Radar Reflectivity Assimilation with the updated WRFDA-4DVAR system

Hongli Wang1, Juanzhen Sun1, Yong-Run Guo1, Xiang-Yu Huang1, and Soichiro Sugimoto2

1 National Center for Atmospheric Research, Boulder, Colorado, USA
2 Central Research Institute of Electrical Power Industry, Tokyo, Japan

Abstract

This research focuses on assimilating radar data into a high-resolution NWP model to improve short-term quantitative precipitation prediction (QPF). The WRFDA-4DVAR (4-dimensional variational data assimilation system of Weather Research and Forecast model) was further developed to assimilate radar reflectivity observations, and a convective case was carried out to assess the performance of the enhanced 4DVAR system. The development includes 1) the inclusion of cloud water and rain water as hydrometeor control variables and their modeling of error covariance; 2) the development of tangent-linear and adjoint of a Kessler warm-rain scheme; 3) a new observation operator for radar reflectivity.

The updated WRFDA-4DVAR system was tested with a convective case occurred on June 13 2002 during IHOP_2002 (International H2O Project) experiment. The impact of radar reflectivity, along with radar radial wind, on analysis and short-term numerical prediction of the IHOP case were assessed. Compared to the control experiment from Eta model 40 km analysis and 3DVAR analysis, radar data assimilation with 4DVAR method initiated the convection system in the first hour and improved the hourly precipitation forecast skill. Assimilating the reflectivity and radial velocity observations in 4DVAR enhanced the moisture and produced a cooling region near the surface in the convective region, which helped to maintain the convective system.
1. Introduction

Improving the short-range quantitative precipitation forecasting (QPF) is a challenge for the numerical weather prediction (NWP) since both a well-formulated mesoscale numerical model and an accurate initial condition are required. For over a decade, high-resolution NWP models have been applied to produce experimental real time numerical forecast in the research community (e.g., Xue et al., 1996; Kain et al., 2008; Weisman et al., 2008). In general, the use of 1-, 2-, and 4-km grids improved the convection forecast over coarser resolutions. Roberts and Lean (2008) found that the convection forecast up to 6 hours using a model of 1-km resolution was the most skillful at all scales than other coarser resolutions, and the biggest improvement occurred for heavier, more localized rain. During the 2007 NOAA Hazardous Weather Testbed (HWT) Spring Experiment, on average, both the 2- and 4-km model forecasts showed substantial improvement over the 12-km NAM in 21- to 33-h forecast period, however, decreasing horizontal grid spacing from 4 to 2 km provided little added value for severe convective QPF in the United States (Schwartz et al., 2010). Weisman (2008) found that the use of 2-km grids improved the structural features of some convective systems over 4-km model, but had little impact on the timing and location of the systems.

The large forecast sensitivity of high-resolution model to initial conditions was reported by several researchers (Reobber et al, 2002; Gallus et al., 2005; Weisman et al., 2008; Xue et al., 2008). A simulation study of the 3 May 1999 Oklahoma tornado outbreak by Roebber et al. (2002) suggested that 2-h forecast errors could be related to resolvable-scale observational errors in the initial upstream conditions. Weisman et al. (2008) found that larger forecast sensitivities arose, when varying the initial conditions for the forecasts (e.g., initializing with the RUC versus Eta) than varying physical parameterization schemes. Xue et al. (2008) found that when mature
pre-frontal convection existed, both 2 and 4 km control forecasts with radar data significantly outperformed that of the 4 km forecast without radar data. The forecasts captured many of the structures and behaviors of the observed storms, including the propagation and evolution of two main groups of convection, and the initiation of a secondary line. Xue’s results provided evidence of the importance of both convection-permitting resolution and radar data assimilation for severe weather prediction. Thus a better observation network than what now exists and better methods of including additional information from radar and satellite could be crucial for improving short term convective forecasting (Gallus et al., 2005).

