



AMERICAN METEOROLOGICAL SOCIETY

Monthly Weather Review

EARLY ONLINE RELEASE

This is a preliminary PDF of the author-produced manuscript that has been peer-reviewed and accepted for publication. Since it is being posted so soon after acceptance, it has not yet been copyedited, formatted, or processed by AMS Publications. This preliminary version of the manuscript may be downloaded, distributed, and cited, but please be aware that there will be visual differences and possibly some content differences between this version and the final published version.

The DOI for this manuscript is doi: 10.1175/MWR-D-18-0033.1

The final published version of this manuscript will replace the preliminary version at the above DOI once it is available.

If you would like to cite this EOR in a separate work, please use the following full citation:

Liu, C., M. Xue, and R. Kong, 2018: Direct Assimilation of Radar Reflectivity Data using 3DVAR: Treatment of Hydrometeor Background Errors and OSSE Tests. *Mon. Wea. Rev.* doi:10.1175/MWR-D-18-0033.1, in press.

© 2018 American Meteorological Society



1

2

3

4 **Direct Assimilation of Radar Reflectivity Data using 3DVAR: Treatment of**
5 **Hydrometeor Background Errors and OSSE Tests**

6

7 Chengsi Liu¹, Ming Xue^{1,2} and Rong Kong²

8 ¹Center for Analysis and Prediction of Storms and ²School of Meteorology

9 University of Oklahoma, Norman Oklahoma 73072

10

11

12

13 Submitted to Monthly Weather Review

14 December 2017

15 Revised August, November 2018

16

17 Corresponding Author:

18 Ming Xue

19 Center for Analysis and Prediction of Storms

20 120 David Boren Blvd

21 Norman OK 73072

22

Abstract

Despite the well-known importance of background error covariance in data assimilation, its impact on the assimilation of radar reflectivity within a three-dimensional variational (3DVar) framework has been little studied. In this study, it is shown that unphysical analysis increments of hydrometeors are produced when using vertically homogeneous background error variance. This issue cannot be fully solved by using the so-called hydrometeor classification in the reflectivity observation operator. Alternatively, temperature-dependent background error profiles for hydrometeor control variables are proposed. With such a treatment, the vertical background error profiles are specified to be temperature dependent, allowing for more physical partitioning of radar-observed precipitation information among the liquid and ice hydrometeors.

The 3DVar analyses using our treatment are compared with those using constant background error or “hydrometeor classification” through observing system simulation experiments with a simulated supercell storm. Results show that 1) 3DVar with constant hydrometeor background errors produces unphysical rainwater at the high levels and unphysical snow at the low levels; 2) the hydrometeor classification approach reduces unphysical rainwater and snow at those levels but the analysis increments are still unphysically spread in the vertical by the background error covariance when the vertically invariant background errors are used; 3) the temperature-dependent background error profiles enable physically more reasonable analyses of liquid and ice hydrometeors from reflectivity assimilation.

46 **1. Introduction**

47 Although much progress has been made in the area of convective-scale data
48 assimilation (DA) and numerical weather prediction (NWP), obtaining accurate initial
49 conditions for convective-scale NWP remains a major challenge. A major issue for convective-
50 scale DA is the lack of sufficient and direct observations within thunderstorms and their
51 environment. Radar reflectivity data, which have higher spatial and temporal resolutions than
52 conventional observations and contain hydrometeor information, are the most important data
53 for convective-scale DA. Complex cloud analysis (Hu et al. 2006), ensemble Kalman filter
54 (EnKF) (e.g. Tong and Xue 2005), 3D variational (3DVar) (Gao and Stensrud 2012) DA
55 methods are currently the three main types of methods for assimilating radar reflectivity for
56 convective-scale models that contain ice microphysics, while 4D variational (4DVar) method
57 has mainly be used with the assumptions of warm rain microphysics (e.g., Sun and Crook 1997).

58 Complex cloud analysis schemes adjust thermodynamic and hydrometeor variables
59 based on radar reflectivity data, satellite and surface observations etc. (Albers et al. 1996).
60 Cloud analysis is generally computationally fast and relatively easy to implement, and has been
61 shown to enable the buildup of reasonable 3D cloud structures and precipitation fields and quite
62 effectively alleviate the precipitation spinup problem (Hu et al. 2006; Kain et al. 2010; Sun et
63 al. 2014; Xue et al. 2003). However, cloud analysis algorithms typically rely on empirical
64 relationships that have many uncertainties. Moreover, such methods do not utilize background
65 and observation error information, and hence cannot obtain statistically optimal estimations.

