
HONGLI WANG, JUANZHEN SUN, XIN ZHANG, XIANG-YU HUANG, AND THOMAS AULIGNÉ

National Center for Atmospheric Research,* Boulder, Colorado

(Manuscript received 8 June 2012, in final form 21 December 2012)

ABSTRACT

The major goal of this two-part study is to assimilate radar data into the high-resolution Advanced Research Weather Research and Forecasting Model (ARW-WRF) for the improvement of short-term quantitative precipitation forecasting (QPF) using a four-dimensional variational data assimilation (4D-Var) technique. In Part I the development of a radar data assimilation scheme within the WRF 4D-Var system (WRF 4D-Var) and the preliminary testing of the scheme are described. In Part II the performance of the enhanced WRF 4D-Var system is examined by comparing it with the three-dimensional variational data assimilation system (WRF 3D-Var) for a convective system over the U.S. Great Plains. The WRF 4D-Var radar data assimilation system has been developed with the existing framework of an incremental formulation. The new development for radar data assimilation includes the tangent-linear and adjoint models of a Kessler warm-rain microphysics scheme and the new control variables of cloud water, rainwater, and vertical velocity and their error statistics. An ensemble forecast with 80 members is used to produce background error covariance. The preliminary testing presented in this paper includes single-observation experiments as well as real data assimilation experiments on a squall line with assimilation windows of 5, 15, and 30 min. The results indicate that the system is able to obtain anisotropic multivariate analyses at the convective scale and improve precipitation forecasts. The results also suggest that the incremental approach with successive basic-state updates works well at the convection-permitting scale for radar data assimilation with the selected assimilation windows.

1. Introduction

Four-dimensional variational data assimilation (4D-Var) is an advanced technique to optimally estimate the initial conditions of an atmospheric model based on observations, first-guess background, and a constraining numerical model. Since its introduction into meteorology in the late 1980s using simple models (e.g., Lewis and Derber 1985; Le Dimet and Talagrand 1986; Talagrand and Courtier 1987; Courtier and Talagrand 1987, 1990), several studies have been conducted using complex numerical weather prediction (NWP) models in the 1990s (Navon et al. 1992; Zupanski 1993; Rabier et al. 1993; Zou et al. 1995; Zupanski 1997; Rabier et al. 1998). The promising results obtained from these studies led to the first operational implementation of the 4D-Var technique at the European Centre for Medium-Range Weather Forecasts (ECMWF; Rabier et al. 2000; Mahfouf et al. 2000; Klinker et al. 2000). Following the success of ECMWF, other 4D-Var systems based on operational NWP models were developed and implemented at several operational centers throughout the world [see Rabier (2005) for a review] as their data assimilation systems for global or regional NWP.

The 4D-Var technique was also applied to convective scales for the initialization of high-resolution models by assimilating Doppler radar observations. Wolfsberg (1987), Sun et al. (1991), and Kapitza (1991) were the first to test the technique on the single-Doppler retrieval problem using high-resolution boundary layer models. Sun and Crook (1997, 1998) further developed the 4D-Var technique of Sun et al. (1991) and demonstrated that the state variables in a cloud model with warm-rain microphysics parameterization could be initialized using
a 4D-Var assimilation system, named the Variational Doppler Radar Analysis System (VDRAS), that assimilates single-Doppler radar observations. The system was applied to the initialization and simulation of two observed convective systems using multiple Doppler radars that had only single-Doppler coverage, and promising results were obtained (Sun 2005; Sun and Zhang 2008). VDRAS has been run in real time in several field demonstration programs to produce high-resolution frequent (12–18 min) analyses (Sun and Crook 2001; Crook and Sun 2002), including the recent implementation in Beijing 2008 Field Demonstration Project (B08FDP) organized by the World Weather Research Program (WWRP) in support of severe weather nowcasting during the 2008 summer Olympics (Sun et al. 2010).

Although VDRAS demonstrated good potential of the 4D-Var technique for convective-scale applications, the system has its limitations because of the simplified cloud model that does not include terrain effects and complex physical processes. The ability of the 4D-Var method in convective-scale assimilation and forecasting has yet to be proven using full-blown operational mesoscale models. Developing such a system, although desirable, is quite challenging. The effort required for the development of the adjoint model (ADM) of a complex NWP model is not trivial. Furthermore, the nonlinearity and complexity associated with microphysics and other physics schemes pose an additional challenge. Therefore, it is necessary to seek approaches that can simplify the 4D-Var method when it is used for an operational NWP model.

Courtier et al. (1994) proposed the “incremental approach,” as a strategy for implementing 4D-Var with operational models. This approach uses a tangent linear model (TLM) instead of the full nonlinear model as the forward constraining model in the 4D-Var cost function minimization. The nonlinear basic-state trajectory is updated successively outside the minimization iteration loop whenever the convergence with respect to the TLM is achieved. This approach not only reduces the computational cost of 4D-Var but also improves the conditioning of the cost function because of the linearization of the forward operator. Common practices in operational implementation of incremental 4D-Var are to use a coarser-resolution TLM and ADM to reduce the computational cost, and to allow a simplified representation of physical processes in TLM and ADM within the data assimilation window. The incremental approach is adopted by a number of three-dimensional variational data assimilation (3D-Var) or 4D-Var systems in major operational centers, including the National Centers for Environmental Prediction (NCEP) gridpoint statistical interpolation (GSI) 3D-Var system (Parrish and Derber 1992; Wu et al. 2002), the Met Office Unified Model (MetUM) 3D-Var (Lorenc et al. 2000) and 4D-Var (Rawlins et al. 2007) systems, the Advanced Research Weather Research and Forecasting Model (ARW-WRF; hereafter referred to as WRF) 3D-Var and 4D-Var systems (Barker et al. 2004, 2012; Huang et al. 2009), the Canadian Regional Data Assimilation system (Fillion et al. 2010; Tanguay et al. 2012), and the High-Resolution Limited-Area Model (HIRLAM) 4D-Var system (Gustafsson et al. 2012).