The advanced data assimilation techniques, such as, variational method (VAR) and ensemble Kalman Filter (EnKF) have been implemented to assimilate radar observations to provide the best analysis for the high resolution convective forecast. The three- and four-dimensional VAR (3/4DVAR) data assimilation systems, such as VDRAS (variational Doppler radar analysis system, Sun et al, 1997), Japan Meteorological Agency’s 4DVAR (Honda et al, 2005) and ARPS (Advanced Research and Prediction System) 3DVAR (Gao et al., 2004), have been implemented. These systems with differences in data assimilation implementation and numerical model all demonstrated potentials in severe storm forecast. The potential of the EnKF’s application to the convective-scale was shown by Aksoy et al. (2009, 2010) using real observations using cases from IHOP_2002.

Assimilating radar observations with the 4DVAR at convective scale could be a superior choice to improve the short-range QPF than the 3DVAR. One of the challenges for the convective-scale application of a 3DVAR data assimilation system is its difficulty in defining the background error covariance, especially when the hydrometeor variables are included in the control variables. However, 4DVAR presents an implicit flow dependent background error covariance to propagate the correlation among the variables through the forward integration of the constraining numerical model. As a result, the analysis is more dynamical consistent with the
numerical model and is expected to reduce the model spin-up time. Another advantage of 4DVAR is that all the observations are assimilated at the time they are valid.

The Radar data assimilation capability with WRF three-dimensional variational data assimilation system (WRFDA-3DVAR) has been developed and evaluated in the National Center for Atmospheric Research (NCAR). The scheme and its performance evaluation were reported by Xiao et al. (2005, 2007). To assimilate the radar reflectivity, WRFDA-3DVAR takes the approach to use the total water as the moisture control variable and a partition scheme to split it into water vapor, cloud water and rainwater. Xiao and Sun (2005) applied the WRFDA-3DVAR to the IHOP case of 12-13 June 2002 and showed improved QPF over experiments without radar data assimilation.

WRFDA-4DVAR was developed upon the same framework of WRFDA-3DVAR, but with the extension of the WRF dynamical model and its adjoint (Huang et al. 2009). As such, it uses the same control variable transformations and precondition mapping. When radar data (in particular the reflectivity data) assimilation is concerned, the WRFDA-4DVAR needs further development to enable the capability of assimilating the cloud- and precipitation related observations. In this paper, we describe the new development pertinent to the convective-scale radar data assimilation, including the hydrometeor control variables and their error statistics and the development of the tangent-linear and adjoint components of a warm-rain microphysics. The newly developed 4DVAR is applied to the IHOP case of 12-13 June 2002. Its QPF performance is compared with 3DVAR. The sensitivity of the 4DVAR system with respect to the radar data with and without the reflectivity, and the length of assimilation window is examined. A specific issue concerning the assimilation of radar reflectivity (Z) is whether the observation itself or the derived rain water mixing ratio \( q_r \) should be assimilated. Sun and Crook (1997) examined this issue by comparing experiments that assimilated the two different variables respectively. In
this paper, a detailed analysis of the nonlinearity of the Z-q<sub>r</sub> relation is provided to further investigate the issue.

The structure of this paper is arranged as follows. Section 2 describes issues in radar reflectivity assimilation with WRFDA-3DVAR and provides an analysis on the nonlinearity of the Z-q<sub>r</sub> relation. Section 3 describes the development of WRFDA-4DVAR for radar data assimilation. The results from a radar data assimilation experiment are presented in Section 4. A summary is provided in the final section.

2. Radar reflectivity assimilation in WRFDA-VAR
   a. Brief description of the WRFDA-VAR system

The WRF variational data assimilation system (WRFDA-VAR) includes both 3DVAR and 4DVAR capabilities within a common framework (Barker et al. 2004, Huang et al. 2009). Here, we provide a brief summary of the system. The reader is referred to the above two publications for further details.

The WRFDA-VAR algorithm takes the incremental VAR formulation that is commonly used in operational systems (Courtier et al. 1994; Veersé and Thépaut 1998; Lorenc 2003). The incremental approach is designed to find the analysis increment that minimizes a cost function defined as a function of the analysis increment instead of the analysis itself. In the incremental 4DVAR, the tangent linear and adjoint models that usually derived from a simplified forward model are used in the inner-loop minimization, while the evolution of the background is predicted with the full forward model. The control variables are streamfunction, the unbalanced components of velocity potential, temperature, and surface pressure, and "pseudo" relative humidity. The “unbalanced” control variables are defined as the difference between the full and the “balanced” (or correlated) components of the field with the streamfunction (Barker et al. 2004).
b. The validity of linearized Z-q_r equation
Since the multi-incremental formulation is taken in WRFDA-VAR system, in practice the linear approximation should be valid for finite amplitude of perturbation in each inner loop. In the incremental formulation, the linearized Z-q_r equation is employed as observation operator in the cost function to assimilate radar reflectivity. In this section, the validity of linear Z-q_r relation is analyzed to determine the error characteristic of the linearization.

