weather Prediction (NWP) models has a significant potential to improve the forecast quality.

This thesis presents a heuristic technique for the assimilation of precipitation observations based on the physical notion that, in an ensemble, the ensemble members closer to the observed precipitation are more likely to be closer to the "truth" in their state variables than the other members. This new approach, called "Localized Ensemble Mosaic Assimilation" (LEMA), initializes the new ensemble forecast by relaxing the background members towards an analysis mosaic constructed using, for each grid column, the information from the ensemble member that is locally the closest to the observed precipitation.

A series of data assimilation experiments are carried out to characterize LEMA under different scenarios. Under the ideal conditions where the spread of the ensemble used for the analysis construction captures the forecast uncertainties, LEMA can produce considerable and long-lived improvements in the forecast quality. However, the effectiveness of LEMA is reduced when the ensemble spread underestimates the forecast errors. To overcome this limitation, the spread of the ensemble used to construct the analysis mosaic is expanded by including model states at different times and also states from forecasts initialized at different times. The expansion of the ensemble spread provides a better representation of the forecast uncertainties, resulting in a better LEMA performance.

A shortcoming of LEMA is that relaxing the background members towards a single analysis

severely reduces the ensemble spread in the regions where precipitation observations are assimilated. To remedy this, a new version of LEMA is proposed. This revised version constructs an ensemble of analysis mosaics using additional members close to the observed precipitation, instead of only the closest member. The information in the analysis ensemble is also used to correct the ensemble mean and the spread of the background ensemble. A series of cycled data assimilation experiments using only precipitation observations show that the revised LEMA method produces forecasts of a comparable quality to those produced by the Local Ensemble Transform Kalman Filter approach.

RÉSUMÉ

Malgré que la précipitation soit une quantité difficile à prévoir, celle-ci est bien documentée et observée. De nos jours, plusieurs systèmes de radars au sol offrent des estimés de précipitations sur de vastes régions. Ces estimations de précipitations sont une source d'information importante pour les prévisions de précipitations étant donné qu'elles offrent des données sur les échelles spatiales et temporelles de ces évènements. Par conséquent, l'assimilation des précipitations observées dans les modèles de prévision numérique du temps offre un grand potentiel d'amélioration sur la qualité de prévision.

Cette thèse présente une technique heuristique pour l'assimilation des observations des précipitations, en se basant sur la notion physique que, dans un ensemble, les membres de cet ensemble qui sont plus près de la précipitation observée ont plus de chance d'être plus proche de la "vérité" dans leurs variables d'état que les autres membres. Cette nouvelle approche, appelée "Localized Ensemble Mosaic Assimilation" (LEMA), initialise la nouvelle prévision d'ensemble en relaxant les membres d'ébauche vers une mosaïque d'analyses en attribuant à chaque colonne du modèle l'information du membre de l'ensemble qui est localement le plus proche aux observations de précipitations.

Une série d'expériences sont menées afin de caractériser LEMA sous différents scenarios. Dans des conditions idéales, où l'écart de l'ensemble utilisé pour la construction d'analyses réussit à capturer l'incertitude des prévisions, LEMA peut produire d'importantes améliorations de longue durée sur la qualité de prévision. Cependant, l'efficacité de LEMA est réduite quand l'écart de l'ensemble sous-estime les erreurs de prévision. Afin de surmonter cette limitation, l'écart de l'ensemble utilisé pour la construction de la mosaïque d'analyses est élargi en incluant les états du modèle à différents moments ainsi que les états des prévisions initialisés à différents moments. L'expansion de l'écart de l'ensemble offre une meilleure représentation des incertitudes

réelles de prévision, résultant en une meilleure performance de LEMA.

Une lacune de LEMA est que la relaxation des membres d'ébauche vers une seule analyse réduit visiblement l'écart de l'ensemble dans les régions où les observations de précipitations sont assimilées. Pour y remédier, une nouvelle version de LEMA est proposée. Cette version révisée construit un ensemble de mosaïques d'analyses en utilisant des membres additionnels proches aux observations de précipitations, au lieu de seulement utiliser les membres les plus proches. L'information de l'ensemble d'analyses est aussi utilisée pour corriger la moyenne ainsi que l'écart de l'ensemble d'ébauche. Une série d'expériences d'assimilation de données cyclées en utilisant seulement des observations de précipitations démontrent que la version révisée de LEMA produit une prévision de qualité similaire à celle produite par l'approche du "Local Ensemble Transform Kalman Filter".

CONTRIBUTION OF AUTHORS

Chapters 2 to 4 of this thesis consist of papers published or submitted for publication to *Monthly Weather Review*, a journal of the American Meteorological Society. I conducted the research and drafted all the manuscripts as part of my Ph.D. studies. Prof. Isztar Zawadzki and Prof. M. K. Yau provided supervision of the research as well as some text editing.

STATEMENT OF ORIGINALITY

The following elements of the thesis represent original contributions to knowledge:

- The development of a new heuristic data assimilation method, LEMA, to construct an analysis mosaic by assigning to each column in the model the information from the ensemble member that is locally closest to the precipitation observations.
- The development of a method to increase the effectiveness of LEMA by constructing the analysis using an augmented background ensemble, including model states at different times and states taken from additional ensemble forecast initialized before the background runs.
- The development of a methodology to quantify the effectiveness of the DA to transfer information from the precipitation to the model state variables using the joint probability $p(\Delta\varepsilon_\phi, \Delta MAD)$, where $\Delta\varepsilon_\phi$ is the decrease of the error in a given state variable and ΔMAD is the precipitation forecast error
- The finding that prioritizing the reduction of the precipitation errors in the large scale leads to:
  - A better transmission of information from the observations to the state variables.
  - An increase in the total area with positive gains in the state error.
  - Long-lasting correction of the model trajectory towards the observed atmospheric state
- The extension of the LEMA methodology to cycled DA applications by constructing an ensemble of analysis mosaics using a small group of members closer to the observations instead of using only the closest one. The analysis ensemble is used to recenter the prior

ensemble around the mean of the analysis ensemble and to scale the original background perturbations to match the spread of the analysis mosaics.

ACKNOWLEDGMENTS

This thesis represents not only a great personal achievement but also symbolizes my journey as a Ph.D. student. The journey started six years ago when I met Prof. Isztar Zawadzki in Argentina. I want to express my sincere gratitude to him not only for allowing me to pursue a Ph.D. at McGill under his supervision but also for teaching me an exceptional way of doing science. I will also want to express my gratitude to Prof. Peter Yau for his guidance and support during this entire period.

I would also like to acknowledge all the people that contributed to this thesis with their many discussions and suggestions. I wish to thank Dominik Jacques for giving me the first recommendations about the Ph.D. life and for following my work closely afterwards. Madalina Surcel, Aitor Atencia, and Bernat Puigdomènech Treserras are thanked for the countless discussions and suggestions during my first years at McGill. Valentina Glazatova is thanked for the French translation of the abstract. I would also like to thank the valuable comments and suggestions of Dr. Peter Houtekamer and Prof. Frédéric Fabry for many interesting discussions and constructive comments.

I am also very grateful to the administrative staff at the Department of Atmospheric and Oceanic Sciences, namely, Amna, Lucy, and Manuela, for making our everyday life easier. I would also like to thank the IT Support staff, Michael Havas and Calin Giurgiu, for the innumerable times that they help me to solve all kinds of problems.

As I mentioned at the beginning, this thesis marks the end of a journey shared with great people. I want to thank all of them for all the laughs and enjoyable moments, especially to Paloma, Ting-Chen, Eric, Natalie, Sonja, and Kyle.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

The origins of numerical weather prediction can be traced back to the beginning of the twentieth century when Abbe (1901) and Bjerknes (1904) recognized weather forecasting as an initial-value problem. Specifically, if one knows with sufficient accuracy the state of the atmosphere at a given time, the physical laws governing the atmospheric dynamics can be used to determine how the initial state evolves with time. Consequently, a skillful weather forecast requires a precise estimation of the initial conditions (ICs) and a model that provides a realistic representation of the atmosphere.

Early experiments on Numerical Weather Prediction (NWP) by Richardson (1922) and Charney et al. (1950) specified the ICs by manually interpolating synoptic charts onto the model grids points. This "subjective" methodology was a time-consuming procedure as it requires an experienced meteorologist to construct the synoptic charts. This problem motivated the development of objective schemes to interpolate the observations onto the model grid without human intervention (Panofsky, 1949; Gilchrist and Cressman, 1954; Barnes, 1964). One shortcoming of these early

interpolation methods is that due to sparse coverage, the observational data alone are insufficient to initialize NWP models. To address this limitation, Cressman (1959) and Bergthórsson and Döös (1955) proposed to supplement the insufficient data with prior information of the atmospheric state, obtained from either a climatological estimate or a previous forecast. Progress on the objective methods to estimate the initial conditions continued to advance and emerged as a field of study referred to as "Data Assimilation" (DA, Daley, 1991; Kalnay, 2002).

Most of the current DA techniques for research and operational applications are based on the Bayesian probabilistic theory. The most general form of this theory represents the atmospheric state $\boldsymbol{x}$ as a probability density function $p(\boldsymbol{x})$, which is updated over time using the Bayes theorem by incorporating the information from new observations of the atmospheric state. More rigorously, Bayes theorem states that the probability of an atmospheric state $\boldsymbol{x}$, given a set of observations $\boldsymbol{y}$, can be expressed as:

$$p(\boldsymbol{x}|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x})}{p(\boldsymbol{y})} \tag{1.1}$$

Hence, the DA inverse problem can be defined as estimating the "best" atmospheric state $\boldsymbol{x_a}$, and its uncertainty, that maximize $p(\boldsymbol{x}|\boldsymbol{y})$. Although this general theory has a simple form, simplifications are still required to make it tractable for typical geophysical applications.

Conventional DA methods widely used in numerical NWP centers solve Eq. 1.1 by assuming that the observations errors and prior estimates of the atmospheric state follow Gaussian distributions (Lorenc, 1986; Hamill, 2006). In four-dimensional variational DA methods (4D-Var) using the model as a strong constraint, the problem of maximizing $p(\boldsymbol{x}|\boldsymbol{y})$ is equivalent to minimizing a cost function over a time window from t=0 to t=T. The cost function is represented by the sum of a term that penalizes the distance of the model trajectory to the observations and another term

that penalizes the distance to a first guess $\boldsymbol{x^b}$:

$$J(\boldsymbol{x}) = \Big(\boldsymbol{x} - \boldsymbol{x_b}\Big)^T \boldsymbol{B}^{-1}\Big(\boldsymbol{x} - \boldsymbol{x_b}\Big) + \sum_{t=0}^{T}\Big(\mathcal{H}(\boldsymbol{x}_t) - \boldsymbol{y_t}\Big)^T \boldsymbol{R}^{-1}\Big(\mathcal{H}(\boldsymbol{x}_t) - \boldsymbol{y_t}\Big) \tag{1.2}$$

where $\boldsymbol{x}$ is the model state at the start of the current assimilation cycle, $\mathcal{H}(\ )$ is the observations operator that links the observed quantities and the model state, and $\boldsymbol{B}$ and $\boldsymbol{R}$ are the background and the observations errors covariance matrices, respectively. To obtain the state (analysis) that maximizes $p(\boldsymbol{x}|\boldsymbol{y})$, 4D-Var methods implement an iterative algorithm to minimize the cost function that requires a linearized version of the observations operators and the forecast model (i.e., the tangent linear model and its adjoint model). The 4D-Var method can be simplified by ignoring the temporal dimension and the time at which the observations are made. This particular implementation, called 3D-Var, is less expensive than 4D-Var and requires fewer development efforts since no tangent linear and adjoint models are required to minimize the cost function. Variational methods have been used successfully in NWP for many years (e.g. Rabier et al., 2000; Rawlins et al., 2007; Tanguay et al., 2012), using time series of previous forecasts to estimate the "climatological" background error statistics (Parrish and Derber, 1992). Although this static background error representation is statistically robust, it has a limited ability to quantify the flow-dependent statistics.

Another widely used DA method in NWP is the Ensemble Kalman filter (EnKF). Instead of using a static covariance model, the EnKF uses an ensemble of possible forecasts to approximate the flow-dependent background error statistics (Evensen, 1994; Houtekamer and Mitchell, 2005; Houtekamer and Zhang, 2016). Using the flow-dependent statistics, and without requiring a linearized version of the model and the observations operator, the EnKF produces an analysis ensemble with a mean value that minimizes Eq. 1.2 and an ensemble spread that reflects the uncertainty of the analysis. In parallel to the variational methods, EnKF have also been applied

successfully in operational NWP (e.g. Bonavita et al., 2010; Houtekamer et al., 2014; Schraff et al., 2016).

In recent years, much effort is focused on combining variational and ensemble data assimilation (EDA) methods to improve the estimation of the background-error statistics and the overall efficiency of the DA (Bannister, 2017). For example, pure "EnVar" methods use an ensemble forecast to estimate the time-varying background error covariances to replace the climatological (static) estimates used in 4D-Var or 3D-Var approaches (e.g. Liu et al., 2008; Buehner et al., 2010). Similarly, hybrid variational methods have been proposed that use a weighted combination of the statistically robust static covariances with the ensemble estimates of flow-dependent error statistics (Hamill and Snyder, 2000; Clayton et al., 2013).

The advances in the DA techniques described above are among the factors contributing to the improvements in the forecast quality over the last three decades. Another crucial factor was the increasing number of observations being assimilated (Bauer et al., 2015). The current generation of DA methods typically assimilates $O\left(10^5\right)$ to $O\left(10^6\right)$ observations from different sources, such as ground-based measurements, satellite radiances, atmospheric motion vectors estimated from satellite imagery, and atmospheric soundings. In addition, precipitation estimates obtained from satellites and ground radar networks are also a valuable source of information with significant potential to improve the forecast quality.

Early precipitation assimilation techniques used a "reversed" cumulus parameterization scheme that updates the humidity profiles to force the model precipitation towards the observed values (e.g. Krishnamurti et al., 1984, 1988; Heckley et al., 1990). The humidity profiles were obtained using a modified version of the Kuo parameterization scheme, which assumes that convection is maintained by the large-scale moisture converge. Similarly, Donner (1988) used a simplified version of the Kuo scheme to adjust the humidity and the temperature fields while keeping them consistent with the observed precipitation.

In parallel to the previous studies, another approach known as Latent Heat Nudging (LHN) was introduced for precipitation DA. Instead of diagnosing humidity and temperature directly from precipitation observations, the latent heating profiles in the model are empirically adjusted to better represent the observed precipitation. The LHN method is based on the simple idea that the precipitation rate is approximately proportional to the vertically integrated latent heating released by condensation. Therefore, to obtain a heating profile compatible with the observations, a reference heating profile is scaled by the ratio of the observed precipitation rate to the model precipitation rate. However, this simple technique requires prior knowledge or assumptions about the vertical structure of the latent heating to be scaled. Early work assumed a parabolic shape for the latent heat profiles (Wang and Warner, 1988). A problem with this method is that the idealized profiles may not be consistent with the model parameterization, and therefore may fail to produce the expected precipitation. To reduce these inconsistencies, Turpeinen et al. (1990); Manobianco et al. (1994) proposed an alternative approach that obtains the heating profiles directly from the model. The main advantage is that the heating profiles are compatible with the model parameterizations. However, special care needs to be taken for the observed precipitation regions where the model fails to produce rain (Jones and Macpherson, 1997). Similarly to the LHN, empirical humidity nudging methods have been proposed for the assimilation of precipitation observations. Instead of adjusting the temperature profiles, these methods modify the humidity profiles in proportion to the difference between the observed and the modeled precipitation (Falkovich et al., 2000; Davolio and Buzzi, 2004; Davolio et al., 2017). Despite its simplicity, the nudging methods mentioned above have performed well in operational and research environments, achieving significant improvements in the precipitation forecast quality, especially in the short-term (~6 h, e.g. Stephan et al., 2008; Jacques et al., 2018).

Alternatively, conventional DA methods, like variational and EnKF approaches, have also been used to assimilate precipitation observations. Early efforts using 4D-Var systems showed

the potential benefits of the precipitation assimilation on the forecast quality (e.g. Županski and Mesinger, 1995; Tsuyuki, 1997; Peng and Zou, 2002). However, these studies also pointed out the difficulties concerning the assimilation of precipitation using variational frameworks. For 4D-Var systems, the highly non-linear and often discontinuous nature of moist convective processes complicates the development of the tangent linear and adjoint models for the observations operator and the forecast model required by variational methods. Instead of assimilating rainfall in 4D-Var directly, Marécal and Mahfouf (2002, 2003) presented a two-step approach. First, a one-dimensional variational method is used to retrieve the vertical profiles of temperature and moisture from the observed rainfall. Next, the total column water vapor profiles are assimilated using a 4D-Var system. Although the precipitation DA using the "1D+4D-Var" approach improves the quality of the moisture analyses and forecasts, one shortcoming is its limited impact on the other dynamical fields (temperature, winds, and geopotential), presumably due to their weak correlations with humidity (Lopez and Bauer, 2007). More recently, Lopez (2011) assimilated the NCEP StageIV precipitation data (available over the continental US) directly in a 4D-Var system simultaneously with other available observations. To better satisfy the linearity assumptions needed by the 4D-Var systems, the author used 6-h precipitation accumulations instead of the original hourly-precipitation data. Although the assimilation of the StageIV observations significantly improved the precipitation forecast up to 12 h, the DA impact on other dynamical fields (temperature, moisture, winds, and geopotential) was generally neutral or slightly positive, presumably due to the competition of the StageIV data with other types of observations available over North America during the DA.

Besides the difficulties related to the linear representation of moist processes, another factor that complicates the assimilation of precipitation is the non-Gaussian characteristics of precipitation errors, which violates the Gaussianity assumptions used in variational and EnKF methods (Errico et al., 2007; Bauer et al., 2011). A popular and practical approach to alleviate the non-

Gaussianity issues is by applying a logarithmic transformation to the precipitation values defined as $ln(R+\alpha)$, where R is the precipitation accumulation in mm/h and α is a positive value typically ranging from 0 to 1 used to avoid undefined values in the transformation (e.g. Lopez and Bauer, 2007; Lopez, 2011). Although the logarithmic transformation alleviates the non-Gaussian characteristics of precipitation, it is not necessarily optimal. Lien et al. (2013) proposed another technique by converting the precipitation variable into a Gaussian distribution using a general transformation derived from the climatological distribution of precipitation values. In a follow-up study, Lien et al. (2016) applied this approach to the assimilation of satellite-derived precipitation in a more realistic configuration, using the National Centers for Environmental Prediction (NCEP) Global Forecasting System (GFS) running at low horizontal resolution ($\sim 200km$). Their results showed that the assimilation of precipitation using the Gaussian transformation improves the forecast quality up to 5 days for temperature, winds, and moisture fields (verified against reanalysis data). Similar results were obtained by Kotsuki et al. (2017) using a similar Gaussian transformation for the assimilation of satellite-derived precipitation data into the Nonhydrostatic Icosahedral Atmospheric Model (NICAM), running at 112 km horizontal resolution.

The difficulties mentioned above have, in part, motivated the development of non-Gaussian DA approaches based on Particle Filters (PF, van Leeuwen, 2009) theory to deal with the non-linearities in the atmospheric processes. As in conventional DA methods, the PF approach is based on Bayesian probabilistic theory. However, instead of expressing the prior probability $p(\boldsymbol{x})$ as a Gaussian distribution that depends on departure from the background mean, PF approximates the prior probability using a discrete number of states $\boldsymbol{x}_m; m = 1, 2, ..., N_e$ (called particles) as:

$$p(\boldsymbol{x}) \approx \frac{1}{N_e} \sum_{m=1}^{N_e} \delta(\boldsymbol{x} - \boldsymbol{x}_m) \tag{1.3}$$

where δ is the delta function. Using the Bayes theorem and the above approximation, the posterior

probability density can be expressed as a weighted sum of all the members:

$$p(\boldsymbol{x}|\boldsymbol{y}) = \sum_{m=1}^{N_e} w_m \delta(\boldsymbol{x} - \boldsymbol{x}_m) \tag{1.4}$$

$$w_m = \frac{p(\boldsymbol{y}|\boldsymbol{x_m})}{\sum_{l=1}^{N_e} p(\boldsymbol{y}|\boldsymbol{x}_l)} \tag{1.5}$$

where w_m represents the weight given to the mth member based on its proximity to the observations, and the expression for the conditional probabilities $p(\boldsymbol{y}|\boldsymbol{x})$ is determined by the observations errors distribution.

An appealing property of PF is that they do not impose any restriction on the shape of the observations error distribution. However, to provide a meaningful posterior density estimate, methods that assign a single (or "global") weight to each ensemble member (particle) typically require a large ensemble size. If the size of the ensemble is not large enough, the PF update (Eq. 1.4) will give most of the weight to the ensemble member closer to the observations, resulting in a collapse of the posterior distribution. To avoid this limitation, Poterjoy (2016) introduced a localized version of the PF, the Local Particle Filter (LPF), where the weights in Eq. 1.5 are computed individually for each grid point using the surrounding observations. Recent studies have shown that these methods may have potential for NWP applications, particularly for the assimilation of precipitation observations, producing a comparable or better forecast quality than Gaussian DA methods (Poterjoy et al., 2017, 2019; Buehner and Jacques, 2020).

1.2 Research questions

The objective of this thesis is to explore a new ensemble DA strategy that is conceptually different from variational, EnKF, and diabatic initialization approaches. This new method does not find

an optimal solution (analysis) using parametric assumptions for the background and observations error distributions. Instead, the proposed technique is based on the assumption that in an ensemble, the ensemble members with a smaller error in the precipitation forecast also have a smaller error in their state variables than the other members.

This thesis is based on a collection of scholarly manuscripts that document the exploration of the new DA strategy. Chapter 2 introduces the heuristic method for the assimilation of precipitation observations, the Localized Ensemble Mosaic Assimilation (LEMA). This method constructs the analysis by assembling different column-states from the ensemble members closest to the precipitation observations in the vicinity of each grid column. This chapter uses an idealized experimental setup consisting of a small ensemble of 20 members to validate the heuristic hypothesis used in LEMA and to determine the parameters that optimize the transfer of information from precipitation observations towards the state variables. Chapter 2 also evaluates, in the context of observing system simulating experiments (OSSEs), the impacts of the assimilation of the precipitation observations in the forecast quality.

Chapter 3 considers unfavorable situations where the background ensemble does not fully represent the forecast uncertainties and investigates how this situation degrades the LEMA effectiveness. This chapter also presents and characterizes a technique to augment the size of the ensemble used for the analysis construction by considering additional model states taken from a set of ensemble forecasts initialized at different times. Finally, this chapter explores the potential of the expanded ensemble in real DA experiments in assimilating StageIV precipitation observations.

The results presented in Chapters 2 and 3 represent our first steps in the exploration of LEMA. The experiments presented in those sections are carried out in the context of a single DA step, relaxing all the background members towards a single analysis mosaic. Although the methodology described above is useful to characterize LEMA, the relaxation towards a single

analysis collapses the ensemble after the DA takes place. Chapter 4 presents a new version of LEMA, called Localized Ensemble Mosaic Assimilation Sequence (LEMAS), that constructs an ensemble of analysis mosaics using a small group of members closer to the observations instead of only the closest one. The LEMAS technique avoids the ensemble collapse by re-centering the background ensemble around the mean of the analysis ensemble while scaling the original background perturbations to match the spread of the analysis mosaics. The new LEMAS technique is evaluated and compared with a LETKF approach using a series of cycled ideal and real DA experiments in assimilating precipitation observations. Finally, Chapter 5 summarizes the main findings and discusses possible directions for future work.

Chapter 2

A Heuristic Approach for Precipitation Data Assimilation

This chapter introduces a heuristic data assimilation (DA) technique suitable for the assimilation of radar-derived precipitation observations. The goal of this chapter is to describe the conceptual basis of the heuristic approach and to characterize it using Observing System Simulation Experiments (OSSEs).

This chapter consists of a paper published in the journal Monthly Weather Review: Pérez Hortal, A.A., I. Zawadzki, and M.K. Yau, 2019. A Heuristic Approach for Precipitation Data Assimilation: Characterization Using OSSEs. Mon. Wea. Rev., 147, 3445–3466, https://doi.org/10.1175/MWR-D-19-0034.1.

Superscript numbers indicate published footnotes while superscripts starting with "E" indicate editorial notes used for clarifications on the published manuscript. In addition, the current chapter includes corrections to minor grammatical errors in the published manuscript.

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, Montréal, Québec, Canada

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 towards 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 DA technique is able to produce forecasts with considerable and long-lived error reductions in the fields of precipitation, temperature, humidity, and wind.

2.1 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 non-linear moist physical processes, limiting the effectiveness of variational and EnKF methods that rely 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 towards 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 (PF, 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-weights 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 above mentioned precipitation DA techniques have shown different degrees of success. Generally, they force the precipitation towards 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 2.2, 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 towards 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 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 2.2. In Section 2.3 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 2.4. Finally, in Section 2.5 we present the discussions and conclusions.

2.2 Localized Ensemble Mosaic Assimilation

The Localized Ensemble Mosaic Assimilation (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 anal-

ysis 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.

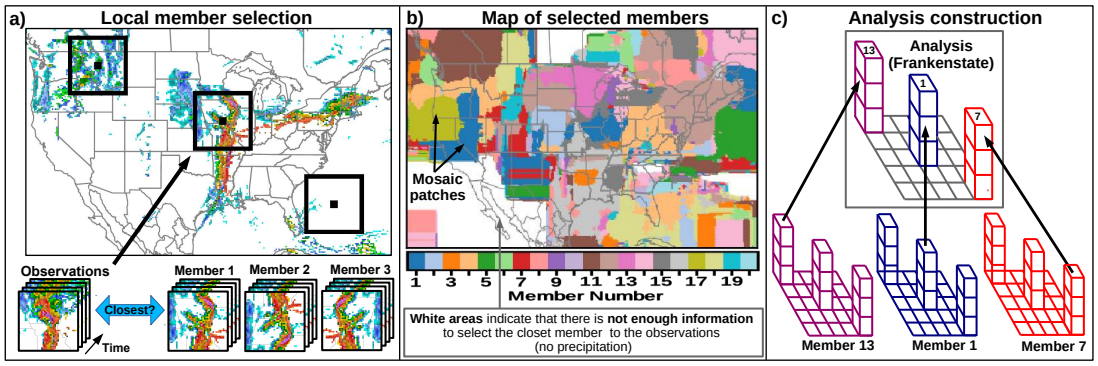

Fig. 2.1: Frankenstate DA method summary. a) Illustration of the local member selection procedure where the black squares denote different observations windows. b) Map of closest members to the precipitation observations used in the construction of the Frankenstate. c) Analysis construction procedure. The analysis is constructed by assembling different columns from the selected members.

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. 2.1a). In this manner, we construct a mosaic of column-states, termed "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. 2.1b and 2.1c). 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 that we applied this DA method.

[1]Name 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 an ensemble forecast).

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 x 41 grid points) and over a $\Delta t = 30\ min$ period preceding the analysis time (see Fig. 2.1a)[E1]. 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 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 $Z[mm^6\ m^{-3}] = 200R^{1.6}$ relation, with R expressed in mm/h. 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[E2].

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

$$MAD_m(i,j) = \frac{1}{N_t N_x N_y} \sum_{t=1}^{N_t} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} \left| Z^{obs}_{t,x,y} - Z^{m}_{t,x,y} \right| \tag{2.1}$$

where Z^{obs} indicates the reflectivity (converted from precipitation observations) and Z^m the mth member reflectivity, both in dBZ units[E3]. The sub-index x and y denotes the x-index and y-index

[E1]The formulation of LEMA described here considers that the observations are available as precipitation intensity values. These values, along with the model rainfall intensity, are transformed into reflectivity units for the MAD computations.

[E2]Corrigendum. The relation that expresses precipitation in a logarithmic scale is the transformation of precipitation intensity values in mm/h to reflectivity values in dBZ units: $dBZ = 10log_{10}(200R^{1.6})$ with R in mm/h.

[E3]For the conversion from precipitation intensity to reflectivity in dBZ, the minimum value allowed for the converted reflectivity is 0 dBZ (equivalent to 0.03 mm/h).

of the horizontal grid point inside the observation windows, while the sub-index 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 minutes, over the 30 minutes 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 $n_{min} = 35$ grid points over the spatiotemporal localization window are used[E4]. 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. 2.1b), 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 towards the Frankenstate. We will refer to these forecasts as "Frankencasts".

The relaxation towards the Frankenstate is done by adding artificial terms to the model's

[E4]A detection threshold of 0.03 mm/h (0 dBZ) is used for the computation of the precipitation coverage.

prognostic equations:

$$\left(\frac{\partial \phi(t)}{\partial t}\right)_{new} = \left(\frac{\partial \phi(t)}{\partial t}\right)_{model} + G\left(\phi^F - \phi(t)\right) \tag{2.2}$$

where $\phi(t)$ indicates a model variable at time t, ϕ^F the Frankenstate[E5], 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 towards the analysis is applied over a time period τ preceding the analysis (pre-forecast nudging). If τ is too small, it forces the model state too strongly towards 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 towards 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

[E5]The omission of the time t in ϕ^F is intentional to indicate that the Frankenstate corresponds to a single state, and it is fixed during the nudging period. For the nudging procedure, the difference $\phi^F - \phi(t)$ is evaluated continuously for each time step since the end goal of this initialization was to bring the model as close as possible to the analysis mosaic at the end of the nudging period while limiting the impacts of the imbalances present in the analysis mosaic.

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 2.A). Consequently, the Frankenstate is constructed only using these four variables.

2.3 Data Assimilation Experiments

2.3.1 The model

The numerical model used in this study is the Weather Research and Forecasting Model (WRF) 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 US and southern Canada. The lateral boundary conditions (LBCs) and initial conditions (ICs) are constructed by downscaling the 1° resolution Global 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 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.

[2]The 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 is accessible from the National Oceanic and Atmospheric Administration (NOAA) Operational Model Archive and Distribution System (NOMADS). More information available in https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-ensemble-forecast-system-gefs.

2.3.2 Observing System Simulation Experiments

The Observing System Simulation Experiments (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 minutes. 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 mid-latitude cyclone over the eastern US 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. 2.2a). In Case B, at 1800 UTC 4 April 2013, three widespread precipitation systems developed around three cyclonic systems over the US, located respectively in the north-west, north-east and south-east regions of the domain (Fig. 2.2b).

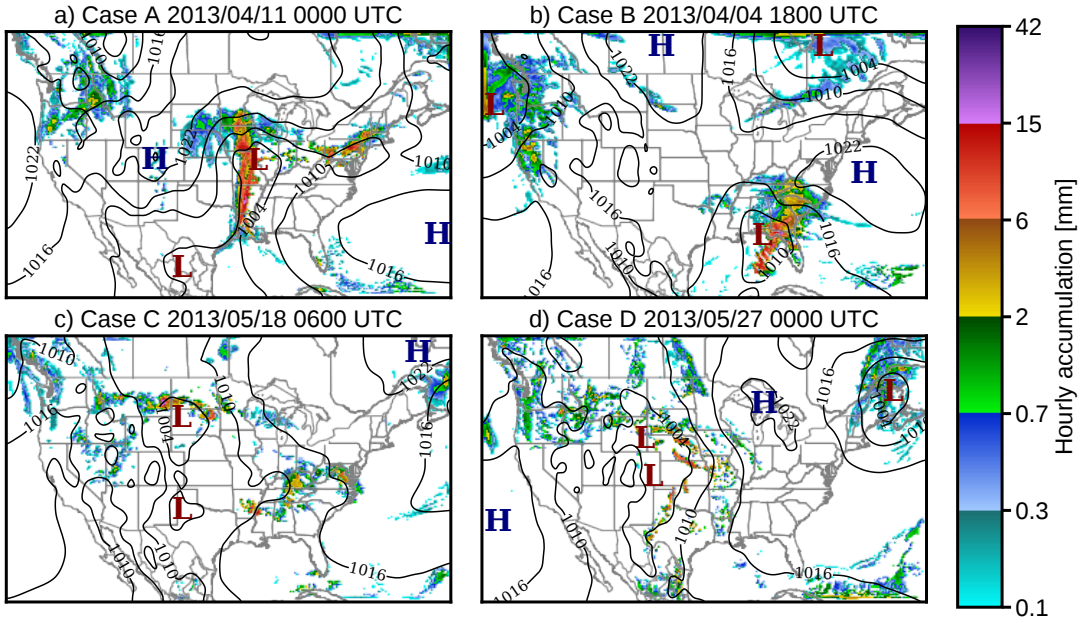

Fig. 2.2: Hourly precipitation for the four study cases. The black contours represent the Sea Level Pressure (SLP).

For the two remaining cases, precipitation was produced by several Mesoscale Convective Systems (MCSs) scattered over the US. Case C occurred at 0600 UTC 18 May 2013 where the precipitation was located in northern and south-eastern US (Fig. 2.2c). For Case D, at 0000 UTC 27 May 2013, several MCSs developed over central and north-western US, along with a cyclonic precipitation system located in the north-eastern part of the domain (Fig. 2.2d).

For each study case, we produce an ensemble forecast of 21 members (0 to 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. Afterwards, 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. 2.3.

For the DA experiment, we use the precipitation observations to construct the Frankenstate 12 h after the WRF model was initialized (end of the spin-up period). Afterwards, each member of the background is relaxed toward this analysis to initialize a 12-h ensemble forecast (Frankencasts). Figure 2.3 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 2.1.

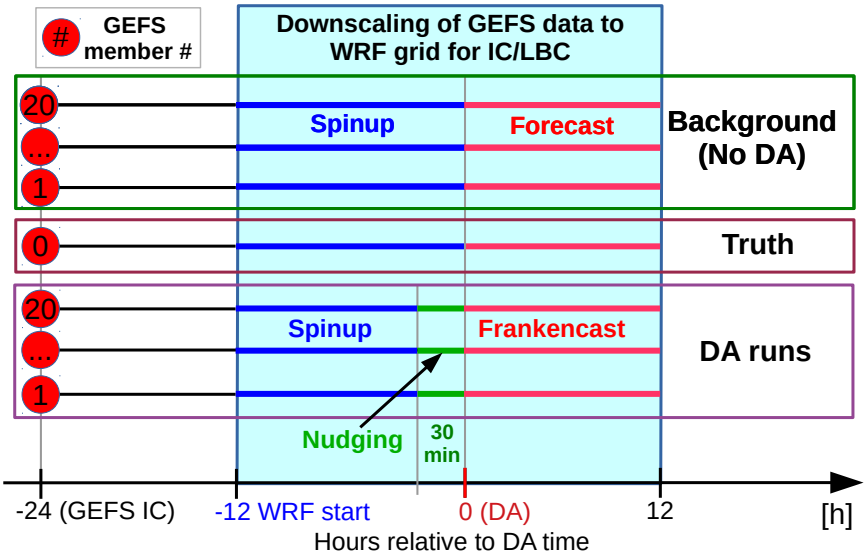

Fig. 2.3: Observing System Simulation Experiments. The horizontal lines indicate the GEFS forecast data initialized from the different analyses (circles) while the numbers inside the circles represent the GEFS member. The simulation period where the model ICs/LBCs are downscaled from the GEFS forecast data is denoted in light blue.

Table 2.1: Precipitation Events

Case	Type	Simulation start date	Frankenstate date	Simulation end date
A	Synoptic front	Apr 10 1200 UTC	Apr 11 0000 UTC	Apr 11 1200 UTC
B	Cyclonic systems	Apr 4 0600 UTC	Apr 04 1800 UTC	Apr 5 0600 UTC
C	MCSs	May 17 1800 UTC	May 18 0600 UTC	May 18 1800 UTC
D	MCSs	May 26 1200 UTC	May 27 0000 UTC	May 27 1200 UTC

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. 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. 2.1) and, in consequence, it may affect the selection of the closest member to the observations. In Appendix 2.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.

[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).

2.4 Results

We first show in Sec. 2.4.1 the characterization of the LEMA method using Case A. This characterization includes the hypothesis testing of the method (Sec. 2.4.1.1), the determination of the optimal observation window size (Sec. 2.4.1.2), and the impacts of the DA in the forecast quality (Sec. 2.4.1.3). In addition, we present the advantages and limitations of initializing the forecast by relaxing (nudging) the background towards the analysis (Sec. 2.4.1.4). In Section 2.4.2 we show the impacts of the DA for Cases B, C, and D.

2.4.1 In-depth study - Case A

2.4.1.1 Hypothesis testing

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

$$\varepsilon_{\phi}^{m} = \sqrt{\frac{\sum_{k}^{N_z}[\phi_m(k) - \phi_t(k)]^2}{N_z}} \;\; ; \;\; \varepsilon_{\phi}^{F} = \sqrt{\frac{\sum_{k}^{N_z}[\phi_F(k) - \phi_t(k)]^2}{N_z}} \tag{2.3}$$

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:

$$Decrease\ in\ RMSE = \Delta\varepsilon_{\phi}^{m} = \varepsilon_{\phi}^{m} - \varepsilon_{\phi}^{F} \tag{2.4}$$

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:

$$Decrease\ in\ MAD = \Delta MAD_m = MAD_m - MAD_F \tag{2.5}$$

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 $p(\Delta\varepsilon_{\phi}, \Delta MAD)$ of decreasing the state variable error by $\Delta\varepsilon_{\phi}$ and the precipitation error by ΔMAD is estimated by the bi-dimensional histogram of the $\Delta\varepsilon_{\phi}$,ΔMAD 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 2.4 shows the joint probability $p(\Delta\varepsilon_\phi, \Delta MAD)$ 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. 2.4), 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.

