Design Strategies of an Hourly Update 3DVAR Data Assimilation System for Improved Convective Forecasting

WENXUE TONG
School of Atmospheric Sciences, Key Laboratory of Meteorological Disaster Ministry of Education, Nanjing University of Information Science and Technology, Nanjing, China

GANG LI
Nanjing University of Information Science and Technology, Nanjing, China

JUANZHEN SUN
National Center for Atmospheric Research, Boulder, Colorado

XIAOWEN TANG
School of Atmospheric Sciences, Nanjing University, Nanjing, China

YING ZHANG
National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 3 March 2016, in final form 22 August 2016)

ABSTRACT

This study examines two strategies for improving the analysis of an hourly update three-dimensional variational data assimilation (3DVAR) system and the subsequent quantitative precipitation forecast (QPF). The first strategy is to assimilate synoptic and radar observations in different steps. This strategy aims to extract both large-scale and convective-scale information from observations typically representing different scales. The second strategy is to add a divergence constraint to the momentum variables in the 3DVAR system. This technique aims at improving the dynamic balance and suppressing noise introduced during the assimilation process. A detailed analysis on how the new techniques impact convective-scale QPF was conducted using a severe storm case over Colorado and Kansas during 8 and 9 August 2008. First, it is demonstrated that, without the new strategies, the QPF initialized with an hourly update analysis performs worse than its 3-hourly counterpart. The implementation of the two-step assimilation and divergence constraint in the hourly update system results in improved QPF throughout most of the 12-h forecast period. The diagnoses of the analysis fields show that the two-step assimilation is able to preserve key convective-scale as well as large-scale structures that are consistent with the development of the real weather system. The divergence constraint is effective in improving the balance between the momentum control variables in the analysis, which leads to less spurious convection and improved QPF scores. The improvements of the new techniques were further verified by eight convective cases in 2014 and shown to be statistically significant.

1. Introduction

Precipitation is one of the high-impact weather phenomena that strongly affects many aspects of our economy and general livelihood. Accurate forecasting of precipitation has long been a challenge to the research and operational communities (Fritsch et al. 1998; Fritsch and Carbone 2004). A precipitation system can exhibit a wide range of spatial and temporal scales depending on the environment. The prediction of small-scale, intermittent, and convection-dominated precipitation is most challenging because of its inherent low predictability, resulting in small tolerance of forecast skill to the timing and location errors of forecasting (Sun et al. 2014). In recent decades, there have been great efforts devoted
to the improvement of quantitative forecasting of convective precipitation, especially through the application of high-resolution numerical weather prediction (NWP) systems. Among many factors that could influence the accuracy of convective precipitation prediction using an NWP system, the specification of a well-balanced and dynamically consistent initial state is critical. This is because the initial state dominates forecasts of model variables and the development of precipitation (Sokol and Zacharov 2012; Sun et al. 2012).

Modern NWP systems typically have dedicated data assimilation (DA) systems to produce an optimal initial field. As a result of the rapidly evolving nature and short predictability of convective systems, rapid updates of current conditions (analyses) are required to improve the prediction and warning of high-impact weather, especially when radar observations are assimilated. Rapid updates are able to add information to the short-term model forecast and reduce the model error. In recent years, many major operational centers have upgraded their regional NWP systems to an update rate of 3 h (Sun et al. 2014). However, an hourly update frequency is highly desired in convective weather forecasts. Previous studies focusing on the rapid cycling of radar data observations (Sun et al. 2010; Hu et al. 2006a,b; Gao and Stensrud 2012) have shown some success in producing analyses with sub-hourly update rates. However, these studies are limited by using a simplified research model, incomplete datasets (exclude conventional data), or a very small research domain. The High Resolution Rapid Refresh (HRRR; Smith et al. 2008) rapid update assimilation and forecasting system is an operational hourly cycled model. However, the radar reflectivity is not assimilated in its three-dimensional variational DA (3DVAR) analysis stage, but rather blended into the initial conditions via a so-called diabatic digital filter initialization (DDFI) technique after the 3DVAR analysis (Weygandt et al. 2008). Therefore, it is still unclear how to effectively assimilate the radar and other large-scale observations with an hourly/subhourly update cycle in an operational configuration.