The nonlinear Z- q_r relation (Sun and Crook, 1997) is presented by:

\[
Z = c_1 + c_2 \cdot \log_{10}(\rho \cdot q_r) \tag{1}
\]

where \(Z\) is the reflectivity (unit of dBZ), \(c_1\) and \(c_2\) is 43.1 and 17.5 respectively, \(q_r\) is the rainwater mixing ratio.

And its linear formulation is,

\[
dZ = c_2 \cdot dq_r / (q_r \cdot a \log(10)) \tag{2}
\]

Given an initial perturbation of rainwater \(dq_r\), the resulting reflectivity can be obtained from equation (1),

\[
Z_{new} = c_1 + c_2 \cdot \log_{10}(\rho \cdot (q_r + dq_r)) \tag{3}
\]

The perturbation of reflectivity caused by \(dq_r\),

\[
dZn = Z_{new} - Z = c_2 \cdot \log_{10}((q_r + dq_r)/q_r) \tag{4}
\]

Let,

\[
k = dq_r / q_r \tag{5}
\]

The physical constraint of \(k\) is \(k \geq -1\)

The linear approximation error (LE) is

\[
LE = dZ - dZn = c_2 \cdot dq_r / (q_r \cdot a \log(10)) - c_2 \cdot \log_{10}((q_r + dq_r)/q_r) = c_2 \cdot k / a \log(10) - c_2 \cdot \log_{10}(1 + k) \tag{6}
\]
From LE-k relation in Fig.1, it is seen that

1. dZ is truncated at the point where k is -1.0, but not dZn.

2. The relation between the linear approximation error and k is nonlinear and asymmetrical. -1<k<0 will lead to larger linear approximation error than 0<k<1.

3. LE is always larger than 0.0, which means dZ always overestimates dZn. Hence when fitting an innovation of radar reflectivity with linear Z-q equation, for the same background rainwater q, an underestimated increment of rainwater dq is obtained compared to the increment derived from nonlinear Z-q equation, resulting in a dry bias in rainwater increment. For example, if there is an innovation of 10.0 dBZ, the k from nonlinear relation is 2.73, whereas from linear relation is 1.32.

4. The linear approximation can have a large error in the case of the background being much wetter than the observation (k is close to -1), which corresponds to the situation when the background forecasted storm is stronger than the observed. The use of the linearized Z-q will result in a negative increment that has a much larger absolute value than that from its nonlinear counterpart, which means the storm is over-corrected and becomes too weak.

Fig.1. Linear perturbation dZ, nonlinear difference dZn, and the LE-k relation
c. Schemes to assimilate reflectivity

The above analysis clearly indicates that the linearization of the $Z$-$q_r$ relation results in significant error in certain circumstances. It implies that when the model forecast has a large deviation from the observation, the optimal solution has difficulty to fit to the observation. To circumvent this problem, we suggest assimilating the retrieved rainwater from radar reflectivity instead of assimilating the reflectivity directly. The benefit will be shown using a single observation by comparing the two approaches.

1) Nonlinear formulation

The cost function that is specifically relevant to the reflectivity can be written as the following,

$$J(q_r) = (q_r - q_{rb})^T B^{-1}(q_r - q_{rb}) + (Z - H(q_r))^T R^{-1}(Z - H(q_r))$$

$$q_{ra} = \min J(q_r)$$

where $B$ and $R$ are the background error and reflectivity observation error respectively. $H$ is the observation operator presented in equation (2).