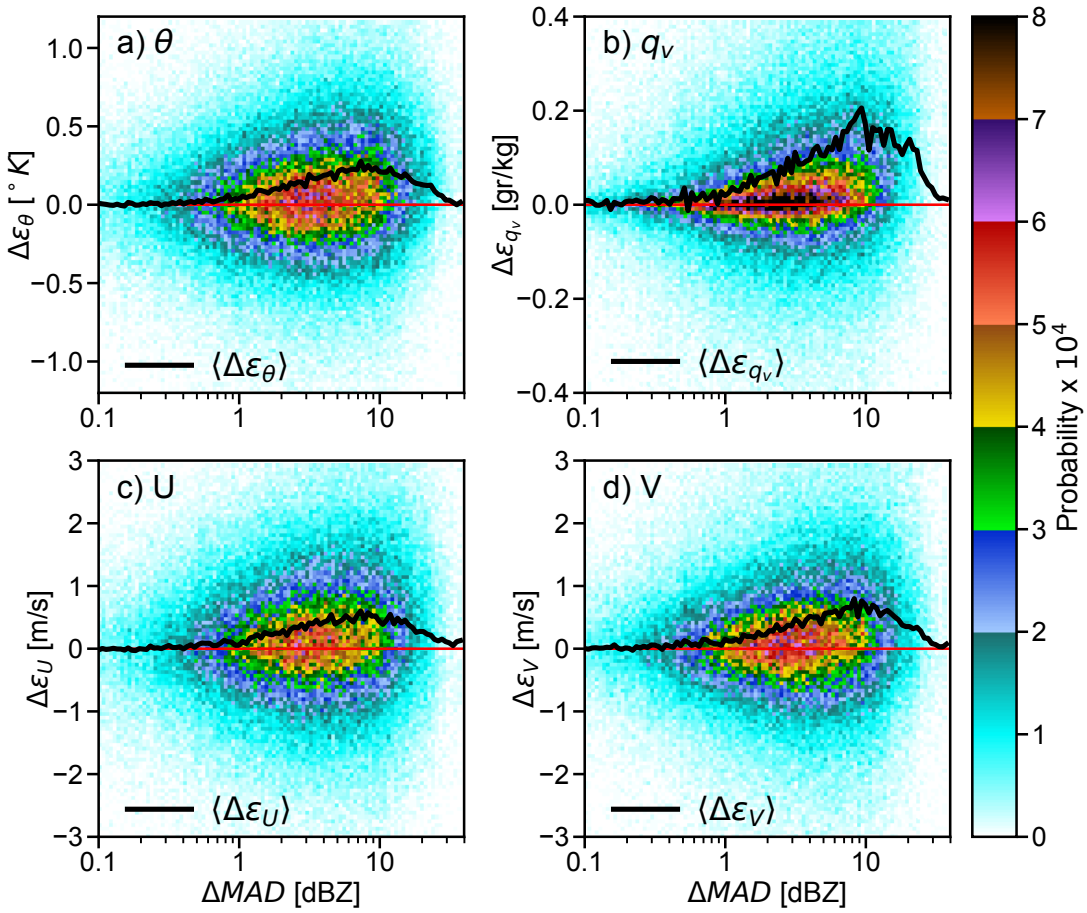

Fig. 2.4: Joint probability of the decrease in the RMSE ($\Delta\varepsilon$) for selected state variables when the error in precipitation is decreased (ΔMAD) by assigning at each grid point the ensemble member with the smallest MAD with respect to the "true" precipitation (20 km observation window). 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: $\langle\Delta\varepsilon_\phi\rangle = \sum_{\Delta\varepsilon_\phi} \Delta\varepsilon_\phi p(\Delta\varepsilon_\phi, \Delta MAD)$. The probability is computed using 100 bins in each axis direction. The values $\langle\Delta\varepsilon_\phi\rangle$ 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 $\Psi = (\vec{U}, \vec{V}, \vec{\theta}, \vec{q_v})$, 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:

$$||\Psi - \Psi_t|| = \varepsilon_\Psi = \sqrt{\frac{1}{N_z}\sum_{z=1}^{N_z}\frac{(\theta-\theta_t)^2}{\sigma_\theta^2} + \frac{(q_v-q_{v_t})^2}{\sigma_{q_v}^2} + \frac{(U-U_t)^2}{\sigma_U^2} + \frac{(V-V_t)^2}{\sigma_V^2}} \tag{2.6}$$

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 Frankenstate for each member "m" and each column is computed as:

$$Decrease\ in\ State\ Error = \Delta\varepsilon_\Psi = ||\Psi_m - \Psi_t|| - ||\Psi_F - \Psi_t|| \tag{2.7}$$

where Ψ_F denotes the Frankenstate column.

Figure 2.5a shows the joint probability $p(\Delta\varepsilon_\Psi, \Delta MAD)$ of decreasing the state error by $\Delta\varepsilon_\Psi$ and the distance to the observations by ΔMAD for a 20 km wide observation window. Similarly, Fig. 2.5b 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. 2.5).

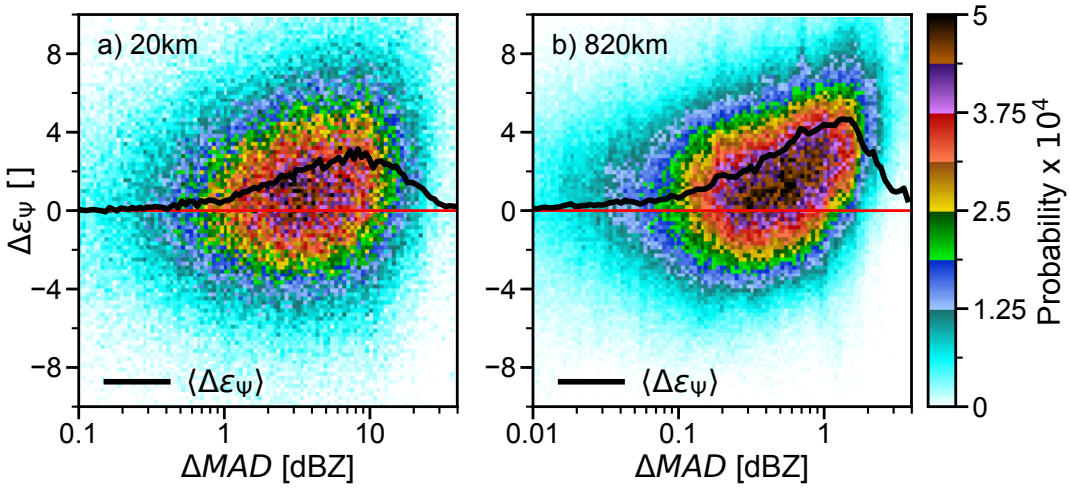

Fig. 2.5: a) Joint probability of the decrease in the State Error ($\Delta\varepsilon_\Psi$) and when the error in precipitation is decreased (ΔMAD) by selecting the ensemble member with the smallest difference to "true" precipitation using a 20 km wide observation window. The probability is computed using 100 bins in each axis direction. b) Same as a) but for an 820 km wide observation windows. The black curve indicates the mean value of the decrease in the state error: $\langle\Delta\varepsilon_\Psi\rangle = \sum_{\Delta\varepsilon_\Psi} \Delta\varepsilon_\Psi p(\Delta\varepsilon_\Psi, \Delta MAD)$. The values $\langle\Delta\varepsilon_\Psi\rangle$ are amplified 100 times to fit the y-scale of the joint probability.

As in Fig. 2.4, Fig. 2.5 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.

2.4.1.2 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. 2.5 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[E6]. The average state gain increased with decreasing MAD resolution (increasing window size). Furthermore, as illustrated in Fig. 2.6, 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 grey area in Fig. 2.6a with respect to Fig. 2.6b)[E7].

Figure 2.7 show the Frankenstate total area as a function of the observation windows width (black line). For the 20 km observation window (single grid point 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 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

[E6] Additionally, smaller ΔMAD values are observed in Fig. 2.5b due to the averaging effects of computing the MAD over a larger area.

[E7] Corrigendum. The text in the parenthesis should read as "the reduction in grey area in Fig. 2.6b with respect to Fig. 2.6a."

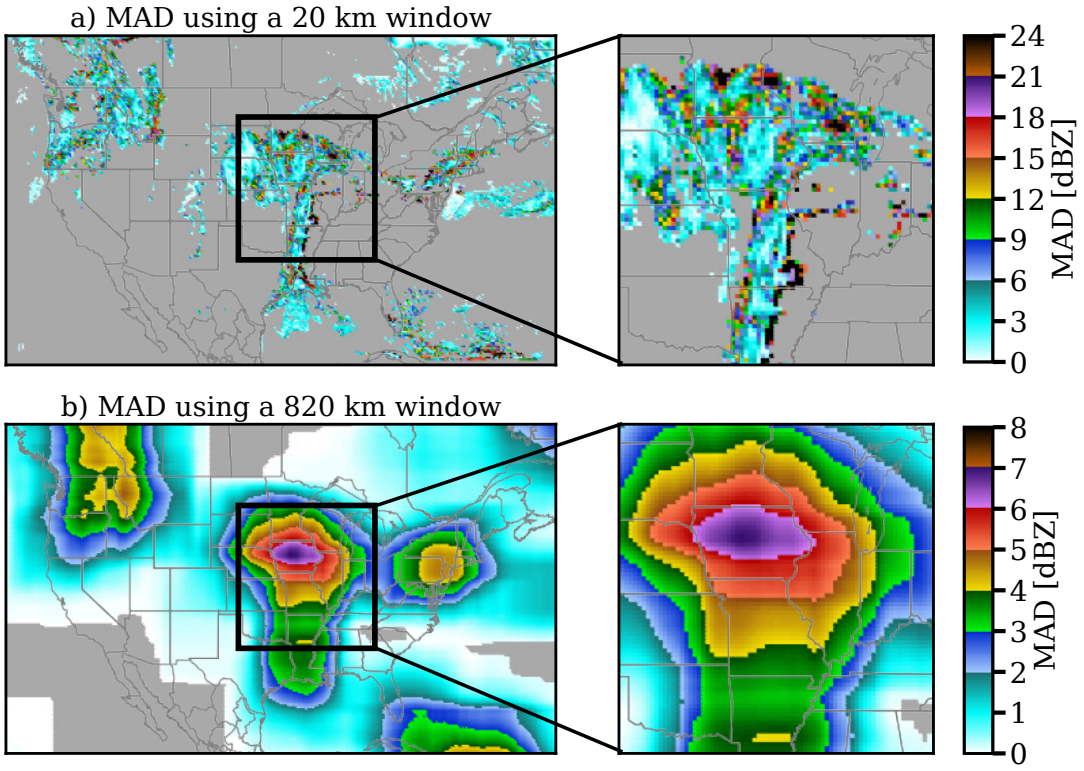

Fig. 2.6: Precipitation distance (MAD) for Member 1 measured using a 20 km (a) and 820 km (b) observation window. The areas where the Frankenstate was left empty are indicated in grey.

positive gains in the state error (red shaded areas in Fig. 2.7). However, it also increases the area with negative impacts (blue shaded region in Fig. 2.7). 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. 2.7). 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

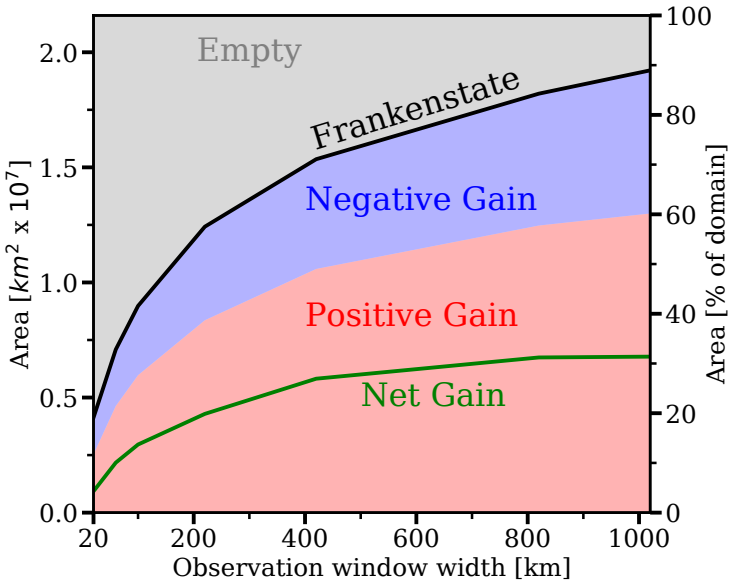

Fig. 2.7: Frankenstate total area as a function of the observation window width (black line). The grey shadow area denotes the areas where the Frankenstate was left empty. The blue and the red shadow areas indicate the area with negative or positive gains, respectively. The green line indicates the net 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 Frankenstate selected that member as the "closest" are excluded because they do not provide any information (0% error reduction by definition).

the domain is directly benefitted from LEMA (Fig. 2.7). Also, it increases the magnitude of the benefit as seen by comparing Figs. 2.5a and 2.5b[E8].

Figure 2.8 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. 2.7. Moreover, the regions with positive and negative gains are different for each ensemble member

[E8]To strengthen this point, the net average improvement over the Frankenstate area is also computed for each experiment as $\overline{\Delta\varepsilon_\Psi} = \sum_{\Delta\varepsilon_\Psi, \Delta MAD} \Delta\varepsilon_\Psi \; p(\Delta\varepsilon_\Psi, \Delta MAD)$. For the experiment using the 820 km window, the net improvement is $\overline{\Delta\varepsilon_\Psi} = 1.7$, 54% larger than the average improvement for the experiment using the 20 km window ($\overline{\Delta\varepsilon_\Psi} = 1.1$).

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 towards 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.

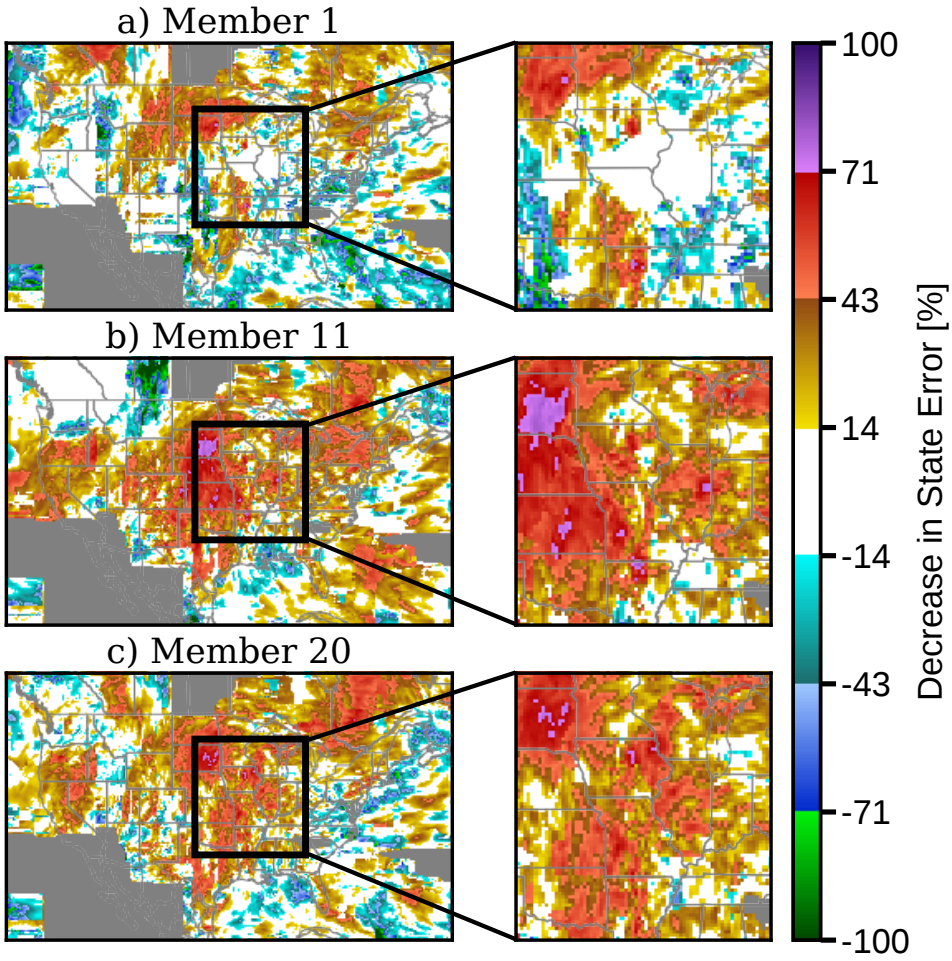

Fig. 2.8: Maps showing the decrease in the state error with respect to the Frankenstate for members 1, 11, and 20. The Frankenstate is constructed using the optimal window size (820 km x 820 km x 30 min).

2.4.1.3 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 towards the Frankenstate. To evaluate the impact in the precipitation forecast quality we compare 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. 2.9a and 2.9b). Consistent improvements are also obtained in the classical contingency scores (Wilks, 1995) such as Equitable Thread Score (ETS; Figs. 2.9c and 2.9d), Probability of Detection (POD; Figs. 2.9e and 2.9f), and False Alarm Ratio (FAR; Figs. 2.9g and 2.9h). The detection threshold used in the contingency score is 0.3 mm/h ($\sim$14 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. 2.10). 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. 2.8).

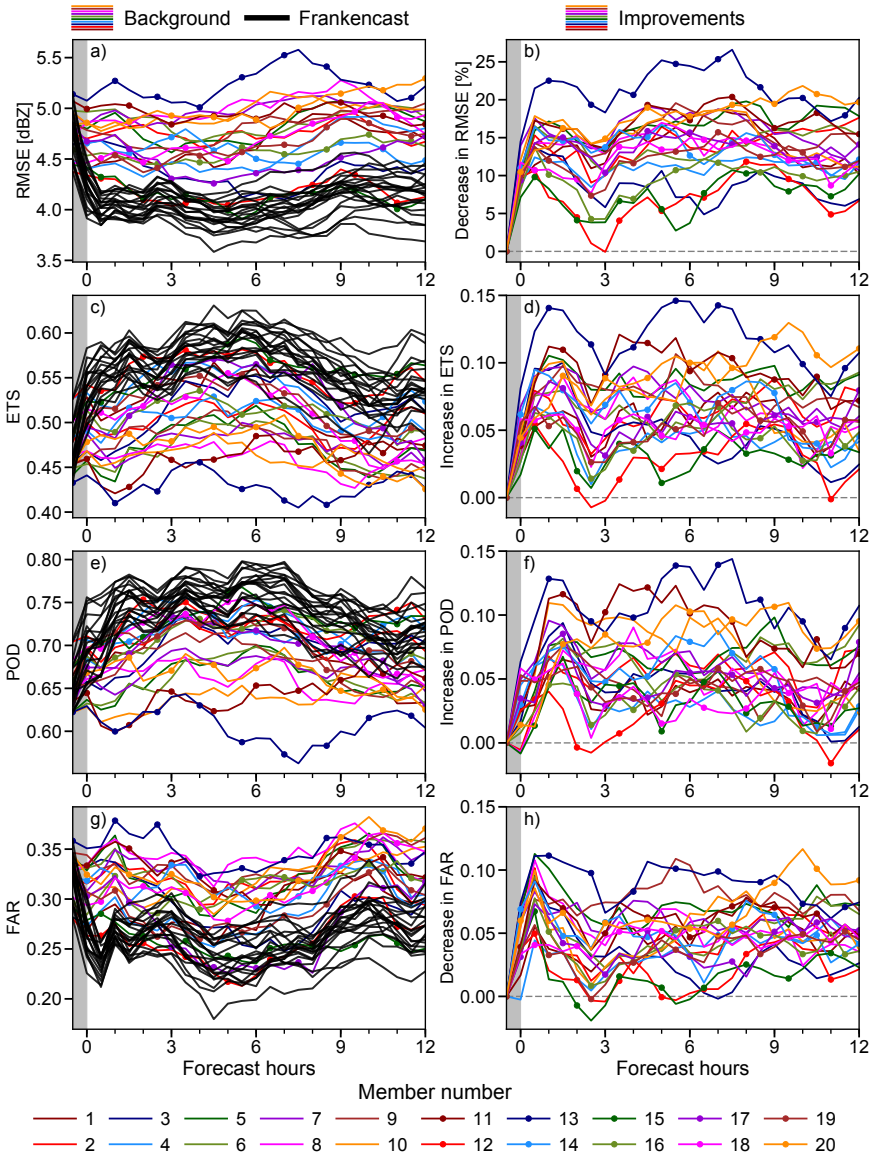

Fig. 2.9: DA impacts on precipitation forecasts for Case A. The left column shows the background (colors) and the Frankencast (black) errors, measured by (a) Reflectivity RMSE, (c) ETS, (e) POD, and (g) FAR. The right panel indicates the error improvement achieved by each Frankencast member, with respect to the corresponding background member. The improvements correspond to: (b) reflectivity relative decrease in RMSE, (d) increase in ETS ($ETS_F^m - ETS_B^m$), (d) 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 sub-indexes "F" and "B" indicate the error for the "Frakenstate" or "Background" and "m" the member number. A threshold of 0.3 mm/h is used for the ETS, POD, and FAR computations. Grey shaded areas indicate the nudging period.

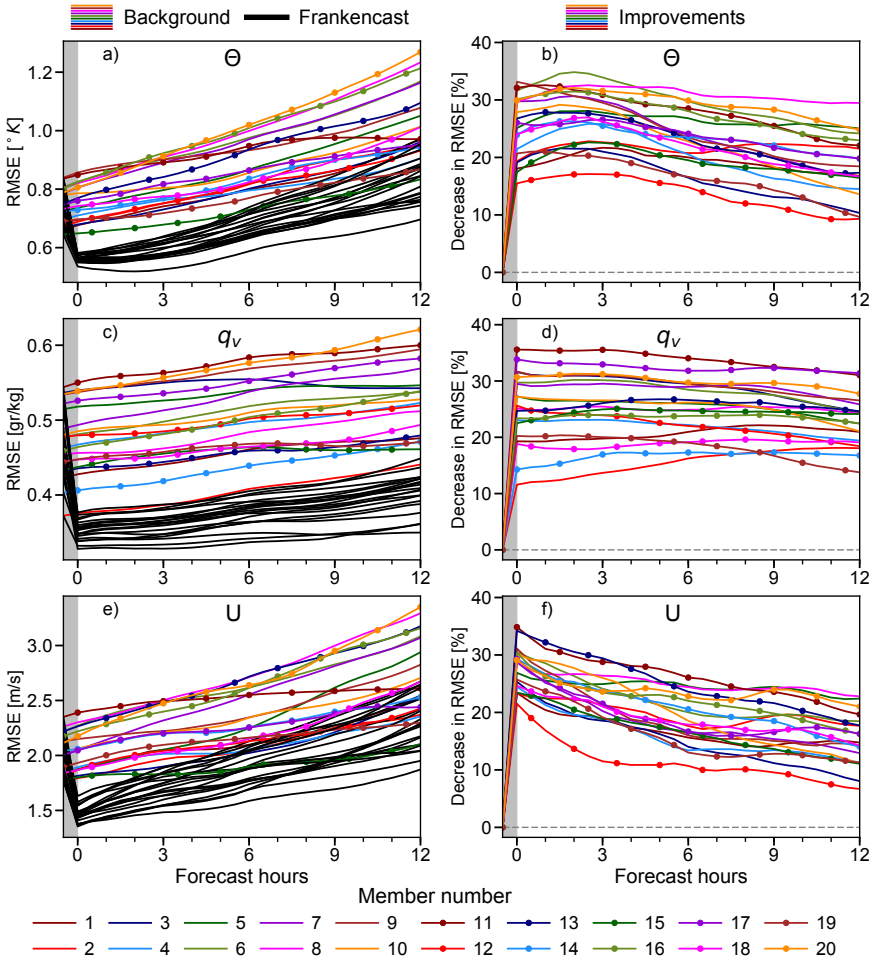

Fig. 2.10: Case A: DA impacts on forecast quality for Potential temperature (a and b), vapor mixing ratio (c, d), and u-wind (e and f). The left panels (a, c, and e) show the RMSE for each variable for each background (colors) and Frankencasts (black) member. The right panels (b, d, and f) display the relative decrease in RMSE for each Frankencast member, with respect to the corresponding background member used in the forecast initialization. Grey shaded area indicates the nudging period.

Figures 2.11a and 2.11b 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 towards 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 Frankenstate'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 towards the analysis, but due to the original model tendencies, the Frankenstate 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):

$$spr_{\phi} = \sqrt{N^{-1} \sum_{i}^{N} \left[(M-1)^{-1} \sum_{m}^{M} \left(\phi_m(i) - \overline{\phi(i)} \right)^2 \right]} \tag{2.8}$$

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 over-dispersive (dispersivity > 1) during the entire forecast period (Fig. 2.11c). After the DA, the dispersivity of the Frankencast ensemble is severely reduced, becoming under-dispersive (dispersivity < 1). Nevertheless, due to the fast growth of the perturbations, the ensemble becomes over-dispersive after 4 h. For vapor mixing ratio and reflectivity (Fig. 2.11d)

the ensemble also becomes under-dispersive 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. 2.11b), compared with the spread increase for u-wind and potential temperature (Fig. 2.11a).

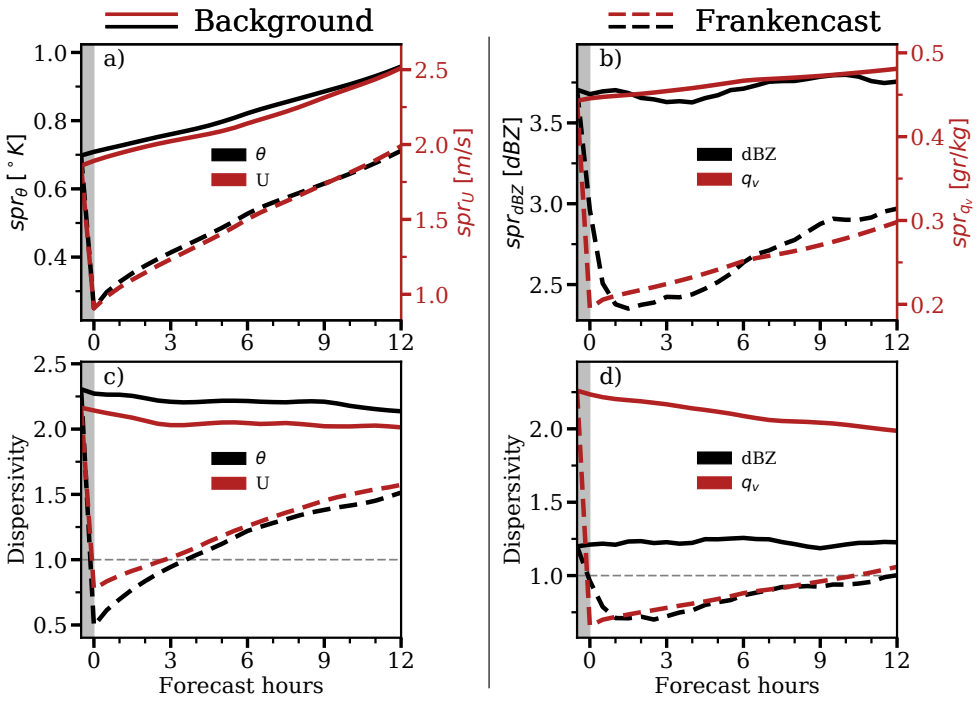

Fig. 2.11: Ensemble spread (first row) and dispersivity (second row) for Case A. The left panels show the results for potential temperature (black) and u-wind (red), while the right panels display the results for reflectivity (black) and vapor mixing ratio (red). In all the plots, the solid lines indicate the Background spread or dispersivity while the dashed ones denote the Frankencast ones.

2.4.1.4 Nudging vs direct insertion forecast initialization

In this section, we show the benefits and the limitations of using nudging to initialize the forecast by 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 towards 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. 2.1b)[E9]. Nonetheless, at the boundaries of those regions where a transition between different selected members occurs (patch boundaries in Fig. 2.1b), 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),

$$S_t = \left\langle \left| \frac{dP_s}{dt} \right| \right\rangle \tag{2.9}$$

where P_s is the surface pressure and $\langle\ \rangle$ denotes the domain average. These pressure perturbations can be 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 towards the Frankenstate are an order of magnitude smaller than the ones produced by the direct insertion initialization (Fig. 2.12). Moreover, in the dynamic initialization, the pressure perturbations have

[E9]Since the Frankenstate is constructed using only four model variables, and not all the prognostic variables, a fully balanced state is not necessarily achieved as mentioned in the text. However, a good "local"balance is expected since each patch in the Frankenstate mosaic uses the information from a single member.

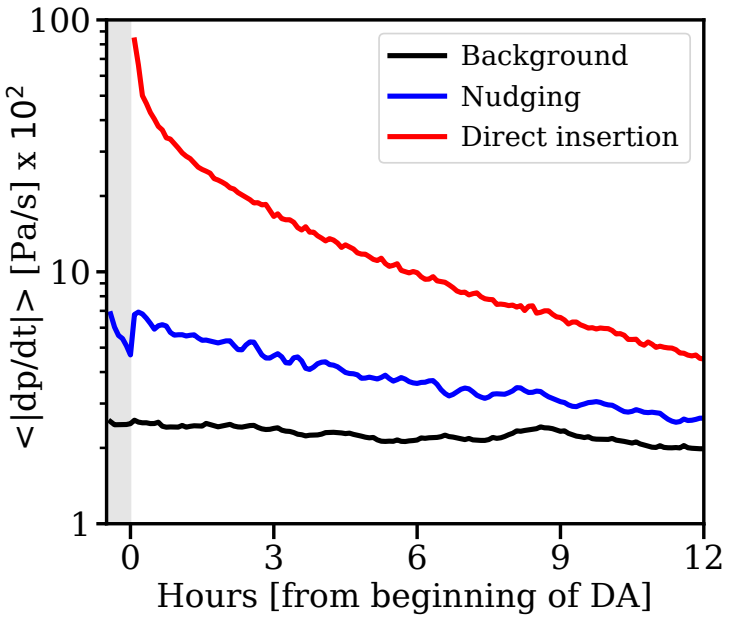

Fig. 2.12: Ensemble mean domain-averaged magnitude of the time derivative of the surface pressure for the background (black line), the forecast dynamically initialized by nudging the background members towards the Frankenstate (Frankencast, blue), and the forecast by direct initialization using the Frankenstate (red).

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. 2.13). 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 towards the Truth values in the four state variables used in the Frankenstate) the maximum attainable reduction in the RMSE is 55-70% for the state variables and 15% for reflectivity (see dot-dashed lines in Fig. 2.13). This is an expected result since only four variables are nudged while the others adjust

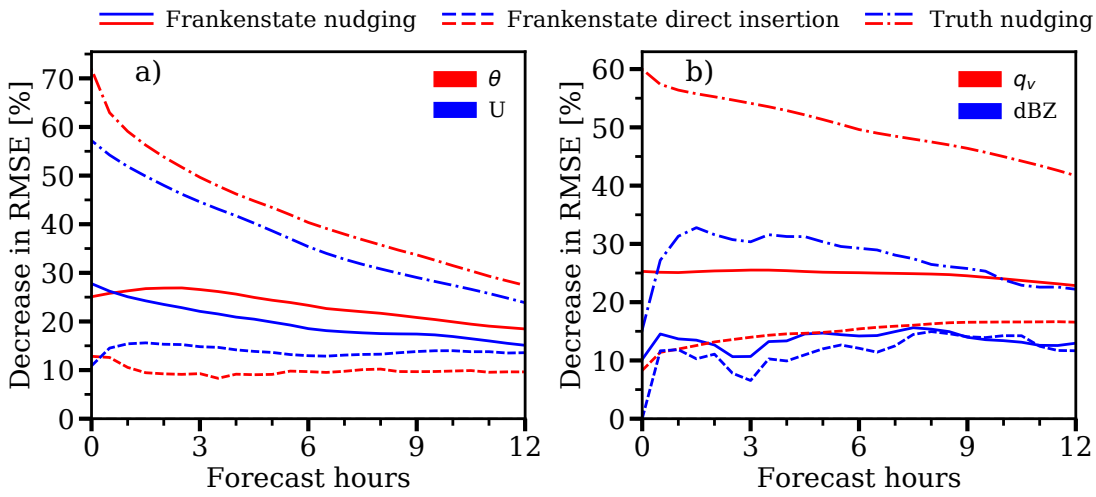

Fig. 2.13: Value of the reduction in the forecast RMSE by different forecast initialization methods, averaged over all the ensemble members, for a) potential temperature θ (red) and u-wind U (blue), and b) vapor mixing ratio q_v (red) and reflectivity dBZ (blue). The solid lines indicate the forecasts initialized by nudging each background member towards 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.

freely to the changes in the state, and also since the artificial forcing needs to compete with the actual model tendencies.

2.4.2 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 2.1).

Figure 2.14 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. 2.14a) which can be associated with the long-lived reduction in the state error (Figs. 2.14d and 2.14g). Similar results were obtained for Case D (Figs. 2.14c, 2.14f, and 2.14i).

For Case C, improvements in reflectivity remain between 5% and 20% over the entire forecast period (Fig. 2.14b). The error reduction in the state variables is similar to the ones obtained for Case B (Figs. 2.14e and 2.14h), 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. 2.14e). Nevertheless, since the error in the other state variables is reduced (Fig. 2.14h for u-wind, similar results for v-wind and vapor mixing ratio), persistent improvements are obtained for that member in precipitation.

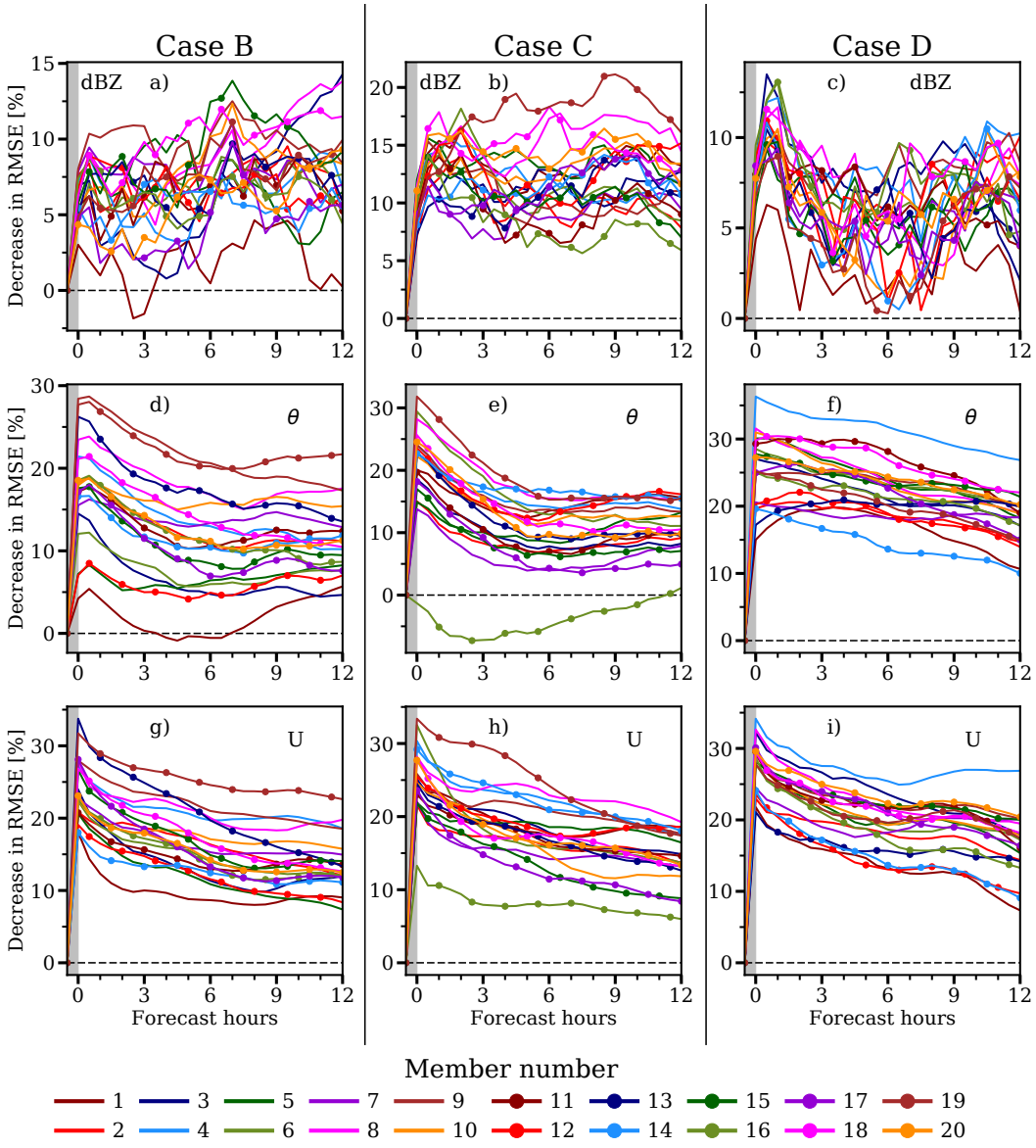

Fig. 2.14: DA impacts on forecasts quality for Case B (first column), Case C (second column), and Case D (third column). The impacts are expressed as the reduction in RMSE for reflectivity (dBZ, first row), potential temperature (θ, second row), and u-wind (U, third row) for each ensemble member. For all plots, the grey shaded area indicates the nudging period.

2.5 Discussions 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 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 towards 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 $p(\Delta\varepsilon_\phi, \Delta MAD)$, where ϕ is a state variable, obtained from model data, such as shown in Fig. 2.4. 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 2.4, the joint pdf has a "noise" component that generates increases in state errors. But LEMA's decrease in state variables' errors is due to the net effect of $p(\Delta\varepsilon_\phi, \Delta MAD)$ 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. 2.15. Here, the graph for potential temperature is shown but similar figures hold for the other three variables. Figure 2.15 shows that the smoothing of MAD from 20 km 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.

From the joint pdf alone (as in Figs. 2.4 and 2.5) 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 $\Delta\varepsilon_\theta$, $\langle\Delta\varepsilon_\theta\rangle = \sum_{\Delta\varepsilon_\theta} \Delta\varepsilon_\theta p(\Delta\varepsilon_\theta, \Delta MAD)$ 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 $\langle\Delta\varepsilon_\theta\rangle/p(\Delta MAD)$ (the green curve) the second effect is eliminated by computing the expectation value of $\Delta\varepsilon$ using the conditional probability $p(\Delta\varepsilon_\theta|\Delta MAD)$ instead of the joint probability. Hence, the green line shows the average relationship $\Delta\varepsilon_\theta = f(\Delta MAD)$ 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

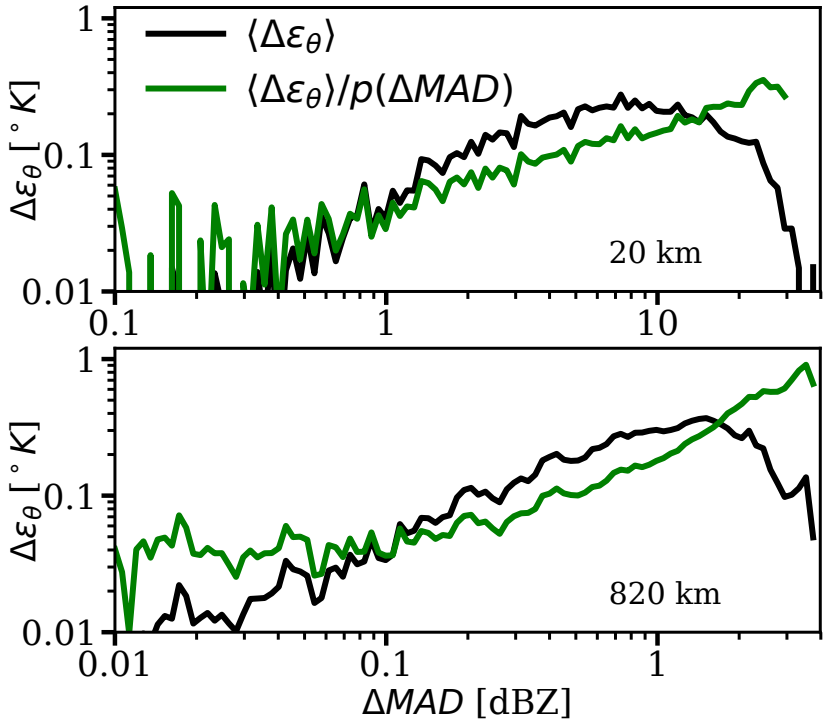

Fig. 2.15: Average transmission of the information from ΔMAD to ε_θ for the 20 km and the 820 km observations windows implicit in the joint probability $p(\Delta\varepsilon_\theta, \Delta MAD)$. The black curve indicates the mean value of the decrease in the RMSE $\Delta\varepsilon$ as a function of ΔMAD: $\langle\Delta\varepsilon_\theta\rangle = \sum_{\Delta\varepsilon_\theta} \Delta\varepsilon_\theta p(\Delta\varepsilon_\theta, \Delta MAD) = \sum_{\Delta\varepsilon_\theta} \Delta\varepsilon_\theta p(\Delta\varepsilon_\theta|\Delta MAD)p(\Delta MAD)$. The green line shows the mean RMSE reduction normalized by the ΔMAD probability: $\langle\Delta\varepsilon_\theta\rangle/p(\Delta MAD)$. The values $\langle\Delta\varepsilon_\theta\rangle$ are amplified 100 times to fit the same y-scale as the normalized expectation values (green line) .

a more effective transmission of the 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 (Durran et al., 2013).

The results summarized in Fig. 2.15 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. 2.15 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 analogues 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.