66 Variational methods can also be used to assimilate reflectivity data; optimal analyses
67 are obtained by minimizing a cost function in which the contributions of the background and
68 observations are weighted by their respective error covariances. Four-dimensional variational
69 (4DVar) method uses the forecast model as a strong constraint in the cost function so that the

70 analysis can be consistent with the forecast model. For the convective scale, the microphysical
71 processes are important and complex ice microphysics schemes are highly nonlinear, which
72 creates many practical issues that are difficult to overcome (Xu 1996), like nonlinearity
73 associated with the highly nonlinear physics. For this reason, existing 4DVar systems that
74 assimilate radar reflectivity data generally assume warm rain microphysics and convert
75 reflectivity to rainwater mixing ratio before assimilation (Sun and Crook 1997; Wang et al.
76 2013b).

77 EnKF uses an ensemble of forecasts to estimate the background error covariances. Not
78 requiring tangent linear or adjoint models of the observation operators, EnKF can more easily
79 use nonlinear observation operators (of, e.g., radar reflectivity) and models with complex ice
80 microphysics. The ability of EnKF to assimilate radar data has been shown by many studies
81 (Aksoy et al. 2009, 2010; Jung et al. 2008a; Jung et al. 2008b; Tong and Xue 2005; Xue et al.
82 2006). While EnKF can estimate a flow-dependent background error covariance, the estimated
83 covariance matrix is often severely affected by the sampling and model errors (Houtekamer
84 and Mitchell 1998; Houtekamer et al. 2009). The hybrid method, which combines the use of
85 ensemble-derived flow-dependent background error covariance with static covariance within a
86 variational framework, requires the capability of variationally assimilating reflectivity data
87 (Hamill and Snyder 2000; Kong et al. 2018).

88 Apart from being needed for constructing a hybrid DA system, using 3DVar to
89 assimilate radar data is also more practical due to its relatively low computational cost. In the
90 MM5 and later WRF 3DVar systems, Xiao et al. (2007) proposed a procedure to directly
91 assimilate reflectivity although only the liquid phase of hydrometeors was considered, to
92 simplify the problem. The more recent implementation of reflectivity assimilation within WRF
93 variational DA systems retrieves hydrometeor mixing ratios prior to assimilation (Wang et al.

94 2013a).

95 When considering ice microphysics in the absence of model constraint (either in the
96 form of a strong constraint as in 4DVar or in the form of ensemble error covariances), 3DVar
97 usually has difficulties in properly partitioning radar-observed precipitation information among
98 multiple hydrometeors, and often produces unphysical analyses of hydrometeors. For example,
99 rainwater can be placed way above the freezing level even without strong updraft while snow
100 may be erroneously analyzed way below the melting level. One approach to avoid this problem
101 is to modify the reflectivity observation operator so that it is temperature-dependent. In Gao
102 and Stensrud (2012), reflectivity is made to be a function of rainwater only below 5 °C level,
103 and a function of snow and hail only above the -5 °C level. Between these two levels, the
104 equivalent reflectivity is made to be a weighted average of the contributions from liquid and
105 ice hydrometeor species, with the weights being linearly dependent on temperature. This
106 approach is called “hydrometeor classification” by Gao and Stensrud (2012).

107 One of the problems with the hydrometeor classification approach is the effective
108 disallowance of hail at levels below 5 °C. On the other hand, when the background error
109 variances of the hydrometeors are assumed to be vertically invariant, as is the case in Gao and
110 Stensrud (2012), unphysical analysis increments of hydrometeors can be produced by their
111 vertical spreading through background error covariance. In addition, with the formula used by
112 Gao and Stensrud (2012), at 0 °C temperature, the contributions of liquid and ice hydrometeors
113 to the equivalent reflectivity are unphysically halved.

114 In ensemble-based DA system, due to the use of flow-dependent background error
115 covariance, the hydrometeors analyses will be physically consistent with forecast model, and
116 the hydrometeor classification approach would not be necessary. When the 3DVar framework
117 is used to construct an ensemble-variational hybrid system (Kong et al. 2018), it is also

118 preferred that the reflectivity operator used in the hybrid system is the same as that used in the
119 coupling EnKF system; in another word, ad hoc modification to the observational operator in
120 the hydrometeor classification approach is undesirable.

121 In this study, we propose an alternative approach to deal with the 3DVar reflectivity
122 assimilation problem discussed above. As opposed to using spatially uniform errors as in Gao
123 and Stensrud (2012), our new approach introduces additional information or constraint to the
124 analysis system by defining temperature-dependent background error covariances for the
125 hydrometeor variables. The impact of the temperature-dependent error profiles on the 3DVar
126 analyses of a supercell storm is examined through observing system simulation experiments
127 (OSSEs), by comparing the results with those using the constant error profiles with or without
128 the hydrometeor classification treatment.