The first challenge of effectively assimilating radar observations in a rapid update system is how to extract information from observations made over different scales. It is known that different observation networks are designed for their corresponding targeted weather systems with different spatial and temporal scales. The operational Doppler radar is optimal for observing convective-scale weather systems. Its importance in numerically predicting convective precipitation has been well recognized (Lilly 1990; Wilson et al. 2010). Despite its ability to resolve the detailed structure of atmospheric convection with high spatial and temporal resolutions, the spatial coverage of radar observations is rather limited compared with conventional observation networks [Global Telecommunications System (GTS)]. Since large-scale atmospheric conditions frequently influence convective systems, assimilating large-scale\(^1\) information is also important in improving the forecast skill of convective systems (Lei et al. 2008; Schenkmans et al. 2011). Conventional observation networks have low spatial and/or temporal resolutions, but are able to resolve large-scale flow patterns and provide important background information for convective systems. It is noted that the variational DA technique has difficulties in obtaining analyses that fit “optimally” to observations representing different scales in one step that minimizes only one cost function (Xie et al. 2011). Therefore, it is important to design a multistep multiscale DA procedure that accounts for the different spatial and temporal scales of different observations. Previous studies (Dong et al. 2011; Xie et al. 2011) have reported promising results in assimilating observations of different scales. However, these studies are limited in the sense that incomplete (exclude radar data) or idealized datasets were used. More studies

\(^1\) Throughout this paper, the term large scale is used to stand for atmospheric motions greater than \(\sim 100\) km, while convective scale is used for those less than \(\sim 100\) km but greater than the scale that can be resolved by a model (which is dependent on the model and its resolution).
are necessary to understand the most effective way of assimilating observations of different scales within the context of operationally rapid updated convective forecasting systems.

Another challenge that arises from the convective-scale DA is the existence and accumulation of noise in a cycled 3DVAR system. The noise is partly attributed to the lack of convective-scale balance between model variables. The imbalances in the analysis could lead to the occurrence of spurious precipitation in the forecast (Huang and Lynch 1993). The problem is more severe in a cycled assimilation–forecast system because the noise and imbalance tends to accumulate in the cycles.

Sun et al. (2012) showed a case in which a 3-hourly cycled 3DVAR system had poor performance because of the erroneously forecasted convective system during the earlier time of the cycles. In some studies, digital filter initialization (DFI) was applied to remove high-frequency components from the initial conditions (Lynch and Huang 1992; Huang and Lynch 1993; Benjamin et al. 2004). However, DFI may also remove real atmospheric features and can only be implemented after the DA process. Lee et al. (2006) used the incremental analysis update (IAU) technique to reduce the impact of the noise in the analysis by gradually adding the increment during the forecast. But the IAU method is unable to completely remove the noise from the analysis.

This study describes and tests two strategies for improving the analysis of convective-scale DA and short-term precipitation forecasts in a 3DVAR system with hourly update cycles. The first strategy is to assimilate conventional observations and radar observations using a two-step procedure to account for their scale differences. The second strategy is to add a dynamical constraint in the cost function to reduce the noise in the 3DVAR analysis. A series of experiments are designed to investigate the impacts of these new strategies to the 3DVAR analysis and QPF using the Advanced Research version of the Weather Research and Forecasting Model (Skamarock and Klemp 2008) and its 3DVAR DA system (Barker et al. 2004). These experiments were conducted using a severe storm case with heavy precipitation that occurred over Colorado and Kansas during 8–9 August 2008. The results of these experiments are verified and analyzed to demonstrate the benefits of the strategies in improving the 3DVAR analysis as well as QPF. To demonstrate the robustness of these strategies, the averaged forecast skill from eight additional summer convective cases over the central plains is also presented.

The paper is organized as follows. In section 2, we describe the design strategies of the hourly rapid update 3DVAR system. The storm case of 8–9 August 2008 and the additional eight cases in 2014 are described in section 3. The experiments are summarized
FIG. 3. Mosaics of radar reflectivity of the lowest PPI scan from the NEXRAD stations in Cheyenne, WY (KCYS); Denver/Boulder, CO (KFTG); Pueblo, CO (KPUX); Dodge City, KS (KDDC); Riverton, WY (KRIW); Grand Junction, CO (KGJX); and Goodland, KS (KGLD) (see Fig. 4 for their locations) at (a) 2200 UTC 8 Aug, and (b) 0000, (c) 0200, (d) 0300, (e) 0500, and (f) 0700 UTC 9 Aug 2008.
in section 4 and their results are presented in section 5. Summary and discussions are given in section 6.

2. Description of the hourly update 3DVAR system

a. General description of WRF 3DVAR system

A developmental version of the WRF data assimilation (WRFDA) 3DVAR system with several refinements for convective-scale radar DA (Wang et al. 2013; Sun et al. 2016) is used in this study. The cost function in the updated 3DVAR system can be expressed in the following form:

$$J = J_b + J_{o}^{GTS} + J_{o}^{radar} + J_c,$$

where $J_b$ stands for the background term, and the two terms with the subscript $o$ stand for observation terms. The data assimilated in this study are conventional observations from the GTS dataset and the reflectivity and radial velocity from the operational WSR-88D weather radar network. As will be shown in the following sections, the GTS data and radar data are assimilated in two separate steps. Therefore, two observation terms are used in Eq. (1) to help our description of the DA procedure. The last term, $J_c$, represents the divergence constraint, which will be discussed in detail later in this section.