Another simplified 4D-Var approach for operational implementation is employed by the Japan Meteorological Agency (JMA), in which JMA’s full nonlinear nonhydrostatic model (JMA-NHM) is used as the forward-constraining model in 4D-Var, while the ADM is developed based on a simplified version of JMA-NHM (Honda et al. 2005). This approach was recently used by Kawabata et al. (2007, 2011) in two studies of convective-scale radar data assimilation.

The basic design of the WRF 3D-Var closely followed that of MetUM, using the incremental approach. WRF 4D-Var was developed by extending the 3D-Var to the time dimension without changing its basic framework. Huang et al. (2009) applied the prototype version of WRF 4D-Var to a hurricane and a cyclone case in which a simplified TLM of WRF was used as the forward model that includes the dry dynamics and large-scale precipitation. They demonstrated that 4D-Var outperformed 3D-Var for the two cases including Typhoon Haitang and a strong cyclone studied with a 6-h assimilation window. However, further development is required before it can be applied to high-resolution radar data assimilation. To assimilate radar reflectivity observations, a microphysics scheme must be included in the forward and adjoint models and microphysical variables must be considered as control variables. Because of the highly nonlinear nature of convective phenomena, a question that to be answered in this study is whether the incremental approach is suitable for convective-scale radar data assimilation.

Besides the progresses in 4D-Var, an ensemble Kalman filter (EnKF) approach has been extensively employed to assimilate radar observations for high resolution convective forecast in the past decade. The potential of the EnKF’s application to the convective scale was shown by assimilating simulated or real radial velocity and reflectivity observations (e.g., Snyder and Zhang 2003; Zhang et al. 2004; Dowell et al. 2004a, 2011; Tung and Xue 2005, 2008; Xue et al. 2006; Jung et al. 2008; Aksoy et al. 2009, 2010). EnKF does not require as much coding effort as 4D-Var and it has been shown to be an effective technique to initialize model prognostic
variables including hydrometeors even with a complex microphysical scheme. However, reliable short-term forecasts from analyses remain a challenge (Snyder and Zhang 2003; Dowell et al. 2004b; Tong and Xue 2005; Aksoy et al. 2010; Yussouf and Stensrud 2010).

The main objectives of the current study are to develop the capability of radar data assimilation for the WRF 4D-Var and examine the feasibility of the incremental 4D-Var for convective-scale radar data assimilation. Previous studies on meso- and convective-scale data assimilation using high-resolution models and 4D-Vars all employed full nonlinear models as forward models (Sun and Crook 1997; Zou et al. 1995; Gao et al. 1998; Kawabata et al. 2007, 2011). Kawabata et al. (2011) briefly discussed an experiment with the incremental approach with a 10-min window and two outer loops and showed controversial results. Recently, the Met Office has been running a prototype high-resolution 4D-Var with the incremental approach in real time since March 2012 (S. Ballard 2012, personal communication). The performance of that system is currently being assessed and it assimilates radar radial velocity but not reflectivity. In this paper, we present an extensive study on the feasibility of the incremental 4D-Var approach to radar observations especially for reflectivity data assimilation using a NWP model at the convection-permitting scale. In Part I, we describe the WRF 4D-Var system with the new developments pertinent to the convective-scale radar data assimilation, including the development of the tangent linear and adjoint models of a warm-rain microphysics, and the extension of new control variables of cloud water, rainwater, and vertical velocity, and their error statistics. Single reflectivityobservation tests of the updated 4D-Var system are carried out to show how the background covariance and model dynamics and physics spread the observation information. The performance of the updated 4D-Var is also assessed by a convective case that occurred in the Great Plains. In Sun and Wang (2013, hereafter Part II), the capability of the 4D-Var radar data assimilation in convective forecasting is examined in detail by comparing the WRF 4D-Var with WRF 3D-Var.

The structure of Part I is arranged as follows. Section 2 describes the WRF 4D-Var system with an emphasis on the radar data assimilation components. Section 3 presents results from single-observation assimilation experiments. The results from real radar data assimilation experiments are presented in section 4. A summary and discussion are provided in the final section.

2. WRF 4D-Var

Huang et al. (2009) described the prototype version of the WRF 4D-Var system and showed its performance on a typhoon and a cyclone analysis and forecast. Recently, the TLM and its ADM of that system were upgraded to match the updated WRF, version 3.3. These upgrades along with some other software enhancements were described in Zhang et al. (2013). To assimilate radar observations, the 4D-Var system is further upgraded through the current study. In this section, we focus on the description of the radar data assimilation components after a general description of the WRF 4D-Var system.

a. General description of WRF 4D-Var

In general terms, the purpose of a 4D-Var data assimilation system is to provide an analysis \(x^a\) of the initial atmospheric state \(x_0\) by minimizing a predefined cost function that includes a model constraint:

\[
J(x_0) = \frac{1}{2}(x_0 - x_0^b)^T B^{-1}(x_0 - x_0^b) + \frac{1}{2} \sum_{k=0}^{K} \{y_k^o - H_k[M_k(x_0)]\}^T R^{-1}\{y_k^o - H_k[M_k(x_0)]\}.
\]

(1)

The background \(x_0^b\) can be provided by a previous forecast. The observations \(y^o\) within a time window are divided into \(K\) discrete subwindows with each subwindow represented by the subscript \(k\); \(M_k\) is the nonlinear forward model used to propagate the initial atmospheric state \(x_0\) to that at the \(k\)th observation time, \(H_k\) is the nonlinear observation operator at time \(k\), and \(B\) and \(R\) are the background and observation error covariance matrices, respectively.