Appendix 2.A: State 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 towards 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 Sec. 2.3.2) except that only one member is used as background (member 1). To characterize the impacts, two simulations are carried out. One of the simulations is considered as the control which produces a forecast from the background ensemble without assimilating any observations. For the second simulation, a forecast is initialized by relaxing the background towards the truth. This experiment is repeated four times, by 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 $\tau = 30min$. The following combinations of variables are used for each experiment:

- Potential temperature (θ)

- Vapor Mixing Ratio (q_v)
- Horizontal winds (U and V)
- 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. 2.4).

Figure A1 shows the decrease in the RMSE for each variable and each experiment. At the end of the nudging period, nudging only θ (Figure A1-a) results in a ~60% improvement on that variable. But, after the nudging took place, most of the improvements are lost within 4 h. Nudging only θ induces small improvements in the horizontal winds (~10%). Similar results occur when we only adjust q_v (Figure A1-b) with only marginal corrections 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 Figure A1-b. Better results are obtained when we nudge the horizontal winds (Figure A1-c). The error reduction at the end of the nudging period is ~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 hours. 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 (Figure A1-d) 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 towards 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.

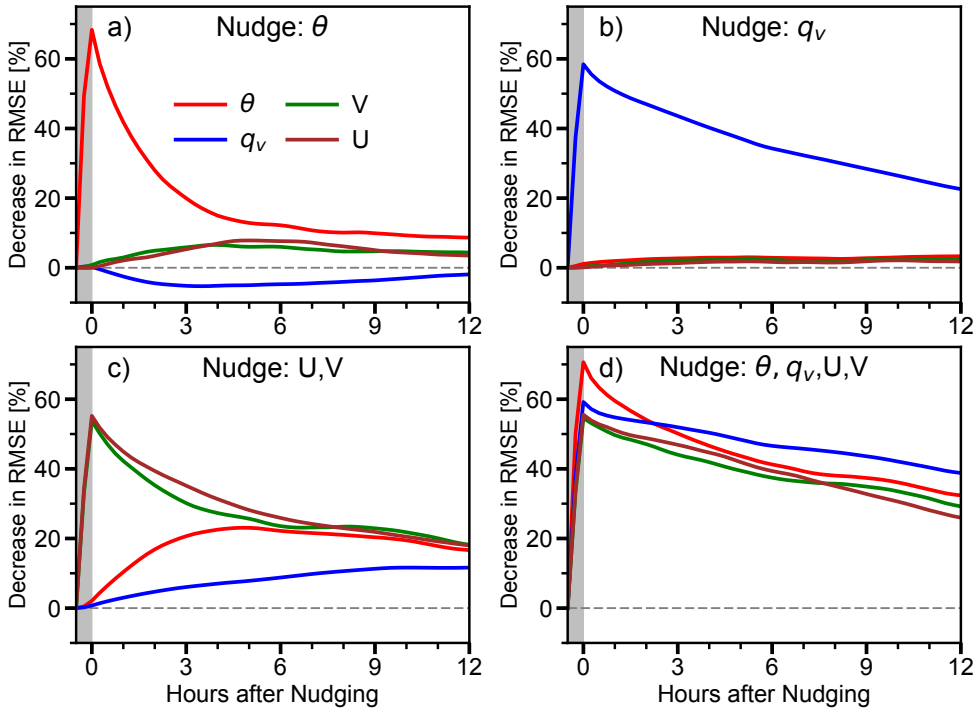

Fig. A1: Nudging variable importance experiments. Impacts on forecast quality using the following combinations of variables during the nudging initialization: a) potential temperature θ, b) vapor mixing ratio q_v, c) horizontal winds (U and V), and d) θ, q_v, U, and V. For all plots, the grey shaded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

Appendix 2.B: Sensitivity of LEMA 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 major ones 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 on range and their structure is scale dependent (Berenguer and Zawadzki, 2009). For lower heights the errors in the Z-R relation dominate, while for elevated observations

the range-dependent errors are dominant.

In LEMA, these errors will affect the calculation of the local distance of each ensemble members to the observations (MAD, Eq. 2.1) 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 $\Delta\varepsilon_\Psi$ using perfect (no errors) and "imperfect" observations (observations + Gaussian noise). The errors were simulated by adding an unbiased Gaussian noise field with σ_{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[4].

Figure B1 shows the expectation values $\Delta\varepsilon_\Psi$ computed using the conditional probability $p(\Delta MAD|\Delta\varepsilon_\Psi)$ for the 820 km observation windows (i.e., the expectation values are normalized by $p(\Delta MAD)$. The expectation values for $\Delta\varepsilon_\Psi$ 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 observation error contribution playing a minor role.

The above results indicate that LEMA is robust with respect to the observation 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 considerably larger than the ones considered here. Under those circumstances, one can simply exclude

[4]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.

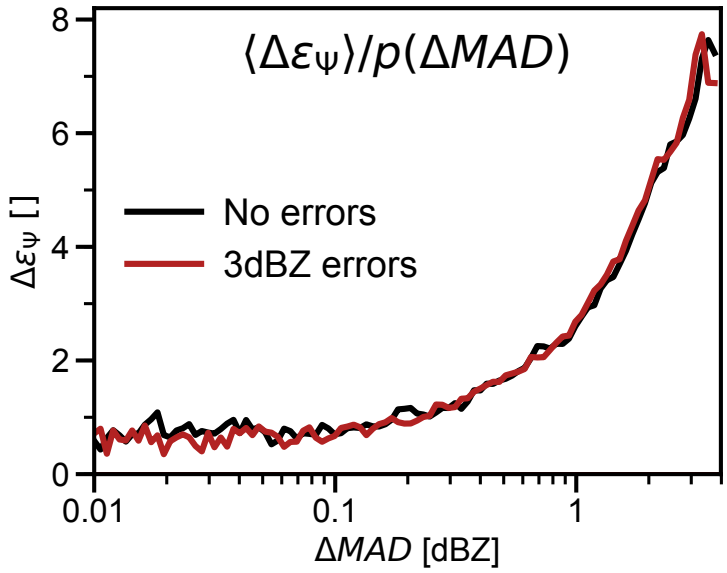

Fig. B1: Expectation values of the decrease in the state error $\Delta\varepsilon_\Psi$ computed using the conditional probability $p(\Delta\varepsilon_\Psi|\Delta MAD)$, where ΔMAD denotes the decrease in the precipitation error. The expectation values are computed as $\langle\Delta\varepsilon_\Psi\rangle/p(\Delta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

the regions with very low-quality observations (Lopez, 2011) or assign quality index (weights 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 observation 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 member 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 observation 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.

Chapter 3

Effect of forecast errors and assimilation of NCEP Stage IV precipitation analyses

In Chapter 2, we presented our first step in the development of the LEMA method. However, the experimental setup used represented very optimistic situations where the background ensemble covered the truth. This chapter investigates how increasing forecast errors, that are not captured by the ensemble spread, degrades the LEMA's effectiveness. Here, we present a method to increase the effectiveness of LEMA by expanding the spread of the background ensemble used to construct the analysis. Finally, this chapter also explores the potential of LEMA using the expanded ensemble in the context of real DA experiments.

This chapter consists of a paper in published in the journal Monthly Weather Review: Pérez Hortal, A. A., I. Zawadzki, and M. K. Yau, 2020: A Heuristic Approach for Precipitation Data Assimilation: Effect of Forecast Errors and Assimilation of NCEP Stage IV Precipitation Analyses. Mon. Wea. Rev., 148, 1629–1651, https://doi.org/10.1175/MWR-D-19-0331.1.

Superscript numbers indicate published footnotes while superscripts starting with "E" indicate

editorial notes used for clarifications on the published manuscript. Also, the current chapter includes corrections to minor grammatical errors in the published manuscript.

A Heuristic Approach for Precipitation Data Assimilation. Effect of forecast errors and assimilation of NCEP Stage IV precipitation analyses

Andrés A. Pérez Hortal, Isztar Zawadzki and M. K. Yau

Department of Atmospheric and Oceanic Sciences, McGill University, Montréal, Québec, Canada

Abstract

Recently, Pérez Hortal et al. (2019) introduced a simple data assimilation (DA) technique named "Localized Ensemble Mosaic Assimilation" (LEMA) for the assimilation of radar-derived precipitation observations. The method constructs an analysis by assigning to each model grid point the information from the ensemble member that is locally closest to the precipitation observations. This study explores the effects of the forecasts errors in the performance of the method using a series of Observing System Simulation Experiments (OSSEs) with different magnitudes of forecast errors employing a small ensemble of 20 members. The ideal experiments show that LEMA is able to produce forecasts with considerable and long-lived error reductions in the fields of precipitation, temperature, humidity, and wind. Nonetheless, the quality of the analysis deteriorates with increasing forecast errors beyond the spread of the ensemble. To overcome this limitation, we expand the spread of the ensemble used to construct the analysis mosaic by considering states at different times and states from forecasts initialized at different times (lagged forecasts). The ideal experiments show that the additional information in the expanded ensemble improves the performance of LEMA, producing larger and long-lived improvements in the state variables and in the precipitation forecast quality.

Finally, the potential of LEMA is explored in real DA experiments using actual StageIV

precipitation observations. When LEMA uses only the background members, the quality of the precipitation forecast shows small or no improvements. However, the expanded ensemble improves the LEMA's effectiveness, producing larger and more persistent improvements in precipitation forecasts.

3.1 Introduction

Although the assimilation of precipitation into Numerical Weather Prediction (NWP) models has a significant potential to improve the forecast quality, assimilating these observations is a particularly challenging task. One of the main reasons is that precipitation results from multivariate non-linear processes that involve many state variables (e.g. temperature, humidity, pressure, winds). As such, obtaining precipitation from the model state is a deterministic forward problem with a unique solution, but the inverse problem is not: a given precipitation value can result from many different meteorological situations (states). Therefore, precipitation observations contain statistical information on each of the state variables related to the precipitation processes.

Conventional variational and Ensemble Kalman Filters (EnKF) data assimilation (DA) methods solve this inverse probabilistic problem assuming Gaussian distributions for the observations errors and prior estimates of the atmospheric states (Lorenc, 1986; Hamill, 2006). Several studies used these Gaussian DA approaches to assimilate precipitation observations over large domains with different degrees of success (Koizumi et al., 2005; Lopez and Bauer, 2007; Lopez, 2011; Kumar et al., 2014; Lien et al., 2016; Kotsuki et al., 2017). Nonetheless, the highly non-linear moist physical processes and the non-Gaussian characteristics of precipitation errors complicate the assimilation of these observations using conventional approaches (Errico et al., 2007; Bauer et al., 2011; Lien et al., 2013).

In addition, simpler and more economical diabatic initialization (nudging) methods are fre-

quently used for precipitation DA. They use theoretical or empirical relationships to modify the trajectory of a numerical model by forcing the model precipitation towards the observed values by adjusting the humidity or temperature profiles (e.g. Falkovich et al., 2000; Davolio and Buzzi, 2004; Davolio et al., 2017; Macpherson, 2001; Bick et al., 2016; Jacques et al., 2018). These methods successfully force the precipitation towards the observed values during the assimilation windows, but most of the positive impacts in precipitation obtained by the DA are quickly forgotten ($\sim$ 3 to 6 hours).

Recent studies explore the potential of non-Gaussian approaches based on Monte Carlo methods, such as Particle Filters (PF, van Leeuwen, 2009). Instead of assuming a Gaussian probability distribution for the model state, PF describe the model probability distribution by a discrete set of model states (ensemble members called particles). The posterior distribution properties are estimated using a weighted combination of all the members, where the members that are closer to the observations receive a higher weight. Poterjoy (2016) introduced a localized implementation of the PF (the Local Particle Filter or LPF) that operates efficiently in high-dimensional systems and has potential for NWP applications. He obtained forecasts with an accuracy similar to those generated using an EnKF system (Poterjoy et al., 2017, 2019), although PF methods do not assume any particular distribution but rely only on the information contained in the ensemble forecast for the estimation of the atmospheric state.

In Pérez Hortal et al. (2019, hereafter PZY19), we introduced a heuristic approach for precipitation assimilation that directly uses the information in the ensemble (without relying on parametric assumptions) to estimate the atmospheric state. The method, called Localized Ensemble Mosaic Assimilation (LEMA), does not assume any particular shapes for the prior probability density (like Gaussianity). Instead, it is based on the intuitive idea that increasing the proximity in precipitation between model and truth will, on average, increase the proximity to the truth in state variables that caused the precipitation in the first place. This idea follows from the joint

probability distributions of precipitation errors and state variables errors observed in the ensemble forecasts. In terms of forecast quality, the Observing System Simulating Experiments (OSSEs) in PZY19 showed that LEMA produces a significant improvement in precipitation forecast that persists up for at least 12 hours after the assimilation time. The improvements in precipitation are associated with a long-lasting reduction in the state variable errors (potential temperature, vapor mixing ratio, u-wind, and v-wind). However, the experiments in PZY19 were performed within ideal conditions: no model errors were considered (i.e. no shortcomings in representing the actual state of the atmosphere), and by construction, the truth was correctly covered by the background ensemble. This is not the case in real ensemble prediction systems where the ensemble spread does not represent the actual forecast uncertainty (Fortin et al., 2014; Zhou et al., 2017). Since LEMA only relies on the information contained in the ensemble forecast to construct the analysis, in order to be effective, the spread of the ensemble members used by LEMA must sample all the possible forecast errors (timing, intensity, and model errors). Therefore, the underdispersivity of the ensemble forecasts will affect the performance of LEMA (as well as for any other methods that exclusively and strongly rely on the information in the ensemble).

The present study extends the OSSEs in PZY19 by considering situations that are likely to occur in real DA cases where the reality may be found partially or fully outside the spread of the background ensemble (underdispersive ensembles). As in PZY19, we extensively use OSSEs despite their limited usefulness for data assimilation in an operational context, where a large discrepancy exists between forecasted and observed precipitation, that is typically beyond the differences in the context of OSSEs. However, an important component of our objectives is to understand how the information contained in ensemble forecasts propagates into state variables and the factors that may correct model trajectory over longer lead-times, including the optimal scale of DA (PZY19) and ways to alleviate the underdispersivity of the ensemble. For these objectives, OSSEs are essential.

Another objective of these experiments is to investigate how increasing forecast errors, that are not captured by the ensemble spread, affects LEMA's effectiveness. In particular, we will show that the quality of the analysis constructed by LEMA deteriorates with increasing forecast errors that are beyond the ensemble spread (underdispersive ensembles). A further objective is to present and characterize our first attempt to expand the ensemble spread by considering model states at different times and also states from forecasts initialized at different times (lagged forecasts).

Finally, we will explore the potential of LEMA with the expanded ensemble in DA experiments using actual StageIV precipitation observations, considering situations where the forecast errors exceed the errors captured by the spread of background ensemble. For these unfavorable situations, when LEMA uses only the background members, the quality of the precipitation forecast shows small or no improvements. Nonetheless, we will show that the expanded ensemble provides a better representation of the actual forecast uncertainties. This improves LEMA's effectiveness in producing larger and more persistent improvements in precipitation forecasts.

The paper is organized as follows. The ensemble forecasts for the nine meteorological situations used in the OSSEs and the real DA experiments are described in Section 3.2. In Section 3.3 we recapitulate the LEMA method with additional details. In Section 3.4 we evaluate the influence of forecast errors on the analysis quality using idealized OSSEs, showing the benefits of using the Expanded Ensemble in LEMA. In Section 3.5 we present the results of the DA experiments with actual precipitation observations and present the benefits of using the Expanded Ensemble. Section 3.6 presents a discussion on the properties of the expanded ensemble and its relevance to other DA methods. Finally, Section 3.7 contains a summary and conclusions.

3.2 Description of the experimental setup

3.2.1 Ensemble Forecasts of the Precipitation Events

For this study, we consider 9 precipitation events, named Cases A to I. Five of the nine cases (A, B, E, F, G) are widespread precipitation events driven by cyclonic systems and with varied degrees of organization. In the rest of the cases (C, D, H, and I), precipitation was produced by several Mesoscale Convective Systems (MCSs) scattered over the US. A brief description of the precipitation events is shown in Table 3.1.

Case A is selected to characterize the effects of forecast errors on LEMA's performance in the context of OSSEs. Afterwards, all the nine cases are used to explore the potential of the method in a real DA application using actual StageIV observations.

Our experiments employ the same NWP model and configuration as in PZY19. The numerical model is the Weather Research and Forecasting Model (WRF) with the Advanced Research WRF (ARW) dynamic solver (WRF-ARW) version 3.7.1 (Skamarock and Klemp, 2008). The model domain covers the continental US and southern Canada. The simulations are carried out using 41 vertical levels and 20 km horizontal grid-spacing. The lateral boundary conditions (LBCs) and initial conditions (ICs) are downscaled from the Global Ensemble Forecast System (GEFS) data (1° resolution) to the WRF grid.

Table 3.1: Ensemble forecasts for each precipitation event

Case	Type	Analysis date	Name	Simulation start date
A	Synoptic front	Apr 11 2013 0000 UTC	En-6	Apr 10 1800 UTC
			En-12	Apr 10 1200 UTC
			En-18	Apr 10 0600 UTC
B	Cyclonic systems	Apr 04 2013 1800 UTC	En-6	Apr 04 1200 UTC
			En-12	Apr 04 0600 UTC
			En-18	Apr 04 0000 UTC
C	MCSs	May 18 2013 0600 UTC	En-6	May 18 0000 UTC
			En-12	May 17 1800 UTC
			En-18	May 17 1200 UTC
D	MCSs	May 27 2013 0000 UTC	En-6	May 26 1800 UTC
			En-12	May 26 1200 UTC
			En-18	May 26 0600 UTC
E	Synoptic front	Dec 28 2015 0600 UTC	En-6	Dec 28 0000 UTC
			En-12	Dec 27 1800 UTC
			En-18	Dec 27 1200 UTC
F	Synoptic front	Oct 14 0600 UTC	En-6	Oct 14 0000 UTC
			En-12	Oct 13 1800 UTC
			En-18	Oct 13 1200 UTC
G	Synoptic front	Aug 15 2016 1800 UTC	En-6	Aug 15 1200 UTC
			En-12	Aug 15 0600 UTC
			En-18	Aug 15 0000 UTC
H	MCSs	Jun 04 2014 0600 UTC	En-6	Jun 04 0000 UTC
			En-12	Jun 03 1800 UTC
			En-18	Jun 03 1200 UTC
I	MCSs	Apr 19 2012 1800 UTC	En-6	Apr 19 1200 UTC
			En-12	Apr 19 0600 UTC
			En-18	Apr 19 0000 UTC

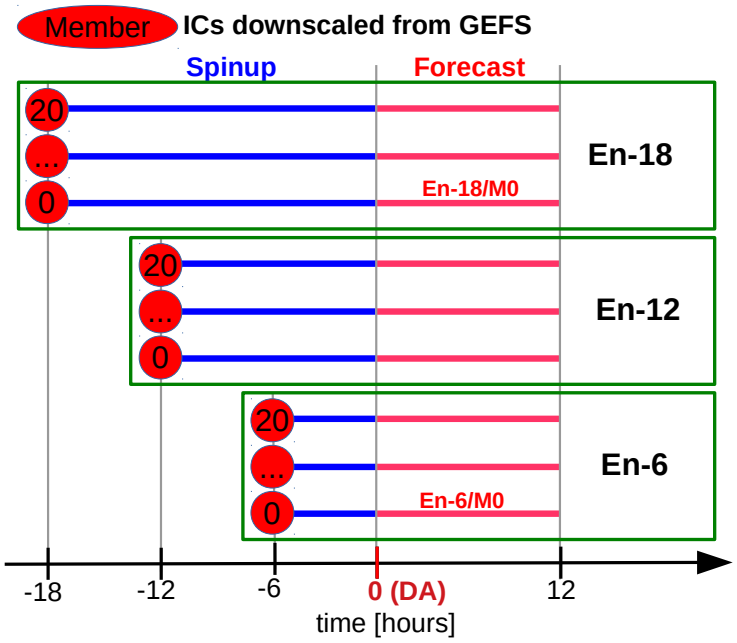

Fig. 3.1: Lagged-forecast scheme used to produce the ensemble forecasts. The red circles indicate the ICs downscaled from GEFS data while the numbers inside the circles denote the GEFS member. The blue lines indicate the spin-up period while the red ones show the forecast period. Data assimilation is performed at t=0 h and a 12 h forecast is run for each member.

The main physics options used in the experiments are the WRF single moment 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.

For each case, we produce three ensemble forecasts of 21 members initialized at 6 h, 12 h, and 18 h before the DA is performed. These three ensembles will be denoted by *En-6*, *En-12* and *En-18* respectively. The ICs/LBCs for each of the ensemble forecast are downscaled from GEFS forecast data initialized at the same time as the WRF forecast initialization (Fig. 3.1). A detailed description of the ensemble forecasts for each case is shown in Table 3.1.

These ensemble forecasts are used to construct different idealized OSSEs as well as DA experiments with actual StageIV observations.

3.2.2 Observing System Simulating Experiments

In PZY19, the experimental setup used was optimal: the truth was covered by the background ensemble (symmetrical spread around the truth), and no model or observations errors were considered. Before considering assimilating actual precipitation observations, we should asses the consequence of forecast errors. Here, we will extend the previous study by considering more significant forecast errors. In these OSSEs, one model run is considered the "true" atmosphere and a different set of runs is considered as the background ensemble. The different forecast errors will be simulated by using different time lags between the initialization time of the truth and the background runs.

The main objective of the OSSEs is to understand how information contained in ensemble forecasts propagates into state variables for different situations with increasing forecast errors that are not captured by the ensemble spread. A second objective is to present and characterize our first attempt to expand the ensemble spread by considering model states at different times and also states from forecasts initialized at different times (lagged forecasts). To that end, we will select Case A for our analysis using an idealized setup. We carry out two OSSEs using the same background ensemble, but each experiment considers a different member as the truth. For easier comparison with the real DA experiments in Section 3.5, for all the idealized experiments, a complete set of hourly precipitation accumulation observations are simulated from the member considered as the truth. The synthetic hourly-precipitation observations are available every hour in the same domain as the StageIV observations used in the real DA experiments (observations located east of 105°W). For simplicity, no model or observations errors are considered.

In the first experiment, we consider the *En-6* runs with the member 0 as the truth run and

the rest of the members as background. This represents an optimal situation where the truth is located at the center of the background ensemble. For this experiment, there is no initialization lag between the truth and the backgrounds runs. We denote this experiment by $\tau = 0h$, where τ indicates the initialization lag.

In a second experiment, to simulate large forecast errors, we use the same background as in $\tau = 0h$, but the "truth" run is taken to be the member 0 from the ensemble initialized 18 h before the analysis time (*En-18/M0*). For this experiment, named $\tau = 12h$, there is a 12 h initialization lag between the truth and the background.

As we will show in Section 3.4.1, the background errors for each experiment increase with the increasing initialization lag between the truth and the background runs. In this section, we also explore the benefits of expanding the background ensemble for the analysis construction. For this, we perform a third experiment, $\tau = 12h/Xpd$ (*Xpd* indicates the expanded ensemble), similar to the $\tau = 12h$ experiment but constructing the analysis using the members from the background ensemble (member 1-20 from *En-6*) as well as all the members from the *En-12* ensemble forecast. In addition, we further increase the ensemble by considering states at different times surrounding the analysis time 0, ±1 h, ±2 h. It is important to note that the expanded ensemble is used only during the analysis construction. The new ensemble forecast is still produced by relaxing each background member towards the analysis.

A summary of all the OSSEs for a given case is shown in Table 3.2a.

Table 3.2: Experiments type summary

a) Observing System Simulating Experiments

Name	Truth	Background	Expanded ensemble
$\tau = 0h$	En-6/M0	En-6/M1-20	No
$\tau = 12h$	En-18/M0	En-6/M1-20	No
$\tau = 12h/Xpd$	En-18/M0	En-6/M1-20	background+En-12

M# denotes the "member #" while M#$_1$-#$_2$ denotes members #$_1$ to #$_2$.

τ indicates the initialization lag between the background and the "truth" runs.

"Xpd" denotes the experiments using the Expanded ensemble in LEMA

b) StageIV DA Experiments

Name	Truth	Background	Expanded ensemble
StageIV/En-6	StageIV (East)	En-6	No
StageIV/Xpd	StageIV (East)	En-6	En-6+En-12+En-18

3.2.3 StageIV DA experiments

For the real DA, we assimilate the Stage IV hourly precipitation, a precipitation product that combines precipitation estimates from about 150 radars and about 5500 hourly rain gauge measurements over the continental United States (Baldwin and Mitchell, 1997; Lin and Mitchell, 2005). Over the Rocky Mountains area, the precipitation estimates are less reliable due to ground clutters, radar beam blockage, sparser radar coverage and poorer rain-gauge representativity. Therefore, all the observations located west of 105°W are discarded as in Lopez (2011).

The original precipitation data are available in a polar-stereographic grid with a 4 km grid spacing. To interpolate the observations to the model grid (20 km grid spacing), we first smooth the Stage IV data by running a 5 x 5 grid points moving-average window. This averaging matches the effective observations resolution with the model resolution and removes the small scales that are not represented by the model. Finally, the observations are interpolated onto the model grid

using bilinear interpolation.

For the real DA with actual StageIV precipitation observations, we use the nine precipitation events described in Table 3.1. For each precipitation event, we conduct two experiments: one using only the background to construct the Frankenstate, and the other using an expanded ensemble. For both experiments, we use as background ensemble the 21 members in the *En-6* runs described in Section 3.2.1[E10]. The analysis is constructed 6 h after the background runs are initialized using the StageIV precipitation observations.

In the first experiment (*StageIV/En-6*), we construct the Frankenstate using only the information in the background ensemble at the analysis time. In the second experiment (*StageIV/Xpd*), we construct the analysis using an expanded ensemble that includes the background ensemble *En-6*, and two other ensemble forecasts initialized at 12 h and 18 h before the analysis time (*En-12* and *En-18* runs, respectively) described in Sec. 3.2.1. The ensemble is expanded even further by considering states at different times t surrounding the analysis time (t_a), where $t = t_a + \Delta t$, with $\Delta t = 0, \pm 1h, \pm 2h$ as described in Section 3.3. This procedure expands the ensemble used to construct the analysis from the 21 members used in the first set of experiments to 315 members (21 members x 3 ensemble forecasts x 5 times).

A summary of the real DA experiments for a given case is shown in Table 3.2b.

3.3 Revisiting LEMA

The Localized Ensemble Mosaic Assimilation (LEMA) creates an analysis using the information in the background ensemble by assigning to each vertical column in the model, the vertical profile of the state variables (u and v wind, temperature, humidity) from the ensemble member that is

[E10]Although the two experiments construct the analysis mosaics using a different ensemble of states (only En-6 versus the expanded ensemble), each new ensemble forecast is initialized by relaxing the same background ensemble towards the different analyses.

locally the closest to the precipitation observations (here, hourly accumulation). This analysis mosaic constructed using the selected members will be referred to as "Frankenstate".

The local error in precipitation (local proximity) is measured by the Mean Absolute Difference (MAD) between the forecasted hourly precipitation and the observed values, computed over a large square-windows of $\Delta x = 820\ km$ width centered in the column. In PZY19, the MAD values are computed using precipitation intensity (in mm/h) expressed in dBZ values, transformed using the Marshal-Palmer relation. Although the same mathematical transformation can be used for hourly precipitation (in mm), the resulting units are not equivalent to dBZ, although they can be considered as "pseudo-reflectivity." However, to avoid confusion with the reflectivity units, in this study we express the hourly precipitation in dBR units ($dBR = 10log_{10}R$, with R in mm)[E11].

Therefore, for a given member "m", the distance to the observations around the (i,j) horizontal grid point is:

$$MAD_m(i,j) = \frac{1}{N_x N_y} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} \left| dBR_{x,y}^{obs} - dBR_{x,y}^{m} \right| \tag{3.1}$$

where dBR^{obs} and dBR^{m} indicates the precipitation observations and the precipitation values corresponding to the m^{th} member, respectively. The sub-index x and y denotes the x-index and y-index of the horizontal grid point inside the observation windows. The summation limits, N_x and N_y, denotes the total number of horizontal grid points in the square window in each direction.

For each horizontal grid point, the ensemble member with the lowest MAD value is considered the "locally closest" one to the observations. To ensure that only members with MAD values strictly greater than zero are used in the closest member selection, during the selection process only members with a minimum precipitation coverage of $n_{min} = 35$ grid points over the localization window are used. If over the observations window, the observed precipitation or no background

[E11] For the dBR transformation, the lowest value allowed for precipitation is 0.03 mm/h. If the precipitation values are lower than this threshold, the $dBR_{min} = 10log_{10}(0.03)$ value is assigned to the transformation.

member exceeds the minimum coverage n_{min}, no information is assigned to that analysis column.

The new ICs for the ensemble forecast is produced by gradually forcing each member of the background ensemble towards the Frankenstate over a 30 min period preceding the analysis (nudging). With this procedure, the possible effects of sharp boundaries in the mosaic of closest members is attenuated[1]. The relaxation towards the Frankenstate is done by adding artificial terms to the model's prognostic equations:

$$\left(\frac{\partial \phi(t)}{\partial t}\right)_{new} = \left(\frac{\partial \phi(t)}{\partial t}\right)_{model} + G\left(\phi^F - \phi(t)\right) \tag{3.2}$$

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 respect to other model processes. During the initialization, the nudging is only applied over the columns where a local closest member is found. On all the other grid points, we let the background to evolve without any artificial forcing. The forecasts initialized from the resulting ICs will be referred to as "Frankencasts".

Since LEMA relies on the information contained in the ensemble forecast only, in order to be effective, the spread of the ensemble members used in the Frankenstate construction must be large enough to cover the reality. In other words, the ensemble used in the analysis construction should sample all the possible forecast errors (timing, intensity, and model errors). Since these forecast errors may not be captured by the background ensemble, we modify the original LEMA approach (PZY19) by expanding background ensemble used to construct the analysis. The ensemble expanded by considering model states from forecasts initialized at previous times (time-lagged ensemble forecasts, see Hoffman and Kalnay, 1983; Dalcher et al., 1988; Van Den Dool and Rukhovets, 1994; Lu et al., 2007). Additionally, to further increase the sample of possible

[1]Strictly speaking, nudging is not necessary. A filtering of high frequencies can also be used or gravity waves generated by the discontinuities can simply be allowed to dissipate). However, the nudging procedure sets the stage for a future exploration in combining LEMA with Latent Heat Nudging.

forecast errors, the ensemble is also expanded by considering, for each available member, states at different times t around the analysis time (t_a), where $t = t_a + \Delta t$, with $\Delta t = 0, \pm 1h, \pm 2h$. The final goal of this expansion is increasing the spread of the background ensemble to cover the "truth".

Let us first revise the fundamental elements of LEMA, considering the experiment $\tau = 0h$ with the truth located at the center of the background ensemble for Case A. We will consider the joint probability of the errors in potential temperature (ε_θ) for a model column and the local precipitation error (MAD, eq. 3.1) around that column. Only errors in potential temperature are shown, but similar results hold for the other variables. The joint probability is estimated from the background ensemble by the bi-dimensional histogram of the MAD, ε_θ pairs (columns), sampled over all the members and the entire domain, considering points with $MAD > 0$.

Figure 3.2a shows the joint probability of error in potential temperature and precipitation $p(\varepsilon_\theta, MAD)$ when MAD is computed using an 820 km localization windows. The solid curve is the expected value of the conditional probability $p(\varepsilon_\theta | MAD)$, representing the average relationship $\varepsilon_\theta = f(MAD)$ per grid point. We can see that, on the average, when MAD decreases, the error in state variables steadily decreases as well. The joint probability can also be interpreted as the potential for improvement in potential temperature as MAD is reduced. This is the conceptual basis for LEMA: we assign to each grid point the vertical profile of the state variables (θ, q_v, U, V) of the ensemble member with the smallest MAD. This creates a mosaic of states (called Frankenstate) that, on the average, has smaller errors in the state variables.

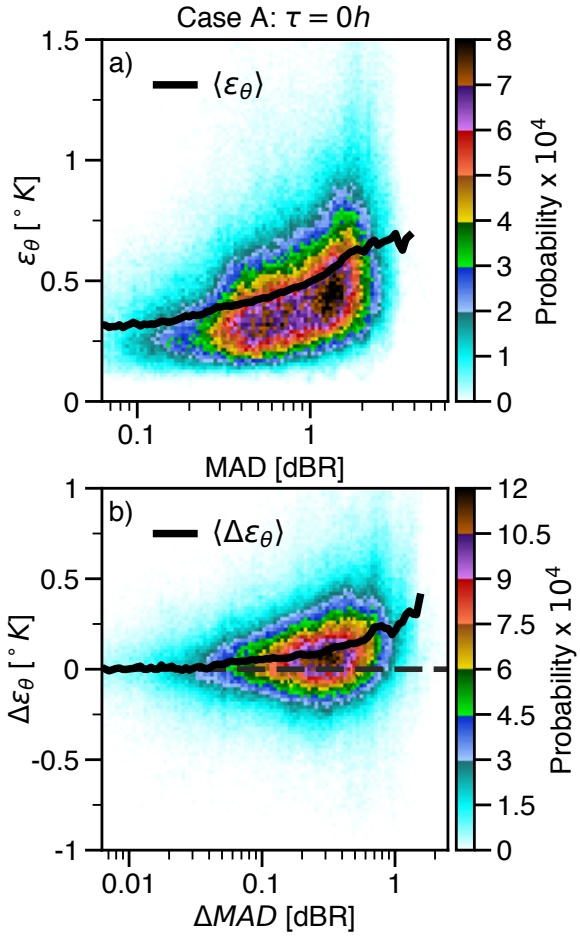

Fig. 3.2: Case A, experiment $\tau = 0h$. The first row shows the joint probability $p(\varepsilon_\theta, MAD)$ of given RMSE in potential temperature ε_θ and an error in precipitation MAD while the second row shows the joint probability $p(\Delta\varepsilon_\theta, \Delta MAD)$ of decrease in errors when the member with lowest MAD is selected. The black curves indicate the expectation values with respect to the conditional probability. $\langle \varepsilon_\theta \rangle = \sum_{\varepsilon_\theta} \varepsilon_\theta p(\varepsilon_\theta | MAD) = \sum_{\Delta\varepsilon_\theta} \varepsilon_\theta p(\varepsilon_\theta, MAD)/p(MAD)$. $\langle \Delta\varepsilon_\theta \rangle = \sum_{\Delta\varepsilon_\theta} \Delta\varepsilon_\theta p(\Delta\varepsilon_\theta | \Delta MAD)$.

Figure 3.2b shows the joint probability of the errors decrease[2] with respect to the Frankenstate (members with lowest MAD). The average gain in the ε_θ is approximately zero for small MAD gains, and increases as the reduction in MAD increases. However, there are grid points where the

[2]The decrease in the errors is formulated in such a way that positive values represent a positive impact in the analysis or forecast quality.

errors of these variables increase even when the MAD decreases. This a result of the stochasticity in the $\varepsilon_\theta = f(MAD)$ relationship. Given this degree of stochasticity, we may ask ourselves whether the member closest to the true precipitation is actually the best at maximizing the gain in state variables. Figure 3.3 shows the gain in state variables as a function of the proximity to true precipitation of the n^{th} closest member. The results are shown for θ and q_v, but similar results are for U and V winds (not shown). Depending on the variable considered, 3 (for U, V and θ) to 6 (for q_v) members close to the truth in precipitation are as effective as the closest one to reduce the error in state variables. Hence, this stochasticity in the joint probability of errors (Fig. 3.2a) has a beneficial aspect on LEMA: it provides a degree of tolerance to observations errors (Appendix B, PZY19).

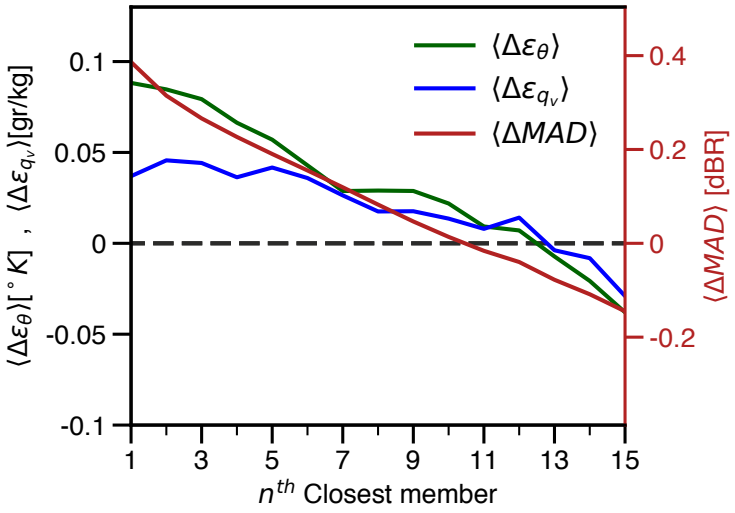

Fig. 3.3: Case A, experiment $\tau = 0h$. Decrease in error (gain) in potential temperature ($\Delta\varepsilon_\theta$, green line) and vapor mixing ratio ($\Delta\varepsilon_{qv}$, blue line) as a function of the proximity to the true precipitation of the n^{th} closest member. The red lines indicate the decrease in the precipitation error ΔMAD when the n^{th} closest member is selected.

Additionally, the adjustment of the atmospheric state will depend on the measure of the distance between model and observations used by LEMA and the optimization criteria used. In our study, we use the local MAD as a distance metric because it does not emphasize localized

high values of errors. Using MAD as a distance measure, LEMA is optimized (by selecting the localization windows size) to maximize the transfer of information from precipitation to the state variables and for correcting the model trajectory in a way that remains closer to the observations for a longer lead-time. In PZY19, we showed that the large observation window has a double benefit. It extends the area of influence of the observations as a larger portion of the domain is affected by LEMA and it also improves the transmission of information from precipitation to the state variables. The physical reason for this is simple: small scale errors are determined by small scale features (precipitation cells). As has been shown by Germann and Zawadzki (2002) and Surcel et al. (2015) and others, small scales are short-lived (in agreement with the general correspondence between time and space scales of atmospheric phenomena, Orlanski, 1975). This suggests that, on the average, these precipitation scales do not contain information relevant for a long-lasting correction of model trajectory. In other words, for data assimilation aiming at improvements lasting over 10 hours or more at regional scales, the relevant information is not where it is intensely raining now but in what general region significant precipitation now occurs.

To further emphasize the last point, let us consider the scale dependence of forecast in the $\tau = 0h$ experiment described in Sec. 3.2, after assimilating synthetic observations using a 20 km and an 820 km localization windows. Figure 3.4 shows 12 hours forecasts after precipitation was assimilated using MAD at 20 km and 820 km. The diagram shows the relative decrease in error in θ and q_v at various scale intervals (similar results hold for U and V winds but not shown)[3]. When the Frankenstate is constructed using MAD at 20 km resolution, the analysis is constructed over the precipitation area getting closer to the detailed structure of the observed precipitation. On the other hand, using the large-scale component of MAD to correct model trajectory leads to an improvement in small scales ($\lambda < 120\ km$) as well as in the large scales, with the gain increasing

[3]The scale decomposition by intervals is done using the Discrete Cosine Transform (DCT, Denis et al., 2002) and a 10^{th} order Butterworth band-pass filter. The DCT is equivalent to the fast Fourier transform (FFT), but it eliminates the problems associated with discontinuities at the boundaries of the domain.

with scale.