129 This paper is organized as follows: In section 2, the proposed temperature-dependent
130 background error profiles for the hydrometeor variables are described. The design of three
131 3DVar experiments, that use, respectively, the proposed background error profiles, constant
132 error profiles with and without the hydrometeor classification, is presented in section 3. In
133 section 4, we compare the OSSE results from the three experiments. Conclusions are given in
134 section 5.

135 **2. Temperature-dependent background error profiles**

136 In the Advance Regional Prediction System (ARPS) 3DVar system (Gao et al. 2004;
137 Xue et al. 2003) , the default background error variances for hydrometeor mixing ratios are
138 assumed to be homogeneous. With such error variances, unphysical analysis increments, such
139 as rainwater at the upper levels and snow far below the freezing level, can be produced. To
140 avoid the unphysical hydrometeor analysis increments, the so-called “hydrometeor
141 classification” algorithm was proposed in Gao and Stensrud (2012), in which the original

142 reflectivity observation operator based on Lin et al. (1983) was modified to be temperature-
 143 dependent.

144 The reflectivity factor Z can be simulated based on model mixing ratios of rainwater
 145 (q_r), snow (q_s) and hail (q_h), following Tong and Xue (2005):

$$146 \quad Z = 10 \log_{10}(Z_e), \quad (1)$$

$$147 \quad Z_e = Z_{er}(q_r) + Z_{es}(q_s) + Z_{eh}(q_h), \quad (2)$$

148 where Z_{er} , Z_{es} and Z_{eh} are the contributions to equivalent reflectivity (Z_e) by rainwater, snow
 149 and hail, respectively.

150 The equivalent reflectivity factors contributed by rainwater is calculated, based on Smith
 151 et al. (1975), from

$$152 \quad Z_{er} = \frac{10^{18} \times 720 (\rho q_r)^{1.75}}{\pi^{1.75} N_r^{0.75} \rho_r^{1.75}} = 3.63 \times 10^9 \times (\rho q_r)^{1.75}, \quad (3)$$

153 where $\rho_r = 1000 \text{ kg} \cdot \text{m}^{-3}$ is the density of rainwater and ρ (in $\text{kg} \cdot \text{m}^{-3}$) is the density of
 154 air; $N_r = 8.0 \times 10^6 \text{ m}^{-4}$ is the intercept parameter in the assumed Marshall-Palmer
 155 exponential raindrop size distribution.

156 If the temperature is less than 0°C , the component of reflectivity for dry snow is

$$157 \quad Z_{es} = \frac{10^{18} \times 720 K_i^2 \rho_s^{0.25} (\rho q_s)^{1.75}}{\pi^{1.75} K_r^2 N_s^{0.75} \rho_i^2} = 9.80 \times 10^8 \times (\rho q_s)^{1.75}. \quad (4)$$

158 Here $\rho_s = 100 \text{ kg} \cdot \text{m}^{-3}$ is the density of snow and $\rho_i = 917 \text{ kg} \cdot \text{m}^{-3}$ is the density of
 159 ice, $N_s = 3.0 \times 10^6 \text{ m}^{-4}$ is the intercept parameter for snow, and $K_i^2 = 0.176$ is the
 160 dielectric factor for ice while $K_r^2 = 0.93$ is the same for water. Wet snow, which occurs at
 161 temperatures higher than 0°C , is treated in a similar way as rainwater, and the equivalent
 162 reflectivity factor-mixing ratio relation is

$$163 \quad Z_{es} = \frac{10^{18} \times 720 (\rho q_s)^{1.75}}{\pi^{1.75} N_s^{0.75} \rho_s^{1.75}} = 4.26 \times 10^{11} \times (\rho q_s)^{1.75}. \quad (5)$$

164 For hail, the wet hail formulation of Smith et al. (1975) is used:

$$165 \quad Z_{eh} = \left(\frac{10^{18} \times 720}{\pi^{1.75} N_h^{0.75} \rho_h^{1.75}} \right)^{0.95} (\rho q_h)^{1.6625} = 4.33 \times 10^{10} \times (\rho q_h)^{1.66}, \quad (6)$$

166 where $\rho_h = 913 \text{ kg} \cdot \text{m}^{-3}$ is the density of hail. The hail intercept parameter of $N_h = 4.0 \times$
 167 10^4 m^{-4} is used, which is slightly larger than that used in Smith et al. (1975), implying more
 168 occurrences of hail at smaller sizes.