The incremental approach minimizes a cost function defined in terms of the analysis increment relative to a nonlinear basic state and the linearized observation and forward operators:

\[
J^{(n)}[\delta x^{(n)}_0] = \frac{1}{2}(\delta x_0^{(n)} - [x_0^b - x_0^{(n-1)}])^T B^{-1}(\delta x_0^{(n)} - [x_0^b - x_0^{(n-1)}]) + \frac{1}{2} \sum_{k=0}^{K} \{H_k M_k [x_0^{(n-1)}] \delta x_0^{(n)} - d_k^{(n-1)}\}^T R^{-1}\{H_k M_k [x_0^{(n-1)}] \delta x_0^{(n)} - d_k^{(n-1)}\},
\]

(2)
where \( n = 1, \ldots, N \) is the index of an outer loop iteration that successively updates the nonlinear basic state and \( \mathbf{H}_k \) and \( \mathbf{M}_k \) are the TLM of \( H_k \) and \( M_k \), respectively. The \( x_0^{(n-1)} \) is the estimate of the atmospheric state from the previous loop [it is equal to the background for the first outer loop \( n = 1 \)]. The innovation

\[
d_k^{(n-1)} = y_k - H_k[M_k(x_0^{(n-1)})] \tag{3}
\]

is the observation departure from its observed equivalent at time \( k \). The linearization assumption

\[
H_k[M_k(x_0^{(n)})] \approx H_k[M_k(x_0^{(n-1)})] + H_k[M_k(x_0^{(n-1)})]\delta x_0^{(n)} \tag{4}
\]

is applied to derive Eq. (2) from Eq. (1). When the increment \( \delta x_0^{(n)} \) at the \( n \)th outer loop is obtained, the estimate of the atmospheric state is updated by \( x_0^{(n)} = x_0^{(n-1)} + \delta x_0^{(n)} \) and used to produce the first-guess trajectory \( M_k[x_0^{(n)}] \) for the next outer loop \( n + 1 \). The basic assumption of the incremental approach is that the solution of the cost function [Eq. (2)] gradually approaches that of the nonlinear cost function [Eq. (1)] given enough outer-loop iterations.

Assuming the background error covariance matrix is given by \( \mathbf{B} = \mathbf{U} \mathbf{U}^T \) and a control variable transform by \( \delta x_0 = \mathbf{U} \mathbf{v} \), the background term of the cost function [Eq. (2)] is simplified to \( 1/2(\sum_{i=1}^{n} v_i^2)(\sum_{i=1}^{n} v_i) \), thereby avoiding the difficulty in computing the inverse of \( \mathbf{B} \). The control variable transform \( \delta x_0 = \mathbf{U} \mathbf{v} \) is implemented through a series of operations \( \mathbf{U} = \mathbf{U}_p \mathbf{U}_h \mathbf{U}_p \) (Barker et al. 2004). The horizontal transform \( \mathbf{U}_h \) is performed using a recursive filter (Hayden and Purser 1995). The vertical transform \( \mathbf{U}_p \) applies an empirical orthogonal function (EOF) decomposition on the vertical component of the background error covariance. The physical transform \( \mathbf{U}_p \) converts the increment in control variable space to analysis variable space. The control variables of the WRF 4D-Var system will be described later in this section.

b. Radar observation operators

The first step in assimilating radar radial velocity and reflectivity observations is to modify the cost function [Eq. (2)] such that it includes the observation terms for these two variables. To do so, the linearized observation operators for radial velocity and reflectivity [denoted by \( \mathbf{H}_k \) in Eq. (2)] are required. The radial velocity operator \( V_r \) is formulated with the 3D wind field \( (u, v, w) \), the hydrometeor fall speed \( V_r \), and the distance \( D \) between the location of a data point and the radar antenna:

\[
V_r = \frac{1}{D} [(x_d - x_r)u + (y_d - y_r)v + (z_d - z_r)(w - V_r)], \tag{5}
\]

\[
D = [(x_d - x_r)^2 + (y_d - y_r)^2 + (z_d - z_r)^2]^{1/2}, \tag{6}
\]

where \( (x_d, y_d, z_d) \) represents the location of the observation point and \( (x_r, y_r, z_r) \) represents the location of the radar station; \( V_r \) is calculated from the rainwater mixing ratio with a height correction following Sun and Crook (1997). The relation in Eq. (5) is linear except that \( V_r \) is nonlinearly dependent on \( q_r \), which needs to be linearized to obtain the linear observation operator for \( V_r \).