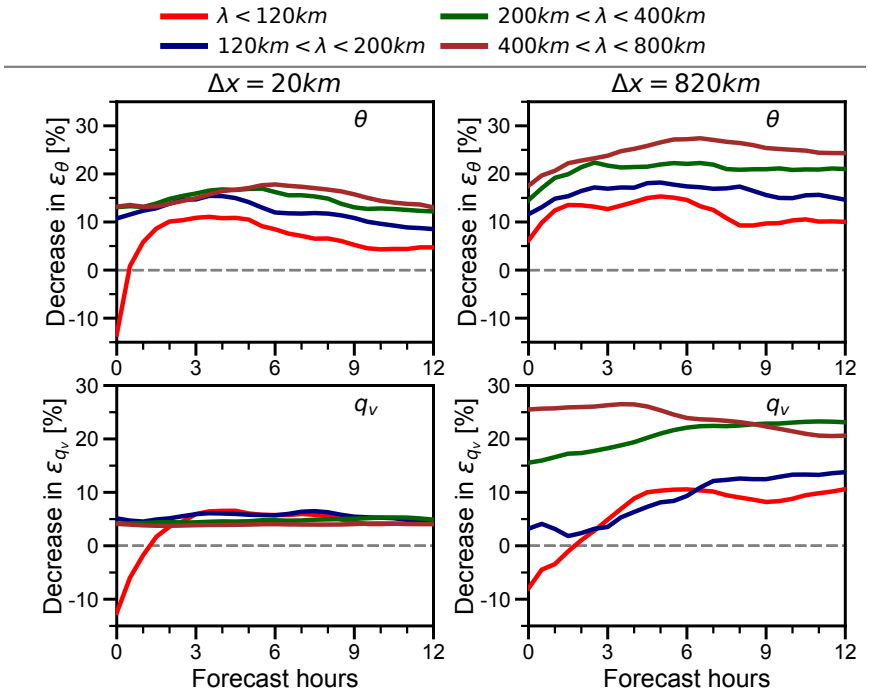

Fig. 3.4: Spatial-scale decomposition of DA impacts for Case A's $\tau = 0h$ experiment using a 20 km localization windows (left panels) and a 820 km windows (right panels). The DA impacts are measured by the relative improvements in the RMSE (ε) for potential temperature θ and vapor mixing ratio q_v, computed considering different spatial-scale intervals. The lines denote the decrease in RMSE averaged over all the ensemble members.

3.4 Influence of forecast errors in the quality of the analysis

3.4.1 Effect of Forecast Errors

Hitherto, we considered a situation with a perfect model and the truth located at the center of the background ensemble. Figure 3.5a compares, at the time when DA takes place, the precipitation probability in the background ensemble with the truth precipitation for the $\tau = 0h$ experiment (Case A). In this optimistic situation, the background ensemble captures remarkably well the

"truth" precipitation pattern, with the regions with a high probability of precipitation (purple colors) being well collocated with the "truth" precipitation (black contours). This corresponds to a situation where the forecast errors are small, or in other terms, when observations fall within the ensemble forecast generated by perturbations of initial conditions. We know that in general this is not the case in real situations: ensemble forecasts are underdispersive[E12]. Hence, before considering assimilating actual precipitation observation, in this section we will assess the consequences of larger discrepancies between the forecast and the "truth" in the context OSSEs. Although the OSSEs may be greatly optimistic compared to real situations, they allow us to understand how LEMA works under different scenarios. In particular, we want to address how the potential for improvements in the state variables is affected by the magnitude of the forecast errors. A similar analyses is a challenging task in real DA cases where we only have a partial knowledge of the actual atmospheric state.

For this, let us consider the same background ensemble as before but the "truth" is taken from a run of the same meteorological situation but initialized 12 h before the background initialization. This corresponds to the $\tau = 12h$ experiment described in Sec. 3.2.2. The larger initialization lag (12 h) results in more substantial differences between the background and the "truth" precipitation patterns, with a significant mismatch between the areas with a high probability of precipitation in the background and the truth precipitation pattern (Fig. 3.5b). This indicates a less optimistic scenario where the truth is only partially covered by the background ensemble.

[E12]Corrigendum. This statement incorrectly describes underdispersion as a systematic problem of all the operational ensemble forecast systems. However, some operational forecasts have a good overall match between the errors and the ensemble spread. Therefore, to better reflect the current state of the art, this statement should mention that the ensemble forecasts may underestimate the actual uncertainties for particular situations (e.g. Haiden et al., 2019).

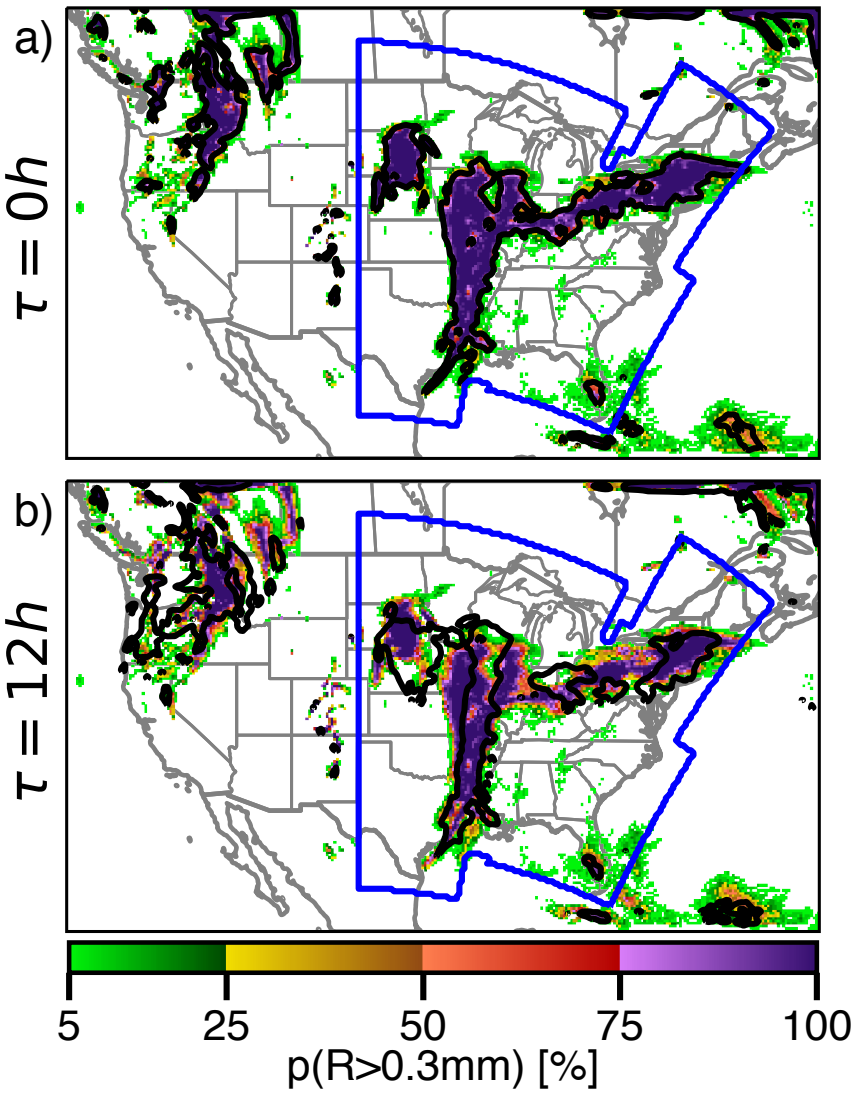

Fig. 3.5: Case A OSSEs overview at 2013/04/11 0000 UTC, when DA is performed. The first row corresponds to the $\tau = 0h$ experiment, while the second to the $\tau = 12h$ experiment. The color-plots depict the precipitation probability in the background ensemble (hourly-accumulation values greater than 0.3 mm). The black contours show the synthetic precipitation observations while the blue contour indicates the area where the observations are available for the DA.

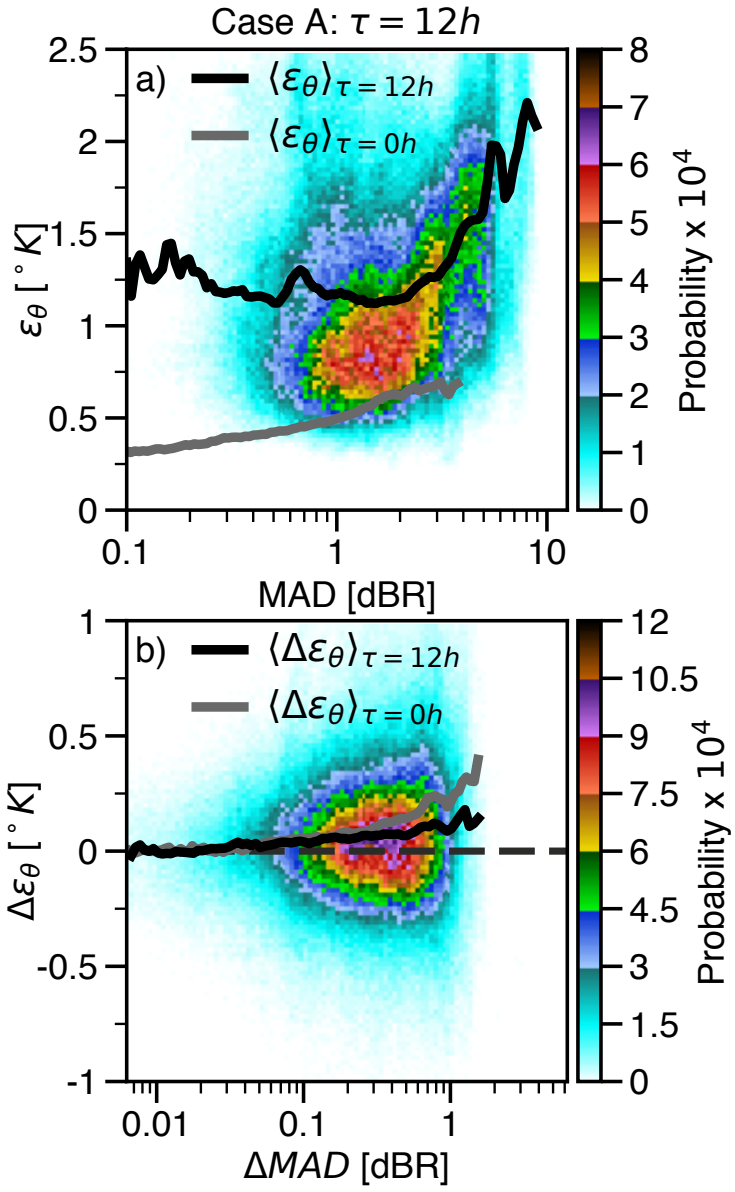

Fig. 3.6: As in Fig. 3.2 but for experiment $\tau = 12h$.

Figure 3.6a shows the joint probability of the decrease in the errors for this experiment. For comparison, the curve of the expected values from the $\tau = 0h$ experiment is shown here in grey. We notice that the errors in potential temperature have increased appreciably (by $\sim 0.7°K$). For small values of MAD, the expected value of ε_θ lost its dependence with MAD. This indicates that in regions with small MAD values, LEMA will not decrease the errors in state variables. For larger values of MAD, the average relationship between errors is maintained, and for the extreme large MAD the increase in ε_θ with increase of MAD is even stronger than the $\tau = 0h$ experiment. In this way, $p(\varepsilon_\theta, MAD)$ is a diagnostic tool for assessing the effect of forecast errors on the analysis. Figure 3.6b indicates that the improvements in the error of potential temperature is very close to the one with small forecast error ($\tau = 0h$ experiment). In addition, large MAD reductions (ΔMAD) are limited to values smaller than 2 dBR (Fig. 3.6b). Since the initial error was larger (Figure 3.6a) the relative gain is quite small. However, given the decrease in error in Figure 3.6b and considering that we adjust four state variables, in the OSSE context we should expect LEMA to still be effective over the 12 h forecast but with lower skill. We will confirm this towards the end of this section.

Figure 3.7 shows the impacts of DA for Case A on the Root Mean Square Error (RMSE) in potential temperature θ and vapor mixing ratio q_v, for the $\tau = 0h$ and $\tau = 12h$ experiments. The RMSE is computed over the entire domain for all the levels located in the troposphere. Similar results are obtained for U and V wind (not shown). As a result of the larger initialization lag between the background and the truth, the background forecast errors are larger in the $\tau = 12h$ than the $\tau = 0h$ runs (black and red lines in Figs. 3.7a and 3.7c). When the synthetic precipitation observations are assimilated, we obtain a persistent reduction in the errors for the state variables for both experiments (green and blue lines). Nevertheless, the error reduction is more important, in both absolute and in relative terms, for the $\tau = 0h$ experiment (smaller forecast errors). For the $\tau = 0h$ experiments, the ensemble-averaged decrease in RMSE for θ, q_v, U, and V are $\sim 12\%$ while

for the $\tau = 12h$ runs, the ensemble-averaged error reduction is 3-6%, remaining approximately constant over the entire 12 h forecast period.

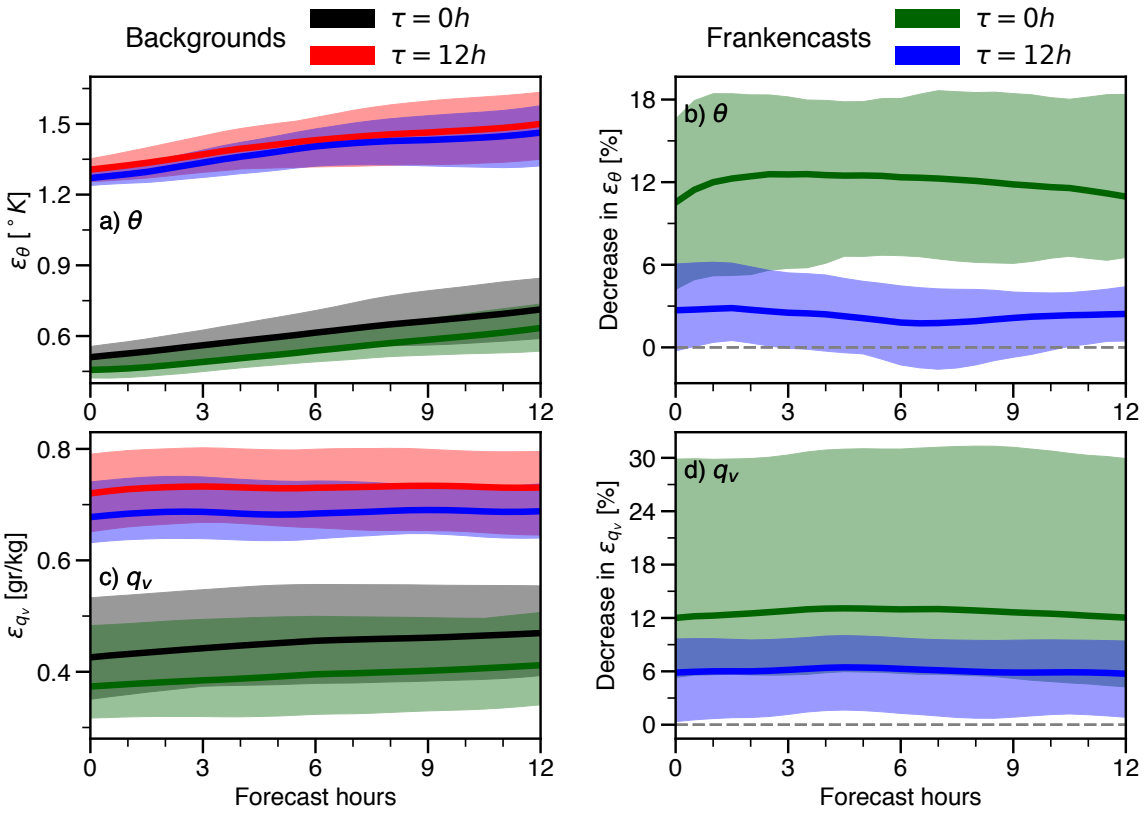

Fig. 3.7: Impacts of DA for Case A OSSEs on the entire domain, measured by the RMSE (ε) for potential temperature θ and vapor mixing ratio q_v. Similar results are for U and V winds. The left panels show the background and the Frankencast RMSE for each variable while the right panels display the relative decrease in RMSE for each Frankencast member. The thick lines denote the RMSE, and the decrease in the RMSE values averaged over all the ensemble members. The shading indicates the corresponding error variability over the ensemble members. Black and red colors indicate the Background errors for the $\tau = 0h$ and $\tau = 12h$ experiments while green and blue colors show Frankencast errors (right) and improvements (left) for the same experiments. The RMSE is computed over the entire domain for all the levels that are located in the troposphere.

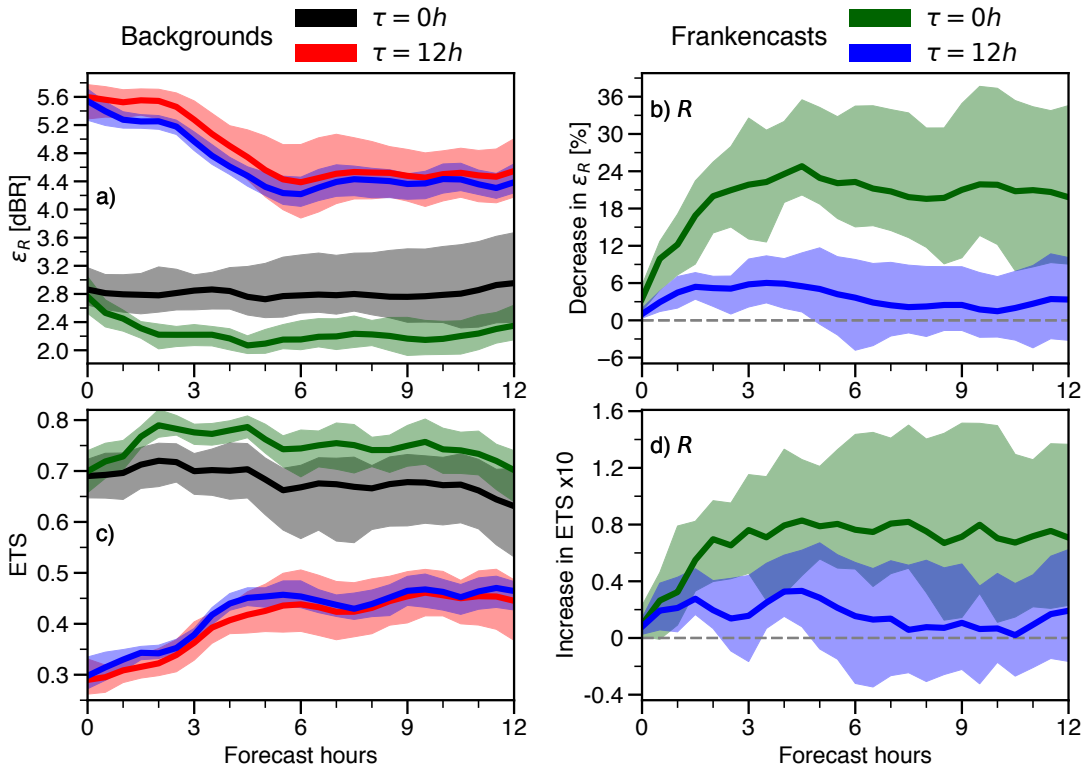

Fig. 3.8: Impacts of DA on precipitation forecasts for the Case A OSSEs computed over the observations domain. The left column shows the background and the Frankencasts errors measured by (a) RMSE in dBR units,ε_R, and (c) ETS. The right panel indicates the error improvement achieved by each Frankencast member, with respect to the corresponding background member. The increase in ETS for the m^{th} member is computed as $ETS_m^F - ETS_m^B$, where "F" and "B" indicate the error for the "Frankenstate" or "background" respectively. The ETS is computed using a 0.3 mm detection threshold. Black and red colors indicate the Background errors for the $\tau = 0h$ and $\tau = 12h$ experiments while green and blue colors show Frankencast errors (right) and improvements (left) for the same experiments. The thick lines denote the ensemble-averaged values, while the shading shows the corresponding error variability over the ensemble members.

The hourly-accumulated precipitation errors are computed against the assimilated observations available only over the eastern United States (blue contour in Fig. 3.5). As for the improvements in the state variables, Fig. 3.8 shows that for the experiment where the background ensemble correctly samples the "truth" ($\tau = 0h$), the DA produces considerable and long-lived improvements. For the $\tau = 12h$ experiment, where the members in the background ensemble are asymmetrically

distributed around the "truth", the DA still produces persistent improvements, but of much smaller magnitude.

3.4.2 Expanded ensemble

In the preceding section, we showed that increasing forecast errors in the background reduce the performance of LEMA. This is not a surprising result since LEMA relies only on the information contained in the ensemble forecast, and for LEMA to be effective, the spread of the ensemble forecast must be enough to reflect its uncertainties and include model states close to the "truth". On the contrary, if all the ensemble members in the background have large errors, the constructed Frankenstate will also have large errors.

The background ensembles used so far have been generated by the perturbations to initial conditions inherited from the GEFS data used to create the ICs/LBCs. In this section, we will use OSSEs to explore the benefits of expanding the ensembles by including different forecast of the same meteorological situation. Concretely, we expand the background ensemble by considering states at different times and also other ensemble forecasts initialized at different times (lagged forecasts).

To that end, we perform a similar experiment to $\tau = 12h$ but constructing the Frankenstate using an expanded ensemble consisting of the background ensemble (*En-6*) and all members from the ensemble *En-12*. In addition, we further increase the ensemble by considering states at different times surrounding the analysis time 0, ±1 h, ±2 h. This results in an expanded ensemble with 200 members (instead of the original 20 members). This corresponds to the $\tau = 12h/Xpd$ experiment described in Sec. 3.2.2.

For the $\tau = 12h/Xpd$ experiment, the additional members in the expanded ensemble produce a small improvement in the proportionality between MAD and ε_θ for small MAD values (Fig. 3.9a). However, in the expanded ensemble, the pool of candidates to be selected for the Frankenstate is

much larger. Hence, there is a greater probability of choosing a members with low MAD values, allowing bigger ΔMAD gains, that leads to larger improvements in the potential temperature errors ($\Delta\varepsilon_\theta$, Fig. 3.9b).

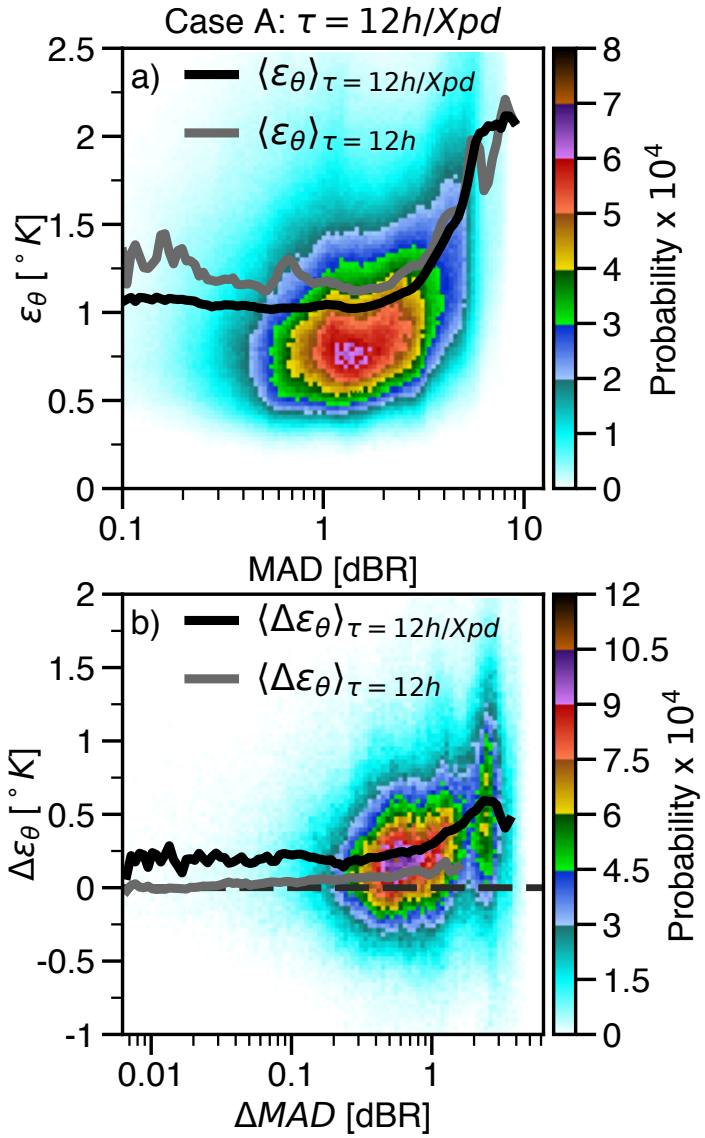

Fig. 3.9: As in Fig. 3.6a but for experiment $\tau = 12h/Xpd$.

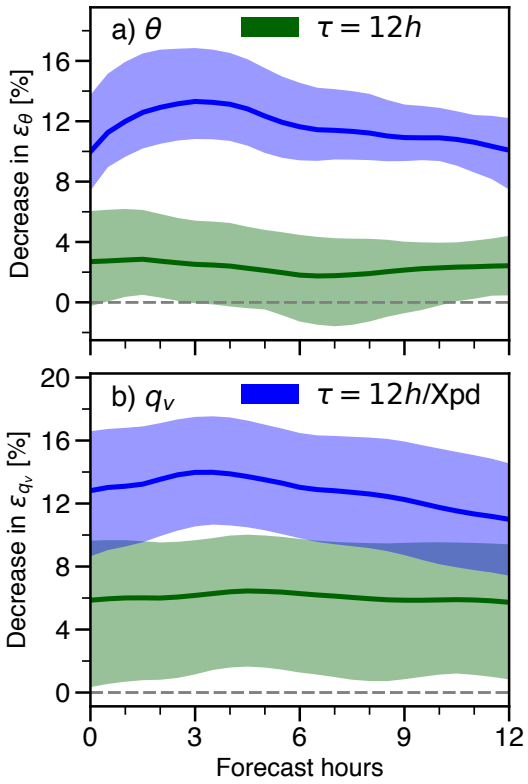

Fig. 3.10: DA impacts on state variables RMSE (ε) for $\tau = 12h$ (green) and $\tau = 12h/Xpd$ (blue) experiments (Case A OSSE). The RMSE is computed over the entire domain. The thick lines denote the ensemble-averaged values, while the shading shows the corresponding error variability over the ensemble members.

Figure 3.10 shows the DA impacts on the RMSE for the state variables when we use the expanded ensemble to construct the Frankenstate. Results are only shown for potential temperature and vapor mixing ratio, but the results are similar for U and V winds. When the Frankenstate is constructed using only the background members *En-6*, LEMA produces only small improvements that persist over the entire 12 h forecast period. However, better results are obtained when the Frankenstate is constructed with the expanded background ensemble ($\tau = 12h/Xpd$). The state variables show a more significant reduction in the RMSE that persist over the entire forecast period (Fig. 3.10). Similar results are obtained for the precipitation forecast quality, computed

over the observations domain (Fig. 3.11). The use of the expanded ensemble to construct the analysis results in a substantial improvement in precipitation forecast quality during the entire forecast period.

Since this procedure allows additional information to be available to LEMA, which translates into better forecast quality, we will consider the expanded ensemble as an integral component of the LEMA technique. We will see in the next section that the concept of Expanded Ensemble is crucial in the assimilation of actual observations.

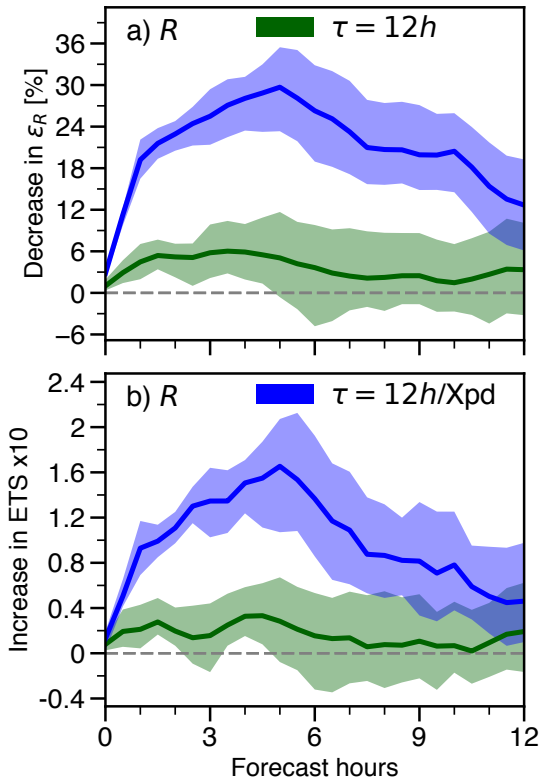

Fig. 3.11: DA impacts on precipitation forecast for $\tau = 12h$ (green) and $\tau = 12h/Xpd$ (blue) experiments (Case A OSSE), computed over the eastern StageIV domain. The same precipitation metrics as in Fig. 3.8 are shown.

3.5 StageIV Data Assimilation Experiments

In this section we present the results of the DA experiments with actual StageIV observations described in Sec. 3.2.3. Similar to the analysis of the DA impacts in the OSSEs, we evaluate LEMA using the RMSE of the hourly precipitation (in dBR units) computed against the StageIV observations. Results for the ETS are not shown since LEMA only produced marginal improvements (< 0.05). However, we will show next that considerable improvements are observed in RMSE, indicating that LEMA is more efficient in correcting the intensity of precipitation. This is expected, since LEMA uses MAD to measure the proximity to the observations that is sensitive to the precipitation intensity while ETS is only sensitive to the precipitation presence.

Let us first show the DA impacts in the forecast quality for the nine DA experiments carried out. Figures 3.12 and 3.13 show the DA impact in the precipitation RMSE when the expanded ensemble is used. To illustrate the benefits of using the expanded ensemble in real DA situations, Figure 3.12 also shows the DA impacts for cases A, B, C, and D when only the background ensemble is used to construct the analysis.

For Case A, when the Frankenstate is constructed using only the background ensemble (*StageIV/En-6*), small improvements in the precipitation RMSE persist for the entire 12 h forecast (Figs. 3.12-b and 3.12-c). Nonetheless, higher gains are obtained when the analysis is constructed using the expanded ensemble (*StageIV/Xpd*).

For Case B, using only the background (*StageIV/En-6*) produces small improvements in RMSE that persist over the entire forecast period (blue line in Figs. 3.12-e and 3.12-f). When LEMA uses the expanded ensemble to construct the analysis (*StageIV/Xpd*), a sharp decrease in the precipitation error is present during the first 2 forecast hours (green line). These improvements are almost lost from 2 to 6 h, and reappear again over the last 6 h.

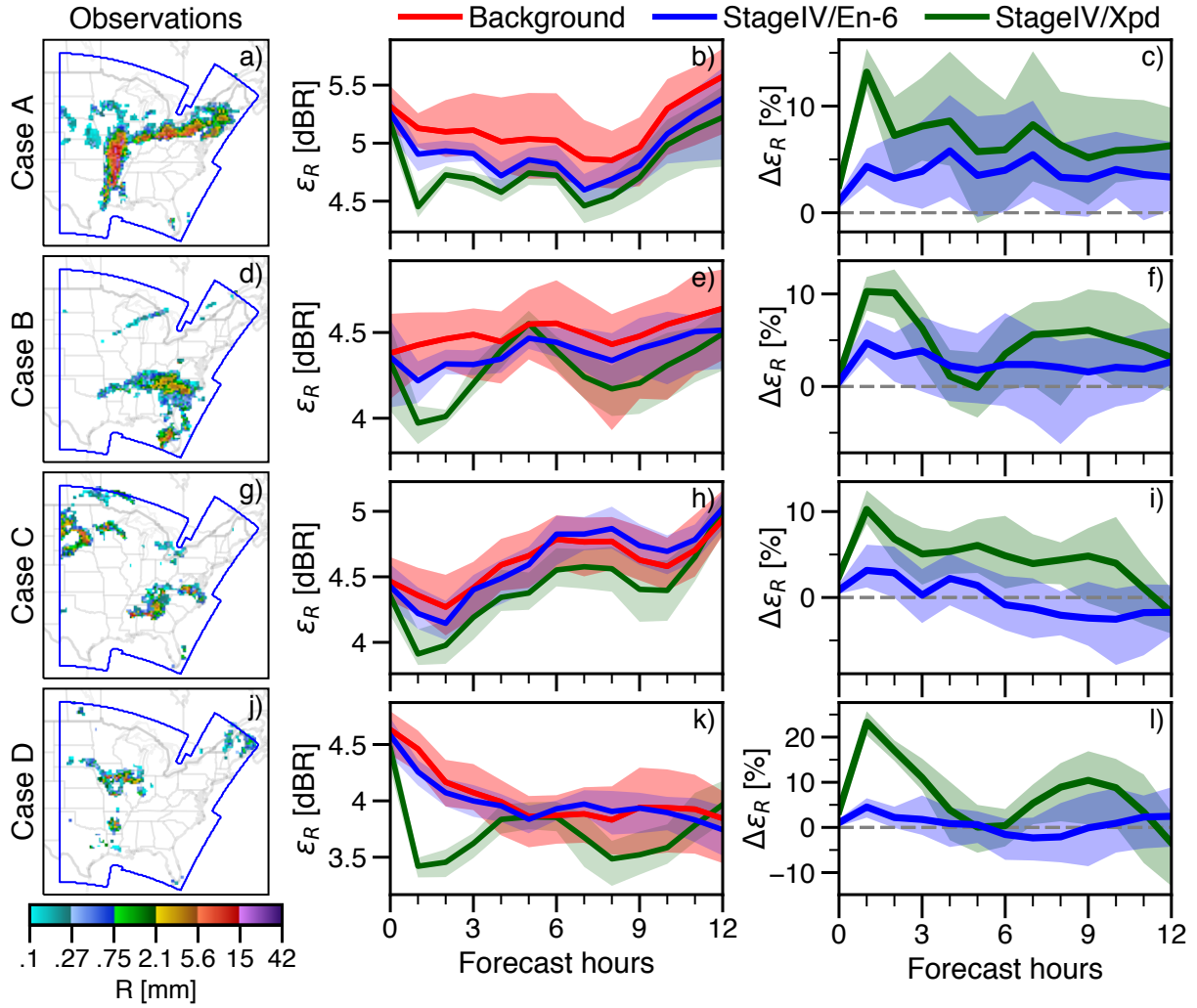

Fig. 3.12: Impacts of DA on precipitation forecasts quality for Case A (first row), Case B (second row), Case C (third row), and Case D (fourth row) for StageIV DA experiments. The first column shows the observed precipitation (StageIV) when the DA takes place. The second column shows the hourly-accumulation forecast error measured by RMSE (in dBR units). The third column indicates the decrease in the RMSE ($\Delta\varepsilon_R$) when the DA is applied (improvements). The background errors are shown in red while blue and green colors display the *StageIV/En-6* and *StageIV/Xpd* errors and improvements. The lines denote the ensemble-averaged values while the shaded area indicates the variability around the mean.

For Cases C and D precipitation was produced by several Mesoscale Convective Systems (MCSs) scattered over the US and is more localized than in Cases A and B (widespread precipitation). For Case C, in the *StageIV/En-6* experiment, the assimilation of the observations does not improve the forecast quality. Only when the expanded ensemble is used (*StageIV/Xpd*) does LEMA produce long-lived improvements on the precipitation RMSE (Figs. 3.12-h and 3.12-i). Case D shows similar results as Case B, with the DA produces small improvements in *StageIV/En-6* experiment and a sharp decrease in the precipitation error during the first 2 forecast hours (green line in Figs. 3.12-k and 3.12-l) in the *StageIV/Xpd* experiment. As in Case B, in the *StageIV/Xpd* experiment there is a loss of skill during 3 to 6 h of lead time that returns towards the end of the forecast period.

The previous results show that in real DA experiments, if only the background ensemble is used by LEMA, the quality of the precipitation forecast shows small or no improvements. However, when the expanded ensemble is used, LEMA produces larger improvements in precipitation forecasts. To confirm the potential of LEMA over a larger sample of cases, in Fig. 3.13 we show the DA improvements for five additional situations. Cases E and F are widespread precipitation events (Figs. 3.13-a and 3.13-d). Although in Case G precipitation is also widespread, it is mostly produced by a large number of small-scale convective systems (Fig. 3.13-g). Lastly, Case H and I correspond to 2 Mesoscale Convective Systems (MCS) precipitation events over central US (Figs. 3.13-j and 3.13-m).

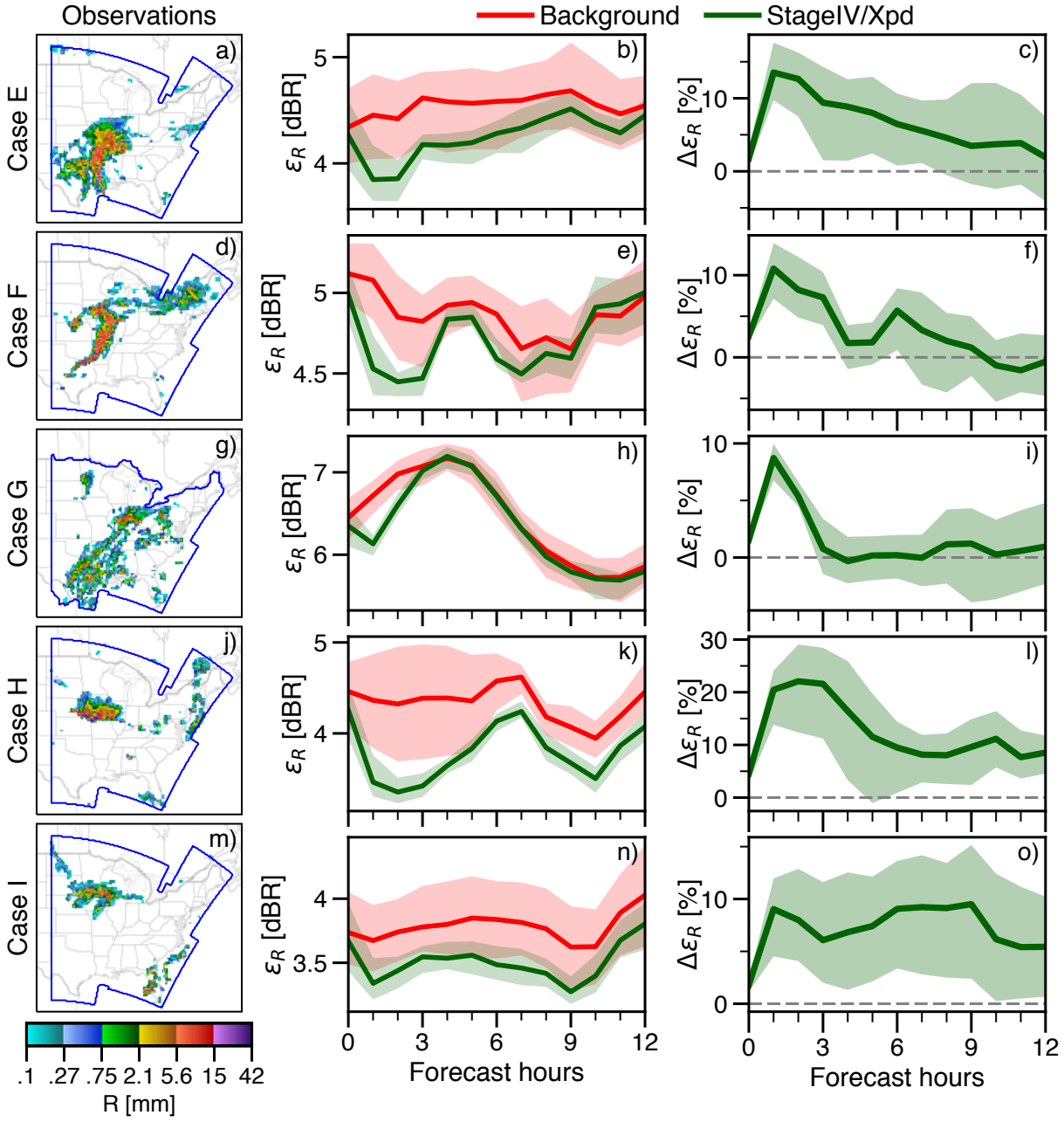

Fig. 3.13: Same as Fig. 3.12, but for for Case E (first row), Case F (second row), Case G (third row), Case H (fourth row), and Case I (last row).