2). Incremental formulation

$$\delta q_r = q_r - q_{rb}, \quad \delta Z = Z - H(q_{rb}), \quad H^\prime = \frac{\partial H}{\partial q_r} |_{q_r = q_{rb}}$$

$$J(\delta q_r) = (\delta q_r)^T B^{-1}(\delta q_r) + (Z - H(q_{rb} + \delta q_r))^T R^{-1}(Z - H(q_{rb} + \delta q_r))$$

$$\approx (\delta q_r)^T B^{-1}(\delta q_r) + (H^\prime \delta q_r - \delta Z)^T R^{-1}(H^\prime \delta q_r - \delta Z)$$

$$\delta q_{ra} = \min J(\delta q_r)$$

$$q_{ra} = q_{rb} + \delta q_{ra}$$

3). Assimilating retrieved rainwater

If the rainwater is assimilated, the cost function is written as

$$J(q_r) = (q_r - q_{rb})^T B^{-1}(q_r - q_{rb}) + (q_r - H^{-1}(Z))^T R^{-1}(q_r - H^{-1}(Z))$$

$$q_{ra} = \min J(q_r)$$
where $H^{-1}$ is the inverse of observation operator $H$ presented in equation (2).

The above equation is a linear formulation that excludes the nonlinear term in equation (7). Its minimum is easily achieved by an iteration method. The observation error $R_o$ in equation (15) is estimated by the relation between the perturbation rainwater and perturbation radar reflectivity in equation (2).

**d. The results of the single observation test**

When it is assumed that there is only one grid point, the equations (8), (13) and (16) can be easily solved. The three schemes of assimilating reflectivity with the nonlinear operator, with the linearized operator, and assimilating the derived rainwater can be compared. In this comparison, the variance of the background error $B$ and reflectivity observation error $R$ are assumed to be $1 \text{ g kg}^{-1}$ and $5 \text{ dBZ}$, respectively. The analysis increments of the three schemes are shown in Fig. 2. It is shown that the scheme of assimilating retrieved rainwater gives the similar result to the nonlinear scheme while the scheme of assimilating reflectivity with the linearized operator has obvious larger departure from the nonlinear scheme.
Fig. 2. Analysis increment of rainwater mixing ratio for schemes by equations (8) (a), (13) (b) and (16) (c).

3. Recent development of WRFDA-4DVAR

The current WRFDA-4DVAR has no hydrometeor control variables directly related to radar reflectivity and no microphysics in tangent-linear and adjoint models to propagate the perturbations of cloud water and rain water. Before the reflectivity data can be assimilated by the WRFDA-4DVAR, it is necessary to develop the system such that it can be applied to the convective scale data assimilation. In this section, we describe the recent development of WRFDA-4DVAR that is related to the assimilation of radar reflectivity.

a. Hydrometeor control variables

In a variational data assimilation system, hydrometeor control variables are required to provide the moisture and hydrometeor analysis. For a 4DVAR system that considers warm rain microphysics, the possible choices are total liquid water, cloud water, and rain water. Two of these three variables are needed because the third one can be derived. Following the MM5-4DVAR (Zou et al., 1998) and RAMDAS (Regional Atmospheric Modeling Data Assimilation System) systems (Zupanski et al., 2005), the model prognostic variables of cloud water and rain water are used as the new microphysical control variables. A recursive filter is used to model the horizontal auto-covariance to spread the increment horizontally. And Empirical Orthogonal Functions (EOFs) are used to model the vertical correlations. The correlations with
other variables are determined by the WRF tangent-linear and adjoint models.

b. Tangent-linear and adjoint of a Kessler scheme

The Radar reflectivity mainly provides the rain water information. Hence we start from a Kessler warm-rain scheme in the Advanced Research WRF (WRF-ARW) (Skamarock et al., 2005). This scheme is a warm cloud scheme that includes water vapor, cloud water, and rain water. 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 tangent-linear and adjoint models were first developed with the automatic differentiation tool, Tapenade (Hascoet and Pascual, 2004), then followed by a manual modification. The original Kessler scheme was reorganized and modified for the components that have high degree of nonlinearity to avoid producing singular values. The derivation of the adjoint of physical processes with on/off switches in Kessler scheme follows that of Zou (1993). That is, the switching times are kept the same as in the basic state. The results of the correctness test for the tangent-linear and adjoint of the Kessler scheme followed the method in Zou’s paper. Based on the analysis in the last section, we assimilate retrieved rainwater converted from the radar reflectivity.