The reflectivity operator follows that of Sun and Crook (1997) and was also used by Xiao et al. (2007) for WRF 3D-Var. It is expressed as

\[
Z = c_1 + c_2 \log_{10}(\rho q_r), \tag{7}
\]

where \( Z \) is the reflectivity factor (dBZ); \( c_1 \) and \( c_2 \) are constants with the value of 43.1 and 17.5, respectively; \( \rho \) is the air density (kg m\(^{-3}\)); and \( q_r \) is the rainwater mixing ratio (g kg\(^{-1}\)). The linearized formulation of Eq. (7) is

\[
dZ = c_2 q_r \times a \log(10). \tag{8}
\]

Equation (8) was used as the observation operator \( \mathbf{H}_k \) to assimilate reflectivity in the WRF 3D-Var developed by Xiao et al. (2007). However, Wang et al. (2013) showed that the linearized equation [Eq. (8)] can result in a dry bias in rainwater analysis. Thus they proposed to indirectly assimilate the derived rainwater mixing ratio \( q_r \) from Eq. (7) to avoid the problems. This approach was also used in Sun and Crook (1997). In this study, we assimilate the derived rainwater from radar reflectivity following Wang et al. (2013).

c. Tangent linear and adjoint models of the Kessler microphysics scheme

The assimilation of reflectivity requires an extension of the TLM and ADM to include a microphysics scheme. In this study, we use the standard Kessler warm-rain scheme in WRF (Skamarock et al. 2008) because it is relatively simple yet captures the major physical processes for warm-season convection. Wang et al. (2012) showed that a warm-rain liquid-only microphysics scheme can be used as an acceptable substitute for the more complex one as long as a short time window (less than 1 h) is used. The WRF Kessler microphysics scheme is composed of the three microphysical variables of water vapor, cloud water, and rainwater. The microphysical processes included are the production, fall, and evaporation of rain, the accretion and autoconversion of cloud water, and the production of cloud water from condensation.
The standard Kessler scheme was modified before the TLM and ADM were developed. Sun and Crook (1997) found that the high degree of nonlinearity associated with microphysics could cause serious problems for the minimization if no modification was made. Following Sun and Crook (1997), large derivatives of the rainwater evaporation and the rainwater fall speed with respect to rainwater mixing ratio were eliminated by setting a threshold for rainwater mixing ratio. Simulations were performed to confirm that the modifications of the Kessler scheme had a negligible effect.

The development of the TLM and ADM were done with the aid of the automatic differentiation tool, Tapenade (Hascoet and Pascual 2004). Tapenade was developed and maintained by the researchers of the tropics Project at the Institut National de Recherche en Informatique et Automatique (INRIA) in France. We used Tapenade to generate the initial FORTRAN codes of the TLM and ADM. A manual process was then followed to improve the codes and to ensure their correctness. The derivation of the adjoint of the physical processes with on/off switches follows that of Zou et al. (1993) by keeping the switching times the same as in the basic state.

The minimization of the incremental 4D-Var is very sensitive to the accuracy of TLM and ADM. Therefore, a necessary and important step in the development of a 4D-Var system is to check the correctness of TLM and ADM. The WRF 4D-Var tangent linear and adjoint check procedure (Zhang et al. 2013) follows Navon et al. (1992). Readers are referred to Zhang et al. (2013) for more details about WRF TLM and ADJ checks. To verify the correctness of the WRF TLM with the added Kessler microphysics scheme, we compared the solution of the nonlinear perturbation with that of TLM. The ratio between nonlinear and tangent linear solution is defined as

\[
\Psi(\alpha) = \frac{\|M(x_0 + \alpha h) - M(x_0)\|}{\|\alpha h\|}, \quad \lim_{\alpha \to 0} \Psi(\alpha) = 1, \quad (9)
\]

where \(\|\|\) denotes norm of a vector. For values of \(\alpha\) that are small but not too close to the machine zero, the ratio \(\Psi(\alpha)\) should be close to 1. For a squall-line simulation with an integration time of 5, 15, and 30 min, respectively, the results of \(\Psi(\alpha)\) are shown in Table 1 for \(\alpha\) ranging from \(10^{-1}\) to \(10^{-12}\). For most values of \(\alpha\), \(\Psi(\alpha)\) in the three TLM integrations are close to 1.0, indicating the TLM is valid. It is also seen that the shorter the integration time is, the better the TLM approximation.

The accuracy of the ADM with the added Kessler scheme was verified against the TLM based on the definition of the adjoint operator:

\[
\langle y, Lx \rangle = \langle L^* y, x \rangle, \quad (10)
\]

where \(x\) stands for the TLM input and \(y\) its output. The symbols \(L\) and \(L^*\) stand for TLM and ADM operators, respectively. The adjoint check was done both for the Kessler microphysics scheme and for the entire WRF ADM as well; both passed the check by holding the equality in Eq. (10) up to the 64-bit machine accuracy.

### Control variables and background error statistics

The standard control variables (CV; Barker et al. 2004) for WRF 3D-Var and 4D-Var are the streamfunction; the unbalanced components of velocity potential, temperature, and surface pressure; and “pseudo” relative humidity. The new microphysical control variables added via the Kessler microphysics scheme are cloud water and rainwater mixing ratios. Because the vertical motion plays an important role in a convective system, it is also added as one of the control variables. As a result, the total number of control variables is eight in the new 4D-Var system for radar data assimilation. It should be
noted that only autocorrelation is considered for relative humidity, cloud water mixing ratio, rainwater mixing ratio, and vertical velocity. Their multivariate correlations are obtained in 4D-Var through the NWP model.

The method for estimating the background error statistics is the same for the new and existing control variables. A recursive filter is used to model the horizontal autocovariance, which is assumed to be spatially homogeneous and isotropic. The vertical error correlations are modeled using the diagnosed EOFs from an estimated background error covariance. In this study, the ensemble method (Fisher 2003) instead of the National Meteorological Center (NMC) method (Parrish and Derber 1992) is applied to better describe background error covariance of the new control variables. An ensemble including 80-member 3-h forecasts with a 4-km grid spacing is used to compute the length scale and vertical correlation EOFs using the WRF 3D-Var tool GEN_BE (Barker et al. 2004). The 3-h rather than longer forecasts were used because Sun et al. (2012) found that WRF forecasts initialized by 40-km Eta Model analysis spun up precipitation within 3 h in a retrospective study using one-week data from the International H2O Project (IHOP_2002). Initial conditions of the ensemble are produced by adding perturbations into the ensemble mean which is provided by Eta Model 40-km analysis. The ensemble perturbations are generated by sampling a static background error covariance (Sun et al. 2012) whose time period covers the case in this study. The boundary conditions are provided by the Eta Model 40-km analysis as well and they are not perturbed in this study.