In four of the five cases, the improvements in the precipitation forecast persist up to 9 to 12 h (Cases E, F, H, and I). However, Case G (third row) shows the worst performance of LEMA where the improvements are quickly lost 3 h after the DA takes place. Let us examine this case in more detail. Figure 3.14 shows the precipitation field at 12 h forecast time without DA (upper left) and with DA (upper right, Frankencast forecast). The lower panels indicate the StageIV fields and the reduction in the precipitation error by the DA. Although the impact of the DA as measured by RMSE is nil (see Fig. 3.13-i), we point out that within the red ellipse in Fig. 3.14a, the background forecast indicates a well-organized line of precipitation that does not resemble the observed cellular structure (Fig. 3.14c). However, the forecast with DA correctly indicates a cellular structure. In addition, the most intense precipitation region (magenta circle) shows a marked DA quantitative improvement. The precipitation within the black circle in the north-west boundary of the observations domain indicates that the change is rather random, probably as a result of the inflow of information from regions without DA. Thus, some improvements are present but not captured fully by the measure we use. The poor performance can perhaps be attributed to the very fragmented precipitation pattern, and hence very limited long-term information on the structure of small scales, and our emphasis on large scale and long duration of the effect of assimilation. In cases like this quantitative improvement by LEMA should not be expected over long lead times.

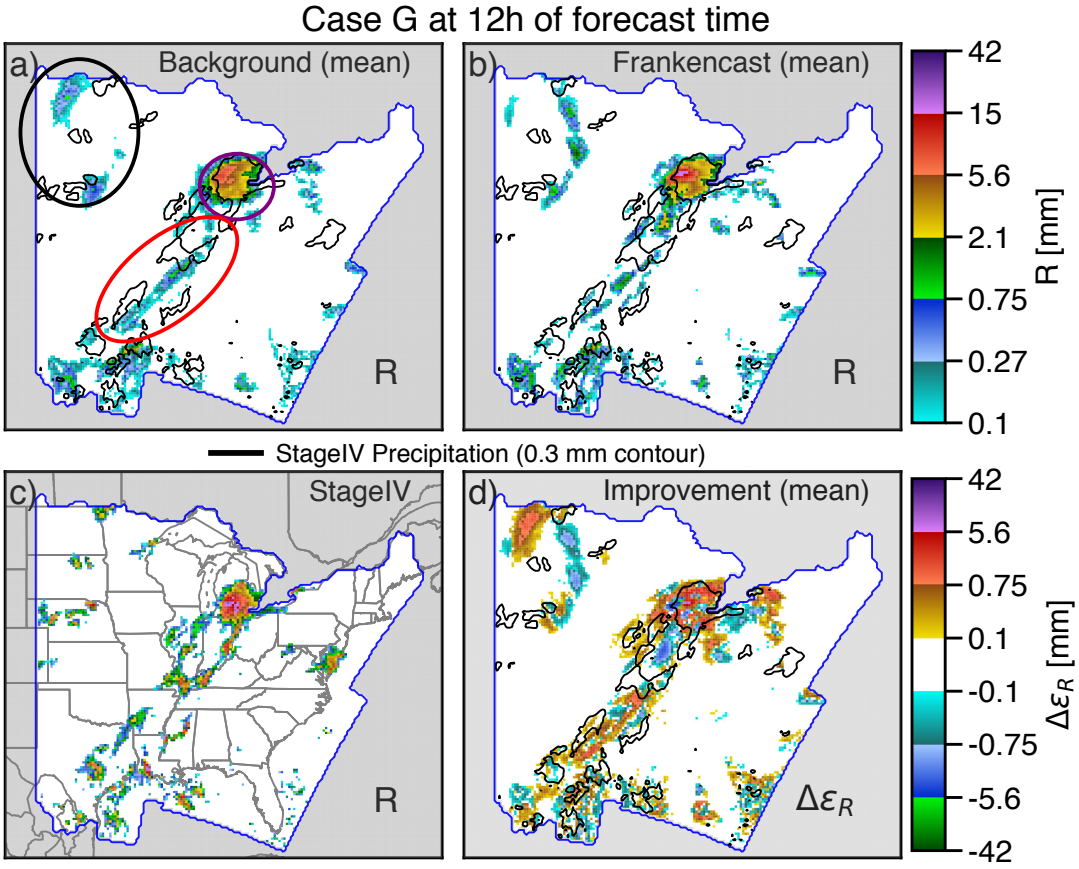

Fig. 3.14: DA impacts of precipitation fields for Case G *StageIV/Xpd* experiment at 12 h of lead time. The upper left and right panels show the ensemble-averaged precipitation for the runs without DA (background) and with DA (Frankencast), respectively. The lower panels indicate the StageIV fields while the right bottom panel shows the ensemble-averaged reduction in the precipitation error by the DA with respect to the background run. The black contour shows the StageIV precipitation (R=0.3mm) observations while the blue contour indicates the area where the observations are available.

However, if we consider the average DA impact on precipitation over the nine cases, the assimilation results in long-lived positive impacts that persist up to 12 h of lead time (Fig. 3.15). These are encouraging results since the persistent improvements are obtained only with a single DA cycle and using an unfavorable setup, consisting of 3 small-sized and underdispersive ensembles (discussed in Sec. 3.6).

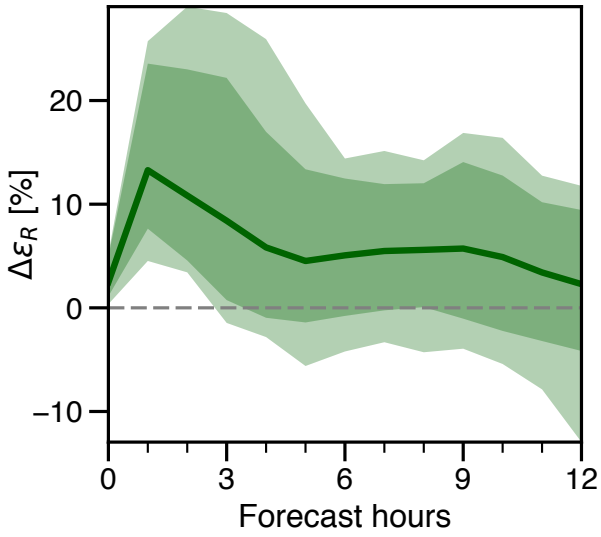

Fig. 3.15: Decrease of the precipitation RMSE for the nine cases for the *StageIV/Xpd* experiment. The green line indicates the average improvements while the dark-green shaded region indicates the 95% confidence intervals. The full variability over the mean is indicated by the light-green shaded area.

Let us now investigate in more details the impacts of LEMA on the precipitation patterns. Hence, for the rest of this section, we will carry out our analysis only in a subset of four of the nine cases, namely Cases A,B, C, and D. Figure 3.16 shows the observed and the forecasted precipitation patterns at 1 h of lead time for the four *StageIV/Xpd* experiments. Only member 0 is shown but the results are similar for other members. For Case A (first row), the background forecast (second column) shows a good general agreement to the observed precipitation. When the observations are assimilated, the precipitation pattern becomes more similar to the observed one

(third column). To better identify the positive (or negative) impacts of LEMA in the precipitation distribution, the fourth column in Fig. 3.16 shows the reduction in the precipitation errors obtained by the DA (positive gains: red and purple colors, negative gains: blue and green colors). For Case A, LEMA is able to produce considerable improvements in the precipitation intensity over most of the frontal precipitation region.

Similarly to Case A, Case B (second row) also shows a good general agreement with the observed precipitation, except in the south-eastern part of the observation domain where precipitation is poorly represented. The DA is able to correct part the overestimation in the south-eastern part of the domain (see red and purple colors in Fig. 3.16-g).

For Cases C and D (third and fourth row), the precipitation patterns are more localized than the other 2 cases and the position errors are important. For Case C, the improvements are not as easy to identify as the other cases. However, Fig. 3.12 shows that there is a 10% improvement in the RMSE. Since the precipitation distance metric used by LEMA was optimized to reduce the large-scale precipitation errors, the improvements in precipitation occur only on average over a particular region. We will discuss this in more detail shortly. For Case D, a large part of the improvements are due to the suppression of the overestimated precipitation in the western part of the domain (indicated by a red arrow in Figs. 3.16-n and 3.16-p).

Figure 3.16 also shows that LEMA is more effective in correcting large-scale errors in the precipitation than small-scale errors. This result is consistent with the analysis being constructed by selecting the members with the lower large-scale errors in precipitation (large MAD window). The use of a large observation windows was motivated by the idea that decreasing errors at the large scales may have benefit at smaller scales (Durran et al., 2013), leading to long-lived improvements in precipitation. Additionally, the large observation window is more effective to transfer information from precipitation to the state variables over a large region (PZY19). Furthermore, the reduction in large-scale errors is also a result of the stochastic relation between precipitation

and the state variables. In LEMA, the transmission of the information from observations to the state variables is represented by the joint probability of a reduction in the precipitation and state variables errors. Fig. 3.9b shows that, albeit there is an overall net reduction in the state errors, there is large variability around those mean values.

However, if we would want to put emphasis on short-term forecast, the distance metric used in LEMA and the observations windows can be optimized to effectively correct the small-scale precipitation errors. For example, one can use ETS over a smaller localization window to measure the local distance, or a weighted average of small-scale and large-scale MAD. Another alternative that may be promising is to combine LEMA, that focuses on large-scale precipitation errors, with Latent Heat Nudging, that takes into account the precipitation errors at the model resolution.

In our DA experiments, the extended ensemble was constructed by considering model states from different time-lagged forecasts, as well as model states at different times. This results in an ensemble with 315 members (21 member x 3 ensemble forecasts x 5 times). All those members are candidates that may be used to construct the analysis. From that pool of candidates, for the analysis construction, LEMA only uses the members that better adjust to the observations. Additionally, the selected members can be used to assess the relative importance of the additional sources of information (time-lagged forecasts or models times) to the Frankenstate construction.

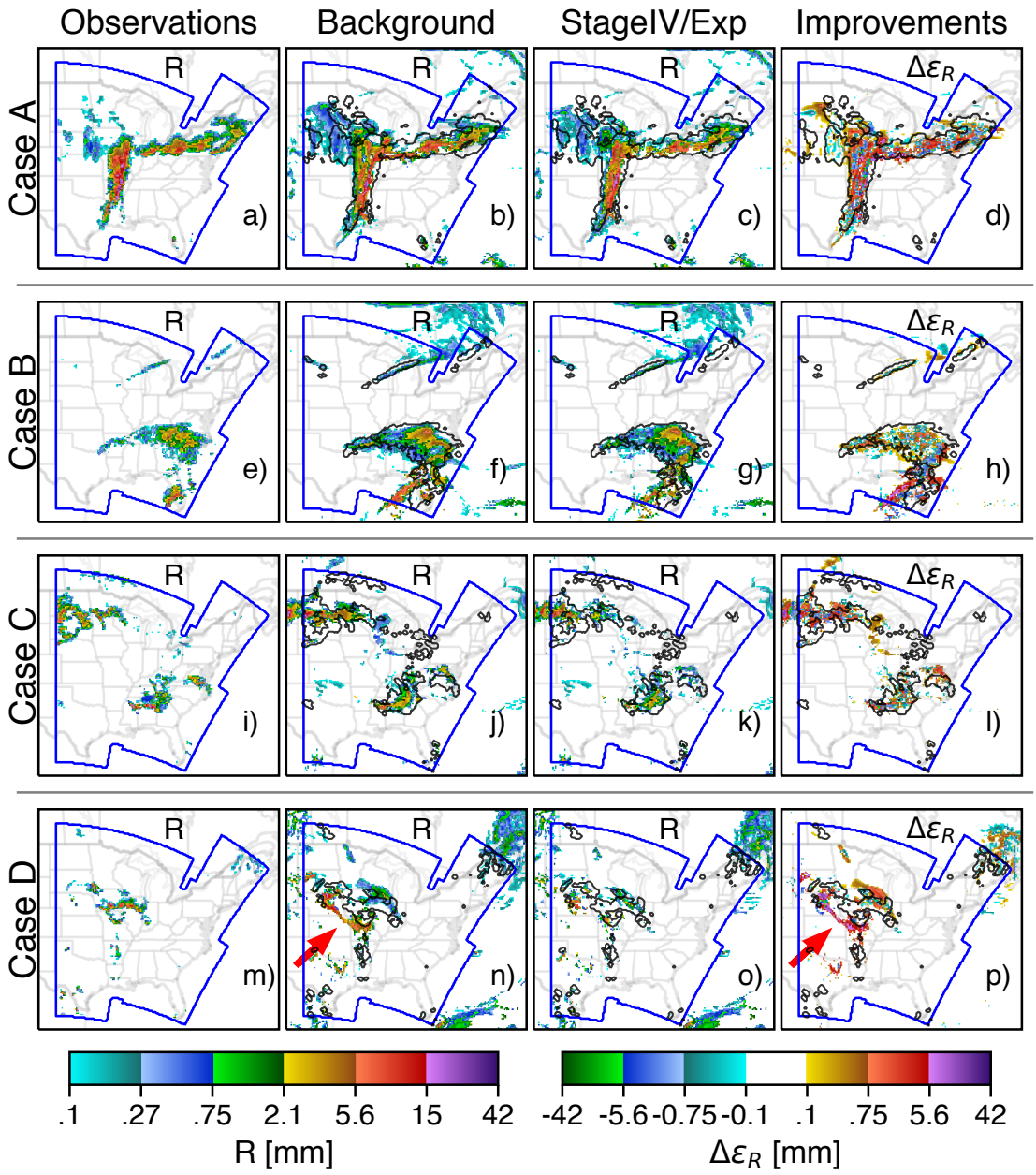

Fig. 3.16: Observed and forecasted hourly precipitation for the *StageIV/Xpd* DA experiments at t=1 h of lead time. The precipitation forecasts without DA (background, second column) and with DA (third column) are shown for Case A (first row), Case B (second row), Case C (third row), and Case D (fourth row). The fourth column shows the decrease in the precipitation error (Δ_R) by the DA at each grid point (difference between background and LEMA forecast). The black lines show the StageIV precipitation (R=0.3mm) observations while the blue contour indicates the area where the observations are available.

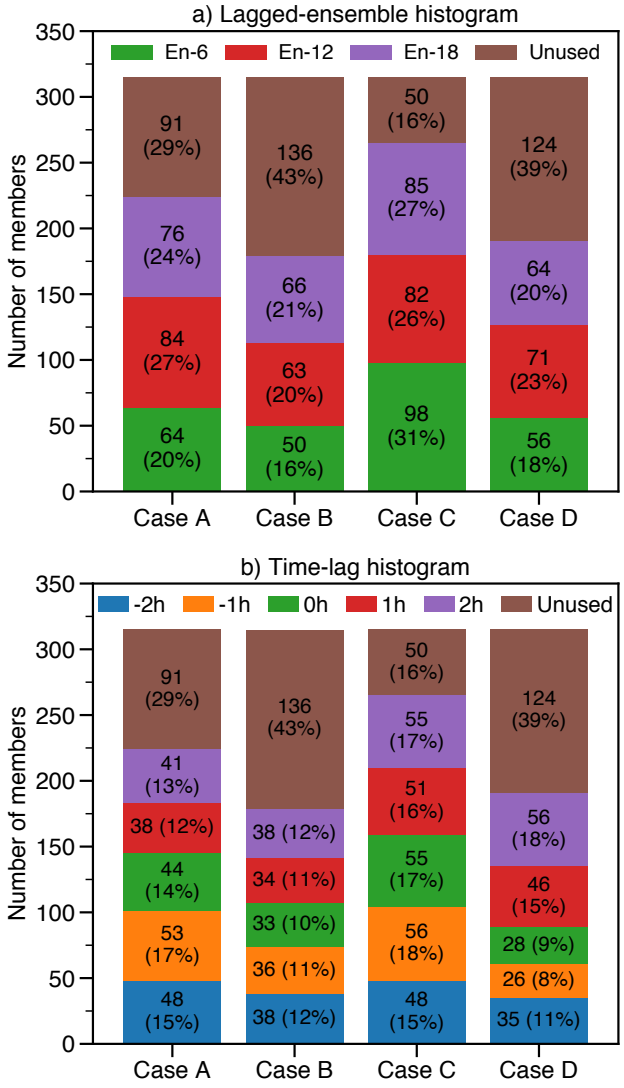

Fig. 3.17: Histogram of selected members for the $StageIV|Xpd$ experiments, for Cases A to D. The upper panel (a) shows the number of members selected from each of the lagged-forecasts runs. The lower panel (b) displays the model times (time lag with respect to the analysis time) selected from the Expanded ensemble. The numbers in each block in the bar plots indicate the numbers of members selected for the corresponding source (lagged-ensemble or model time).

Figure 3.17a shows the number of grid points chosen from each lagged ensemble-forecast for the subset of four cases analyzed. Most the 315 members and all the lagged-forecasts ensembles are used in the Frankenstate construction, with contributions from the background ensemble (*En-*

6) that are similar in magnitude to the additional lagged-forecasts ensembles *En-12, En-18*. In addition to the lagged forecasts, the expanded ensemble includes different model forecast times around the analysis time to further expand the ensemble. Figure 3.17b shows the models times (time lag with respect to the analysis time) in the Expanded Ensemble. For the cases shown, all the model times included in the expanded ensemble contribute to the analysis. That is, LEMA using the expanded ensemble corrects errors originally present in the background ensemble (0 h time lag).

3.6 Discussion on the Expanded Ensemble properties

LEMA exclusively relies upon the information contained in the ensemble forecast, as represented by the joint probabilities of errors such as in Fig. 3.2. However, in practice, ensembles are notoriously underdispersive (e.g. Zhou et al., 2017), which means that the ensemble does not represent all the forecast uncertainties[E13]. Possible causes include the underestimation of the errors in the initial conditions, non-optimal perturbations that do not capture the growth of forecast errors, and model errors. Hence, forecast errors can be seen as a manifestation of the model and initial conditions errors not taken into account in the generation of the ensembles.

In Sec. 3.4 we used the joint probabilities $p(\Delta MAD, \Delta\varepsilon_\theta)$ to evaluate the effectiveness of LEMA for a particular situation (e.g. Fig. 3.2b and 3.6b). When the ensemble members used for the analysis construction capture the forecast errors, a decrease in the precipitation error MAD (by construction) results in a considerable decrease in the state variables errors. However, the latter approach is only suitable for idealized situations where the truth state is fully known. In a real case, the only information available is the forecast errors with respect to the observations (MAD).

[E13] Although it is stated that "... ensembles are notoriously underdispersive ..." a better choice of words would have been "ensembles may be underdispersive", since underdispersion is not necessarily a systematic problem of all the operational ensemble forecast systems (e.g. Haiden et al., 2019).

In this case, to measure the extent to which the ensemble used in LEMA capture the actual forecast errors, we can use the ensemble-averaged distributions of precipitation errors, p(MAD).

Figure 3.18 compares the spectrum of errors captured by the background spread (in green and red) to the p(MAD) of the actual errors with respect to the observations (orange curves) at 20 km resolution (grid scale)[E14]. As shown, the forecast errors exceed the errors captured by the spread of ensemble forecasts En-6 (green), indicating that the background ensemble does not capture the actual uncertainties given by the orange curves, *StageIV/En-6*. With the Expanded Ensemble (red curves) the forecast errors captured by the new ensemble spread, *Spread/Xpd* (considering all possible members as "truth") the distribution of errors approaches the actual forecast errors blue dashed curves, StageIV/Xpd. Clearly, the Expanded Ensemble provides a better representation of the actual uncertainties, improving the performance of LEMA (see Fig. 3.12).

In our experiments, the perturbations in the ICs are inherited from the GEFS data used to create the ICs and LBCs. The GEFS data used to create the ICs/LBCs for the WRF runs never exceeded the 30 h of lead time, where the GEFS dispersivity remains close to 1 for the 500 hPa geopotential height field (Zhou et al., 2017). However, for precipitation the forecast error probabilities, p(MAD), with respect to the precipitation observations (Stage IV) are quantitatively and qualitatively different from p(MAD) captured by the ensemble spread (*Spread/En-6*). These differences between OSSE and StageIV experiments in the p(MAD) hold for different cases, representing

[E14]For a given grid point (x, y) and member m, the MAD with respect to the StageIV observations is computed as $MAD_m(x,y) = \left|dBR^{obs}_{x,y} - dBR^{m}_{x,y}\right|$, where $dBR^{obs}_{x,y}$s and $dBR^{m}_{x,y}$ denote the observed and the modeled precipitation in dBR units, respectively. The probability p(MAD) is obtained by computing the histogram of all the MAD values calculated for every member and every grid point where the model or the observations have non-zero precipitation (>0.03mm). Therefore, this way of computing the p(MAD) with respect to the StageIV measures the distribution of the precipitation forecast errors. In addition, to measure the errors captured by the ensemble spread, the p(MAD) is computed using the same procedure as before but considers each ensemble member as a possible "truth" instead of the StageIV observations. Consequently, the different p(MAD) obtained for each possible "truth" provides a measure of the errors distributions captured by the ensemble spread.

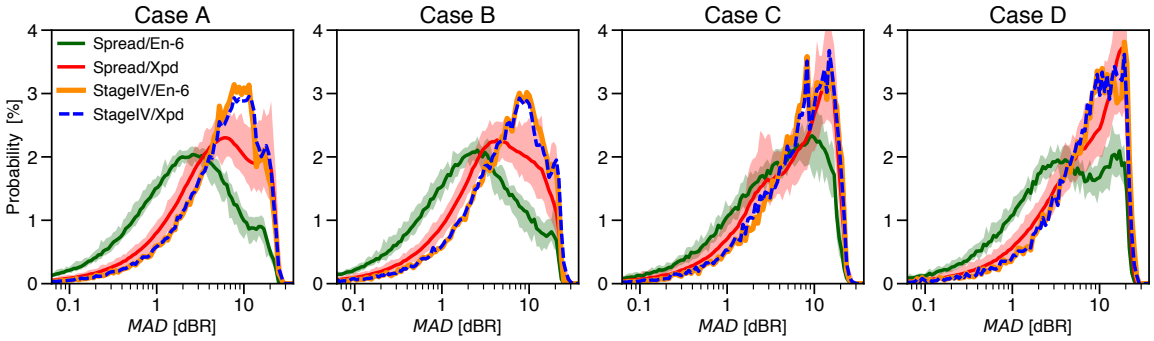

Fig. 3.18: Probability of precipitations errors $p(MAD)$ for the background ensemble (*En-6*) and the Expanded Ensemble (*Xpd*). The MAD is computed with respect to the StageIV observations (*StageIV/* prefix, blue and orange lines) and with respect to every possible members in the ensemble (*Spread/* prefix, red and green lines). The probability is computed from the histogram (using 100 bins) of MAD values over every grid point where $MAD > 0$ and over all the ensemble members. The thick line denotes the average values over all the truths considered, while the shaded area the variability around the mean.

different meteorological situations and initial conditions (ICs). This difference could be partly (or mostly) due to large model errors (particularly errors in the convective parameterization), a common element to all the runs[E15]. Since the model errors are not explicitly taken into account, LEMA will try to compensate for those errors through the new ICs (analysis) by choosing the models state variables generating precipitation closer to the observations and an improvement in the precipitation forecast after the DA takes place. If this is the case, the forecast may diverge from the observations. This could be the cause of the temporary loss of skill observed for Case B and D in the StageIV DA experiments (see second and forth row in Fig. 3.12).

Perturbing the model's initial and boundary conditions, choosing different parameterizations included in the model formulation, are the usual methods to generate the spread of ensemble

[E15]In the previous analysis, the StageIV observations are considered a proxy for the truth without considering possible observations errors. Therefore, another element that can explain the difference between the p(MAD) captured by the ensemble spread (*Spread/En-6*) and the p(MAD) of the errors with respect to the observations could be the presence of systematic errors in the precipitation observations, where the large precipitation values may be overestimated.

forecasts. In the OSSEs as well as in the assimilation of StageIV data we used some forecast errors as a proxy for model and initial condition errors to justify the expansion of the ensemble. A similar idea was used in a different context (Berner et al., 2015), where forecast errors are used as reflecting model errors. Hence, they can measure how different model-error schemes represent the actual uncertainties due to model errors. We show here that the initialization time of a limited area model is an aspect of uncertainty that can also be considered as an element of model perturbation that helps to expand the ensemble spread. Thus, the concept of the expanded ensemble, which we now consider as an integral part of LEMA, can be seen as a heuristic expansion of model perturbations. In an operational setup, where the ensemble forecasts are run periodically (e.g. every 6 h), this expansion of the model perturbations does not require additional model runs, since often the lagged forecasts are already available. Hence, the expanded ensemble is a computationally inexpensive alternative to the generation of additional model forecast perturbations.

In Sec. 3.5 we showed that when the Expanded Ensemble is used, all the lagged-forecast initialized at different times and different model times contribute to the Frankenstate (Fig. 3.17). This suggests that for precipitation forecasts, the different initialization and model times have a similar probability to represent the actual atmospheric state. That is, the members of the Expanded Ensemble appear to be as equiprobable with respect to precipitation errors as the members of the original ensemble generated by perturbations of the initial conditions. We have shown that the Expanded Ensemble improves the coverage of the reality (dispersivity). To further emphasize this, Fig. 3.19 shows that the dispersivity of background (*En-6*) and the Expanded Ensemble (*Xpd*), with respect to the StageIV observations[E16]. The correct dispersivity in the ensemble results when

[E16]In the discussion presented in this paragraph, the precipitation observations are considered as a proxy for the "truth" atmospheric state. Hence, the dispersivity calculations presented here are used to measure how well the ensemble covers the StageIV observations.

the RMSE of the ensemble mean matches the ensemble spread (Fortin et al., 2014). For the cases shown, the background ensemble (red curve) are underdispersive (dispersivity ~ 0.5-0.6) and the positive impacts of LEMA are limited (Fig. 3.12). The Expanded Ensemble considering all the candidate members (*Xpd/Prior*) have a much better dispersivity (closer to one) as a result of a reduction in the RMSE of the ensemble mean (not shown). When we only keep the members selected by LEMA (*Xpd/Posterior*), the dispersivity is improved even further. This is an expected result since the members with larger errors are discarded, reducing the distance of the ensemble mean to the "truth".

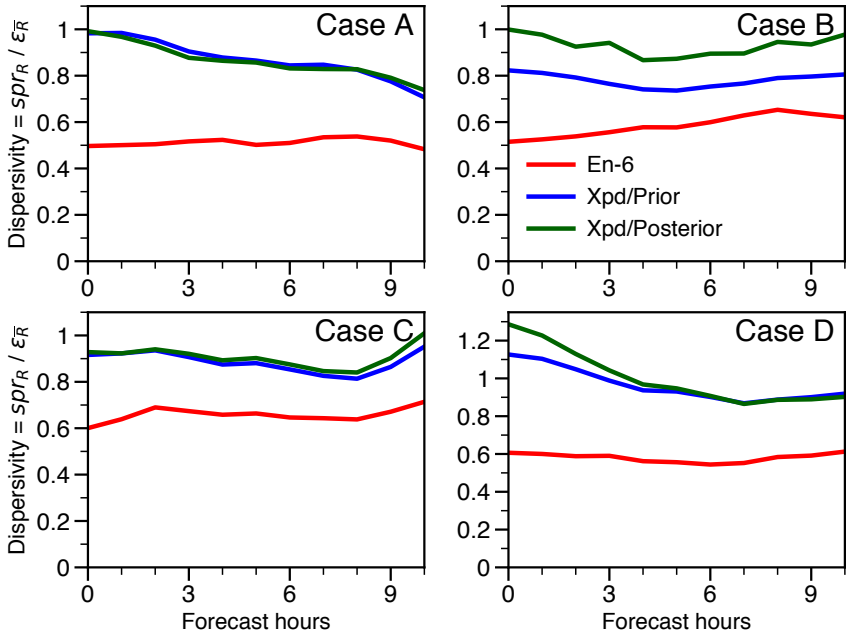

Fig. 3.19: Dispersivity computed with respect to the StageIV observations for Cases A to D. For each case the dispersivity is shown for the background ensemble used in the real DA experiments (*En-6*, red curve), the Expanded Ensemble considering all the possible candidates (*Xpd/Prior*, blue curve), the Expanded Ensemble considering only the members selected by LEMA (*Xpd/Posterior*, green curve). The dispersivity for the hourly accumulation (in dBR units) is computed as the ratio of the ensemble spread (spr_R) and the RMSE of the ensemble mean ($\varepsilon_{\overline{R}}$).

3.7 Summary and Conclusions

In PZY19 we introduced a heuristic DA method for precipitation assimilation. The method, called Localized Ensemble Mosaic Assimilation (LEMA), uses the information from an ensemble of model runs in a direct manner to construct a mosaic analysis by selecting the ensemble members that locally have the smallest precipitation error (MAD) measured at large scales. The present study extends the idealized experiments in PZY19 by considering less optimistic situations that are likely to occur in real DA situations with significant forecast errors and underdispersive ensembles.

The OSSEs with different characteristics of forecasts errors show that, when a small number of ensemble members is used, the impacts of DA in the forecast quality deteriorate rapidly with increasing forecast errors beyond the spread of the ensemble. This is an expected result since LEMA relies on the information contained in the ensemble forecast only. If the ensemble does not cover the true state and all the ensemble members in the background have large errors, the constructed Frankenstate will also have large errors. To partially remedy this problem, we present a method to expand the spread of the background ensemble used by LEMA with the goal of improving the representation of the actual forecast uncertainties by considering additional lagged-forecasts ensembles and model states at different times. With OSSEs, we found that the additional information in the expanded ensemble improves the efficacy of LEMA, produces larger and long-lived improvements in the state variables and in the precipitation forecast quality. Although the OSSEs allow us to understand how the forecast errors affect the information transfer from precipitation to the state variables, they are over-optimistic compared to real situations.

Therefore, we also explore the potential of a LEMA in real DA experiments using StageIV precipitation observations. As a proof of concept, we consider unfavorable situations where the forecast errors were not captured by the spread of background ensemble. For these unfavorable

situations, when LEMA uses only the background members, the quality of the precipitation forecast shows small or no improvements. However, when the Frankenstate is constructed with the expanded ensemble, LEMA produces larger improvements in precipitation forecasts. We consider these results encouraging since the expanded ensemble was constructed using 3 small-sized and underdispersive ensembles of 21 members each. However, this setup is only a proof-of-concept and more initialization lags are not harmful to LEMA, even if it produces an overdispersive ensemble. The member selection in LEMA will filter out the members that are far away from the truth and only use the information available in the members with the smallest large-scale MAD, improving the dispersivity of the expanded ensemble.

To understand the origins of the better effectiveness in LEMA, in Section 3.6 we explored the properties of the expanded ensemble. In our real DA experiments, the forecast errors exceed the errors captured by the spread of the background ensembles, indicating that they do not represent the actual uncertainties. We showed that the expansion of the ensemble spread captures part of the uncertainties not taken into account in the original background ensemble, which results in a better performance of LEMA. However, despite the larger spread and corrected dispersivity of the expanded ensemble, not all of the forecast errors are taken into account. Figure 3.18 shows that the expanded ensemble does not fully correct the distribution of precipitation errors, especially for large MAD values. This difference could be partly explained by significant model errors not being explicitly taken into account in the expansion of the ensemble. Hence, to better capture the errors, the background ensemble can be further expanded by taking into account model errors. This can be done, for example, by adding perturbations to the Kain-Fritsch convective parameterization.

Additionally, the expanded ensemble concept is not only relevant to LEMA. It can be applied directly to any assimilation method that relies on the information contained in the ensemble to improve their performance when the background ensemble does not represent the actual forecast uncertainties or the size of the ensemble is small. For example, the expanded ensemble can

be used in Particle Filters (PF) to improve the representation of the model probability density function (pdf). Although its potential is yet to be determined, the expanded ensemble can also be used to increase the ensemble size in EnKF, in particular when small-sized ensembles are used.

The main findings in this study are relevant to other approaches based on Monte Carlo methods that are similar to LEMA, such as the Local Particle Filter (LPF) introduced in Poterjoy (2016). The analysis mosaic in LEMA has similarity to analysis mean obtained by LPF algorithm when the posterior distribution is strongly narrowed down (collapsing of the weights in the algorithm into a single member). When this collapse occurs, the mosaic of selected members resembles the one obtained in LEMA by selecting the member closest to the observations. In this particular case, the analysis mean produced by the narrow posterior distribution appears to perform better in a deterministic forecast (Mark Buehner, private communication).

In this work and PZY19, the OSSEs are used to understand how the information contained in ensemble forecasts propagates into state variables and to determine the optimal parameters for LEMA. However, the OSSEs can also be used as part of the DA procedure for a flow-dependent optimization of the algorithm's parameters. For example, idealized experiments can be used before the analysis construction to determine the size of the window most adequate for each meteorological situation. This is part of our future work.

Finally, one limitation of LEMA not considered as yet is that all the background ensemble members are relaxed towards the single analysis (the Frankenstate). As we showed in PZY19, the ensemble spread is reduced, resulting in a very underdispersive ensemble at the beginning of the forecast. This limits the cycling of LEMA to long time intervals, of the order of 9 h, at which time the ensemble spread in the forecast is restored. A too short cycling period, Δt, of precipitation assimilation is not necessarily desirable because precipitation slowly decorrelates in time and hence information update is slow (particularly at large scales). The cycling period of 9 h is probably too long. However, it is possible to shorten the cycling period by adding to

LEMA another heuristic step: nudging the ensemble towards the Frankenstate while maintaining the original (or adjusted) background's spread. This is part of next steps.

Chapter 4

Extending LEMA to cycled DA applications.

In the previous chapter, we showed that LEMA has potential for real DA applications. Using a single DA cycle and an unfavorable setup, consisting of 3 small-sized and underdispersive ensembles, LEMA produces improvements in the precipitation forecasts that persist up to 12 h of the forecast. This chapter presents a new version of LEMA, called Localized Ensemble Mosaic Ensemble Assimilation Sequence (LEMAS), suitable for cycled DA. Next, we explore the potential of LEMAS in a cycled DA framework using ideal and real DA experiments in assimilating precipitation observations.

This chapter consists of a manuscript submitted to the journal Monthly Weather Review: Pérez Hortal, A.A., I. Zawadzki, and M.K. Yau. A Sequential Non-Gaussian Approach for Precipitation Data Assimilation.

A Sequential Non-Gaussian Approach for Precipitation Data Assimilation

Andrés A. Pérez Hortal, Isztar Zawadzki and M. K. Yau

Department of Atmospheric and Oceanic Sciences, McGill University, Montréal, Québec, Canada

Abstract

In two recent studies, the authors presented a new data assimilation (DA) method for precipitation observations that does not require Gaussianity or linearity assumptions. The method, called Localized Ensemble Mosaic Assimilation (LEMA), initializes the new ensemble forecast by relaxing the background ensemble (prior) towards a single analysis composed of different column states taken from the ensemble members with the lowest error in the precipitation forecast. However, a limitation of the LEMA is that relaxing the background ensemble towards that analysis severely reduces the spread of the ensemble, thus, limiting its usefulness for cycled DA applications. This study presents a new version of LEMA, called Localized Ensemble Mosaic Assimilation Sequence (LEMAS), suitable for cycled DA operations. LEMAS constructs an ensemble of analysis mosaics using a small group of members closer to the observations instead of only the closest one. The new ensemble forecast is then initialized by recentering the prior ensemble around the mean of the analysis ensemble while scaling the original background perturbations to match the spread of the analysis mosaics. A series of ideal and real DA experiments are used to evaluate the potential of LEMAS for the assimilation of hourly accumulation observations. A comparison of LEMAS with the Local Ensemble Transform Kalman Filter (LETKF) using idealized experiments shows that LEMAS produces similar or slightly better forecast quality than the LETKF in temperature, water vapor, winds, and precipitation. Extending this comparison to real DA experiments assimilating StageIV precipitation observations shows that both methods

produce precipitation forecasts of comparable quality.

4.1 Introduction

Although precipitation is a difficult quantity to forecast, it is also a well-observed quantity. Over the past decades, several networks of ground-based radars provide precipitation estimates over several continents (e.g. Crum and Alberty, 1993; Huuskonen et al., 2014; Min et al., 2019). Since these dense observation networks provide detailed information on the spatial and temporal scales of the precipitation events, their effective assimilation has a significant potential to improve the quality of precipitation forecasts. Consequently, several studies have explored the benefits of assimilating precipitation observations using conventional DA assimilation approaches, like variational and Ensemble Kalman Filter (EnKF) techniques (Lopez, 2011; Kumar et al., 2014; Lien et al., 2016; Kotsuki et al., 2017), simpler humidity, and latent heat nudging (LHN) approaches (Davolio et al., 2017; Jacques et al., 2018), as well as non-Gaussian approaches based on Particle Filters (PF) (Poterjoy et al., 2019; Buehner and Jacques, 2020).

Of the methods mentioned above, the LHN is a simple DA approach that has been used operationally for Numerical Weather Prediction (NWP) in many centers over the last two decades (e.g. Macpherson, 2001; Simonin et al., 2017; Stephan et al., 2008; Jacques et al., 2018). This method uses the precipitation observations to adjust the model buoyancy through temperature and humidity changes to bring the model precipitation close to the observed values. Despite its simplicity, the assimilation of precipitation observations by the LHN successfully brings the model precipitation close to the observed values right after the DA. However, most of the positive impacts in the precipitation forecast are commonly lost after ~3 h of the forecast (e.g. Jacques et al., 2018).

Alternatively, conventional DA methods, like variational and EnKF approaches, have also

been recently used to assimilate precipitation observations with encouraging results (Kumar et al., 2014; Lien et al., 2016; Kotsuki et al., 2017). However, these studies also point out the difficulties related to the assimilation precipitation observations due to the highly non-linear nature and non-Gaussian error statistics of precipitation processes that do not necessarily satisfy the Gaussianity assumptions needed by conventional DA methods (Bauer et al., 2011). Nonetheless, recent studies by Lien et al. (2013, 2016) show that the Gaussianity issues can be alleviated by applying a Gaussian transformation to the precipitation observations, constructed from its climatological distribution.

The difficulties related to the non-linear nature of the precipitation process have motivated the exploration of alternative DA methods that do not require Gaussian statistics, like the localized Particle Filters (LPF, Poterjoy et al., 2019). Instead of assuming a Gaussian distribution for the prior and the posterior estimates, the basic idea of the LPF is to describe the model probability distribution by a limited number of model states. Using similar assumptions as in the LPF, Buehner and Jacques (2020) proposed a deterministic non-Gaussian approach targeted to the assimilation of precipitation observation. Results show that these non-Gaussian DA methods produce forecasts with comparable or better quality than those produced by conventional Gaussian DA methods (Poterjoy et al., 2017, 2019; Buehner and Jacques, 2020).

Recently, Pérez Hortal et al. (2019, hereafter PZY19) and Pérez Hortal et al. (2020, hereafter PZY20) introduced the Localized Ensemble Mosaic Assimilation (LEMA) method, a simple non-Gaussian DA technique for precipitation data assimilation that can improve the precipitation forecast quality up to 12 h of lead time. Although the results presented in PZY19 and PZY20 are encouraging, the studies were performed in the context of a single DA step, where all the background members were nudged towards a single analysis mosaic constructed only using the ensemble members that are locally the closest to the precipitation observations. A shortcoming of our previous studies is that relaxing the background ensemble towards the analysis mosaic

severely reduces the ensemble spread and produces underdispersive forecasts.