The updated WRFDA 3DVAR system adopted the $u$ and $v$ wind components as the momentum control variables in replacement of the streamfunction and velocity potential, which were used in most previous studies. The use of $u$ and $v$ winds allows closer fitting to high-resolution observations (e.g., radar radial velocity) and results in improved precipitation prediction (Sun et al. 2016). The same momentum control variables have been used in other variational DA systems with a primary goal of improving the convective-scale analysis (Zou et al. 1995; Sun and Crook 1997; Gao et al. 1999). Other control variables in the 3DVAR system include temperature $T$, surface pressure $P_s$, and pseudo–relative humidity $^2$ (RH). The radar radial velocity assimilation follows the procedure described in Xiao and Sun (2007). An indirect assimilation scheme (Wang et al. 2013) is used for the radar reflectivity in which the retrieved rainwater mixing ratio is assimilated instead of reflectivity.

b. The divergence constraint

An important improvement in this study is the adding of a divergence constraint to the updated WRFDA 3DVAR system. Sun et al. (2016) demonstrated that the use of $u$ and $v$ as control variables was superior to employing the streamfunction and velocity potential within the context of convective-scale DA and precipitation forecasts even without the cross-variable correlation. However, it is suggested that the addition of the cross-variable correlation via statistical analysis or physical constraint would further improve the analysis by reducing the analysis noise. Since it is difficult to obtain meaningful climatological correlation at the convective scale, the physical constraint method has drawn more attention recently. Vendrasco et al. (2015) showed that a physical constraint added to the 3DVAR cost function using large-scale analyses resulted in significant improvement of their convective-scale 3DVAR analysis and hence the precipitation forecast. Li et al. (2012)

---

2 Pseudo–relative humidity is defined as the water vapor mixing ratio divided by its saturated counterpart in the background field.
demonstrated that a dynamical constraint based on the steady-state momentum equation improved the forecast of a tropical cyclone. In this study, a divergence constraint is added to the cost function through the \( J_c \) term in Eq. (1):

\[
J_c = \frac{1}{2} R_{\text{div}} \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right)^2 ,
\]

(2)

Where \( R_{\text{div}} \) is a weighting coefficient that controls the relative importance of this term. Since only the

<table>
<thead>
<tr>
<th>Start time</th>
<th>Radar cycle frequency (h)</th>
<th>GTS/radar Length scale (km)</th>
<th>Variance scale</th>
<th>Divergence constraint?</th>
</tr>
</thead>
<tbody>
<tr>
<td>3HCYC</td>
<td>1200 UTC 8 Aug</td>
<td>3</td>
<td>120/120</td>
<td>D/D</td>
</tr>
<tr>
<td>1HCYC</td>
<td>1200 UTC 8 Aug</td>
<td>1</td>
<td>120/120</td>
<td>D/D</td>
</tr>
<tr>
<td>2STEP</td>
<td>1200 UTC 8 Aug</td>
<td>1</td>
<td>120/30</td>
<td>D/2D</td>
</tr>
<tr>
<td>2STEP00</td>
<td>0000 UTC 9 Aug</td>
<td>—</td>
<td>120/30</td>
<td>2D/2D</td>
</tr>
<tr>
<td>1STEP00_S</td>
<td>0000 UTC 9 Aug</td>
<td>—</td>
<td>120/120</td>
<td>D/D</td>
</tr>
<tr>
<td>1STEP00_L</td>
<td>0000 UTC 9 Aug</td>
<td>—</td>
<td>120/30</td>
<td>2D/2D</td>
</tr>
<tr>
<td>2STEP_DIV</td>
<td>1200 UTC 8 Aug</td>
<td>1</td>
<td>120/30</td>
<td>D/2D</td>
</tr>
</tbody>
</table>

TABLE 3. Assimilation configurations of the experiments. The letter (D) represents the value from the default BES.

Fig. 5. FSSs of hourly accumulated precipitation with thresholds of (a) 1 and (b) 5 mm for experiments 3HCYC and 1HCYC initialized at 0000 UTC 9 Aug 2008, and the corresponding biases with thresholds of (c) 1 and (d) 5 mm.
horizontal divergence is used, this constraint is strictly zero horizontal mass divergence. The main effect of this weak dynamical constraint is to couple the $u$ and $v$ wind components for the assimilation of radar radial velocity data and enforce smoothness where no radar observations are available. Since it is a weak constraint, zero mass divergence is not strictly satisfied. The weighting coefficient determines the amount of adjustment to the wind components and is empirically determined with the principle that the small-scale unrealistic divergence is suppressed while the large-scale divergence is maintained. A similar concept was used Gao et al. (1999) in radar wind retrieval. Hu et al. (2006b) implemented a similar constraint in the 3DVAR system to assimilate radar observations with very short assimilation periods (10 min). This study will focus on its impact in the hourly update WRFDA 3DVAR system with the purpose for employing it in more general applications.