3. Single-observation experiments

a. Experimental setup

Single-observation assimilation experiments are conducted to show background covariance structures in the updated 4D-Var system. The experiments will demonstrate how the background error covariance $B$ and model dynamics and physics $M$ spread the observation information. The single-observation tests are performed using the simulation of a midlatitude squall-line case that occurred in the U.S. Great Plains on 13 June 2002 during IHOP_2002 (Weckwerth et al. 2004). The appearance of this convective case is presented by the stage-IV hourly accumulated precipitation analysis (Lin and Mitchell 2005) at 0000 and 0600 UTC 13 June in Fig. 1 on the experiment domain. Both the single-observation experiments and real data experiments that will be presented in the next section use the same experimental domain shown in Fig. 1. It has a single, fixed grid with a horizontal resolution of 4 km and $151 \times 118 \times 31$ grid points. The following physics options are used in the simulation and outer loop update of the 4D-Var experiments: the Rapid Radiative Transfer Model (RRTM) longwave radiation, Dudhia shortwave radiation, Yonsei University (YSU) PBL schemes, Noah land surface model, and the Kessler microphysics. The description of the above schemes can be found in the WRF-ARW technical report (Skamarock et al. 2008). The TLM and ADM only have a Kessler microphysics and a simple vertical diffusion scheme. The TLM and ADM are
integrated with a 4-km resolution and a time step of 20 s. The lateral boundary control in WRF 4D-Var is not used in this study. A WRF forecast is initialized at 2100 UTC 12 June 2002. The 40-km Eta Model analyses provide both initial conditions and boundary conditions. For the single-observation tests, data assimilation is carried out between 0000 and 0010 UTC with a 10-min assimilation window. Figure 2 shows the forecast fields on the 13th model level (midtroposphere) at the end of the assimilation window (0010 UTC). A line of convection (Fig. 2e) is simulated with the corresponding features of convergence and updrafts/downdrafts (Fig. 2a), positive temperature perturbation (Fig. 2b), increased moisture (Fig. 2c), and cloud water (Fig. 2d). The convection generally occurs along the surface pressure isolines.

Since the microphysics TLM and ADM influence the analysis mainly through the assimilation of reflectivity data, we focus single-observation tests on the reflectivity observation. Two experiments are carried out by changing the value of reflectivity on a grid point located at (36.7°N, 97.3°W) at 13th model level (4000 m above sea level), the horizontal location of which is marked by the black dot in Fig. 2. The first single-observation test assumes the observed reflectivity is larger than the background predicted value (hereafter SOT-1). In addition to assessing how the updated 4D-Var system spreads observation information, this experiment is to see if the 4D-Var system can intensify the convection system by the reflectivity assimilation. The second experiment assumes the observed reflectivity is smaller than the background predicted value (hereafter SOT-2; i.e., the background overpredicted the convection system). In this experiment, it is also expected that the 4D-Var can weaken the convection system by fitting to the observation. A reflectivity observation with the values of 58 and 15 dBZ for SOT-1 and SOT-2, respectively, is used at the end of data assimilation window. The dBZ values are equivalent to 7.1 and 0.02 g kg⁻¹ rainwater mixing ratio, respectively, using Eq. (7) with $c_1$ setting to 43.1 and $c_2$ to 17.4 following Sun and Crook (1997). The background rainwater is 4.6 g kg⁻¹ at the assumed single-observation location and its equivalent simulated reflectivity is 55 dBZ.
b. Background error statistics

Figure 3 shows the first vertical EOFs for the seven control variables with its percentage to total variance and horizontal length scale shown on the top-right corner of each plot. The EOFs of each control variable are found by computing the eigenvectors of the corresponding control variable's vertical background error covariance matrix. The first eigenvector is important because it represents a pattern that can explain the largest percentage of variance to total variance, whereas the horizontal length scale describes to what extent the horizontal background covariance can spread observation information. The streamfunction, relative humidity, vertical velocity, and rainwater have the same features of a single peak with a large contribution to the total variance (at least 36%). For the streamfunction and unbalanced velocity potential, the eigenvectors have maximum values at about 250 hPa (about the 20th model level), which implies large forecasting uncertainty in the high-level jet around this level (Figs. 3a,b). The leading velocity potential eigenvector indicates a negative correlation in divergent wind between the upper and middle troposphere (Fig. 3b). The first mode of the unbalanced temperature peaks at lower levels, indicating large vertical background variances near the surface. The maximum forecasting uncertainty of the vertical velocity peaks at the midlevels, and is at the correlation transition region of the divergent wind.

The relative humidity variable and hydrometeor variables of cloud water and rainwater display the same features as that of a monotonic correlation (all positive or negative) along the entire vertical depth. The peaks of the eigenvectors of all variables are above 720 hPa (10th model level), suggesting that the large background variance occurs on the high levels. For cloud water, rainwater, and vertical velocity, the first eigenvector accounts for 57.1%, 80.7%, and 60.1% of the total variance, respectively, suggesting that the first EOFs give good representations of their vertical covariance.