4. Radar data assimilation experiment

a. Case description

A squall line that occurred in Oklahoma and Kansas on 12–13 June 2002 during IHOP_2002 is chosen for this study. This case was documented by 12 ground-based Doppler radars in Oklahoma, Kansas, and Texas. Detailed analyses of the case were conducted by Weckwerth et al. (2004) and Wakimoto et al. (2004). This study focuses on a northeast-to-southwest-oriented squall line from the Kansas and Oklahoma border to the Texas Panhandle, initiated at around 2100 UTC 12 June 2002. Isolated
convections formed from the Kansas and Oklahoma border to the Texas Panhandle along a dryline at 2100 UTC 12 June. They gradually strengthened along most of the dryline, and a severe storm developed near a triple point, formed by the dryline and a cold air outflow. At 0000 UTC 13 June, a squall-line structure was well developed in the reflectivity field, with the intense convection in the triple-point area near the Kansas and Oklahoma border. It then moved southeastward and dissipated after 0900 UTC 13 June 2002.

b. Experimental design

The objectives of the experiments are to examine the performance of the Doppler radar data assimilation capability in WRFDA-4DVAR, the impact of multiple–Doppler radar data assimilation on short-range QPF, and the reasons leading the better precipitation prediction. Six numerical experiments are conducted to examine the sensitivity of the data assimilation with respect to some options in the system. All data assimilation experiments begin at 0000 UTC 1300 June 2002. One domain with horizontal resolution of 4 km is adopted in data assimilation and forecast experiments.

Table 1. Experiments and descriptions (RV stands for radial velocity and RF for reflectivity)

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BG</td>
<td>ETA 40 km analysis</td>
</tr>
<tr>
<td>FG</td>
<td>Assimilate RF by 3DVAR</td>
</tr>
<tr>
<td>3DVAR-R2</td>
<td>Assimilate <strong>RV and RF</strong> by 3DVAR</td>
</tr>
<tr>
<td>4DVAR-R2-T10</td>
<td>Assimilate RV + RF with <strong>10</strong> minute time window</td>
</tr>
<tr>
<td>4DVAR-R2-T15</td>
<td>Assimilate RV + RF with <strong>15</strong> minute time window</td>
</tr>
<tr>
<td>4DVAR-RV-T15</td>
<td>Assimilate RV with <strong>15</strong> minute time window</td>
</tr>
</tbody>
</table>

Table 1 lists these experiments and gives a brief description of each experiment. In these experiments, we only assimilate the radial velocity in the rain region defined by the reflectivity values larger than or equal to 20 dBZ. The method of Doppler radar data quality control and other preprocessing procedures are the same as that described in Xiao et al. (2005, 2007). The experiment BG is a WRF simulation with initial and
boundary condition interpolated from Eta model 40 km analysis. To examine the benefit of 4DVAR method, two 3DVAR experiments FG and 3DVAR-R2 are carried out. The analysis of FG is used as first guess for the 4DVAR experiments. Three 4DVAR experiments are conducted to assess the gain of the 4DVAR method, and to test the impact of time window and reflectivity observations.

The WRF-ARW model is used as the forecast model. The WRF is the next generation mesoscale model designed to serve both operational forecasting and atmospheric research applications. The model uses a third order Runge-Kutta time integration, third to fifth order advection operators, and split-explicit fast wave integration conserving both mass and energy. The model physics options include the rapid radiative transfer model (RRTM) longwave radiation (Mlawer et al., 1997), Dudhia shortwave radiation (Dudhia, 1989), Yonsei University (YSU) PBL schemes (Hong et al., 2006), and single-moment 6-class microphysics scheme (WSM6) (Hong et al., 2006) in all the experiments.

The performance of the 4DVAR and 3DVAR are compared in terms of the precipitation forecast. The overall precipitation forecast improvement is estimated by comparing the precipitation threat scores. The threat score is computed against Stage IV rainfall data in the model domain.

c. Results

The threat scores of 5-mm of hourly precipitation amounts are shown in Figure 3. From Figure 3a, it is seen that even with the 3DVAR, radar observations improve the precipitation forecast up to 5 hours. Assimilating both RV (radial velocity) and RF (reflectivity) gives the better results. The 4DVAR experiment yields significantly better forecasts than the 3DVAR experiment. Figure 3b shows that TSs for the 4DVAR experiments 4DVAR-R2-T10 and 4DVAR-R2-T15 are much higher than that for 3DVAR, indicating the significant positive impact of 4DVAR method. The longer time window of 15 minutes improves the precipitation prediction slightly
compared to the time window of 10 minutes. Comparing TS of the experiment 4DVAR-RV-T15 with the experiment 4DVAR-R2-T15, the impact of RF is positive and significant up to 6 hours.