c. Design strategies of a two-step procedure

A two-step assimilation procedure with different background error length scales, variance scales, and cycle frequencies is introduced in this study to optimally incorporate information from the GTS and radar observations. Since the variance and length scales determine the magnitude and spread of the analysis increment introduced by the observations, different values of variance and different length scales take into account the representative scales of different observations. Similar to the sequential assimilation procedure used in Xie et al. (2011), larger (smaller) scale data are assimilated in the first (second) step. Specifically, only the GTS data (without $J^{radar}_o$) are assimilated in the first step with the length ($\sim$120 km) and variance scales obtained using the National Meteorological Center (NMC, now known as NCEP) method [hereafter referred to as default background error statistics (BES)]. The analysis after the GTS assimilation is used as the background for the second step, in which only the radar data (without $J^{GTS}_o$) are assimilated with the length scale reduced by a factor of 0.25 and the variance scale doubled compared with those in the first step. The change in the length and variance scales in the second step is based on the consideration that the background field has larger uncertainties at the convective scale. The specific values used above are determined based on a series of trial-and-error experiments. The two-step procedure also takes into account the differences of the temporal frequencies between the GTS and radar observations by using different cycle frequencies. As shown in Table 1, the GTS observations have temporal frequencies of 3 h or above while the radar observations have a frequency less than 1 h. Therefore, the GTS observations are assimilated every 3 h, and the radar observations are assimilated every 1 h.

**FIG. 6.** The first hour ($t = 1$ h) forecasts of hourly accumulated precipitation (color; mm) valid at 0100 UTC 9 Aug 2008 for experiments (b) 3HCYC and (c) 1HCYC, with (a) the stage IV analysis shown for verification.
A schematic diagram of the two-step procedure is shown in Fig. 1 by assuming all cycles are initialized at 1200 UTC. In step 1, only the GTS data are assimilated to produce a 3DVAR analysis (3DV_GTS) using the 0.5° Global Forecast System (GFS) analysis of the National Centers for Environmental Prediction (NCEP) as the background. In step 2, only the radar observations are assimilated to produce another 3DVAR analysis (3DV_RD) using 3DV_GTS as the background. A 0–12-h forecast is then conducted using the 3DV_RD as the initial field. The assimilation of radar data (step 2) is repeated at 1300 and 1400 UTC using the WRF forecast at $t = 1$ h as the background. A similar two-step DA and forecast cycle is repeated every 3 h with the only difference being that the WRF forecast ($t = 1$ h) initialized at the previous hour is used as the background for the GTS data assimilation in step 1.

3. Experiment dataset

This study is part of a National Center for Atmospheric Research (NCAR) cross-laboratory project that aims to improve short-term high-resolution QPF for flash flood prediction. A severe convective system that produced a flash flood in Denver’s Cherry Creek on 8–9 August 2008 was chosen to show the details of the hourly update 3DVAR system. Figure 2 shows the overall weather pattern of this event. At 1200 UTC 8 August 2008, there was a synoptic-scale ridge over the mountain states (Wyoming and Colorado) at 500 hPa (Fig. 2a). Under the ridge, there
was a short-wave trough located in the northwest of Colorado, covering the state. At the lower level (Fig. 2b), the southeasterly flow in eastern Colorado and western Kansas is favorable for the development of the convective system through the advection of moisture from the south.

The evolution of the convective system is illustrated in Fig. 3 with radar reflectivity observations. At 2200 UTC 8 August 2008, the system was initiated in the southeast corner of Wyoming. It then organized into a squall line along a thin convergence line between 0000 and 0200 UTC. The system further propagated southeastward into Colorado with a northeast–southwest orientation. The intense, localized storm that caused the flash flood (marked by the arrow in Fig. 3c) can be seen at 0200 UTC (2000 local time). By 0500 UTC, a new convective line was initiated in northwestern Kansas and moved southeastward as the old system in Colorado weakened.

In addition to the verification and detailed analysis based on the above case, eight convective cases over the central plains of the United States in summer 2014 are used to demonstrate the robustness of the improvement in QPF after the implementation of the new technique. Table 2 lists the dates and main weather characteristics of the eight cases. These cases cover convective systems with different locations, sizes, and durations, indicating the potential applicability of the new techniques in operations.

4. Data assimilation experiments and verification methods

All assimilation and prediction experiments in this study employ a two-way, two-domain nested grid shown, as shown in Fig. 4. The outer (inner) domain has
211 × 160 (320 × 280) grids with 15- (3-) km horizontal grid spacing and 50 terrain-following levels in the vertical. The initial and lateral boundary conditions are interpolated from the GFS analyses and forecasts, and the BES results were calculated using the NMC method. The GTS data are assimilated on both the outer and inner domains while the radar observations are assimilated only across the inner domain. The radars used in the experiments are shown in Fig. 4. A preprocessing and quality control procedure similar to that in Sun (2005) and Lim and Sun (2010) is employed to process radar observations and specify observation errors. The radar data are thinned to a grid with 3-km spacing by averaging all observations in the same grid. The radar volumes closest to the top of the hour within a ±5-min window are used in the assimilation. Other physics options include the Kain–Fritsch cumulus parameterization scheme (Kain and Fritsch 1993) (only in the outer domain), the Thompson bulk microphysics scheme (Thompson et al. 2008), the Mellor–Yamada–Janjić (MYJ) PBL (Janjić 2002) scheme, the Noah land surface model (Ek et al. 2003), and the RRTMG radiation (Iacono et al. 2008) scheme. Details of the above schemes can be found in the WRF technical report (Skamarock et al. 2008).