This paper represents a sequel to PZY19 and PZY20 by developing a new version of LEMA that mitigates the limitation mentioned above. Instead of constructing a single analysis mosaic using only the closest member to the precipitation observations as in our previous studies, an ensemble of different analysis mosaics is constructed by considering the group of members close to the truth in precipitation. This new version, called Localized Ensemble Mosaic Assimilation Sequence (LEMAS), recenters the background ensemble on the mean of the ensemble of analysis mosaics and uses the spread of the mosaics ensemble to adjust the magnitude of the background perturbations to reflect the potential reduction in the uncertainties produced by the DA.

The goal of this study is to evaluate the new LEMAS technique in a cycled DA framework assimilating only precipitation observations and compare the forecast produced by LEMAS against those produced by the Local Ensemble Transform (LETKF) DA method using a similar configuration. The LETKF approach is chosen due to its similarities with LEMAS since both DA methods construct the analysis using a combination of the background members that are locally closer to the observations. However, these methods rely on different assumptions. The LETKF method assumes Gaussian error statistics and makes a linear approximation for the observations operator that are not required by LEMAS. Instead, LEMAS relies on a discrete number of model states to construct the analysis. To evaluate the new DA method, we first use a set of idealized experiments to compare the LEMAS and LETKF approaches using two different magnitudes of forecast errors. Then, this study also explores the potential of LEMAS in a real DA context by assimilating StageIV precipitation observations.

The article is organized as follows. We describe the LEMAS method in Sec. 4.2 and give a short description of the LETKF algorithm in Sec. 4.3. Section 4.4 describes the study cases, the model, and the experimental setup for the idealized experiments in assimilating synthetic precipitation observations, and the real DA experiments in assimilating StageIV precipitation

observations. Section 4.5 presents the results of the idealized DA experiments while Sec. 4.6 presents the results of the DA experiments using StageIV precipitation observations. Finally, Section 4.7 contains a summary and conclusions.

4.2 Localized Ensemble Mosaic Assimilation Sequence

The DA method presented in this study is a revised version of the LEMA method presented in PZY20 and is suitable for cycled DA applications. The revised version, called Localized Ensemble Mosaic Assimilation Sequence (LEMAS), can be conceptually divided into three parts. The first part involves constructing an ensemble of analysis mosaics composed of different column states selected from the subset of ensemble members closest to the precipitation observations. The second part uses the analysis-mosaic ensemble to recenter the background perturbations around the analysis ensemble mean and adjust the magnitude of the original background perturbations to account for the uncertainties of the updated ensemble. In the last part, the new analysis ensemble is used to initialize the new forecast by gradually relaxing (nudge) the background ensemble to the updated analysis ensemble.

4.2.1 Construction of the analysis mosaics

In PZY19 and PZY20, the physical notion that the proximity in precipitation of two model states is associated with proximity to the state variables was used to develop a simple heuristic method for precipitation DA. However, in these previous studies, the concept of proximity was used in an extreme manner: only the ensemble member that was locally closest to the precipitation observations was used to construct a single analysis. However, in PZY20 (their Fig. 3), it is briefly shown that other members close to the observations can also reduce the error in the state variables.

In this study, we extend the approach used in PZY20 to make use of the N_F members closer to the observations instead of only the closest one. To that end, an ensemble of states is used to construct an analysis composite by taking the values of the potential temperature θ, vapor mixing ratio q_v, and horizontal winds (U and V) from the model column corresponding to the ensemble members with the lowest local precipitation error in the vicinity of the column.

As in PZY20, the local precipitation error is measured using the Mean Absolute Difference (MAD) of the forecasted and the observed hourly precipitation values, considering only the valid observations inside a square window of width W centered on the column. Although the hourly precipitation accumulation values in mm units can be used for the MAD computations, errors related to high precipitation intensities are penalized more than errors at low precipitation values. To partially alleviate this problem by giving a more similar weight to the errors associated with small and large accumulation values, as in PZY20, we apply a logarithmic transformation to the hourly precipitation values by converting them from mm to dBR units as:

$$dBR = \begin{cases} 10log_{10}R, & R \geq 0.03\text{mm} \\ 10log_{10}(0.03), & \text{otherwise} \end{cases} \tag{4.1}$$

where R denotes the hourly accumulation in mm.

Using the transformed precipitation accumulation, the local MAD value associated with a horizontal grid point (i, j) and a member m is computed as:

$$MAD_m(i,j) = \frac{1}{K}\sum_{k=1}^{K}\left|dBR_k^{obs} - dBR_k^m\right| \tag{4.2}$$

where dBR^{obs} and dBR^m denotes precipitation observations and the precipitation values of the m^{th} member, respectively. The summation limit, K, indicates the total number of valid observations in

a square window of width W centered on the column (referred to as "localization" or "observation" window).

After the MAD is computed for each grid point and each member in an ensemble of states of size N_e, the next step is to construct $N_F(< N_e)$ different analysis mosaics $\{\boldsymbol{x}_k^F : k = 1, 2, ..., N_F\}$ by using at each horizontal grid point, the information from the first N_F members of the ensemble with the lowest MAD. Here, $\boldsymbol{x}$ denotes the model state that includes the four prognostic variables used in PZY19: the potential temperature θ, vapor mixing ratio q_v, and horizontal winds (U and V). The kth analysis mosaic $\boldsymbol{x}_k^F$ is constructed by assigning to each model column, the values of the model state $\boldsymbol{x}$ from the ensemble members with the kth smallest local MAD. For the member selection procedure, only members with a precipitation area coverage greater than a 2% threshold over the observation window are considered, as in PZY20. A 0.03mm detection threshold is used for the precipitation presence (hourly accumulation). If no member satisfies this condition for a given column, no values are assigned to that column. Consequently, the analysis is left "empty" and no updates are applied to the background over those grid points.

Since each analysis is constructed using only the information contained in the ensemble of states, using only states from the background ensemble at the analysis time may lead to a poor analysis quality when the ensemble spread underestimates the actual forecast errors (PZY20). Therefore, to increase the probability of finding states that are "locally close" to the actual atmospheric state, the ensemble of states is constructed using an "expanded ensemble" consisting of the model states taken from all members in the background ensemble, and also from additional lagged-forecasts ensembles initialized before the background runs. Additionally, from each member in the ensemble forecasts, the model states lagging the analysis time by 0, ±1, and ±2 hours are included in the expanded ensemble. However, the concept of the expanded ensemble is only used for the construction of the analysis. The actual ensemble forecast is performed by relaxing each member in the background ensemble towards the updated analyses to be described

next.

4.2.2 Recentering and re-scaling of the background perturbations

Once the N_F analyses are constructed, the background states $\{\boldsymbol{x}_m^b : m = 1, 2, ..., N_e\}$ (priors) are updated by recentering them around the ensemble mean of the analysis mosaics $\bar{\boldsymbol{x}}^F$, and rescaling the original background perturbations (at each grid point) to match the ensemble spread of the analysis mosaics σ^F. Although the different N_F analyses can be used directly as an update to the different background members, since $N_F < N_e$, the diversity of states in the posterior ensemble will be reduced compared to the background. Also, preliminary experiments showed that due to the mosaicked nature of the analyses $\boldsymbol{x}^F$, the resulting perturbations $\boldsymbol{x}_m^F - \bar{\boldsymbol{x}}^F$ are spatially uncorrelated which tend to cancel each other, thus reducing the effective ensemble spread. On the other hand, by rescaling the background perturbations to match the posterior spread, the resulting ensemble is expected to have similar spatial correlations and growth properties than the background ensemble. Consequently, this procedure may facilitate the blending of LEMAS with more sophisticated ensemble data assimilation (EDA) systems that deal with the other available observations, since the rescaling of the prior perturbations does not affect the structure of the perturbations obtained by the EDA designed to capture better the growth of the forecast errors with time (e.g. Wei et al., 2008; Buizza et al., 2008).

Based on the above-mentioned considerations, for each ensemble member “m” and model

grid point “i”, the analysis mosaics are constructed and updated according to:

$$\boldsymbol{x}^a_{m,i} = \bar{\boldsymbol{x}}^F_i + \left(\boldsymbol{x}^b_{m,i} - \bar{\boldsymbol{x}}^b_i\right)\left(\frac{\sigma^F_i}{\sigma^b_i}\right) \tag{4.3}$$

$$\bar{\boldsymbol{x}}^F_i = \frac{1}{N_F}\sum_{k=1}^{N_F} \boldsymbol{x}^F_{k,i} \quad ; \quad \bar{\boldsymbol{x}}^b_i = \frac{1}{N_e}\sum_{k=1}^{N_e} \boldsymbol{x}^b_{k,i} \tag{4.4}$$

$$\sigma^F_i = \frac{1}{N_F - 1}\sqrt{\sum_{k=1}^{N_F}\left(\boldsymbol{x}^F_{m,i} - \bar{\boldsymbol{x}}^F_i\right)^2} \tag{4.5}$$

$$\sigma^b_i = \frac{1}{N_e - 1}\sqrt{\sum_{k=1}^{N_e}\left(\boldsymbol{x}^b_{m,i} - \bar{\boldsymbol{x}}^b_i\right)^2} \tag{4.6}$$

where the super-indexes b denote the background and F the analysis mosaics, while $\boldsymbol{x}_{m,i}$ denotes the model state for member “m” at grid point “i”. The prior and the posterior spread at each grid point are denoted by σ^b_i and σ^F_i, respectively. For all the grid points where the analyses $\boldsymbol{x}^F$ were not constructed (left “empty") due to insufficient information, the prior model states remain unchanged.

Since LEMAS constructs a limited number of analysis mosaics (less than the background ensemble size), the resulting analysis ensemble may underestimate the actual uncertainties. To artificially increase the posterior spread to better capture the uncertainties, the Relaxation To Prior Perturbations (RTPP) method of Zhang et al. (2004) is used to relax the analysis perturbations towards the background perturbations as:

$$\left(\boldsymbol{x}^a_{m,i} - \bar{\boldsymbol{x}}^F_i\right) \leftarrow \left(\boldsymbol{x}^a_{m,i} - \bar{\boldsymbol{x}}^F_i\right)(1-\alpha) + \left(\boldsymbol{x}^b_{m,i} - \bar{\boldsymbol{x}}^b_i\right)\alpha \tag{4.7}$$

where α denotes the relaxation factor. By applying the equation above to Eq. 4.3, the final update

equation with the RTPP can be expressed as:

$$\boldsymbol{x}^{a}_{m,i} = \bar{\boldsymbol{x}}^{F}_{i} + \left(\boldsymbol{x}^{b}_{m,i} - \bar{\boldsymbol{x}}^{b}_{i}\right)\left(\frac{\sigma^{F}_{i}}{\sigma^{b}_{i}}(1-\alpha)+\alpha\right) \tag{4.8}$$

4.2.3 Ensemble Forecast Initialization

The new ensemble forecast is initialized by gradually forcing the model states $\boldsymbol{x}^{b}_{m}$ of each background member "m" towards the analysis $\boldsymbol{x}^{a}_{m}$ obtained by Eq. 4.8. The nudging approach is implemented by adding artificial forcing terms in the prognostic equations for the potential temperature θ, vapor mixing ratio q_v, and horizontal winds (U and V), that gradually force the model towards a given analysis state as in PZY20 (their Eq. 2). The nudging is applied for 30 min preceding the analysis time. The forcing is only applied to grid points where precipitation observations are assimilated, leaving the model to evolve without any artificial forcing in the rest of the grid points.

4.3 Local ensemble transform Kalman filter (LETKF)

The Local ensemble transform Kalman filter (LETKF; Hunt et al., 2007) is an efficient version of the EnKF that computes the analysis by minimizing the Kalman cost function over the subspace spanned by the background ensemble perturbations (here referred to as the "ensemble space" or the "perturbations space"). The local implementation of the LETKF computes a separate analysis for each model grid point using a weighted combination of the background ensemble perturbations to estimate the posterior analysis ensemble, where the magnitudes of the weights are determined based on their distance to the observations. This method will be compared to LEMAS because both methods directly use the background ensemble information to construct an independent analysis for each grid point using the observations in the neighborhood of the

grid point. However, an important difference between the methods is that the LETKF relies on Gaussianity assumptions and linear approximations for the observation operators to construct the analysis, while LEMAS does not make any parametric assumption. Also, the LETKF uses the perturbations space to compute the analysis, while LEMAS relies on a discrete number of model states to construct the analysis. Finally, another important difference between LEMAS and the LETKF methods is that the latter can assimilate different types of observations, while LEMAS only deals with precipitation observations.

In this section, the main components of the LETKF technique are described using the same notation as in Sec. 4.2. For additional details, the reader is referred to Hunt et al. (2007). As it was mentioned earlier, the LETKF computes the posterior analysis mean, independently at each grid point, using a linear combination of the deviations of the background ensemble from the ensemble mean. The LETKF analysis mean is computed as:

$$\bar{\boldsymbol{x}}_i^a = \bar{\boldsymbol{x}}_i^b + \sum_{m=1}^{N_e} \bar{\mathrm{w}}_{m,i}(\boldsymbol{x}_{m,i}^b - \bar{\boldsymbol{x}}_i^b) \tag{4.9}$$

where $\boldsymbol{x}$ is used to denote the model state vector (with the state variables θ, q_v, U, and V) at the grid point "i", and $\overline{w}_{m,i}$ denotes the weights given to each member "m", based on its proximity to the precipitation observations.

The local weights $\bar{\mathbf{w}}$ are computed using only the neighboring observations $\boldsymbol{y}^o$ to the grid point "i" as:

$$\bar{\mathbf{w}} = \widetilde{\boldsymbol{P}}^a (\boldsymbol{Y}^b)^T \boldsymbol{R}^{-1} \left[\boldsymbol{y}^o - \overline{\mathcal{H}(\boldsymbol{x}^b)}\right] \tag{4.10}$$

$$\widetilde{\boldsymbol{P}}^a = \left[(N_e - 1)\boldsymbol{I} + (\boldsymbol{Y}^b)^T \boldsymbol{R}^{-1} \boldsymbol{Y}^b\right]^{-1} \tag{4.11}$$

where $\boldsymbol{R}$ is the observations error covariance matrix (here assumed diagonal). $\boldsymbol{Y}^b$ is the matrix

whose mth column is the difference between the mth background member observations $\boldsymbol{y}_m^b = \mathcal{H}(\boldsymbol{x}_m^b)$ and their ensemble-averaged value $\overline{\boldsymbol{y}}^b = (1/N_e)\sum_m^{N_e} \boldsymbol{y}_m^b$. The matrix $\widetilde{\boldsymbol{P}}^a$ in Eq. 4.11 represents the analysis error covariance matrix in the ensemble space (i.e. has a dimension of $N_e \times N_e$).

After the analysis mean is computed, the analysis ensemble is constructed by transforming the background perturbations according to:

$$\boldsymbol{W} = \left[(N_e - 1)\widetilde{\boldsymbol{P}}^a\right]^{-1/2} \tag{4.12}$$

$$\boldsymbol{x}_{m,i}^a = \bar{\boldsymbol{x}}_i^a + \sum_{m=1}^{N_e} W_{m,n}(\boldsymbol{x}_{m,i}^b - \bar{\boldsymbol{x}}_i^b) \tag{4.13}$$

where "m" indicates the ensemble member, and $\boldsymbol{W}$ is a matrix of size $N_e \times N_e$ whose columns contain the coefficients of the linear combination that transforms the background ensemble perturbations to the analysis ensemble perturbations. In addition, the RTPP inflation (Eq. 4.7) is applied to increase the posterior spread in the locations where the observations are assimilated to deal with the possible underestimation of the posterior uncertainties (Zhang et al., 2004).

In addition to the localization implemented by computing the local weights for each grid point considering only the observations in the surroundings of the grid point, the influence of each observation is decreased to zero when the distance to the grid point increases by increasing the error variance (element of the matrix $\boldsymbol{R}$). In the LETKF, this can be implemented by multiplying each element in the $\boldsymbol{R}^{-1}$ by a smooth localization function (Hunt et al., 2007). In this work, the Gaspari–Cohn localization function (Gaspari and Cohn, 1999) is applied to the horizontal scale, with a cutoff radius L, and no localization is applied in the vertical direction. Since no vertical localization is applied, the weights computed in Eq. 4.13 are computed for each horizontal grid point and applied to the corresponding model column. Therefore, this implementation of the LETKF works column-wise in a manner similar to LEMAS.

In this study, the LETKF is used to assimilate hourly-accumulated precipitation observations (denoted by R), transformed to logarithmic units according to $ln(R+0.1)$. This transformation is used instead of the dBR transformation in LEMAS because it was recently used in other LETKF studies in assimilating precipitation observations to alleviate the non-Gaussian characteristics of these observations (e.g. Lien et al., 2016; Kotsuki et al., 2017; Buehner and Jacques, 2020). Following Lien et al. (2013), precipitation observations are assimilated using Eqs. 4.9 and 4.13 only at grid points where a significant proportion of the background members have precipitation values greater than 0.03 mm. In this study, a 75% proportion of the background members is required to assimilate the observations as in Kotsuki et al. (2017)[1].

For the LETKF method, only the members used in the actual ensemble forecast (20 for the ideal experiments, and 21 for the real ones) are used to construct the analysis ensemble. Thus, the concept of the "expanded ensemble" used to construct the analysis ensemble in LEMAS does not applies to the LETKF in this study. One reason for this is that the application of the expanded ensemble to the LETKF analysis update equations is not straightforward. In particular, it is not clear which is the best strategy to construct the analysis ensemble of size N_e, using states $\boldsymbol{x}^b$ and weights w computed using an expanded ensemble of size $N_{expanded} > N_e$. Therefore, evaluating the potential of the expanded ensemble in the LETKF approach is left outside the scope of this study. Instead, we focus our efforts on the characterization of the LEMAS technique and explore its comparison with an implementation of the LETKF previously used in the literature for the assimilation of precipitation observations.

Once the analysis ensemble is constructed, the new forecast ensemble is initialized by nudging each background ensemble towards the new analysis ensemble as in the LEMAS approach (Sec. 4.2.3).

[1]The 75% value requires that 15 members over 20 (or 21) must have precipitation. This ratio is the closest approximation to the proportion of 72% value used in Kotsuki et al. (2017).

4.4 Description of the experimental setup

Next, in Sec. 4.4.1, we summarize the elements in the setup similar to PZY20, namely, the study cases, numerical model configuration, and lagged ensemble forecasts used for the construction of the different experiments. Section 4.4.2 describes the Observing System Simulating Experiments (OSSEs) used to evaluate LEMAS and to compare it with the LETKF approach. Finally, Sec. 4.4.3 describes the cycled DA experiments in assimilating StageIV precipitation observations to evaluate the potential of LEMAS in a more realistic scenario.

Table 4.1: Model configuration

Horizontal grid	30 km (180x300)
Vertical levels	41
Computational time-step	1 min
Microphysics	WRF single moment (WSM3, Hong and Lim, 2006)
Planetary boundary layer	Yonsei University scheme (YSU, Hong et al., 2006)
Cumulus convection	Kain-Fritsch (Kain, 2004)
Shortwave radiation	Dudhia Shortwave Scheme (Dudhia, 1989)
Longwave radiation	Rapid Radiative Transfer Model (RRTM Mlawer et al., 1997)
Other parameters	Default WRF options

4.4.1 Model configuration and Lagged Ensemble Forecasts

For the numerical simulations, version 3.7.1 of the Weather Research and Forecasting model with the Advanced Research core (WRF-ARW, Skamarock and Klemp, 2008) is used. The model configuration is shown in Table 4.1. The simulation domain is the same as in PZY20 (contiguous US). However, the analysis of the forecast quality for both ideal and real DA experiments is carried out over the region excluding the Rocky Mountains (red rectangle in Fig. 4.1), since no precipitation observations are assimilated over that region due to lower data quality of the StageIV observations used in the real DA experiments (Sec. 4.4.3).

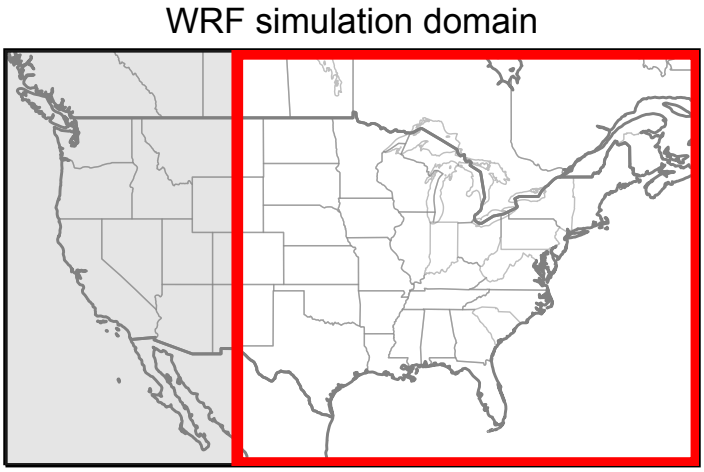

Fig. 4.1: WRF simulation domain. The red rectangle's interior indicates the region of interest where the forecast quality is evaluated in the numerical experiments.

Nine precipitation events described in Table 4.2 are used to evaluate LEMAS and the LETKF. They are named alphabetically, from A to I. Five of the nine cases are widespread precipitation events driven by synoptic fronts while the rest consists of several Mesoscale Convective Systems (MCSs) over the US. Figure 4.2 shows an overview of the observed precipitation patterns for the nine events over the simulation domain.

Table 4.2: Precipitation events

Case	Type	Beginning of DA
A	Synoptic front	0000 UTC 11 Apr 2013
B	Cyclonic systems	1800 UTC 04 Apr 2013
C	MCSs	0600 UTC 18 May 2013
D	MCSs	0000 UTC 27 May 2013
E	Synoptic front	0600 UTC 28 Dec 2015
F	Synoptic front	0600 UTC 14 Oct 2012
G	Synoptic front	1800 UTC 15 Aug 2016
H	MCSs	0600 UTC 04 Jun 2014
I	MCSs	1800 UTC 19 Apr 2012

Adapted from PZY20

For each of the nine cases, we produce three different lagged ensemble forecasts of 21 members each, initialized at 6-hours intervals, 18 h (*En-18*), 12 h (*En-12*), and 6 h (*En-6*) before the beginning of the DA period. The ICs/LBCs for each ensemble are constructed by interpolating the GEFS forecast data (1° grid spacing) into the WRF grid using the global forecast initialized at the same time as the WRF forecasts.

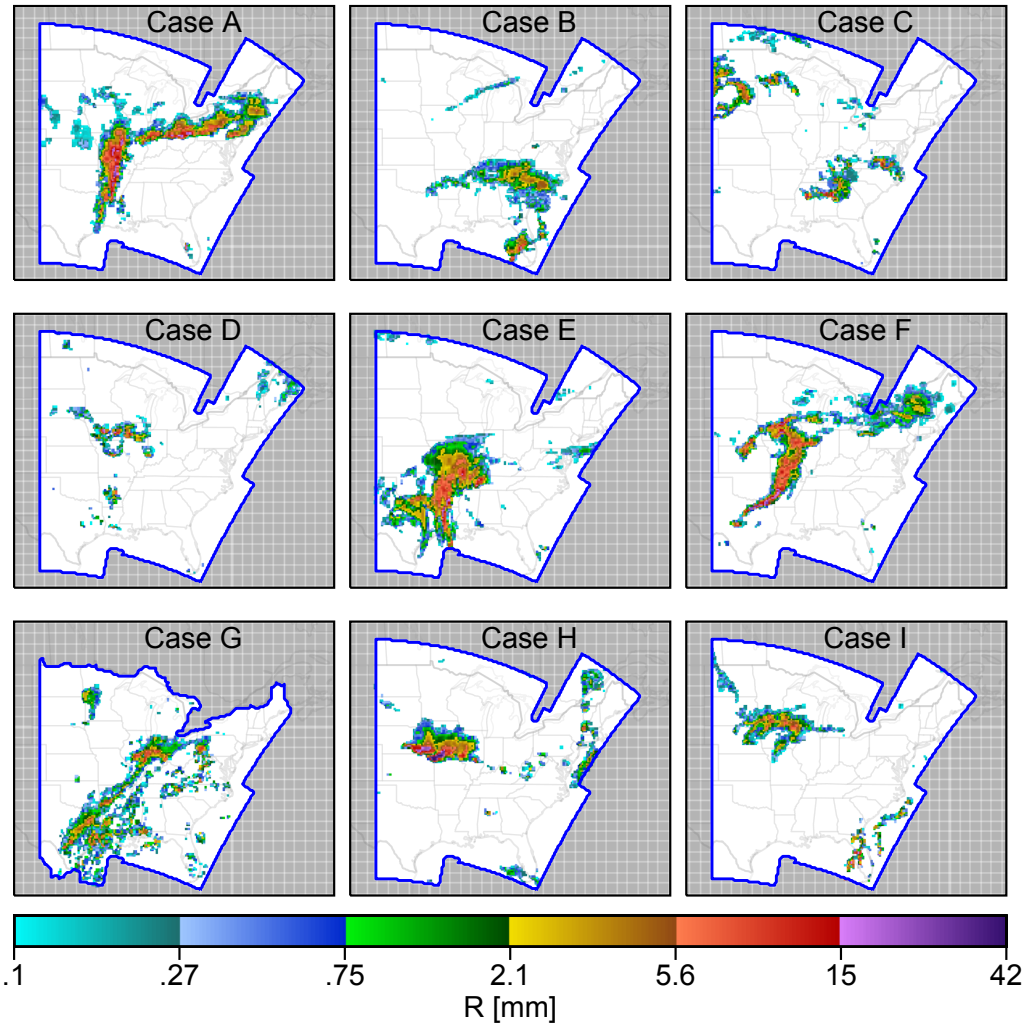

Fig. 4.2: StageIV precipitation overview at the beginning of the assimilation period for the nine cases used in this study (Cases A to I). The blue contour indicates the region where the synthetic and the StageIV observations are available for the assimilation.

4.4.2 Observing System Observations Experiments

For each study case (A to I), two different experiments with different magnitudes of forecast errors are carried out. The OSSEs are constructed by considering all the members in *En-18* ensemble as background runs, excluding member 15 (to be used as the "truth" run). The *En-18* ensemble is chosen as background since they correspond to the lagged-forecast with the larger ensemble spread when the DA takes place. The two different magnitudes of forecast errors are simulated using a similar methodology as in PZY20 by selecting the "truth" run from the same ensemble forecast as the background runs (smaller forecast errors), or from an ensemble forecast initialized six hours after the background (resulting in more significant forecast errors).

For the OSSEs with the smaller forecast errors, the member 15 was randomly chosen from the ensemble *En-18* as the truth. These experiments are denoted by $\tau = 0h$, where τ represents the initialization lag between the background and the truth runs. Since the truth was taken from the same ensemble as the background runs, it is expected that the spread of the background ensemble provides a good representation of the actual forecast errors for this experiment. The ideal experiments simulating larger forecast errors use the same background runs as the $\tau = 0h$ experiments, but the member 15 from the *En-12* ensemble forecast is taken as the truth. For these experiments, denoted by $\tau = 6h$, the truth and the background runs have an initialization lag of six hours. Since the truth and the background runs use the same model configuration, no model errors are simulated in our OSSEs.

For both types of experiments, synthetic hourly precipitation observations are simulated in the area enclosed by the blue contours in Fig. 4.2. This area corresponds to the same region where the StageIV precipitation product is assimilated in the real DA experiments described in Sec. 4.4.3 and it was chosen to make the ideal experiments as similar as possible to the real DA ones. For simplicity, no observations errors are considered. A schematic description of the OSSEs is shown

in Fig 4.3.

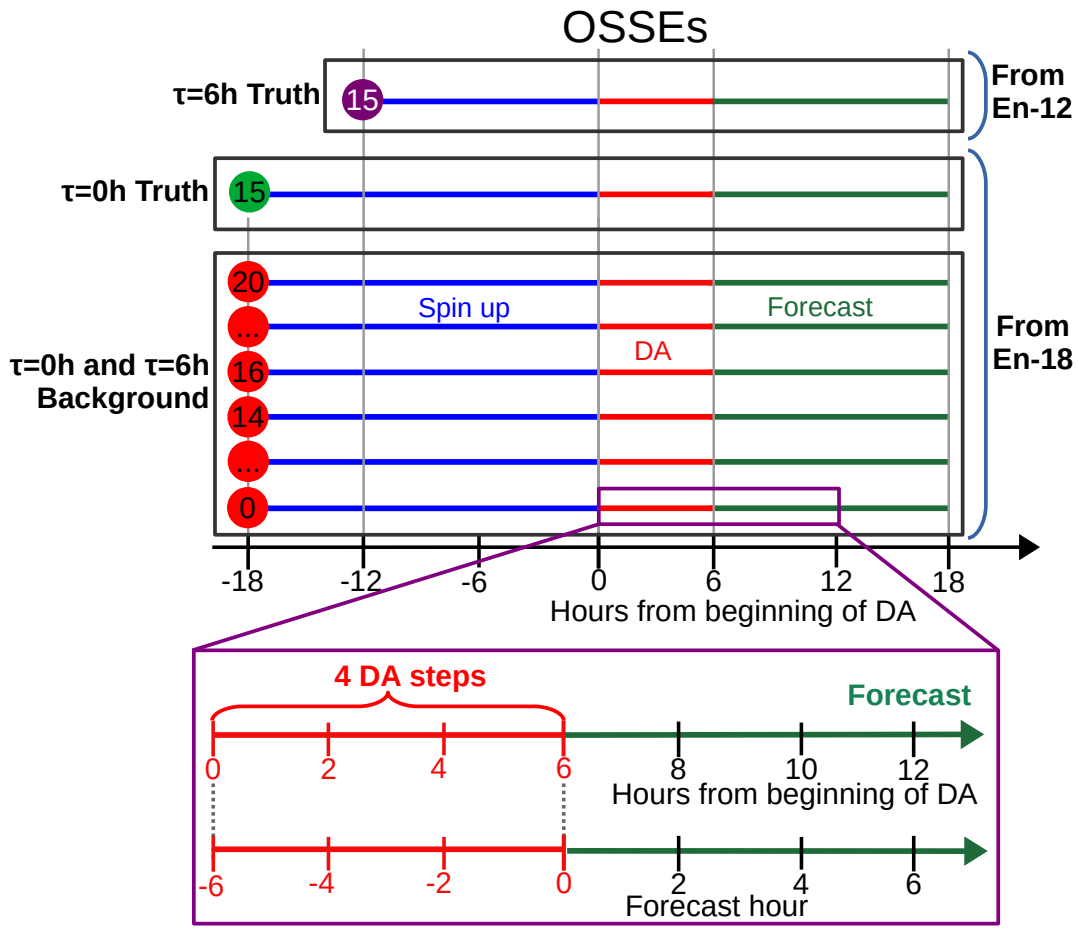

Fig. 4.3: Schematic description of the $\tau = 0h$ and $\tau = 6h$ idealized experiments. Each horizontal line denotes a single forecast run from a particular lagged forecast. The numbers inside the circles indicate the lagged-forecast member number.

A series of cycled DA OSSEs are carried out for each study case using the $\tau = 0h$ and $\tau = 6h$ setups to evaluate LEMAS and compare it with the LETKF method. For all the experiments, synthetic hourly accumulation observations are assimilated every 2 h during a 6 h DA period, followed by a 12 h forecast. The 2 h cycling interval is chosen to allow the precipitation fields to respond to changes in the state variables introduced by the DA, and for those changes to be better reflected in the hourly accumulated precipitation. A summary of all the OSSEs is presented in

Table 4.3a. The results of these experiments are compared against the control runs without DA to address the impacts of the assimilation of precipitation observations.

The parameter values for each DA algorithm were chosen based on a limited number of sensitivity experiments using the $\tau = 0h$ setup (not shown). The optimal parameters were found to give the best compromise between producing a larger error reduction in θ, q_v, U, and V while adjusting the ensemble spread to account for updated uncertainties.

For the LEMAS experiments, the $N_F = 10$ closest members to the precipitation observations are used to construct the analysis ensemble, using an observation window of width W=420 km. The ensemble of N_F analysis mosaics are constructed using an expanded ensemble that includes the model states lagging the analysis time by 0, ±1 h, and ±2 h. For the first DA step, the expanded ensemble includes the model states taken only from the *En-6* runs (without DA). However, in the subsequent DA steps, states available from the recentered ensemble forecast are also included in the expanded ensemble. Therefore, at each DA step, the expanded ensemble contains at least 100 candidates (20 members x 5 model times). This expansion of the ensemble is only used for the analysis construction in LEMAS, and it is not applied to the LETKF. For the LETKF experiments, and observation error $\sigma = 2$ is used for the transformed precipitation observations and a localization radius of L=400 km. Finally, the same inflation factor of $\alpha = 0.7$ is used for the LETKF as well as LEMAS. The parameters in each experiment are also summarized in Table 4.3a

Although it is not shown, preliminary experiments showed that using smaller values for σ or N_F produced similar or slightly better error reductions in the state variables, but they produced a more substantial underestimation of the posterior uncertainties.

4.4.3 StageIV DA Experiments Setup

For each of the nine precipitation events in Table 4.2, we carry out the two cycled DA experiments, described in Table 4.3b, that assimilate actual Stage IV hourly precipitation observations (Baldwin

and Mitchell, 1997; Lin and Mitchell, 2005) using the LEMAS and the LETKF methods with the same parameters configuration as the OSSEs. For the cycled DA experiments assimilating StageIV observations, the 21 members of the *En-6* ensemble are used as first guess (background) since they are lagged forecast initialized closer to the DA period (6h). For these experiments, the 21 members of the *En-6* ensemble are used as background, and the StageIV observations are assimilated every 2 h during a 6 h DA period. After the DA period, a 12 h forecast is performed.

Table 4.3: Summary of the cycled DA experiments

a) Observing System Simulating Experiments

Name	DA method	DA Parameters	Truth	Background
LETKF/τ0	LETKF	$L = 400km$; $\sigma = 2$; $\alpha = 0.7$	En-18/M15	En-18 (excluding M15)
LETKF/τ6			En-12/M15	
LEMAS/τ0	LEMAS	$W = 420km$; $N_F = 10$; $\alpha = 0.7$	En-18/M15	
LEMAS/τ6			En-12/M15	

M# denotes the "member #"

b) StageIV DA Experiments

Name	DA method	DA Parameters	Truth	Background
LETKF/StageIV	LETKF	$L = 400km$; $\sigma = 2$; $\alpha = 0.7$	StageIV	En-6
LEMAS/StageIV	LEMAS	$W = 420km$; $N_F = 10$; $\alpha = 0.7$		

Although StageIV precipitation observations (hourly accumulation) are available over most of the contiguous US, the precipitation observations over the Rocky Mountains are discarded due to their low quality on account of significant radar beam blockage and sparser radar coverage. Therefore, only the precipitation observations located east of 105°W are assimilated (indicated by the area inside the blue contours in Fig. 4.2). Since the StageIV observations are available on a different grid and different grid spacing than the WRF domain grid, the precipitation observations are remapped into the WRF model grid using the same procedure described in PZY20 (Sec. 2.c). The two-step interpolation procedure first smoothes the Stage IV data to match the StageIV

and model grid resolutions, and then a bilinear interpolation is used to remap the smoothed observations to the WRF grid.

For the DA using LEMAS, during the first DA step the analysis mosaics are constructed using an expanded ensemble build-up of the model states lagging the analysis time by 0, ±1 h, and ±2 h, taken from the *En-6* (no DA), *En-12*, and *En-18* ensemble forecasts. In the following DA steps, states from the runs from the ensemble forecast where the DA is applied are also used to construct the expanded ensemble. Therefore, the expanded ensemble for the real DA experiments consists of 315 or more model state candidates (21 members x 5 model times x 3 lagged forecasts, plus the additional states from the ensemble runs with DA).

4.5 Results of the OSSEs

This section presents the results of the idealized DA experiments described in Sec. 4.4.2 comparing the LEMAS and LETKF approaches for two different magnitudes of forecast errors. The impacts of the DA on the forecast quality are evaluated for selected model state variables using the ensemble mean Root Mean Square Errors (RMSE). Additionally, the RMSE is compared with the ensemble spread to evaluate the effects of DA on the representation of the forecast uncertainties by the ensemble forecast. An ensemble forecast that is representative of the actual forecast uncertainties is obtained when the RMSE of the ensemble mean is close to the ensemble spread. That is, the $\overline{spr}/RMSE$ ratio remains close to 1. The ensemble spread is computed following Fortin et al. (2014) as:

$$\overline{spr}_{\boldsymbol{x}} = \sqrt{M^{-1}\sum_{i}^{M}\left[(N_e-1)^{-1}\sum_{m}^{N_e}\left(\boldsymbol{x}_{m,i}-\overline{\boldsymbol{x}_i}\right)^2\right]} \tag{4.14}$$

where x denotes a state variable, "i" a grid point, mth a given ensemble member, and the overline indicates the ensemble mean values. The summation limits N_e and M indicates the ensemble size,

and the total number of grid points over the selected region. The RMSE and spread are computed over the selected area of the domain indicated in Fig. 4.1, considering all the levels located in the troposphere.

Finally, the overall skill for the precipitation forecasts is evaluated for the hourly precipitation of the ensemble mean, using the RMSE of the hourly precipitation (in dBR units) and three categorical contingency scores (Wilks, 1995). Three categorical scores are used to evaluate the forecast skill for precipitation occurrence. The Equitable Thread Score (ETS) measures the skill of a forecast relative to a random forecast, taking values from -1/3 to 1, with 1 indicating a perfect forecast, and values close to 0 or lower no skill. The Probability of Detection (POD) that measures the number of precipitation events that were correctly forecasted, relative to the total number of events (positive and negative precipitation occurrences), with 1 indicating a perfect forecast and 0 no skill. Last, the False Alarm Ratio (FAR) that measures the fraction of the false-positive precipitation forecasts with respect to the total number of positive forecasts, with 0 indicating a perfect score and 1 no skill. A detection threshold of 1 mm (hourly accumulation) is used to compute all the contingency scores.

4.5.1 $\tau = 0h$ experiments

In this section, we present the results from the $\tau = 0h$ OSSEs where the truth is a randomly chosen member taken from the same ensemble runs used as background (i.e., no model initialization lag between the truth and the background).

Figure 4.4 shows the impacts of the DA on the Root Mean Square Error (RMSE) of the ensemble mean LEMAS and the LETKF experiments. A control experiment with no DA (No DA) is also shown for comparison. If we look at the RMSE values averaged over the nine cases (first column of Fig. 4.4), both DA methods produce a similar reduction of the forecast errors that persists over the entire forecast period. However, the magnitudes of improvements for each

case show a large case-to-case variability (the second and third columns in Fig. 4.4). In general, both DA methods reduce the errors in the four variables by 12% or more when precipitation is assimilated. An exception to the mentioned behavior are Cases A and G for LEMAS forecasts, and Cases A and B for the LETKF ones, where a small increase in the RMSE is observed in at least three variables.

Figure 4.5 shows the values ratio of the RMSE of the ensemble mean and the ensemble spread for θ, q_v, U, and V. The results for the control runs without DA are also shown for reference. If we consider the averaged values over all the cases, the control experiments without DA show spread/RMSE ratio that is close to one, with a case-to-case variability ranging from 0.8 to 1.2. Therefore, this feature shows that the background used in the $\tau = 0h$ experiments corresponds to an optimal situation where the ensemble spread well represents the background errors. For the LEMAS and the LETKF experiments, although there is large case-to-case variability (shaded area), the spread/RMSE ratio values averaged over all the cases are only slightly larger than the values of the No DA experiments, remaining close to one during the DA period (-6 to 0 h) and the 12 h forecast period.