In the WRF 4D-Var system, the length scale is defined by a distance at which the correction decreases to $e^{-1/8}$ (0.88; Barker et al. 2003). In this study, all the control variables except for streamfunction and unbalanced temperature have the maximum length scales at first EOF modes. The maximum length scale for the streamfunction is at the 12th EOF with a value about 80 km. The maximum length scale of the unbalanced temperature takes place at second EOF with a value of 12 km. The length scale of cloud water, rainwater, and vertical velocity is about 4 km, which means these convection-driven variables can only influence the nearby grid points.
c. Results

Figure 4 shows the horizontal increments of various model variables from the experiment SOT-1 at the end of the assimilation window. It gives a graphic representation of 4D-Var structure function. The results indicate that in addition to fitting the observation at its observed location (Fig. 4e), the observation influences its surrounding grid points with a flow-dependent pattern. With this error structure, 4D-Var can spread the single reflectivity observation information spatially and result in anisotropic multivariate analysis, meaning the other model variables, such as, water vapor, cloud water, temperature, and surface pressure are adjusted because of the dynamical model constraint even though only the rainwater information is assimilated. The convective system around the observation location is intensified as evidenced by the positive increment of the rainwater (Fig. 4e), along with the increased water vapor and cloud water (Figs. 4c,d). The horizontal wind convergence results in an upward vertical motion that supports the convection (Fig. 4a). The latent heat release due to the conversion of water vapor to cloud water, results in a positive increment of potential temperature (Fig. 4b). A negative surface pressure increment is obtained in response to the upward vertical motion and an intensified low- and midlevel convergence (Fig. 4f). Figure 4 also shows that all of the analysis increments display a southwest-northeast orientation in agreement with that of the simulated convective system (Fig. 2e), indicating the obvious flow-dependent nature of the increments. The cross section of the increment of the rainwater along with that of the vertical velocity from SOT-1 (Fig. 6a) shows that the convection is intensified as suggested by the increases of the rainwater and updraft. Figure 5 depicts the horizontal increments of various
model variables from the experiment SOT-2. Similar to the experiment SOT-1, 4D-Var produces analysis increments that have clear flow-dependent structures. As expected, the convection in this experiment is weakened as shown by the negative increments of the water vapor, cloud water, and rainwater (Figs. 5c,d,e) around the observation location. The divergent outflow is consistent with the weakened vertical motion (Fig. 5a). The negative cloud water and rainwater increments result in weakened latent heating that cause the negative temperature increment (Fig. 5b), which results in the increase of the surface pressure (Fig. 5f). Compared with Fig. 4, the increments in Fig. 5 show less directional orientation, which is possibly because the dissipation process is less controlled by the background dynamical forcing. Figure 6b clearly shows the decrease of the rainwater and vertical velocity over a depth of \(-8\) km.

The above single-observation tests confirm that the updated 4D-Var system with the addition of the Kessler microphysics scheme and the new control variables along with the derived background statistics is able to spread the observation information spatially and result in flow-dependent multivariate convective-scale analysis even with an assimilation window as short as 10 min. In the next section, the system will be tested on a real data case and the sensitivity of the system to the length of the assimilation window will be examined.

4. Real data experiments

The main purpose of the real data experiments presented in this section is to examine the performance of the incremental 4D-Var in a convective flow regime that has a high degree of nonlinearity. Since the nonlinear errors in a convective system grow quickly with forecast time, the assimilation window should be short enough such that the TLM approximation remains valid. A long window may decrease the accuracy of TLM approximation.
and have a negative impact on 4D-Var data assimilation, but on the other hand, a window too short could result in an insufficient balance between the analysis variables because the model integration is not long enough to allow the information to be transferred from the observed variables to the unobserved. In addition, a longer time window also allows more observations to influence analysis, which is another advantage for using a longer window. Radar observations in the United States are typically updated every 5 min or so in a storm mode; a 4D-Var window should be at least 5 min in order to include observations at two time levels. Wang et al. (2012) suggested that a short time window (less than 1 h) be used for 4D-Var to reduce the model error and nonlinearity and thus improve the validity of tangent linear model when a liquid-only microphysics scheme is used. The 5–15-min windows were used for 4D-Var radar data assimilation using VDRAS (Sun and Crook 1997, 1998, 2001; Sun and Zhang 2008).

Three experiments, 4DV_T05, 4DV_T15, and 4DV_T30, are conducted with the assimilation windows of 5, 15, and 30 min, respectively, to investigate the sensitivity of analysis to the length of the assimilation. Radar data are assimilated every 5 min with an observation time slot \( \pm 2.5 \) min. All six radars have 14 elevation scans with the elevation angles of \( 0.5^\circ, 1.5^\circ, 2.4^\circ, 3.3^\circ, 4.3^\circ, 5.2^\circ, 6.2^\circ, 7.5^\circ, 8.7^\circ, 10^\circ, 12^\circ, 14^\circ, 16.7^\circ, \) and \( 19.5^\circ \). The starting time of the three experiments is at 0000 UTC 13 June 2002. The ending time of the three experiments, 4DV_T05, 4DV_T15, and 4DV_T30, are at 0005, 0015, and 0030 UTC, respectively. Only radar radial wind and reflectivity observations are assimilated in the above experiments. The null reflectivity and clear-sky radial wind observations are not used in this study. A radar preprocessing and quality control procedure (Sun 2005; Lim and Sun 2010) is employed to process radar radial wind and reflectivity observations and specify observation errors. Six outer loops are performed in these experiments, that is, the nonlinear model is updated 6 times during the minimization of the cost function. For each outer loop, 20 iterations are carried out for the minimization of the cost function in the inner loop, which amounts to a total number of 120 iterations. With the same total number of iterations, we have conducted experiments with fewer nonlinear model updates by increasing the number of inner loop iterations and found the results were slightly degraded. On the other hand, fewer than 20 inner loop iterations with more outer loop updates are not desirable because the minimization is not given enough time to achieve a reasonable convergence.