169 Based on the consideration that rainwater should generally exist below the freezing level,
 170 and snow and hail should generally exist above the freezing level, the reflectivity operator is
 171 modified to the following form in Gao and Stensrud (2012):

$$172 \quad Z = \begin{cases} 10 \log_{10} Z_{er} & T_b \geq 5^\circ \text{C} \\ 10 \log_{10} (Z_{es} + Z_{eh}) & T_b \leq -5^\circ \text{C} \\ 10 \log_{10} [\alpha Z_{er} + (1 - \alpha)(Z_{es} + Z_{eh})] & -5^\circ \text{C} > T_b > 5^\circ \text{C}, \end{cases} \quad (7)$$

173 where T_b is the background temperature in Celsius, and weighting coefficient a varies linearly
 174 between 0 and 1 as a function of temperature ranging between -5 to 5°C . With this modified
 175 reflectivity observation operator, the sensitivity of reflectivity to snow and hail below the 5°C
 176 level, and that to rainwater above the -5°C level are disallowed. As a result, in the absence
 177 of background error correlation, the assimilation of reflectivity will not generate rainwater
 178 above the -5°C level, or snow and hail below the 5°C level. If homogeneous background error
 179 covariance is used, however, as in the default scheme of ARPS 3DVar, undesirable hydrometers
 180 at these levels will appear. Moreover, the above treatment does not consider the situation of
 181 hail falling below the freezing level.

182 Another problem using the modified observation operator with the homogenous
 183 background error covariance is that the reflectivity data assimilation will be unable to remove
 184 spurious rainwater in the background at the upper levels, and more importantly unable to

185 remove spurious snow and hail at the lower levels when assimilating reflectivity data because
 186 the 3DVar system cannot detect such spurious hydrometeors since they do not appear in the
 187 reflectivity observation operator). As pointed out earlier, in the transition zone between -5 and
 188 $5^{\circ}C$, the equivalent reflectivity contributions from co-existing liquid and ice phases are
 189 excessively reduced, and in fact, to half of the correct values at $0^{\circ}C$.

190 Our proposed solution to the problem is, as mentioned earlier, to define temperature-
 191 dependent background error profiles for the hydrometeor variables. For each of the profiles,
 192 the errors are defined separately in three vertical sections divided by temperature thresholds.
 193 The error values in the top and bottom sections are still constant but in different magnitudes.
 194 The value in the middle sections increases or decreases from the bottom-section value to top-
 195 section value based on a hyperbolic tangent function, according to:

$$196 \quad E = \begin{cases} E_h & T < T_h \\ (E_h + E_l) / 2 - \frac{(E_h - E_l)}{2 \times \tanh(2\alpha)} \tanh \left[\frac{2\alpha}{1-a} (i-a) \right] & T_h \leq T \leq T_l \\ E_l & T > T_l \end{cases}, \quad (8)$$

197 where E is the background error magnitude and T the temperature. E_h and E_l are constant errors
 198 for the high and low levels with temperature smaller than T_h and greater than T_l , respectively. i
 199 is the temperature of certain level in the transition zone and n is the temperature difference
 200 between high level and low level. The variable “ a ” is equal to $(1+n)/2$ and α is a tuning
 201 parameter which can be a value between 0.2 and 5.0, and is set as 1.0 in this paper. In real data
 202 assimilation experiments, the background error profiles can be determined statistically using
 203 data from precipitating columns, e.g., with a form of the NMC method (Parrish and Derber
 204 1992). In our idealized experiments presented in this paper, the error magnitudes at the high
 205 and low levels are estimated by the root mean square errors (RMSEs) with the mean calculated
 206 over all the cycles in the OSSEs. For rainwater, T_h and T_l are set as $-5^{\circ}C$ and $5^{\circ}C$,

207 respectively, while for snow and hail, they are set as -30°C and 5°C , respectively. The lower
208 temperature of the upper threshold helps prevent discontinuity that can appear in the analysis
209 increments of snow and hail. The parameters in Eq. (8) used in this study are given in Table 1
210 and the error profiles are shown in Fig. 1.

211 **3. Experimental design**

212 In this study, a classic supercell storm that occurred on 20th May 1977 in Del City is
213 simulated using the ARPS model (Xue et al. 2001) to serve as a truth simulation of the OSSEs.
214 The domain size is $35\times 35\times 35$, with 2 km horizontal and 0.5 km vertical grid spacings. The
215 environment of the supercell is defined by a modified real sounding while the truth storm
216 simulation is initialized by a thermal bubble. More information about the truth simulation can
217 be found in Tong and Xue (2005).