Table 3 lists the details of each experiment. It is noted that experiments 3HCYC and 1HCYC do not include any of the improvements introduced in this study and serve as the benchmark for other experiments. The two experiments use the same default BES and other configurations, and only differ in their cycle frequencies. The experiment 2STEP is the same as 1HCYC but uses the two-step procedure described in section 2c and Fig. 1. The next three experiments (2STEP00, 1STEP00_S, and 1STEP00_L) are similar to 2STEP but with different combinations of length and variance scale values for the GTS and radar data. The three experiments are not cycled and only perform the 3DVAR DA at 0000 UTC 9 August 2008 using the same background from the GFS analysis. Since the same background is used, the differences in the analysis fields can only be attributed to the different DA procedures. So the three experiments are able to reveal the effects of the different procedures more clearly. Experiment 2STEP_DIV is the same as experiment 2STEP, except that the divergence constraint term is added to the cost function.

Since this study aims to improve short-term precipitation prediction, the neighborhood-based fractions skill score (FSS; Schwartz et al. 2009) is used to verify the skill of precipitation forecasts. The FSS is defined as
where \( P_f \) and \( P_o \) are the forecasted and observed fractional coverages of rainfall that exceed a given threshold within a radius (12 km is used in this study) centered at a grid point, and \( N \) gives the number of grid points in the verification domain. The hourly gridded stage IV precipitation analysis is used as truth for the verification. The FSS ranges from 0 to 1, with 1 (0) signifying a perfect (no skill) forecast. Since FSS is sensitive to bias (Mittermaier and Roberts 2010), the bias is also calculated as

\[
\text{bias} = \frac{a + b}{a + c},
\]

where \( a, b, \) and \( c \) stand for the hitting points, false alarm points, and missing points, respectively. Bias is greater (less) than 1 when the spatial precipitation coverage is overforecasted (underforecasted). In addition to the use of FSS and bias to verify the precipitation forecast, the variance of vertical velocity in nonprecipitation regions is used to examine the noise level in the analysis for some experiments. Moreover, verifications against radiosonde observations, surface observations, and aircraft reports in terms of the root-mean-square errors (RMSEs) of wind, temperature, and humidity are also performed for some experiments. All verifications are performed over the 3-km inner domain shown in Fig. 4.

5. Results

a. Comparison of the 3- and 1-h cycle control experiments

Before presenting the impact of the new strategies described in section 2, we first examine the performance of the hourly cycled configuration relative to its 3-hourly counterpart without using the new strategies. The FSS scores for the 12-h forecasts from 3HCYC and 1HCYC initialized at 0000 UTC 9 August 2008 are shown in Figs. 5a and 5b. Experiment 3HCYC has higher scores than those of 1HCYC for both the 1- and 5-mm thresholds from the 1st hour to the 10th hour. Figures 5c and 5d show that experiment 3HCYC has a lower bias than is found for 1HCYC for most hours for the 1- and 5-mm thresholds. The spatial distributions of predicted precipitation at forecast time \( t = 1 \) h verified against the stage IV data are shown in Fig. 6. Experiment 3HCYC was able to capture the major rainband located around the northern border of Colorado while the same rainband is disorganized in 1HCYC. Both

![Fig. 11. RMSEs of the 3DVAR analyses from experiments 1HCYC and 2STEP compared with (a) radiosondes, (b) METARs, and (c) aircraft reports.](image)

![Fig. 12. As in Fig. 8b, but for experiments 2STEP and 1HCYC.](image)
experiments show different amounts of false precipitation in the southeast part of Colorado. Since the observations assimilated at 0000 UTC are the same for both experiments, the differences in QPF must have been caused by the differences in the background. The quality of the two backgrounds at 0000 UTC for the two experiments (3-h forecast starting from 2100 UTC for 3HCYC and 1-h forecast starting from 2300 UTC for 1HCYC) is evaluated using the radiosonde, surface, and aircraft observations from the GTS dataset. It is shown (Fig. 7) that experiment 3HCYC has lower RMSEs than 1HCYC in all verified variables, indicating the potential for a better forecast.