For these three experiments, the background is generated from 40-km Eta Model analysis at 0000 UTC 13 June 2002, which is the starting time of the assimilation window for each of these three experiments. The background is also regarded as the first guess in the first outer loop (i.e., forecast from the background provides the first trajectory for WRF TLM and ADM). Forecasts up to 0600 UTC are run after the 4D-Var minimizations in all three experiments. Note that the WRF single-moment 5-class microphysics scheme (WSM5; Hong...
Fig. 7. Hourly accumulated precipitation (mm) valid at (left) 0100 UTC and (right) 0300 UTC. Forecasts at these times from the experiments (c),(d) NoDA; (e),(f) 4DV_T05; and (g),(h) 4DV_T30 are compared with (a),(b) the stage-IV precipitation analyses.
et al. 2004) that has ice and snow physics is used during the forecast period rather than the Kessler warm-rain scheme as in the 4D-Var assimilation period. Both the analysis and forecast are conducted on the domain shown in Fig. 1a, in which the locations of the six WSD-88D radars that are assimilated in the 4D-Var experiments are also indicated.

The sensitivity of the analysis to the assimilation length is assessed by the quantitative precipitation forecasting (QPF) skill. In this study, we use both the neighborhood-based fractions skill score (FSS; Roberts and Lean 2008) and the precipitation bias to verify the QPF skill. The stage-IV 4-km hourly precipitation analysis is used as “observations” in the verification. Before showing the statistical precipitation skills, in Fig. 7, the hourly precipitation fields from 4DV_T05 (Figs. 7c,d) and 4DV_T30 (Figs. 7g,h) at $t = 1$ and $t = 3$ h are compared with each other and with the stage-IV analysis (Figs. 7a,b). The precipitation forecasts from the “cold start” run, NoDA, which is initialized from the first-guess fields (e.g., interpolated from the 40-km Eta Model analysis), are shown in Figs. 7c,d. It is clear that the forecasts

![FSS graphs](image-url)
from NoDA miss a large part of the convective system, resulting in low FSS in Fig. 8. In contrast, the two 4D-Var experiments predict the convective lines that agree better with the stage-IV analysis. At $t = 1$ h, 4DV_T05 captures all of the six major storm clusters A–F (marked in Fig. 7a), but some of them are overpredicted with higher values of precipitation accumulation than the stage-IV analysis. The problem of the overprediction
worsens at $t = 3$ h. In comparison to 4DV_T05, 4DV_T30 captures the three clusters A, B, and D and misses most part of the other three at $t = 1$ h (Fig. 7g). It is also noted that the forecasted squall line at $t = 3$ h in both 4DV_T05 and 4DV_T30 appears to be slower (farther north) than observed. However, the forecast at $t = 3$ h has a closer agreement with the stage-IV analysis as can be seen subjectively from Fig. 7h.

The FSSs in Fig. 8 show that all the 4D-Var experiments have much higher scores than the cold start experiment, indicating the significant impact of the radar data assimilation. The low score in the cold start experiment is because it takes time to develop convection and convective-scale dynamics when the high-resolution detail is spinning up from the lower-resolution analysis data used (Lean et al. 2008). Figure 8 also indicates that overall the three 4D-Var experiments do not show significant differences in FSS except that a slightly lower FSS from 4DV_T05 is noticeable for the threshold of 5 mm. However, when the precipitation bias is plotted, we have found that the two experiments with shorter assimilation lengths have higher bias, as demonstrated in Fig. 9. The 30-min assimilation window produces weaker precipitation analysis as shown in Fig. 9 by the lower than 1 bias values at $t = 1$ h. Nevertheless, the precipitation forecast at $t > 1$ h shows less bias than the experiments that use shorter window lengths.

The QPF results shown above suggest that the 4D-Var radar data assimilation is capable of initializing the WRF with a single 4D-Var assimilation window and producing encouraging convective forecasting. Next, we will provide a detailed examination on whether the incremental 4D-Var has a stable behavior with respect to the multiple basic-state updates. The reduction on the cost function is first shown in Fig. 10 with respect to the 126 iterations (21 inner loop iterations for each of the 6 outer loops). Note that the values of the cost function for the three experiments are different at the first iteration because of the difference of the assimilation length. Figure 10 also suggests that the longer the assimilation length, the larger reduction of the cost function is produced. For experiment 4DV_T30, the cost function decreases steadily in each inner loops but a jump occurs when a new outer loop starts. In contrast, the other two experiments (4DV_T05 and 4DV_T15) experience smaller jumps. This jump happens because in the inner loop $H_kM_k(x^{(n-1)}_k)\delta x^{(n)}_k - \{y^{(n)}_k - H_k[M_k(x^{(n-1)}_k)]\}$ is used to calculate the cost function [Eq. (2)], whereas $\{y^{(n)}_k - H_k[M_k(x^{(n-1)}_k + \delta x^{(n)}_k)]\}$ is used to calculate the value of the cost function [Eq. (1)] at the start of the outer loop. If the simulated observations from the updated first guess $H_k[M_k(x^{(n-1)}_k + \delta x^{(n)}_k)]$ can be well represented by its tangent linear assumption $H_k[M_k(x^{(n-1)}_k)] + H_kM_k(x^{(n-1)}_k)\delta x^{(n)}_k$, the jump should not be significant, as in the case of 4DV_T05. When examining the analysis fields, it is found that the analysis after the first outer loop from 4DV_T30 triggers convection and its associated nonlinear regime, resulting in the apparent jump in the beginning of the second outer loop. For 4DV_T15, the jump does not occur until the beginning of the third outer loop when convection is first initiated. We have shown earlier that 4DV_T30 yields reasonable forecasts despite the apparent discontinuity caused by the linearization.