Let us now analyze the impacts of the DA on the ensemble mean hourly precipitation forecasts. Figure 4.6, shows the ensemble-mean precipitation errors measured by the RMSE (in dBR units), and three categorical scores (ETS, FAR, and POD). If we look at the average over the nine cases, both LEMAS and the LETKF produce comparable improvements of approximately 10% that persist over the entire forecast period (first and second columns in Fig. 4.6), with LEMAS obtaining slightly larger improvements than the LETKF (~2%). To look into more detail the case-to-case differences between each DA method, we show in the third column of Fig. 4.6 the LEMAS−LETKF improvements differences for each case. During the DA window, no DA method outperforms the other for a majority of the cases. Similar behavior is observed for the FAR over the entire forecast period. However, between 3 to 9 h of the forecast, for seven or more cases

(77%), LEMAS produces a larger error reduction than the LETKF for the RMSE, ETS, and POD.

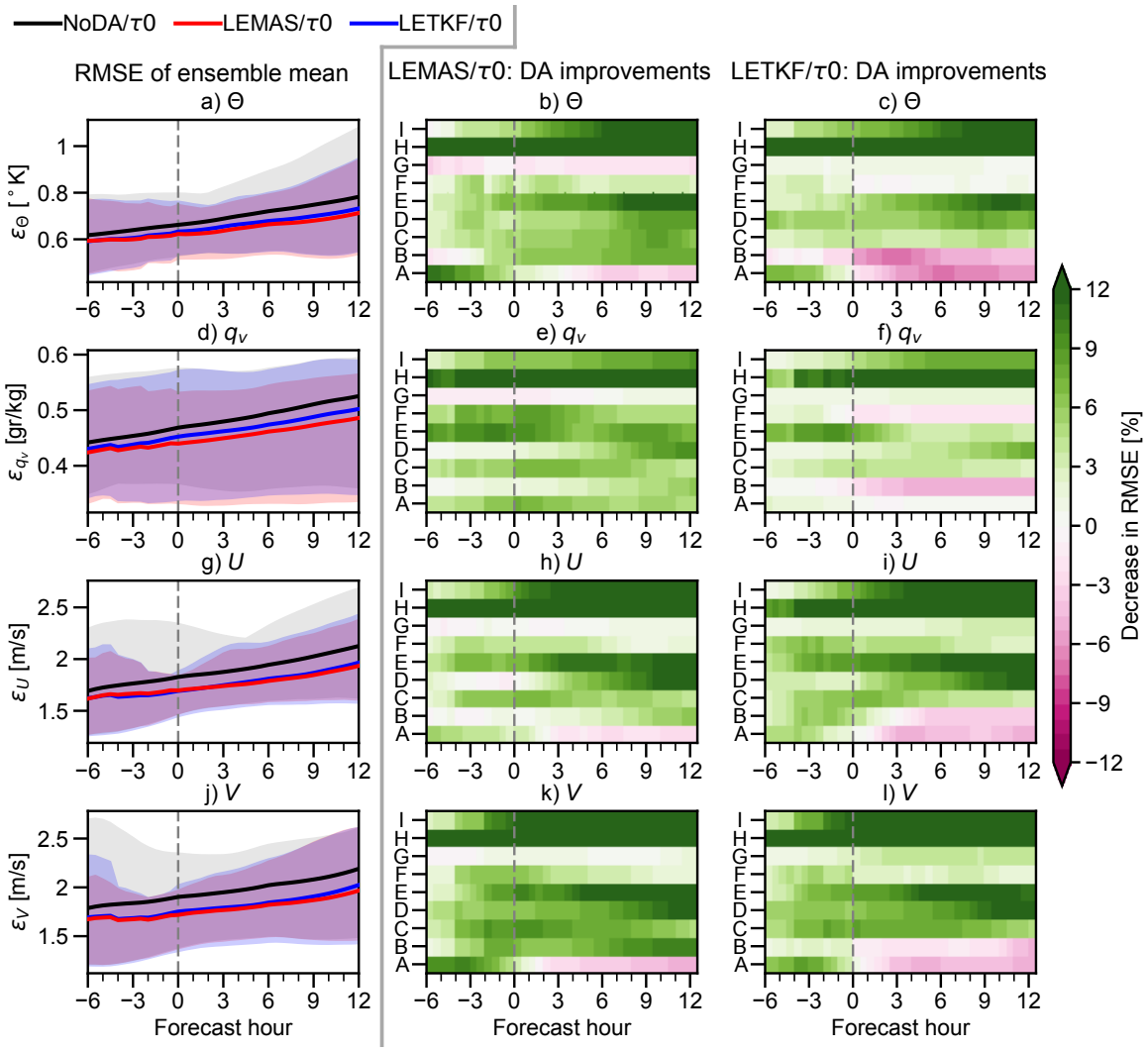

Fig. 4.4: $\tau = 0h$ OSSEs. Impacts of DA on the RMSE of the ensemble mean on the potential temperature θ (first row), vapor mixing ratio q_v (second row), U wind (third row), and V wind (last row). The first column displays the RMSE values averaged over the nine cases (thick lines) for the forecasts with no assimilation (NoDA, black), and the forecasts produced by LEMAS (red) and the LETKF (blue). The shaded area indicates the full case-to-case variability of the RMSE. The second and third column show the relative decrease in the RMSE produced by LEMAS (second column) and by the LETKF (third column). The decrease in the RMSE is measured with respect to the runs with no DA. Positive values indicate improvements (reduction in the error). The vertical dashed line denotes the end of the DA period.

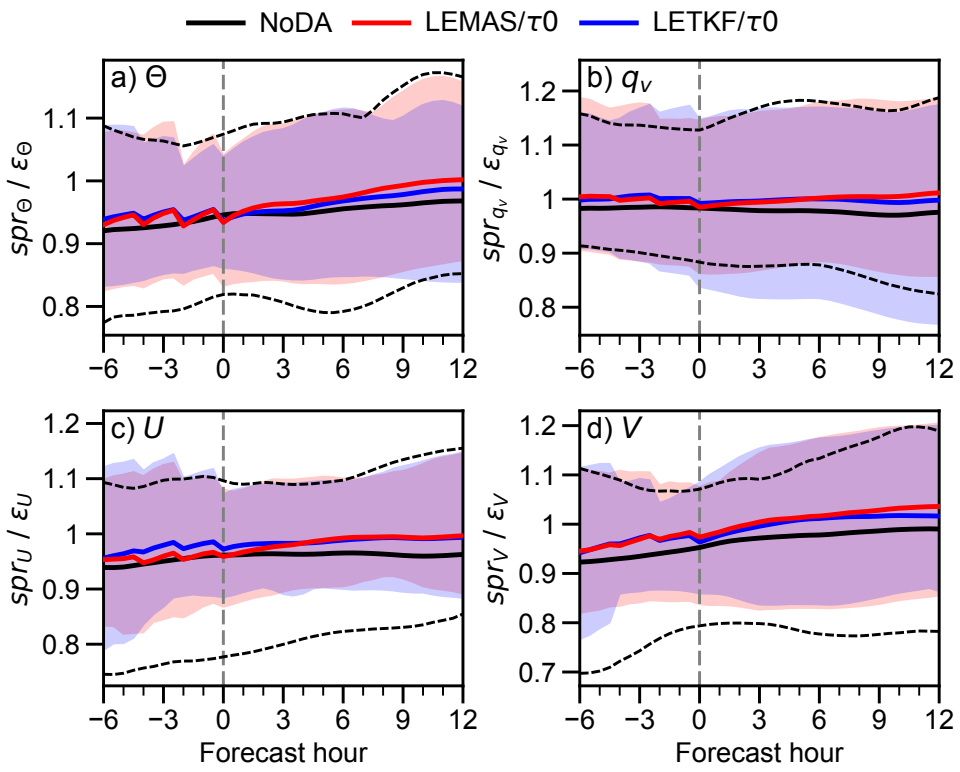

Fig. 4.5: $\tau = 0h$ OSSEs. Ratio of the ensemble spread $\overline{spr}$ and the $RMSE$ of the ensemble mean for potential temperature θ (a), vapor mixing ratio q_v (b), U wind (c), and V wind (d). The $\overline{spr}/RMSE$ ratio is displayed for the forecasts with no assimilation (black), and the forecasts produced by LEMAS (red) and the LETKF (blue). The thick lines indicate the values averaged over the nine study cases, while the shaded area indicates the full case-case variability for the LEMAS (red) and the LETKF (blue) experiments. The dashed black lines indicate the full case-to-case variability for the runs without DA (NoDA).

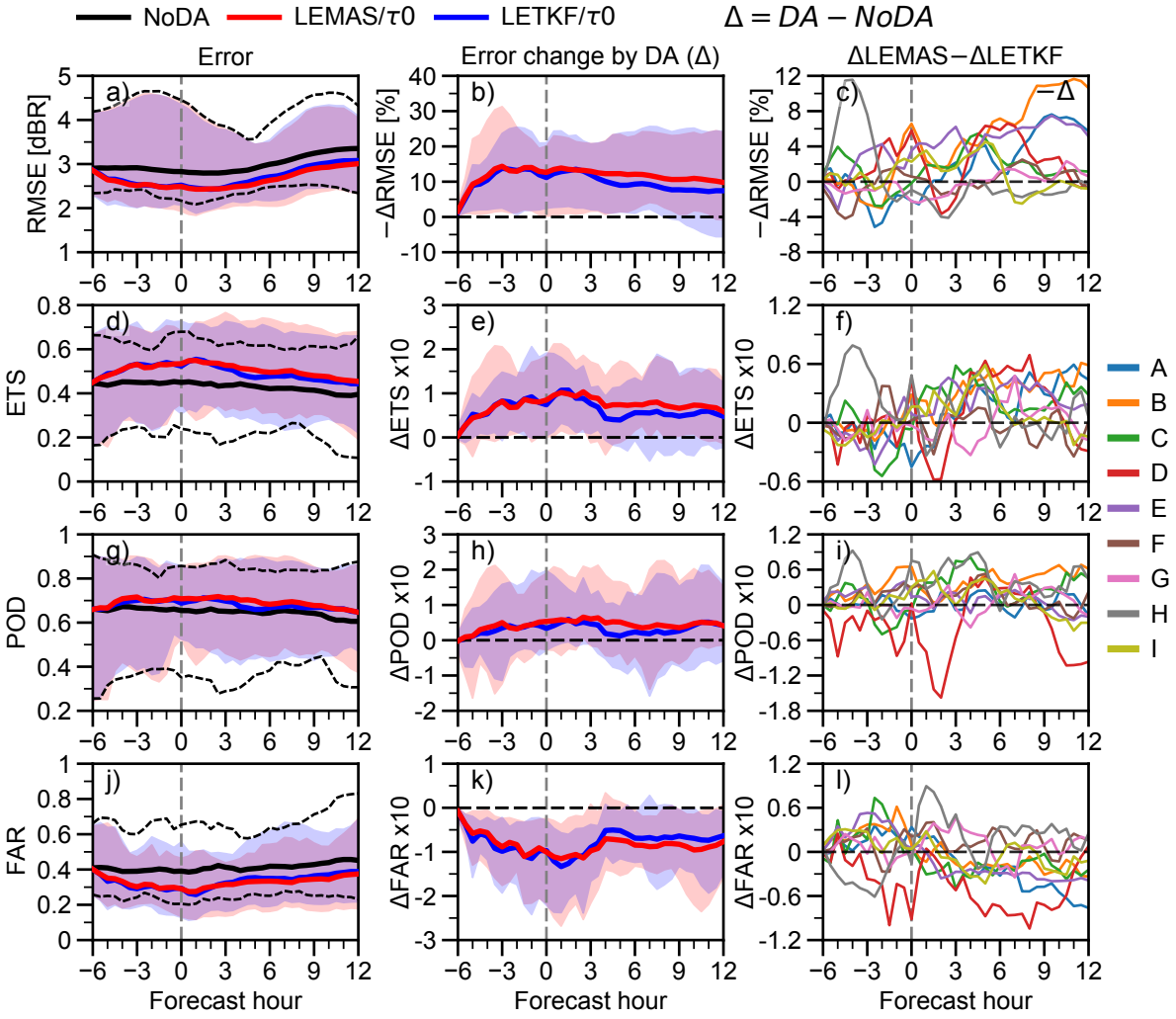

Fig. 4.6: $\tau = 0h$ OSSEs. Impacts of DA on the ensemble mean precipitation forecast measured by RMSE (first row), the ETS (second row), the POD (third row), and the FAR (last row). The errors are computed using the hourly accumulated precipitation, and a threshold of 1 mm for the contingency scores. The first column displays the absolute errors averaged over the nine cases (thick lines) for the control runs without DA (NoDA) and the LEMAS (red) and the LETKF (blue) experiments. The shaded area indicates the full case-to-case variability for the LEMAS and the LETKF forecasts, while the dashed line the variability for the control runs. The second column indicates the error change by the DA with respect to the NoDA runs (denoted by Δ). The change in the RMSE has its sign reversed to make positive values indicate improvements, as in Fig 4.4. The third column shows the difference in the error change of the LEMAS runs against the LETKF ones for each case (A to I).

4.5.2 $\tau = 6h$ experiments

In the previous section, we show that for the $\tau = 0h$ experiments, the LEMAS approach produces a similar quality forecast than those obtained by the LETKF. By construction, the previous experiments have an average spread-to-skill ratio close to 1. In this section, we extend the previous analysis to the $\tau = 6h$ experiments that, as we will show next, represent a less optimistic situation where the forecast errors are larger, and the ensemble spread considerably underestimate the forecast errors. Although current operational forecasting systems typically have a good overall match between spread and error over time, for particular situations, the ensemble spread may considerably underestimate the uncertainties (see, for example, Haiden et al., 2019, their Fig. 10 and 11). Therefore, in this section, we will explore how LEMAS performs under this less favorable situation and evaluate its potential to correct more significant forecast errors. As in the previous section, the LEMAS approach is also compared with the LETKF.

Figure 4.7 shows the impacts of the DA for the $\tau = 6h$ OSSEs on the temperature, humidity, and horizontal winds errors. Let us first focus our attention on the first column of the figure showing the absolute RMSE. The errors of the control runs with no DA for the $\tau = 0h$ experiments are also shown as a reference (dashed lines). Overall, the magnitude of simulated forecast errors in the $\tau = 6h$ experiments are approximately 30-40% larger than the $\tau = 0h$ ones. For the $\tau = 6h$ experiments, both DA methods produce small but long-lived improvements on the RMSE with respect to the runs without DA, with LEMAS performing slightly better than the LETKF. If we look at the relative error decrease produced by the DA for each study case, shown in the second and third columns in Fig. 4.7, the magnitude of the improvements produced by LEMAS are generally larger than the LETKF ones in most of the study cases.

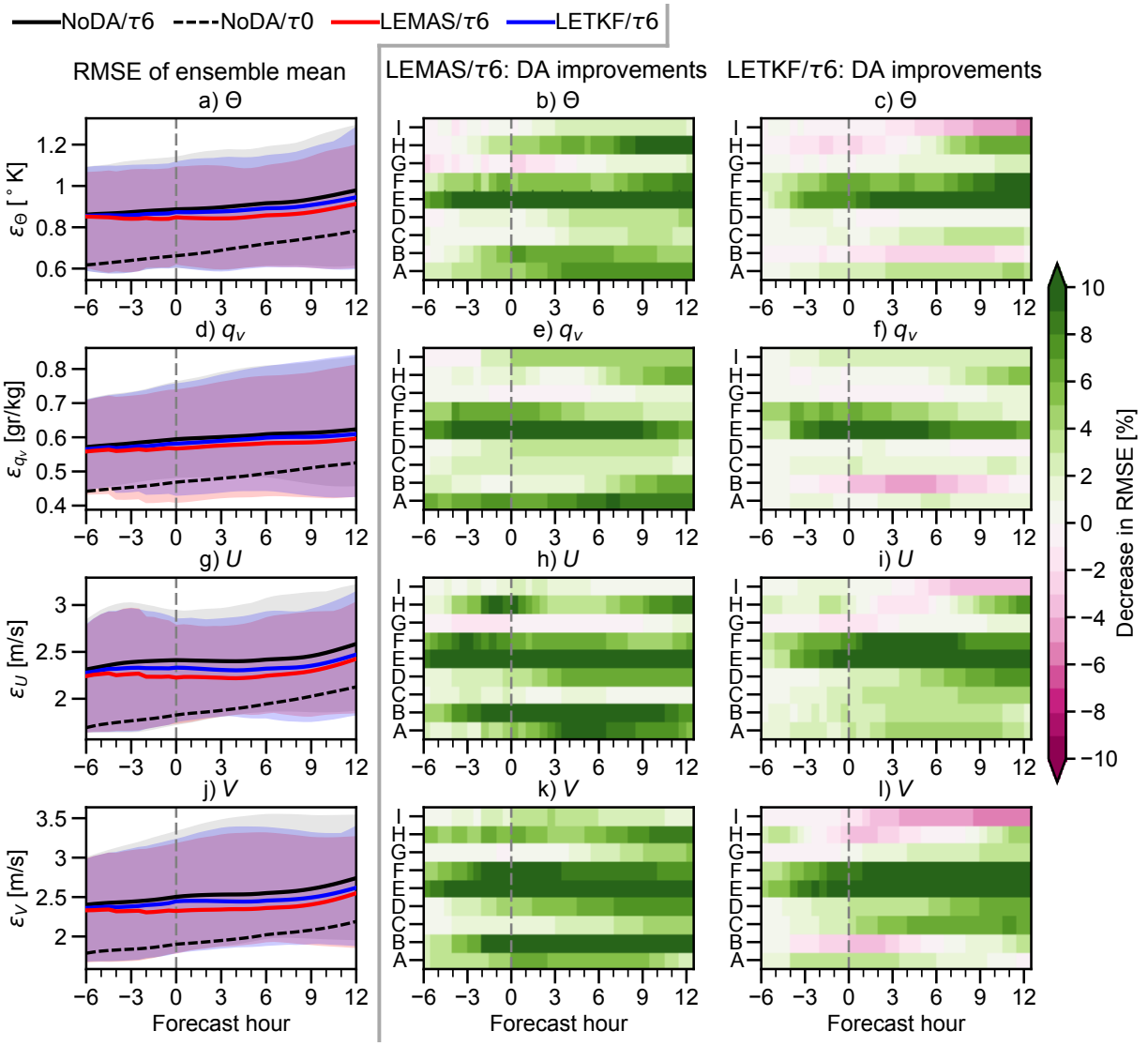

Fig. 4.7: As Fig. 4.4 but for $\tau = 6h$ OSSEs. In the first column, the errors of the control runs without DA for the $\tau = 0h$ experiments are also show as reference (NoDA/τ0, dashed lines).

To explore into more detail the case-to-case differences in the performance of each DA method, we show in Fig. 4.8 the differences between the second and the third column of Fig. 4.7. In other words, the case-by-case difference between the error decrease in the LEMAS experiments against the LETKF ones. For the vapor mixing ratio (Fig. 4.8b), LEMAS produces a similar or larger error reduction than the LETKF for all the 9 cases over the entire forecast period, and in 7 of the

9 cases for temperature, and horizontal winds (Fig. 4.8a, c, and d, respectively).

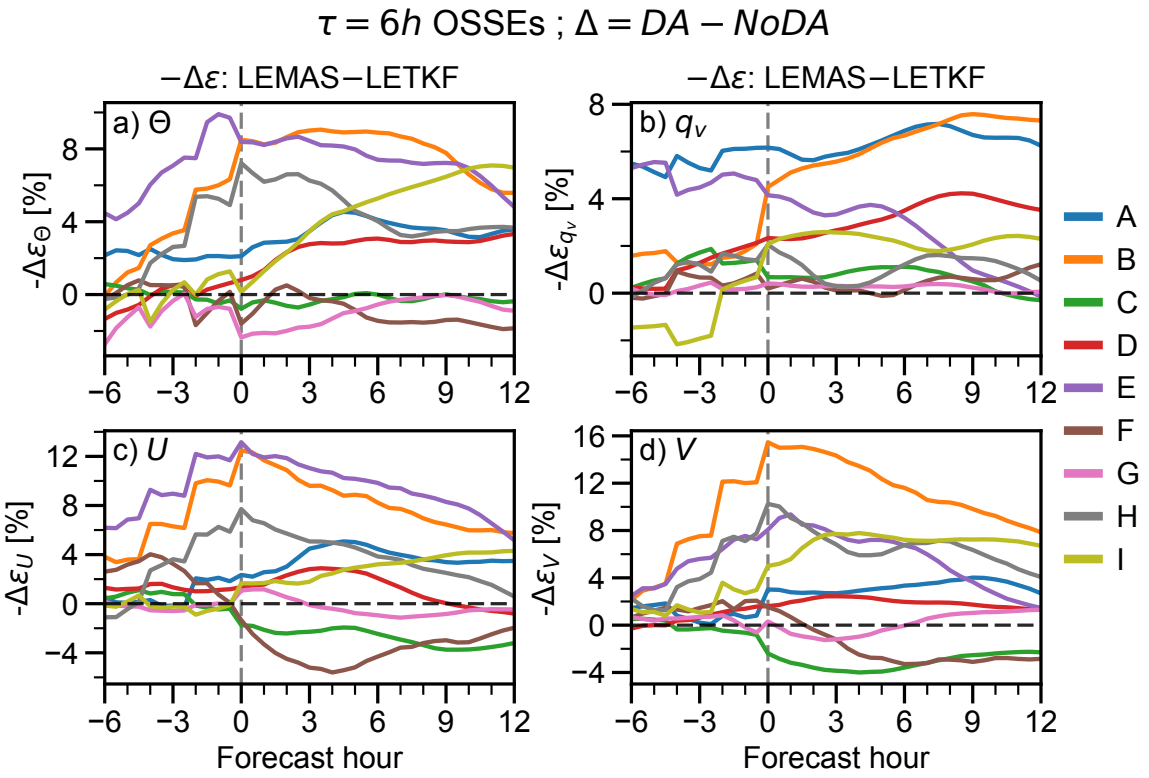

Fig. 4.8: $\tau = 6h$ OSSEs difference between the decrease in ensemble mean RMSE in the LEMAS experiments against the LETKF ones for each study case (A to I). The change in the RMSE by the DA (denoted by Δ), measured with respect to the NoDA runs, has its sign reversed to make positive values indicate improvements. Hence, positive values in the plots indicate larger improvements by LEMAS with respect to the LETKF.

Figure 4.9 shows the ensemble spread to RMSE ratio for temperature, vapor, and winds for the $\tau = 6h$ experiments. At the beginning of the DA period (-6 h), the spread/RMSE ratio before the DA (NoDA run) has an average value of ~0.7, with values ranging from 0.5 to 0.9. The low spread/RMSE ratio indicates that for these experiments, the forecast errors are underestimated by the ensemble spread. In general, the assimilation of the precipitation observations using both DA methods does not improve the spread/RMSE ratio with respect to the control run without DA. These results suggest that for the experiments with larger forecast errors, the assimilation of

precipitation alone is not enough to improve significantly the representation of the uncertainties by the ensemble (bring the spread/RMSE ratio close to 1). However, maintaining a similar spread/RMSE ratio indicates that the reduction in the RMSE at each DA step is accompanied by a similar reduction in the ensemble spread.

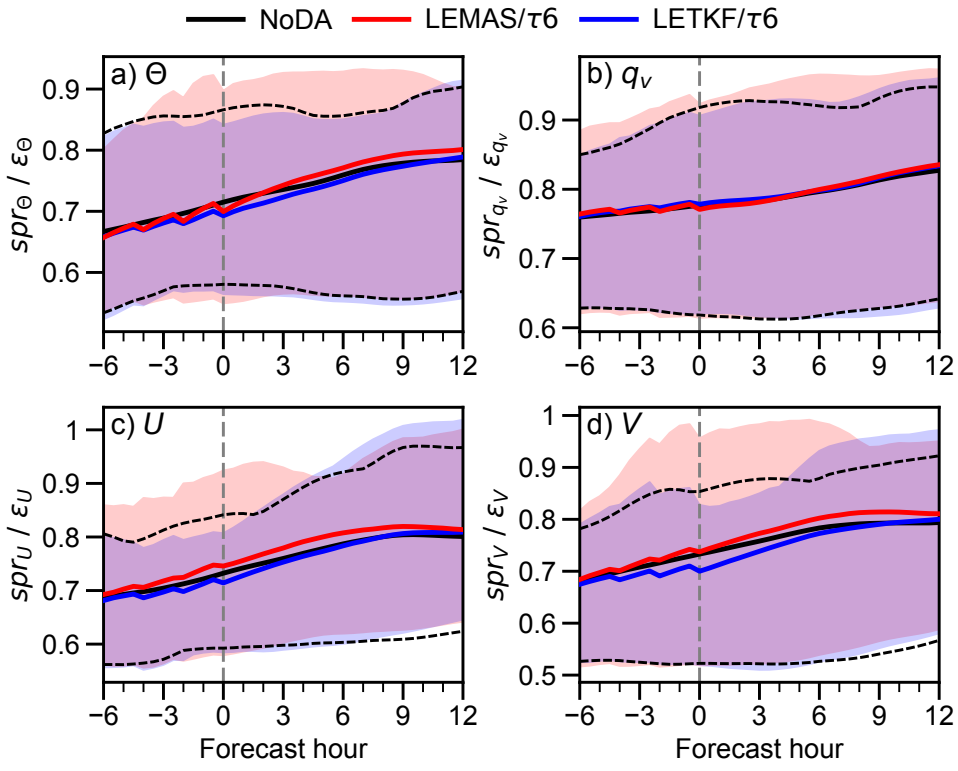

Fig. 4.9: As Fig. 4.5 but for $\tau = 6h$ OSSEs.

Figure 4.10 shows the effects of the DA in the precipitation forecast quality for the $\tau = 6h$ experiment. If we look at the average improvements over all the cases (first and second columns), the LEMAS forecasts produce a similar or slightly better precipitation forecast than the LETKF ones for all the scores. These results are in agreement with the larger positive impacts of LEMAS on temperature, humidity, and winds (Fig. 4.7). Let us now look at the case-to-case differences between the LEMAS improvements with those obtained by LETKF (third column in Fig. 4.10). For the RMSE, LEMAS produces comparable or larger error reductions than the LETKF for all

the cases, except for case G during the 0 to 6 h the forecast. For the categorical scores (Fig. 4.10f, i, and l), during the DA period (-6 to 0 h), no DA method outperforms the other over a large number of cases. From 6 to 12 hours of lead time, LEMAS tends to produce slightly better results than the LETKF on the ETS and FAR for almost all the cases. For the POD, the benefits of LEMAS over the LETKF are significant ($\geq 7/9$ of the cases) only during the first 6 hours of the forecast.

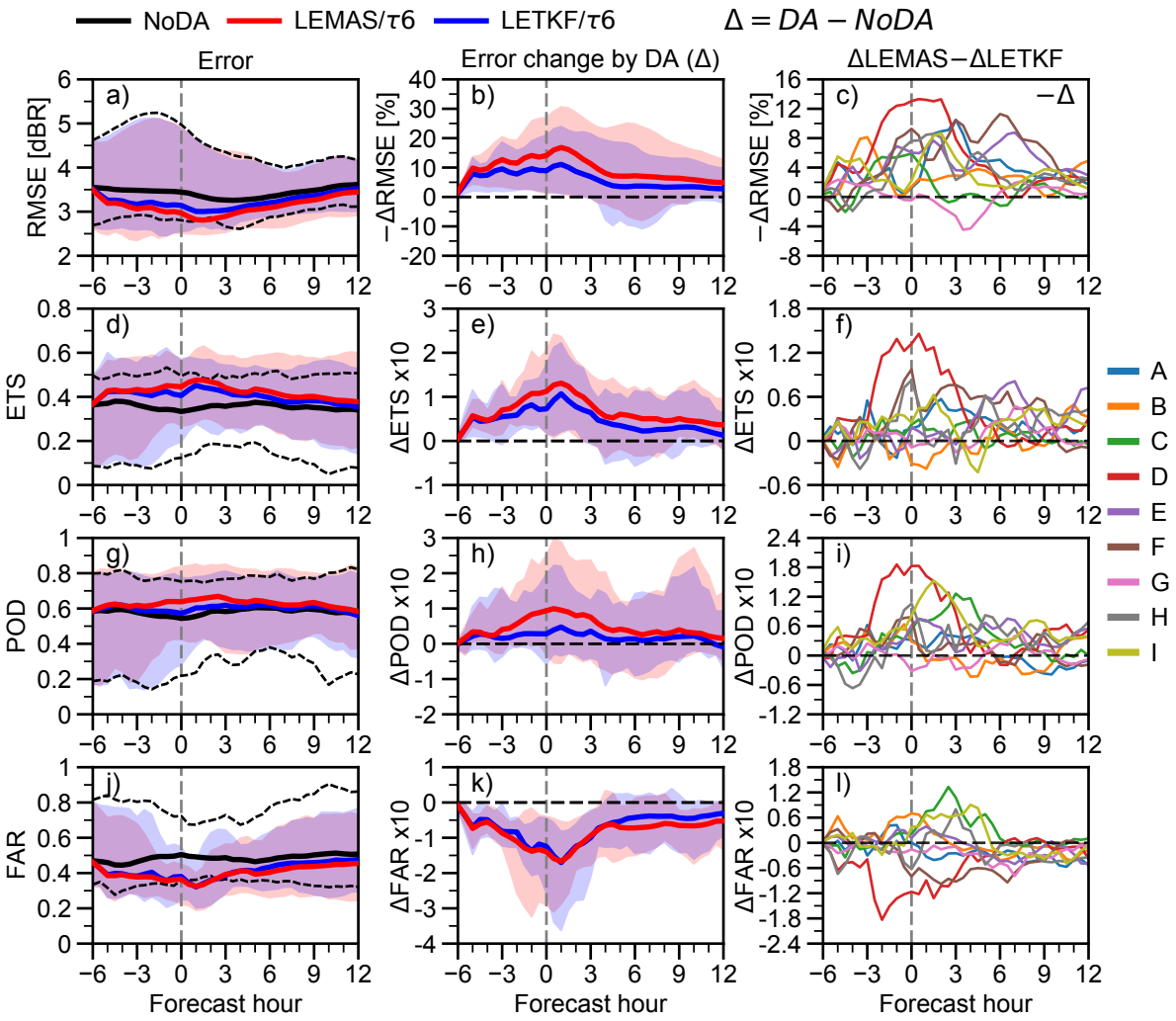

Fig. 4.10: As Fig. 4.6 but for $\tau = 6h$ OSSEs.

4.6 Results for the StageIV DA experiments

In this section, we evaluate LEMAS in a real DA situation by assimilating StageIV precipitation observations and comparing its performance with the LETKF approach in the OSSEs. As in the previous section, the forecasts produced by LEMAS are compared against those produced by a LETKF approach. For a detailed description of the experiments discussed in this section, the reader is referred to Sec. 4.4.3.

Figure 4.11 shows the impacts of the DA on the ensemble mean precipitation forecasts for the LEMAS/StageIV and the LETKF/StageIV experiments. The same scores used to analyze the ideal experiments (Sec. 4.5) are used, computed over the region where the StageIV precipitations are available (denoted by the blue contour in Fig. 4.2). The results for the control run without DA are also shown as a reference. On the average, both DA approaches produce forecasts of similar quality, improving all the scores up to 9 h of the forecast, except for the POD that is not improved by the DA. In particular, for the RMSE, during the first 3 h of the forecast, LEMAS produces a $\sim 5\%$ larger improvement than the LETKF (Fig. 4.11b). Although it is not shown, these larger improvements by LEMAS are present in 7 of the 9 cases. Interestingly, the difficulties to improve the POD score were not observed in the idealized experiments shown in Sec. 4.5, where small but significant improvements in the POD are present (e.g. Fig. 4.6 and 4.10). One explanation for this different behavior is the tendency of the model to overestimate the precipitation occurrence with respect to the observed values, which can maintain higher POD values.

To explore this in more detail, in Fig. 4.12a, we show the Frequency Bias Index score (FBI, Wilks, 1995), computed for the ensemble mean precipitation using the same detection threshold as in the other categorical scores. The FBI measures the frequency of precipitation occurrence relative to the observed occurrence. The FBI ideal score is 1, where values higher than one indicate an overestimation of the forecasted precipitation occurrence, while a value less than

one indicates the opposite. On average, the forecasts without DA overestimate the precipitation occurrence by 75% to 100%(1.75 < FBI < 2) on most of the cases, and even higher for cases B and C (Fig. 4.12a). When precipitation is assimilated by both DA methods, the overestimation of the precipitation occurrence is reduced, reaching FBI values close to 1 by the end of the DA

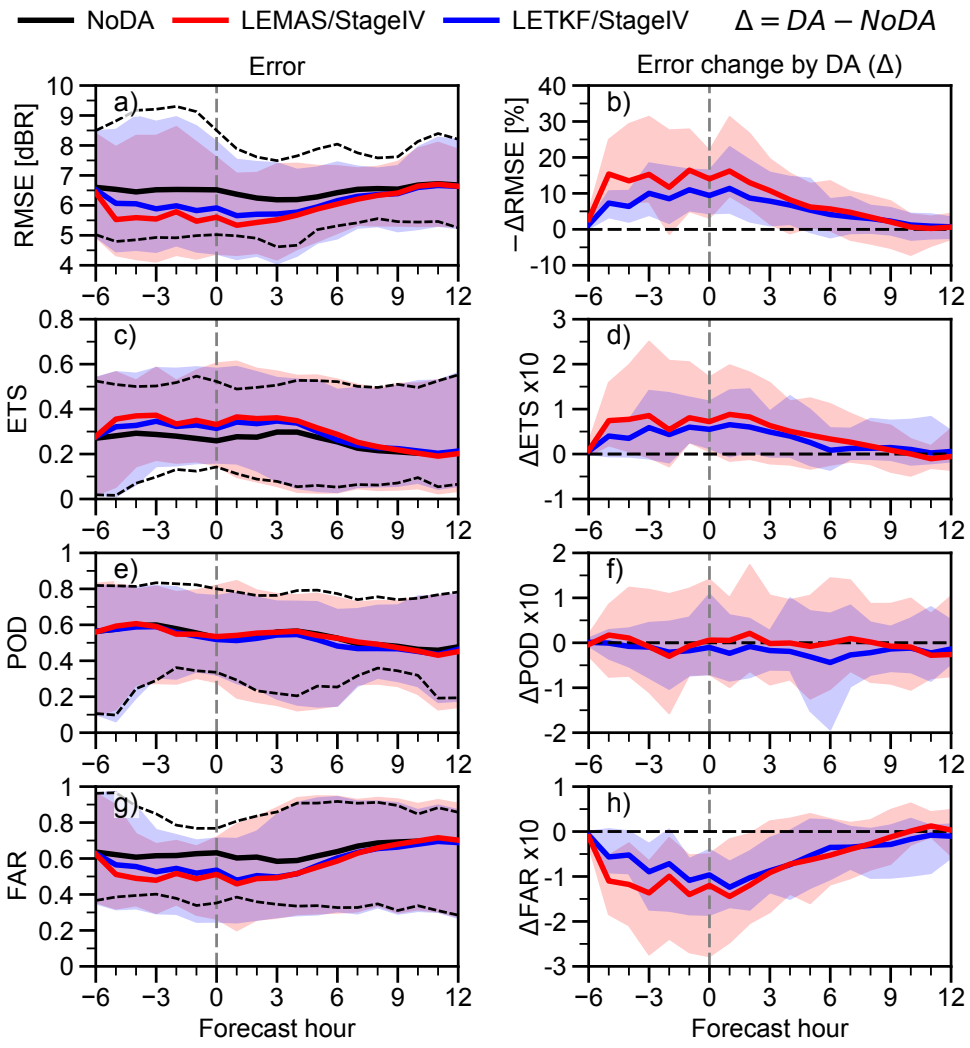

Fig. 4.11: Impacts of DA on the precipitation forecast quality for DA experiments in assimilating StageIV observations. The forecast errors are measured by the RMSE (first row), ETS (second row), POD (third row), and FAR (last row). The same description for the first and second columns in Fig. 4.6 applies.

period (Fig. 4.12b). Similarly to the other error metrics shown in Fig. 4.11 , the positive impacts of the DA are quickly lost after 9 hours the forecast.

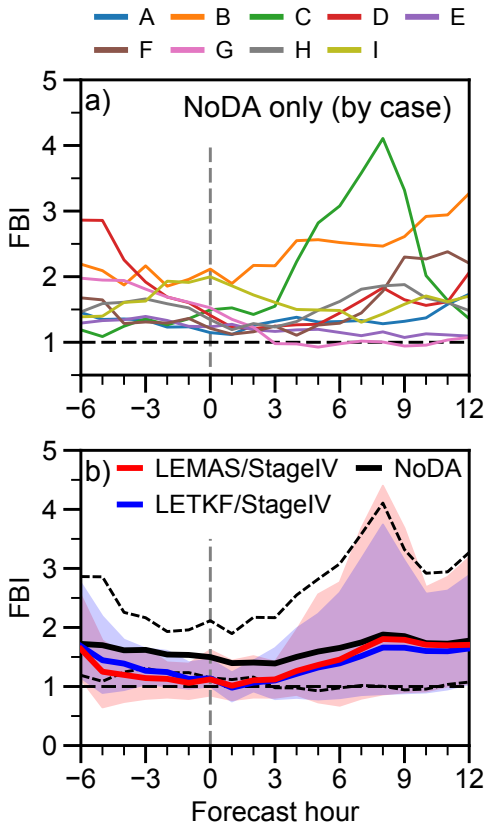

Fig. 4.12: Frequency Bias Index (FBI) for the StageIV DA experiments. a) The FBI for the control experiments without DA, shown by case. b) Impacts of DA on the precipitation FB) for DA experiments assimilating StageIV observations. The thick lines indicate the FBI values averaged over the nine cases while the shaded area indicates the full case-to-case variability for the LEMAS and the LETKF forecasts, while the dashed line the variability for the control runs.

Next, let us investigate the effects of the DA on the precipitation patterns in more detail, focusing on two of the nine precipitation events, namely cases A and D. These two cases are selected to illustrate how both DA methods perform under very different weather situations. Case A corresponds to a widespread precipitation system driven by a synoptic front, while in Case D precipitation is produced by two MCSs, having a more localized and intense nature than the

precipitation system in Case A.

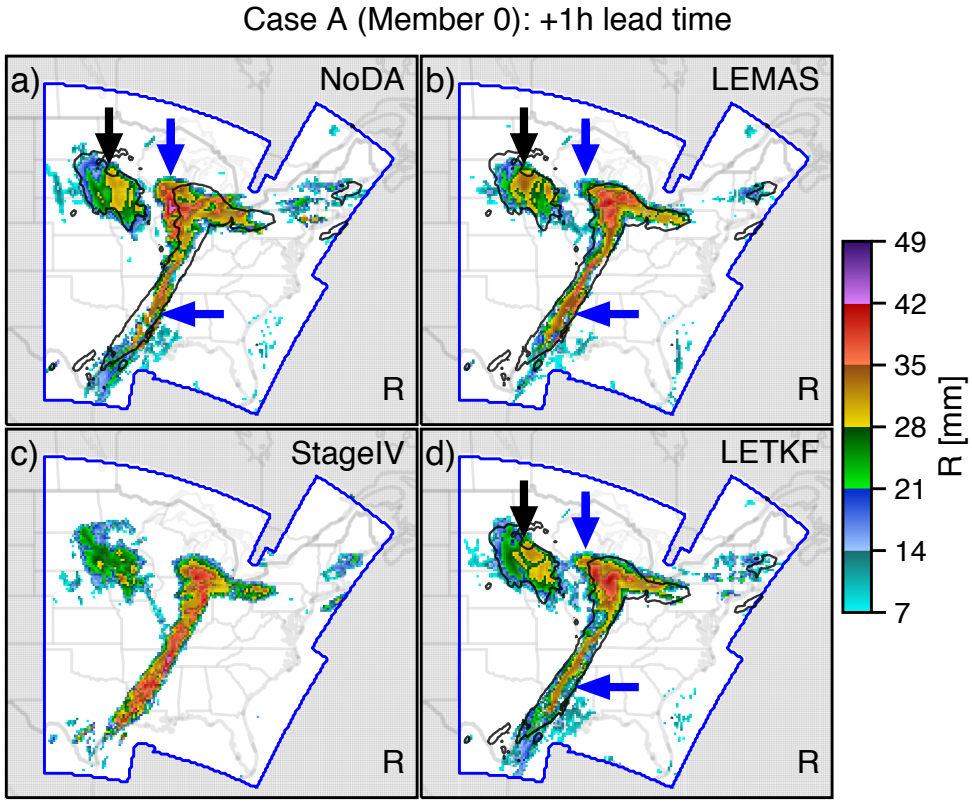

Fig. 4.13: Precipitation forecast for Case A experiments in assimilating StageIV observations. Results are shown for Member 0. All the panels show the hourly accumulated precipitation at 1 h of forecast time (1 h after the end of the DA period). The first column shows the control experiment results without DA (a) and the StageIV hourly accumulation observations (c). The second column displays the precipitation for the LEMAS (b) and the LETKF (d) experiments. The black contours indicate the StageIV precipitation (1 mm), while the region inside the blue contour denotes the area where the StageIV observations are available.