The inferior background at 0000 UTC from experiment 1HCYC may be partially attributed to the accumulation of noise in the cycle process, because the hourly cycle has less time to achieve a balanced state between the model variables. It is hypothesized that the accumulation of noise is most evident in regions where no radar observations are available to correct the noise, and the noise could manifest itself as large nonphysical updrafts and downdrafts in the nonprecipitation region caused by unrealistic convergence and divergence. To verify this hypothesis, the vertical velocity variance over the regions without precipitation is shown in Fig. 8. The nonprecipitation region (Fig. 8a) is defined where the reflectivity value is lower than 5 dB\(\text{Z}\) in the reflectivity composite generated using all radar volumes from 1200 UTC 8 August to 0000 UTC 9 August 2008. The smaller threshold of 5 dB\(\text{Z}\), compared to the typical value of 12–15 dB\(\text{Z}\) employed in other studies, is used to exclude any
real updrafts/downdrafts associated with convection. Figure 8b shows that variance from experiment 1HCYC over the nonprecipitation region is greater compared to that from 3HCYC at most vertical levels.

The above analysis suggests that the accuracy of the analysis and subsequent forecast can be degraded by small-scale noise. The noise could come from different sources, including spurious gravity waves (oscillations), imbalances in the analysis, and model errors in physical processes, etc. In the following section, we will show that the strategies described in section 2 can improve the analysis and precipitation forecast by reducing the error/noise in the analysis.

b. The impact of the two-step assimilation

Figures 9a and 9b show a comparison of the FSSs for the 12-h forecasts from experiments 1HCYC and 2STEP. Experiment 2STEP has higher scores for most hours for the 1-mm threshold. The improvement in the 5-mm threshold is not as consistent as that of the 1-mm threshold and is only significant at certain forecast hours (the 1st, 2nd, 7th, 8th, 9th, and 12th hours). Meanwhile, as shown in Figs. 9c and 9d, experiment 2STEP has lower biases during the first 10 (8) hours for the 1-mm (5 mm) threshold. The verification of predicted hourly precipitation at $t = 1$ h (Fig. 10) shows that experiment 2STEP not only captures the major rainband in northern Colorado but also shows less false precipitation in southeastern Colorado. The 3DVAR analysis fields at 0000 UTC 9 August 2008 from experiments 2STEP and 1HCYC are evaluated against the GTS observations in the same way as that in experiments 3HCYC and 1HCYC. Figure 11 shows that the analysis fields from experiment 2STEP all have lower RMSEs,
indicating a better analysis than that from 1HCYC. The comparison of the vertical velocity variance in the non-precipitation region (Fig. 12) shows that experiment 2STEP has smaller vertical velocity variances at all levels, particularly in the upper levels, indicating that the two-step assimilation is effective in reducing noise.

To understand the impact of the two-step procedure in a clear way, the FSSs and bias scores for the 12-h forecasts from experiments 2STEP00, 1STEP00_L, and 1STEP00_S are shown in Fig. 13. Figures 13a and 13b show that 1STEP00_L has the lowest score, except for the 4–6-h period for the 5-mm threshold. Experiment 1STEP00_S has a high initial score, but the score drops quickly between t = 1 and 6 h, especially for the 5-mm threshold. Experiment 2STEP00 outperforms the other experiments except for the first hour for the 1-mm threshold. The three experiments all show relatively lower FSSs between 6 and 9 h because that was a transitional period when the first convective system gradually dissipated and the second convective system began to develop. The bias of the three experiments (Figs. 13c,d) shows that experiment 1STEP00_L has the highest bias for most hours, and experiment 1STEP00_S underpredicted the bias for the 5-mm threshold except during four hours (the 1st, 7th, 8th, and 9th hours). The spatial distributions of the predicted hourly precipitation at t = 2 and 6 h are compared with the stage IV data. At t = 2 h (Fig. 14), all three experiments capture the major convective band and the smaller storm that produced the Denver flash flood. The rainband in 1STEP00_L (Fig. 14c) is weaker and displaced slightly to the east. The forecasts of experiments 2STEP00 (Fig. 14b) and 1STEP00_S (Fig. 14d) improve the location of the rainband. At t = 6 h (Fig. 15), experiments 2STEP00 (Fig. 15b) and 1STEP00_L (Fig. 15c) captured the first rainband. The size of the rainband in 2STEP00 is closer to that in the observations, and there is less false precipitation in 2STEP00 compared with 1STEP00_L.
FIG. 16. The $u$, $v$ winds (vectors; m s$^{-1}$) at the 6th model level and $w$ wind (color; m s$^{-1}$) at the 12th model level at 0000 UTC 9 Aug 2008 for (a) the analysis of 2STEP00, (b) the increment of 2STEP00, (c) the analysis of 1STEP00_L, (d) the increment of 1STEP00_L, (e) the analysis of 1STEP00_S, and (f) the increment of 1STEP00_S.
despite some false alarms in southwest Nebraska. The forecast rainband is much weaker in 1STEP00_S (Fig. 15d).