To show the contributions of different outer loop updates to the analyses, we compare the analysis increments in 4DV_T30 after the first outer loop at 850 hPa in Fig. 11 with those after the sixth outer loop in Fig. 12. It is found that the analysis increments from the first outer loop account for a large part of the total analysis increments, but differences from those after the sixth outer loop are identifiable. The magnitude of the $u$ and $v$ increments is increased (cf. Fig. 12a with Fig. 11a and Fig. 12b with Fig. 11b). The intensities as well as patterns of some of the cold pools in Fig. 12c are adjusted from those in Fig. 11c. The most significant differences appear in the relative humidity field. In addition to the mainly positive moisture increment shown in Fig. 11d, a more small-scale negative analysis increment is found in Fig. 12d. The iterations beyond the first outer loop add small-scale and small-amplitude details to the analysis as shown by Fig. 13 in which the differences of the analyses between fifth and sixth outer loops are plotted.

To evaluate whether these fine structure adjustments to the analysis are physically meaningful, forecasts after each outer loop are conducted and their QPF skills are compared in Fig. 14. It is shown that in general the FSS steadily increases with the increase of the outer loop number. After the third loop it changes very little for the 1-mm threshold and only has small adjustment for the
5-mm threshold relative to the first three loops. The results indicate that three outer loops are enough to produce skillful precipitation forecasts for the case studied and the incremental 4D-Var results in a converged precipitation forecast.

5. Summary and discussion

The main objectives of this study are developing the capability of radar data assimilation for WRF 4D-Var and examining the feasibility of the incremental 4D-Var for convective-scale radar data assimilation. The WRF 4D-Var is further developed to include the tangent linear and adjoint models of a Kessler warm-rain microphysics scheme and new control variables of cloud water, rainwater, and vertical velocity for the assimilation of reflectivity. The background error statistics are derived using an ensemble method. Although the motivation of the new development is for radar data assimilation, the updated WRF 4D-Var system can be applied to assimilate other types of cloud and rain-affected observations as long as their corresponding observation operators are developed.

Single radar reflectivity observation assimilation experiments are carried out to show how the reflectivity information is spread spatially and propagated to other
variables in the updated WRF 4D-Var system. The results show that the system is not only able to spread the information from a single reflectivity observation to its surrounding grid points anisotropically but also transfer that information to other variables, producing multivariate analysis owing to the use of the TLM dynamical constraint. The assimilation of a reflectivity observation can effectively intensify (weaken) the convection system if it is underpredicted (overpredicted) in the background, which induces the dynamical responses of other model variables.

The results from real radar data assimilation experiments indicate that the updated WRF 4D-Var with multiple outer loops produces reasonable convective analyses and forecasts by assimilating radial wind and reflectivity observations with window lengths of 5, 10, and 30 min. The cost functions are reduced steadily with the iterative updates of the basic state. Although the longer window of 30 min shows temporary jumps at the beginning of each update, the 4D-Var minimization behaves properly, yielding reduced value of the cost function than the previous outer loop. It is also shown that the analyses yield precipitation forecasts with stabilized skills after three outer loops. The assimilation experiments with the different window lengths produce forecasts with similar skill scores when measured by FSS. However, the 30-min experiment produces the best forecasts in term of bias.

The preliminary results presented in this paper show the updated WRF 4D-Var with the incremental

Fig. 12. As in Fig. 11, but for the final analysis increments after the sixth outer loop.
formulation works well at the convection-permitting scale for radar data assimilation. In Part II, we will further evaluate the system by comparing it with the WRF 3D-Var. More detailed diagnosis of the results will be conducted to show the reason for the improved capability of 4D-Var in convective-scale analysis and forecasting.

The next step of our 4D-Var developments will be to design a procedure allowing the WRF 4D-Var system runs with continuous cycles. The challenge lies in how to make a smooth transition between the simplified physics in the 4D-Var and in the full nonlinear forecast model. Another future project is the development of a more sophisticated microphysics scheme that includes ice and snow physics to broaden the application of the WRF 4D-Var into winter-season storms in which ice and snow processes play an important role. The development of the adjoint model of more complicated microphysics for convective-scale data assimilation could be a challenge. Difficulties are not only in writing the adjoint code, but also obtaining a tangent linear model with acceptable precision at the convective scale. Past studies (Honda and Yamada 2007; Amerault et al. 2008) showed controversial results when a microphysics scheme with ice and snow processes were used in high-resolution 4D-Var data assimilation. Further tests of the current 4D-Var radar data assimilation system on more cases with different characteristics are also required to give a comprehensive evaluation of the system.

Fig. 13. As in Fig. 11, but for differences of the analyses between the fifth and sixth outer loops.
Acknowledgments. This work was supported by the National Science Foundation’s U. S. Weather Research Program (USWRP) and the Air Force Weather Agency. The authors thank Yong-Ruo Guo for constructive discussions. We also thank Dr. Ming Xue and another three anonymous reviewers for their valuable comments that helped improved this paper. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES


——, and ——, 1990: Variational assimilation of meteorological observations with the direct and adjoint shallow water equations. Tellus, 42A, 531–549.


Fisher, M., 2003: Background error covariance modelling. Proc. ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean, Reading, United Kingdom, ECMWF, 45–63.