Figure 4.13 shows the forecasted precipitation for the Case A at 1 h of lead time (after the end of the DA period). The precipitation forecasts are only shown for member 0, but qualitatively similar results are obtained for the other members. In general, the $NoDA$ run (Fig. 4.13a) produces a precipitation forecast that is close to the observed values (Fig. 4.13c). For example, overall, the shape of the precipitation is generally captured by the control run. However, there are

significant positional errors in the precipitation pattern in the regions indicated by the blue arrows. The assimilation of precipitation using both LEMAS (Fig. 4.13b) and the LETKF (Fig. 4.13d) methods corrects most of the positional errors, producing a precipitation pattern that is close to the observed one.

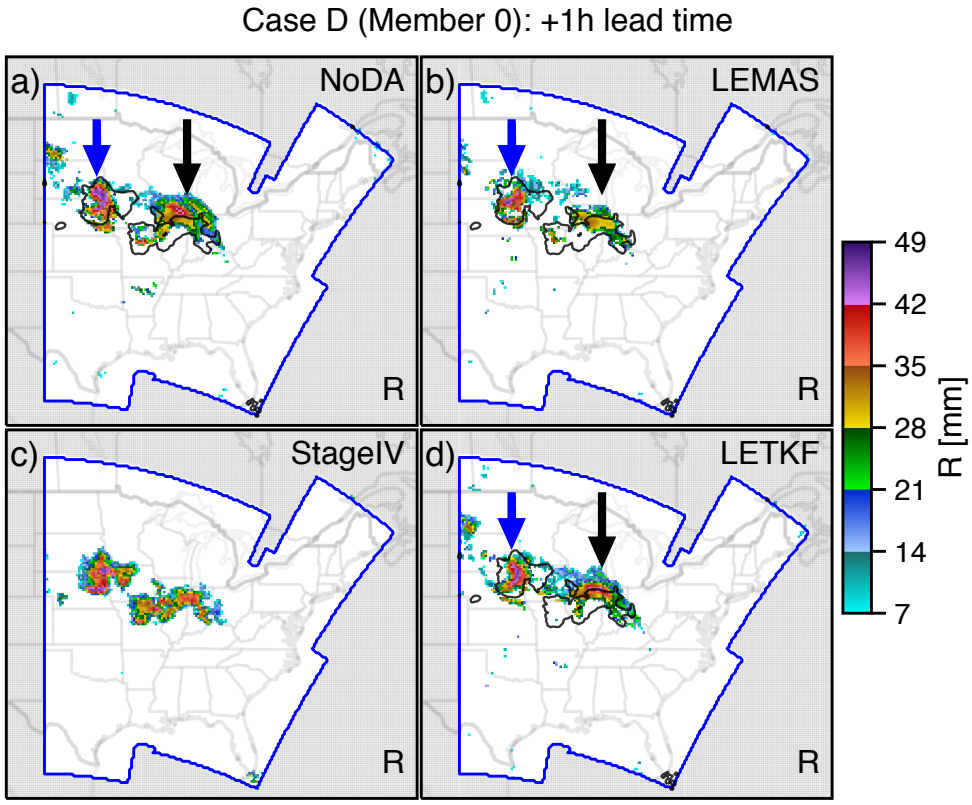

Fig. 4.14: As in Fig. 4.13, but for Case D.

Similar to Fig. 4.13, Fig. 4.14 shows the forecasted precipitation, but for the Case D. For this case, the forecast without DA (Fig. 4.14a) produces a precipitation forecast that is similar to the observed precipitation (Fig. 4.14b). Although the two observed convective systems are present in the control forecast, they have a different structure and location than the observed precipitation systems. For example, the convective system indicated with a blue arrow in Fig. 4.14a is approximately at the correct location, by it has a different shape and a slightly smaller spatial

extent than the observed system. While the other convective system, pointed at by the black arrow, is displaced north with respect to the observed pattern. The impacts of the assimilation of the precipitation observations using either LEMAS (Fig. 4.14b) or the LETKF (Fig. 4.14d) are more evident on the eastern convective system than the western one. Both precipitation forecasts produced by LEMAS and the LETKF suppress an important fraction of the overestimated convection (indicated by the black arrows), with LEMAS producing a less intense precipitation system.

4.7 Conclusions

This paper represents a sequel to the PZY19 and PZY20 studies that introduced the Localized Ensemble Mosaic Assimilation (LEMA) method for precipitation observations. Here, we present a new version of LEMA, called Localized Ensemble Mosaic Assimilation Sequence (LEMAS), that constructs an ensemble of analysis mosaics using the members that are locally closer to the precipitation observations. The new ensemble forecast is then initialized by recentering the prior ensemble around the mean of the analysis ensemble while scaling the original background perturbations to match the spread of the analysis mosaics.

A series of cycled idealized DA experiments for nine different precipitation events are used to evaluate the performance of LEMAS for two different magnitudes of simulated forecast errors. For the idealized experiment with the smallest forecast errors, LEMAS produces similar quality forecasts than the ones produced by the LETKF for temperature, water vapor, winds forecasts, and precipitation. For the experiment with more significant forecast errors, LEMAS produced better forecasts for most precipitation events.

In addition to the idealized experiments, the potential of LEMAS is also explored in the context of real cycled DA experiments assimilating StageIV precipitation observations. For

these experiments, LEMAS produces precipitation forecasts with comparable quality to the ones produced by the LETKF approach. However, contrary to the idealized experiments, the DA improvements were lost 9 h after the end of the DA. A possible explanation for this limited persistence of the improvements is the tendency of the model to overestimate the precipitation occurrence, indicating the presence of significant model errors.

Although LEMAS is a simple method for precipitation DA that does not rely on any Gaussianity linearity assumptions, it can only handle precipitation observations. Therefore, DA strategy is needed to combine LEMAS with conventional ensemble DA (EDA) methods that assimilate the rest of the available observations. A simple strategy could be combining both DA methods using a "hybrid gain" approach (Penny, 2014; Bonavita et al., 2015), where the EDA analysis ensemble is recentered on a new analysis mean that is a linear combination of the LEMAS analysis mean $\bar{\boldsymbol{x}}^F$ and the EDA analysis mean $\bar{\boldsymbol{x}}^a_{EDA}$ as:

$$\bar{\boldsymbol{x}}^a = \beta \bar{\boldsymbol{x}}^a_{EDA} + (1-\beta)\bar{\boldsymbol{x}}^a_{LEMAS} \tag{4.15}$$

where β is a tunable parameter controlling the relative weight given to the LETKF analysis with respect to the LEMAS analysis, reflecting the confidence on each analysis. This hybrid approach can also be extended towards the spread of the LEMAS analysis mosaics. An alternative approach was proposed by Buehner and Jacques (2020), where a non-Gaussian DA method (like LEMAS) is first applied to assimilate precipitation observations using the updated ensemble as background state for the EDA. However, Buehner and Jacques (2020) shows that this DA strategy may degrade the forecast quality. The potential of these or other strategies to combine LEMAS (or other non-Gaussian methods) with conventional DA approaches will be explored in future work.

Data availability statement. The WPS and ARW-WRF model (version 3.7.1) codes are publicly available at https://www2.mmm.ucar.edu/wrf/users/download/get_source.html. The GEFS data used to create the WRF LBCs/ICs is accessible at https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-ensemble-forecast-system-gefs. Since the model outputs of the experiments presented in this study are too large (>20 TB) to be made available online, researchers wishing to obtain model for collaborative research purposes, as well as the code implementing the LEMAS method, should contact the corresponding author. The LETKF code used in this study is based on the open source code released by Takemasa Miyoshi that is available at https://github.com/takemasa-miyoshi/letkf.

Acknowledgments. We acknowledge the valuable comments and suggestions of Prof. Frédéric Fabry, Dr. Peter Houtekamer, and two anonymous reviewers that helped us to improve the manuscript. The research reported here has been supported by the NSERC/Hydro-Quebec Industrial Research Chair program.

Chapter 5

Summary, conclusions, and ideas for future work

5.1 Summary and Conclusions

This thesis explores a new DA strategy that approaches the DA problem from a heuristic point of view, based on the intuitive idea that model states that are closer to the observed precipitation (a proxy of the "truth") are also closer to the truth state in the state variables that caused precipitation in the first place. The new heuristic DA method named "Localized Ensemble Mosaic Assimilation" (LEMA) does not assume Gaussian statistics or linearity for the analysis construction. Instead, it relies on the information in an ensemble forecast and the mentioned physical notion to construct the analysis. In more rigorous terms, LEMA constructs an analysis mosaic where each grid column contains the column-state from the ensemble member that is locally closer to the observed precipitation. For each column, the local proximity is measured using the Mean Absolute Difference (MAD) between the modeled and the observed precipitation computed over a rectangular window centered at the column. Using Observing System Simulation Experiments (OSSEs),

in Chapter 2, the notion that smaller errors in the precipitation forecast are associated with a lower error in the state variables is expressed in terms of the joint probabilities $p(\Delta\varepsilon_\phi, \Delta MAD)$ involving the decrease of the error in a given state variable by $\Delta\varepsilon_\phi$ and the precipitation error ΔMAD. In the context of a single DA step, sensitivity studies using increasing sizes of the observation window showed that using a large observation window (820 km) improves the transmission of the information from precipitation towards the state variables and increases the area over which the analysis can be constructed. In terms of forecast quality, the idealized experiments showed that LEMA produces a long-lasting reduction in the state variable and the precipitation errors.

However, the experiments in Chapter 2 were performed within an optimistic setup, with the truth being correctly covered by the background ensemble. Chapter 3 extends the idealized experiments to less optimistic circumstances that may occur in real DA situations, where the forecast errors are large, and the background ensemble may not fully capture the reality for particular situations. The OSSEs with different characteristics of forecasts errors show that the effectiveness of LEMA decay when the ensemble used to construct the analysis does not cover the truth state. To alleviate this problem, the effectiveness of LEMA is improved by expanding the ensemble used for the construction of the analysis. Instead of constructing the analysis mosaic using only the background ensemble, the number of states available to LEMA is augmented by considering additional model states taken from different lagged-forecast ensembles and model times. Using idealized and real DA experiments (assimilating actual StageIV precipitation observations), it is shown that the additional information in the expanded ensemble improves the effectiveness of LEMA, which translates to a better forecast quality in precipitation and the state variables. An interesting aspect of the expanded ensemble concept is that it can be applied to other data assimilation methods relying on the information contained in the ensemble.

The exploration carried out in Chapters 2 and 3 showed the LEMA may have potential for real DA applications. However, that exploration was done in the context of a single DA step and by

initializing the forecast by relaxing all background ensemble members towards the analysis mosaic (Frankenstate). This initialization procedure severely reduces the ensemble spread, producing an underdispersive ensemble forecast that does not represent the forecast uncertainties. To avoid the collapse of the ensemble, Chapter 4 presents a new version of LEMA, called Localized Ensemble Mosaic Assimilation Sequence (LEMAS), that constructs an ensemble of Frankenstates using a small group of members that is close to the observed precipitation. This analysis ensemble is used to translate the background ensemble towards the ensemble mean of the analysis mosaics and adjust the magnitude of the background perturbations towards the spread of the analysis ensemble. The new LEMAS methodology was evaluated and compared against a LETKF approach, using idealized and real DA experiments in assimilating precipitation observations. These experiments shows that LEMAS produces a comparable or better forecast quality than the forecast produced by the LETKF.

5.2 Ideas for future work

This thesis explored the potential of LEMA and LEMAS using experiments in assimilating precipitation observations in isolation from other types of observations. A natural direction for future work is to explore blending strategies to combine the LEMAS approach with conventional DA methods that assimilate other available observations (like aircraft, satellite, soundings, among others). In the Conclusions section of Chapter 4, two potential blending strategies were discussed. One strategy is to apply LEMAS to correct large errors on the background ensemble, and then use the LEMAS ensemble as background for the assimilation of observations other than precipitation. The other proposed alternative is to construct an analysis ensemble using a weighted combination of the ensemble obtained by LEMAS and the one obtained using a conventional DA method. Finally, another blending strategy could be using a two-step procedure similar to the "1D+4D-Var"

approach proposed by Marécal and Mahfouf (2002, 2003); Lopez and Bauer (2007). In the "1D+4D-Var" method, a one-dimensional variational method is used to retrieve the vertical profiles of temperature and moisture from the observed rainfall. Then the vertical profiles are assimilated as "pseudo-observation" using a 4D-Var system with the rest of the observations. This two-step procedure can be adapted to combine LEMAS and conventional DA methods by using the ensemble mean of the analysis mosaics obtained by LEMAS as a "pseudo-observation", that is then assimilated by the conventional DA method with the rest of the observations. For this approach, the spread of the LEMAS ensemble (representing the uncertainties in the analysis) can be used to model the pseudo-observation error.

In addition, the heuristic DA approach presented here can be extended to other types of observations with complex non-linear observations operators, like for example, satellite microwave observations. Some of the microwave radiances observations are highly sensitive to the clouds and precipitation, that translates into highly non-linear relationship between the observed radiances and the model state variables. These non-linearities result in significant deviations from the Gaussianity assumptions used in conventional DA methods, making their assimilation challenging (e.g. Geer et al., 2018; Migliorini and Candy, 2019). Therefore, it would be interesting to explore the potential benefits of using a non-Gaussian approach like LEMAS for this observation type. To that end, extending LEMAS to satellite observations would require a precise radiative transfer model (observation operator) that relates the model state to the observations, and a suitable metric to measure the distance to the observations.

THE END

Bibliography

Abbe, C., 1901: THE PHYSICAL BASIS OF LONG-RANGE WEATHER FORECASTS. *Monthly Weather Review*, **29 (12)**, 551–561, doi:10.1175/1520-0493(1901)29[551c:TPBOLW]2.0.CO;2.

Anthes, R. A., 1974: Data Assimilation and Initialization of Hurricane Prediction Models. *Journal of the Atmospheric Sciences*, **31 (3)**, 702–719, doi:10.1175/1520-0469(1974)031<0702:DAAIOH>2.0.CO;2.

Baldwin, M., and K. Mitchell, 1997: The NCEP hourly multisensor U.S. precipitation analysis. *Preprints, 11th Conf. on Numerical Weather Prediction*, Norfolk, VA, Amer. Meteor. Soc.

Bannister, R. N., 2017: A review of operational methods of variational and ensemble-variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, **143 (703)**, 607–633, doi:10.1002/qj.2982.

Barnes, S. L., 1964: A Technique for Maximizing Details in Numerical Weather Map Analysis. *Journal of Applied Meteorology*, **3 (4)**, 396–409, doi:10.1175/1520-0450(1964)003<0396:ATFMDI>2.0.CO;2.

Bauer, P., G. Ohring, C. Kummerow, and T. Auligne, 2011: Assimilating Satellite Observations of Clouds and Precipitation into NWP Models. *Bulletin of the American Meteorological Society*, **92 (6)**, ES25–ES28, doi:10.1175/2011BAMS3182.1.

Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525 (7567)**, 47–55.

Berenguer, M., and I. Zawadzki, 2008: A Study of the Error Covariance Matrix of Radar Rainfall Estimates in Stratiform Rain. *Weather and Forecasting*, **23 (6)**, 1085–1101, doi:10.1175/2008WAF2222134.1.

Berenguer, M., and I. Zawadzki, 2009: A Study of the Error Covariance Matrix of Radar Rainfall Estimates in Stratiform Rain. Part II: Scale Dependence. *Weather and Forecasting*, **24 (3)**, 800–811, doi:10.1175/2008WAF2222210.1.

Bergthórsson, P., and B. R. Döös, 1955: Numerical weather map analysis1. *Tellus*, **7 (3)**, 329–340, doi:10.1111/j.2153-3490.1955.tb01170.x.

Berner, J., K. R. Fossell, S.-Y. Ha, J. P. Hacker, and C. Snyder, 2015: Increasing the Skill of Probabilistic Forecasts: Understanding Performance Improvements from Model-Error Representations. *Monthly Weather Review*, **143 (4)**, 1295–1320, doi:10.1175/MWR-D-14-00091.1.

Bick, T., and Coauthors, 2016: Assimilation of 3D radar reflectivities with an ensemble Kalman filter on the convective scale. *Quarterly Journal of the Royal Meteorological Society*, **142 (696)**, 1490–1504, doi:10.1002/qj.2751.

Bjerknes, V., 1904: Das Problem der Wettervorhers-age, betrachtet vom Standpunkte der Mechanik und der Physik (The problem of weather prediction, considered from the viewpoints of mechanics and physics). *Meteor. Z.*, **21**, 1–7, translated and edited by Volken E. and S. Bronnimann. Meteorol. Z. 18 (2009), 663–667.

Bonavita, M., M. Hamrud, and L. Isaksen, 2015: EnKF and Hybrid Gain Ensemble Data Assimilation. Part II: EnKF and Hybrid Gain Results. *Monthly Weather Review*, **143 (12)**, 4865–4882, doi:10.1175/MWR-D-15-0071.1.

Bonavita, M., L. Torrisi, and F. Marcucci, 2010: Ensemble data assimilation with the CNMCA regional forecasting system. *Quarterly Journal of the Royal Meteorological Society*, **136 (646)**, 132–145, doi:10.1002/qj.553.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010: Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter for Global Deterministic NWP. Part I: Description and Single-Observation Experiments. *Monthly Weather Review*, **138 (5)**, 1550–1566, doi:10.1175/2009MWR3157.1.

Buehner, M., and D. Jacques, 2020: Non-Gaussian Deterministic Assimilation of Radar-Derived Precipitation Accumulations. *Monthly Weather Review*, **148 (2)**, 783–808, doi:10.1175/MWR-D-19-0199.1.

Buizza, R., M. Leutbecher, and L. Isaksen, 2008: Potential use of an ensemble of analyses in the ECMWF Ensemble Prediction System. *Quarterly Journal of the Royal Meteorological Society*, **134 (637)**, 2051–2066, doi:10.1002/qj.346.

Charney, J. G., R. Fjörtoft, and J. Von Neumann, 1950: Numerical Integration of the Barotropic Vorticity Equation. *Tellus*, **2 (4)**, 237–254, doi:10.1111/j.2153-3490.1950.tb00336.x.

Clark, A. J., and Coauthors, 2011: Probabilistic Precipitation Forecast Skill as a Function of Ensemble Size and Spatial Scale in a Convection-Allowing Ensemble. *Monthly Weather Review*, **139 (5)**, 1410–1418, doi:10.1175/2010MWR3624.1.

Clayton, A. M., A. C. Lorenc, and D. M. Barker, 2013: Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Quarterly Journal of the Royal Meteorological Society*, **139 (675)**, 1445–1461, doi:10.1002/qj.2054.

Cressman, G. P., 1959: AN OPERATIONAL OBJECTIVE ANALYSIS SYSTEM. *Monthly Weather Review*, **87 (10)**, 367–374, doi:10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2.

Crum, T. D., and R. L. Alberty, 1993: The WSR-88D and the WSR-88D Operational Support Facility. *Bulletin of the American Meteorological Society*, **74 (9)**, 1669–1688, doi:10.1175/1520-0477(1993)074<1669:TWATWO>2.0.CO;2.

Dalcher, A., E. Kalnay, and R. N. Hoffman, 1988: Medium Range Lagged Average Forecasts. *Monthly Weather Review*, **116 (2)**, 402–416, doi:10.1175/1520-0493(1988)116<0402:MRLAF>2.0.CO;2.

Daley, R., 1991: *Atmospheric data analysis*. Cambridge atmospheric and space science series ; 2, Cambridge University Press, Cambridge ;.

Davolio, S., and A. Buzzi, 2004: A Nudging Scheme for the Assimilation of Precipitation Data into a Mesoscale Model. *Weather and Forecasting*, **19 (5)**, 855–871, doi:10.1175/1520-0434(2004)019<0855:ANSFTA>2.0.CO;2.

Davolio, S., F. Silvestro, and T. Gastaldo, 2017: Impact of Rainfall Assimilation on High-Resolution Hydrometeorological Forecasts over Liguria, Italy. *Journal of Hydrometeorology*, **18 (10)**, 2659–2680, doi:10.1175/JHM-D-17-0073.1.

De Maesschalck, R., D. Jouan-Rimbaud, and D. L. Massart, 2000: The mahalanobis distance. *Chemometrics and intelligent laboratory systems*, **50 (1)**, 1–18.

Denis, B., J. Côté, and R. Laprise, 2002: Spectral Decomposition of Two-Dimensional Atmospheric Fields on Limited-Area Domains Using the Discrete Cosine Transform (DCT). *Monthly Weather Review*, **130 (7)**, 1812–1829, doi:10.1175/1520-0493(2002)130<1812:SDOTDA>2.0.CO;2.

Donner, L. J., 1988: An Initialization for Cumulus Convection in Numerical Weather Prediction Models. *Monthly Weather Review*, **116 (2)**, 377–385, doi:10.1175/1520-0493(1988)116<0377:AIFCCI>2.0.CO;2.

Dudhia, J., 1989: Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model. *Journal of the Atmospheric Sciences*, **46 (20)**, 3077–3107, doi:10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.

Durran, D. R., P. A. Reinecke, and J. D. Doyle, 2013: Large-Scale Errors and Mesoscale Predictability in Pacific Northwest Snowstorms. *Journal of the Atmospheric Sciences*, **70 (5)**, 1470–1487, doi: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. *Journal of the Atmospheric Sciences*, **64 (11)**, 3785–3798, doi:10.1175/2006JAS2044.1.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, **99 (C5)**, 10 143–10 162, doi:10.1029/94JC00572.

Falkovich, A., E. Kalnay, S. Lord, and M. B. Mathur, 2000: A new method of observed rainfall assimilation in forecast models. *Journal of applied meteorology*, **39 (8)**, 1282–1298, doi: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? *Journal of Hydrometeorology*, **15 (4)**, 1708–1713, doi:10.1175/JHM-D-14-0008.1.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, **125 (554)**, 723–757, doi:10.1002/qj.49712555417.

Geer, A. J., and Coauthors, 2018: All-sky satellite data assimilation at operational weather forecasting centres. *Quarterly Journal of the Royal Meteorological Society*, **144 (713)**, 1191–1217, doi:10.1002/qj.3202.

Germann, U., and I. Zawadzki, 2002: Scale-Dependence of the Predictability of Precipitation from Continental Radar Images. Part I: Description of the Methodology. *Monthly Weather Review*, **130 (12)**, 2859–2873, doi:10.1175/1520-0493(2002)130<2859:SDOTPO>2.0.CO;2.

Gilchrist, B., and G. P. Cressman, 1954: An experiment in objective analysis. *Tellus*, **6 (4)**, 309–318, doi:10.1111/j.2153-3490.1954.tb01126.x.

Haiden, T., M. Janousek, F. Vitart, L. Ferranti, and F. Prates, 2019: Evaluation of ECMWF forecasts, including the 2019 upgrade. Tech. Rep. 853. doi:10.21957/mlvapkke, URL https://www.ecmwf.int/node/19277.

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.

Hamill, T. M., and C. Snyder, 2000: A Hybrid Ensemble Kalman Filter–3D Variational Analysis Scheme. *Monthly Weather Review*, **128 (8)**, 2905–2919, doi:10.1175/1520-0493(2000)128<2905:AHEKFV>2.0.CO;2.

Heckley, W. A., G. Kelly, and M. Tiedtke, 1990: On the Use of Satellite-Derived Heating Rates for Data Assimilation within the Tropics. *Monthly Weather Review*, **118 (9)**, 1743–1757, doi:10.1175/1520-0493(1990)118<1743:OTUOSD>2.0.CO;2.

Hoffman, R. N., and E. Kalnay, 1983: Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus A*, **35A (2)**, 100–118, doi:10.1111/j.1600-0870.1983.tb00189.x.

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.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes. *Monthly Weather Review*, **134 (9)**, 2318–2341, doi:10.1175/MWR3199.1.

Houtekamer, P. L., X. Deng, H. L. Mitchell, S.-J. Baek, and N. Gagnon, 2014: Higher Resolution in an Operational Ensemble Kalman Filter. *Monthly Weather Review*, **142 (3)**, 1143–1162, doi:10.1175/MWR-D-13-00138.1.

Houtekamer, P. L., and H. L. Mitchell, 2005: Ensemble Kalman filtering. *Quarterly Journal of the Royal Meteorological Society*, **131 (613)**, 3269–3289, doi:10.1256/qj.05.135.

Houtekamer, P. L., and F. Zhang, 2016: Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation. *Monthly Weather Review*, **144 (12)**, 4489–4532, doi:10.1175/MWR-D-15-0440.1.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, **230 (1)**, 112 – 126, doi:https://doi.org/10.1016/j.physd.2006.11.008.

Huuskonen, A., E. Saltikoff, and I. Holleman, 2014: The Operational Weather Radar Network in Europe. *Bulletin of the American Meteorological Society*, **95 (6)**, 897–907, doi:10.1175/BAMS-D-12-00216.1.

Jacques, D., D. Michelson, J.-F. Caron, and L. Fillion, 2018: Latent Heat Nudging in the Canadian Regional Deterministic Prediction System. *Monthly Weather Review*, **146 (12)**, 3995–4014, doi: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. *Monthly Weather Review*, **142 (3)**, 1053–1073, doi: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. *Meteorological Applications*, **4 (3)**, 269–277, doi:10.1017/S1350482797000522.

Joss, J., A. Waldvogel, and C. G. Collier, 1990: *Precipitation Measurement and Hydrology*, 577–606. American Meteorological Society, Boston, MA, doi:10.1007/978-1-935704-15-7_39.

Kain, J. S., 2004: The Kain-Fritsch Convective Parameterization: An Update. *Journal of Applied Meteorology*, **43 (1)**, 170–181, doi:10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.

Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, doi:10.1017/CBO9780511802270.

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, doi: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). *Journal of Geophysical Research: Atmospheres*, **122 (2)**, 631–650, doi:10.1002/2016JD025355, 2016JD025355.

Krishnamurti, T., K. Ingles, S. Cocke, T. Kitade, and R. Pasch, 1984: Details of Low Latitude Medium Range Numerical Weather Prediction Using a Global Spectral Model. *Journal of the Meteorological Society of Japan. Ser. II*, **62 (4)**, 613–649, doi:10.2151/jmsj1965.62.4_613.

Krishnamurti, T. N., H. S. Bedi, W. Heckley, and K. Ingles, 1988: Reduction of the Spinup Time for Evaporation and Precipitation in a Spectral Model. *Monthly Weather Review*, **116 (4)**, 907–920, doi:10.1175/1520-0493(1988)116<0907:ROTSTF>2.0.CO;2.

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. *Journal of Geophysical Research: Atmospheres*, **119 (5)**, 2017–2031, doi:10.1002/2013JD020005.

Lee, G. W., A. W. Seed, and I. Zawadzki, 2007: Modeling the Variability of Drop Size Distributions in Space and Time. *Journal of Applied Meteorology and Climatology*, **46 (6)**, 742–756, doi:10.1175/JAM2505.1.

Lien, G.-Y., E. Kalnay, and T. Miyoshi, 2013: Effective assimilation of global precipitation: simulation experiments. *Tellus A: Dynamic Meteorology and Oceanography*, **65 (1)**, 19 915, doi:10.3402/tellusa.v65i0.19915.

Lien, G.-Y., T. Miyoshi, and E. Kalnay, 2016: Assimilation of TRMM Multisatellite Precipitation Analysis with a Low-Resolution NCEP Global Forecast System. *Monthly Weather Review*, **144 (2)**, 643–661, doi:10.1175/MWR-D-15-0149.1.

Lin, Y., and K. E. Mitchell, 2005: The NCEP stage II/IV hourly precipitation analyses: Development and applications. *19th Conf. Hydrology*, San Diego, CA, Amer. Meteor. Soc., 1.2, [Available online at https://ams.confex.com/ams/Annual2005/techprogram/paper_83847.htm.].

Liu, C., Q. Xiao, and B. Wang, 2008: An Ensemble-Based Four-Dimensional Variational Data Assimilation Scheme. Part I: Technical Formulation and Preliminary Test. *Monthly Weather Review*, **136 (9)**, 3363–3373, doi:10.1175/2008MWR2312.1.

Lopez, P., 2011: Direct 4D-Var Assimilation of NCEP Stage IV Radar and Gauge Precipitation Data at ECMWF. *Monthly Weather Review*, **139 (7)**, 2098–2116, doi:10.1175/2010MWR3565.1.

Lopez, P., and P. Bauer, 2007: "1D+4DVAR" Assimilation of NCEP Stage-IV Radar and Gauge Hourly Precipitation Data at ECMWF. *Monthly Weather Review*, **135 (7)**, 2506–2524, doi:10.1175/MWR3409.1.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, **112 (474)**, 1177–1194, doi:10.1002/qj.49711247414.

Lu, C., H. Yuan, B. E. Schwartz, and S. G. Benjamin, 2007: Short-Range Numerical Weather Prediction Using Time-Lagged Ensembles. *Weather and Forecasting*, **22 (3)**, 580–595, doi:10.1175/WAF999.1.

Macpherson, B., 2001: Operational experience with assimilation of rainfall datain the met office mesoscale model. *Meteorology and Atmospheric Physics*, **76 (1-2)**, 3–8, doi:10.1007/s007030170035.

Manobianco, J., S. Koch, V. M. Karyampudi, and A. J. Negri, 1994: The Impact of Assimilating Satellite-Derived Precipitation Rates on Numerical Simulations of the ERICA IOP 4 Cyclone. *Monthly Weather Review*, **122 (2)**, 341–365, doi:10.1175/1520-0493(1994)122<0341:TIOASD>2.0.CO;2.

Marécal, V., and J.-F. Mahfouf, 2002: Four-Dimensional Variational Assimilation of Total Column Water Vapor in Rainy Areas. *Monthly Weather Review*, **130 (1)**, 43–58, doi:10.1175/1520-0493(2002)130<0043:FDVAOT>2.0.CO;2.

Marécal, V., and J.-F. Mahfouf, 2003: Experiments on 4D-Var assimilation of rainfall data using an incremental formulation. *Quarterly Journal of the Royal Meteorological Society*, **129 (594)**, 3137–3160, doi:10.1256/qj.02.120.

Migliorini, S., and B. Candy, 2019: All-sky satellite data assimilation of microwave temperature sounding channels at the met office. *Quarterly Journal of the Royal Meteorological Society*, **145 (719)**, 867–883, doi:10.1002/qj.3470.

Min, C., S. Chen, J. J. Gourley, H. Chen, A. Zhang, Y. Huang, and C. Huang, 2019: Coverage of China New Generation Weather Radar Network. *Advances in Meteorology*, **2019**, doi:10.1155/2019/5789358.

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. *Journal of Geophysical Research: Atmospheres*, **102 (D14)**, 16 663–16 682, doi:10.1029/97JD00237.

Orlanski, I., 1975: A Rational Subdivision of Scales for Atmospheric Processes. *Bulletin of the American Meteorological Society*, **56 (5)**, 527–530.

Panofsky, R. A., 1949: OBJECTIVE WEATHER-MAP ANALYSIS. *Journal of Meteorology*, **6 (6)**, 386–392, doi:10.1175/1520-0469(1949)006<0386:OWMA>2.0.CO;2.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. *Monthly Weather Review*, **120 (8)**, 1747–1763, doi:10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2.

Peng, S. Q., and X. Zou, 2002: Assimilation of NCEP multi-sensor hourly rainfall data using 4D-Var approach: A case study of the squall line on April 5, 1999. *Meteorology and Atmospheric Physics*, **81 (3)**, 237–255, doi:10.1007/s00703-002-0545-y.

Penny, S. G., 2014: The Hybrid Local Ensemble Transform Kalman Filter. *Monthly Weather Review*, **142 (6)**, 2139–2149, doi:10.1175/MWR-D-13-00131.1.

Pérez Hortal, A. A., I. Zawadzki, and M. K. Yau, 2019: A Heuristic Approach for Precipitation Data Assimilation: Characterization Using OSSEs. *Monthly Weather Review*, **147 (9)**, 3445–3466, doi:10.1175/MWR-D-19-0034.1.

Pérez Hortal, A. A., I. Zawadzki, and M. K. Yau, 2020: A Heuristic Approach for Precipitation Data Assimilation: Effect of Forecast Errors and Assimilation of NCEP Stage IV Precipitation Analyses. *Monthly Weather Review*, **148 (4)**, 1629–1651, doi:10.1175/MWR-D-19-0331.1.

Poterjoy, J., 2016: A Localized Particle Filter for High-Dimensional Nonlinear Systems. *Monthly Weather Review*, **144 (1)**, 59–76, doi:10.1175/MWR-D-15-0163.1.

Poterjoy, J., R. A. Sobash, and J. L. Anderson, 2017: Convective-Scale Data Assimilation for the Weather Research and Forecasting Model Using the Local Particle Filter. *Monthly Weather Review*, **145 (5)**, 1897–1918, doi:10.1175/MWR-D-16-0298.1.

Poterjoy, J., L. Wicker, and M. Buehner, 2019: Progress toward the Application of a Localized Particle Filter for Numerical Weather Prediction. *Monthly Weather Review*, **147 (4)**, 1107–1126, doi:10.1175/MWR-D-17-0344.1.

Rabier, F., H. Järvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Quarterly Journal of the Royal Meteorological Society*, **126 (564)**, 1143–1170, doi:10.1002/qj.49712656415.

Rawlins, F., S. P. Ballard, K. J. Bovis, A. M. Clayton, D. Li, G. W. Inverarity, A. C. Lorenc, and T. J. Payne, 2007: The Met Office global four-dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, **133 (623)**, 347–362, doi:10.1002/qj.32.

Richardson, L. F., 1922: *Weather prediction by numerical process*. Cambridge University Press.

Schraff, C., H. Reich, A. Rhodin, A. Schomburg, K. Stephan, A. Periáñez, and R. Potthast, 2016: Kilometre-scale ensemble data assimilation for the COSMO model (KENDA). *Quarterly Journal of the Royal Meteorological Society*, **142 (696)**, 1453–1472, doi:10.1002/qj.2748.

Simonin, D., C. Pierce, N. Roberts, S. P. Ballard, and Z. Li, 2017: Performance of Met Office hourly cycling NWP-based nowcasting for precipitation forecasts. *Quarterly Journal of the Royal Meteorological Society*, **143 (708)**, 2862–2873, doi:10.1002/qj.3136.

Skamarock, W. C., and J. B. Klemp, 2008: A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *Journal of Computational Physics*, **227 (7)**, 3465–3485, doi:10.1016/j.jcp.2007.01.037.

Snyder, C., T. Bengtsson, P. Bickel, and J. Anderson, 2008: Obstacles to High-Dimensional Particle Filtering. *Monthly Weather Review*, **136 (12)**, 4629–4640, doi: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. *Monthly Weather Review*, **118 (6)**, 1250–1277, doi: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. *Quarterly Journal of the Royal Meteorological Society*, **134 (634)**, 1315–1326, doi:10.1002/qj.269.

Surcel, M., I. Zawadzki, and M. K. Yau, 2015: A Study on the Scale Dependence of the Predictability of Precipitation Patterns. *Journal of the Atmospheric Sciences*, **72 (1)**, 216–235, doi:10.1175/JAS-D-14-0071.1.

Tanguay, M., L. Fillion, E. Lapalme, and M. Lajoie, 2012: Four-Dimensional Variational Data Assimilation for the Canadian Regional Deterministic Prediction System. *Monthly Weather Review*, **140 (5)**, 1517–1538, doi:10.1175/MWR-D-11-00160.1.

Tsuyuki, T., 1997: Variational Data Assimilation in the Tropics Using Precipitation Data. Part III: Assimilation of SSM/I Precipitation Rates. *Monthly Weather Review*, **125 (7)**, 1447–1464, doi:10.1175/1520-0493(1997)125<1447:VDAITT>2.0.CO;2.

Turpeinen, O. M., L. Garand, R. Benoit, and M. Roch, 1990: Diabatic initialization of the Canadian Regional Finite-Element (RFE) Model Using Satellite Data. Part I: Methodology and Application to a Winter Storm. *Monthly Weather Review*, **118 (7)**, 1381–1395, doi:10.1175/1520-0493(1990)118<1381:DIOTCR>2.0.CO;2.

Van Den Dool, H. M., and L. Rukhovets, 1994: On the Weights for an Ensemble-Averaged 6–10-Day Forecast. *Weather and Forecasting*, **9 (3)**, 457–465, doi:10.1175/1520-0434(1994)009<0457:OTWFAE>2.0.CO;2.

van Leeuwen, P. J., 2009: Particle Filtering in Geophysical Systems. *Monthly Weather Review*, **137 (12)**, 4089–4114, doi: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. *Monthly Weather Review*, **139 (2)**, 403–423, doi:10.1175/2010MWR3487.1.

Wang, W., and T. T. Warner, 1988: Use of Four-Dimensional Data Assimilation by Newtonian Relaxation and Latent-Heat Forcing to Improve a Mesoscale-Model Precipitation Forecast: A Case Study. *Monthly Weather Review*, **116 (12)**, 2593–2613, doi:10.1175/1520-0493(1988)116<2593:UOFDDA>2.0.CO;2.

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 A: Dynamic Meteorology and Oceanography*, **60 (1)**, 62–79, doi:10.1111/j.1600-0870.2007.00273.x.

Wilks, D. S., 1995: *Statistical methods in the Atmospheric Sciences*. Academic Press, 467pp pp.

Zawadzki, I., 1984: Factors affecting the precision of radar measurement of rain. *Preprints, 22nd Int. Conference on radar meteorology*, Zurich, Switzerland,, Amer. Meteor. Soc., 251–256.

Zhang, F., C. Snyder, and J. Sun, 2004: Impacts of Initial Estimate and Observation Availability on Convective-Scale Data Assimilation with an Ensemble Kalman Filter. *Monthly Weather Review*, **132 (5)**, 1238–1253, doi:10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2.

Zhou, X., Y. Zhu, D. Hou, Y. Luo, J. Peng, and R. Wobus, 2017: Performance of the New NCEP Global Ensemble Forecast System in a Parallel Experiment. *Weather and Forecasting*, **32 (5)**, 1989–2004, doi:10.1175/WAF-D-17-0023.1.

Županski, D., and F. Mesinger, 1995: Four-Dimensional Variational Assimilation of Precipitation Data. *Monthly Weather Review*, **123 (4)**, 1112–1127, doi:10.1175/1520-0493(1995)123<1112:FDVAOP>2.0.CO;2.