To understand the different forecast results of the three experiments, Fig. 16 shows the wind fields in the analysis and the increment (analysis minus background) field. Experiment 2STEP00 (Figs. 16a,b) strengthens the confluence and the updrafts associated with the convective rainband (area A). In addition, the southerly flow (area B) advects warm, moist air from the south.

The lower-level confluence, which implies convergence, and moisture advection are favorable for the initiation and development of convection. In contrast to 2STEP00, the vertical velocity in area A (Fig. 16d) is disorganized in 1STEP00_L. The northerly increment in area B also weakens the moisture advection. The differences in the analysis of 1STEP00_L mainly results from the use of the same length and variance scales from the NMC BES for both the radar and GTS data. The information from the radar radial velocity is exaggeratedly spread out,
resulting in unrealistically large wind increments even in the area without radar observations.

Figures 17a–c show the temperature increment near the surface (the first model level above ground) from the three experiments. The 2STEP00 experiment (Fig. 17a) shows two cooling areas (areas A and C) and a warming area (area B). Compared with the radar observations at the same time, the cooling areas are located right behind the convective rainband and are likely the result of the evaporative cooling of precipitation. The small but strong cooling area C near the rainband with a northwesterly outflow is clearly a convective-scale feature associated with the convective system in northeastern Colorado. The large area of warm increments (area B; Fig. 17a) results from the GTS data assimilation with a longer length scale, capturing enhanced warm advection associated with the environment southeasterly flow, compared to 1STEP00_S (Fig. 17c). The cooling and warming dipole is consistent with the analysis of the horizontal wind pattern shown in Fig. 16a and is favorable for the initialization and development of convection. The temperature increments from 1STEP00_L (Fig. 17b) are similar to those of 2STEP00 but with weaker cooling. Both the warming (B) and cooling areas (A and C) in 1STEP00_S (Fig. 17c) become much weaker compared to those in 2STEP_00 and 1STEP00_L. The use of smaller length and variance scales severely limits the spread of temperature increments. Figures 17d–f and 17g–i show the convective available potential energy (CAPE) and the convection inhibition (CIN) fields.

Fig. 18. FSSs of hourly accumulated precipitation with thresholds of (a) 1 and (b) 5 mm for experiments 1HCYC, 2STEP, and 2STEP_DIV initialized at 0000 UTC 9 Aug 2008, and the corresponding bias with thresholds of (c) 1 and (d) 5 mm.
averaged over the first three levels above ground. The CAPE from 1STEP00_S (D in Fig. 17f) is smaller than for the other two experiments (D in Figs. 17d,e), and the CIN from 1STEP00_S (E in Fig. 17i) is higher than in the other two experiments (E in Figs. 17g,h). As a result of a less supportive environment, the convection is not well maintained and dissipates quickly in 1STEP00_S, as shown in Fig. 15d.

The above comparisons suggest that assimilating the GTS and radar observations into a single step is not practical for optimally extracting information from different observations. Worse still, the assimilation of radar data with improper length and variance scales can disturb the synoptic-scale balance in the background field, which could further cause additional noise in the analysis (i.e., Fig. 16d). With the two-step procedure, the large-scale (convective scale) information contained in the GTS (radar) data can be effectively incorporated into the analysis.

c. The impact of divergence constraint

Experiment 2STEP_DIV examines the effect of the divergence constraint on the hourly update DA system. Figures 18a and 18b compare the FSS scores for the 12-h forecasts from experiments 2STEP and 2STEP_DIV. For the 1-mm threshold, experiment 2STEP_DIV shows improvements over 2STEP from the second to the ninth hours. For the 5-mm threshold, experiment 2STEP_DIV outperforms 2STEP from the first hour to the seventh hour. The bias scores from experiment 2STEP_DIV show improvement over 2STEP in the first 6 h (Figs. 18c,d). Figure 19 compares the predicted hourly precipitation of experiments 2STEP_DIV and 2STEP at $t = 5$ h when 2STEP shows little improvement over 1HCYC. Although the old rainband in northeastern Colorado and the newly developed one in northwestern Kansas are both captured by 2STEP (Fig. 19b), there is spurious precipitation to the west of the old rainband (area A) and south of the new one (area B). The addition of the divergence constraint in 2STEP_DIV reduces the magnitude of the false precipitation in the corresponding areas (Fig. 19c).