Figure 4 compares the hourly precipitation forecast at 01Z 13 June 2002 from different experiments with stage IV observation. Consistent with the objective verification scores in figure 3b, the forecast initiated from 40-km Eta analysis almost missed the convections during the first hour. Both 3DVAR and 4DVAR experiment produce precipitation structures that are much more realistic and in better agreement with the observation. The experiment 4DVAR-R2-T10 and 4DVAR-R2-T15 give the better location and intensity forecast especially for the precipitation near the west boundary compared to the 3DVAR experiment.

As the forecast proceeds, the convection in BG begins to develop, but the location does not agree with the observations. The convection system in experiment 3DVAR-R2 moves fast then the observation. While the location and evolution of the convection in the experiment 4DVAR-R2-T10 and 4DVAR-R2-T15 are very close to the observations. Figure 5 shows the hourly precipitation forecast at 06Z 13 June 2002 from different experiments and stage IV observation. Consistent with the objective threat scores in Figure 3b, 4DVAR experiments with both RV and RF assimilation produce the better forecast then experiment BG and 3DVAR-R2.
Fig. 4. Hourly rainfall at 01Z 13 June of Stage IV observations, and forecast from BG, 3DVAR and 4DVAR experiments.

Fig. 5. Hourly rainfall at 06Z 13 June of Stage IV observations, and forecast from BG, 3DVAR and 4DVAR experiments.

Analyses from the 3DVAR and 4DVAR experiments are compared to show the reason leading to the improved precipitation forecast. Figure 6 shows the analysis at 00Z 13 June 2002 for the experiments 3DVAR-R2, 4DVAR-RV-T15 and 4DVAR-R2-T15. It is seen that assimilating RF and RV in 4DVAR enhances the
moisture and produces a cooling region near the surface under the convective system, which is the key for its development and maintenance.

Fig. 6 The analysis at 00Z 13 June 2002 from experiments of a) 3DVAR-R2, b) 4DVAR-RV-T15, and c) 4DVAR-R2-T15 of precipitable water (shaded), wind field at 10 m (vector) and temperature (red contour) at 2 m altitude.

5. Summary

In this study, the WRFDA-4DVAR was further developed to assimilate cloud- and precipitation-related observations. The system was tested with radar reflectivity observations for a mid-latitude squall line occurred during IHOP_2002. The development included 1) the addition of cloud water and rain water as hydrometeor control variables; 2) the development of the tangent-linear and adjoint of a Kessler warm-rain scheme; 3) a new observation operator for radar reflectivity.
A scheme, which assimilates the retrieved rainwater from reflectivity, was suggested to improve the rainwater analysis based on the validity analysis of the linear Z-R equation. It was found that, the validity of the linear Z-q_r equation only depends on the ratio $k$ defined by the rainwater perturbation to the basic state rainwater. The linear approximation error is always larger than 0.0. Hence when fitting an innovation of radar reflectivity using the linearized Z-q_r equation, an underestimated increment of rainwater is obtained compared to results from the nonlinear Z-q_r equation. The experiment with a single observation showed that assimilating the derived rain water was a superior choice than assimilating the reflectivity with the linearized operator.

The impact of radar reflectivity, along with radar radial wind, on analysis and short-term numerical predictions of the IHOP case were assessed. Compared to the forecasts from Eta model 40 km analysis and 3DVAR analysis, radar data assimilation with the 4DVAR method initiated the convection system in the first hour and improved the hourly precipitation forecast skill up to 6 hours, which is the maximum length of the forecast for the case studied with a relative small domain. Assimilating radial velocity and reflectivity in 4DVAR enhanced the moisture and produced a cooling region near the surface under the convective system, which helped to maintain its development.
Reference