Figure 20 shows the verification of the analysis fields from experiments 2STEP and 2STEP_DIV against the GTS observations at 0000 UTC 9 August 2008. The results show that the analyses from experiment 2STEP_DIV consistently have lower RMSEs compared to those of 2STEP. The dynamical effect of the divergence constraint is to couple the horizontal momentum variables through the mass continuity equation. Therefore, it is supposed that we reduce the small-scale noise in the analysis. The vertical velocity variance from the analysis fields of the two experiments in the nonprecipitation...
region are compared in Fig. 21. It is evident that experiment 2STEP_DIV has less variance at most levels, particularly in the upper levels, indicating that the divergence constraint is effective in reducing noise.

d. Multiple cases verification

The analyses presented in the previous sections demonstrated improved QPF after applying the new techniques and showed the details of how such improvements were achieved. To verify whether the improvements in the single case are robust, the same techniques and model configurations are used to rerun eight convective cases over the central plains of the United States during the summer of 2014. For each case, 12-h forecasts were initialized at 1500, 1800, 2100, 0000, and 0300 UTC for experiments 1HCYC, 2STEP, and 2STEP_DIV, and there are 40 sets of 12-h forecasts in total for each experiment.

The averaged FSSs of each experiment over the eight cases are shown in Figs. 22a,b. Experiment 2STEP shows clear improvement over experiment 1HCYC and further improvement is achieved by 2STEP_DIV. The averaged absolute deviation of bias from 1 is also smaller for 2STEP (2STEP_DIV) compared with 1HCYC (2STEP), which verifies that the improvement of FSS is not caused by the increase in bias. To further investigate the statistical significance of the differences of FSS, the $p$ values of a $t$ test between two pairs of experiments (1HCYC versus 2STEP and 2STEP versus 2STEP_DIV) are shown in Fig. 23. The $\alpha$ value (1.68) at the 95% significant level is denoted as the dashed gray line. The $p$ values above this gray line show that the improvement of FSS is statistically significant. The improvement of all forecasts passes the 95% confidence level except for the forecasts at $t = 1$ and 8 h for the 5-mm threshold between experiments 2STEP_DIV and 2STEP.

6. Summary and discussion

One challenge in assimilating convective-scale observations, such as those from Doppler radar, is in producing analyses that can closely fit to the observations while still maintaining a balanced large-scale background. This study implemented two strategies in a 3DVAR radar DA system with an hourly rapid update cycle to address such issues. The first strategy was to run the 3DVAR system
with a two-step assimilation procedure, in which the radar observations were assimilated with modified BES results in a second step after the GTS observations were assimilated. The second strategy was to add a divergence constraint to the cost function to mitigate the small-scale erroneous convergence/divergence and the resultant unrealistic updrafts/downdrafts. A set of experiments was performed on a convective weather system that produced a severe flash flood in Denver, Colorado, on 8–9 August 2008 to examine the impacts of the new strategies on the 3DVAR analyses and subsequent precipitation forecasts. Diagnoses of the experimental results were conducted by evaluating the analysis accuracy against the conventional observations and verifying the precipitation forecast skill. We also used the variance of the vertical velocity in the nonprecipitation region as a measure of the noise level.

We first conducted two experiments to demonstrate that, without these strategies, the hourly cycled 3DVAR produced degraded precipitation forecasts compared to its 3-hourly counterpart as a result of the accumulation of noise in the forecast background. We further showed that the use of the two-step assimilation procedure and the divergence constraint significantly improved the forecast skill of the hourly cycled 3DVAR analysis and forecast. It was found that the two-step assimilation procedure is able to better construct relevant dynamical structures both for the storm environment, such as the warm advection associated with a southeasterly flow, and at the convective scale, such as the low-level convergence and cold pool.
As a result of these improvements in the analysis field, the two-step assimilation substantially improved the skill of QPF over the one-step assimilation for the 12-h forecast range. The divergence constraint was able to reduce the analysis noise in the nonprecipitation region and enhance the dynamical balance in momentum variables. The more balanced initial conditions resulted in improved precipitation forecasts in the first 8–9 h. The new techniques are further applied to eight convective cases in 2014, and the results from 40 sets of 0–12-h forecasts show effective and robust improvements.

Although some encouraging results have been presented in this study, there are still some detailed problems related to multistep assimilation that require in-depth investigation. For example, the surface data, which represent the small-scale and rapid evolution in the boundary layer, were assimilated in the same step as other GTS observations. Studies are being conducted to examine the characteristics of the surface data and the optimal technique for assimilating the high-frequency observations from automated weather stations. Another limitation of the current study is that the background error statistics were adjusted empirically for the second-step radar DA. A more objective method, for example, based on ensemble forecasts, should be developed to estimate the uncertainties of the flow-dependent forecast background. Despite these limitations, this study provides a prototype for operational short-term prediction of local convective weather systems.

Acknowledgments. This work was performed while WT was visiting the National Center for Atmospheric Research (NCAR). WT is grateful for the support provided by the Mesoscale Microscale Meteorology (MMM) and the Research Application Laboratory (RAL) of NCAR during the visit. This work is funded by CINRAD Variational Assimilation Program of China Meteorological Administration Meteorological Observation Center (20110266); NCAR’s Short Term Explicit Prediction (STEP) Program, which is supported by the National Science Foundation funds for the U.S. Weather Research Program (USWRP); the China Special Fund for Meteorological Research in the Public Interest (Grant GYHY201506004); and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

REFERENCES