218 The simulated radial velocity (V_r) and reflectivity (Z) are produced by a pseudo-radar
219 located at the southwest corner of the model domain. The radial velocity operator based on the
220 model velocity is written as

$$221 \quad V_r = u\sin\phi\cos\mu + v\cos\phi\cos\mu + w\sin\mu, \quad (9)$$

222 where μ and ϕ are the elevation and azimuth angles of radar beams, and u , v , and w are the
223 model-simulated velocities interpolated to the scalar grid points.

224 Random noise is added to the simulated V_r and Z observations given by Eqs. (9) and
225 (1), respectively. The noise is sampled from Gaussian distributions with zero mean and
226 standard deviations of 1 m s^{-1} for V_r and 3 dBZ for Z . Both radial velocity and reflectivity
227 observations are assimilated by the ARPS 3DVar at 5 minute intervals in the one-hour DA
228 window. The first analysis is performed at the 25 minutes of the truth simulation. Similar to
229 Kong et al. (2018), the initial background is the ensemble mean of the 5-minute ensemble

230 forecasts initiated from the ensemble perturbations valid at the 20-minute of the truth
231 simulation. The background error decorrelation scales are 3.5 km in the horizontal and 1.1 km
232 in the vertical for all analysis variables, based on the optimal scales obtained in Kong et al.
233 (2018) for similarly configured 3DVar OSSEs. To help suppress spurious reflectivity in the
234 background, clear-air reflectivity is also assimilated (Tong and Xue 2005).

235 Three experiments are conducted using 3DVar with (1) the invariant background error
236 profiles (CTRL), (2) the invariant background error profiles and hydrometeor classification in
237 the reflectivity observation operator (3DVarHC), and (3) the new temperature-dependent
238 background error profile (3DVarTD). Another experiment, 3DVarHCS, which is the same as
239 3DVarHC, except that a much smaller vertical background error decorrelation scale (100 m) is
240 used. The experiment results are presented in the next section.

241 **4. Experimental results**

242 *a. Results of the first analysis*

243 *1) Comparison of total reflectivity analyses*

244 Figure 2 shows the reflectivity and wind vector fields in an east-west vertical cross section
245 through the maximum vertical velocity of the simulated truth storm at the time of first analysis
246 (25 minutes), from the truth simulation, background and three DA experiments. Compared with
247 the simulated truth, the reflectivity in the background is weak and has a wide horizontal
248 coverage unlike that of a deep convective cell. Also, there is some spurious reflectivity away
249 from the main storm. After the first assimilation of radar data, the analyzed reflectivity fields
250 in CTRL, 3DVarHC, and 3DVarTD are much closer to the truth, and the spurious reflectivity
251 in the east of the main storm is partly suppressed. However, the intensity of the reflectivity core
252 is much underestimated in all three analyses, and the vertical motion in the main storm region

253 is still mostly downward. Some new spurious reflectivity is also produced in supposedly clear
254 air regions. It is a known problem of direct variational assimilation of reflectivity that large
255 gradients of cost function can be produced in areas of small background hydrometeor mixing
256 ratios, which prevents efficient minimization convergence (Sun and Crook 1997). Figure 2 also
257 shows that the vertical velocity is not properly analyzed within the updraft region. In the ARPS
258 3DVAR system whose static background error does not contain cross covariance between wind
259 and hydrometeor fields, the wind field can only be updated by the radial velocity data. We have
260 found that when using hydrometeor mixing ratios as the control variables and when analyzing
261 reflectivity and radial velocity data simultaneously, the assimilation of radial velocity data is
262 ineffective, and special treatments are needed to circumvent this problem which will be discussed
263 in a separate paper. Because of the focus on reflectivity assimilation in this paper, this issue is
264 unimportant for this study. Through the assimilation cycles, rather accurate vertical velocity does
265 develop later on (not shown).

266 Figure 2 also indicates that the reflectivity analyses from CTRL, 3DVarHC, and 3DVarTD
267 are generally similar. Since both the “hydrometeor classification” and temperature-dependent
268 background error profiles aim to better partition the radar-observed precipitation information
269 among the hydrometeors, and the cost function minimization tries to minimize the simulated
270 reflectivity from observed values, the analyzed reflectivity is expected to be relatively close to
271 the observed reflectivity values. The key is in the accuracy of the analyzed hydrometeor fields,
272 which are not directly observed.

273 2). *Comparison of analyzed hydrometeor fields*

274 When multiple hydrometeors species contribute to the reflectivity, and only reflectivity is
275 observed, the problem of directly assimilating reflectivity data for the purpose of determining
276 several hydrometeor variables is under-determined. Therefore even if the analyzed total

277 reflectivity fits well to its observations, the individual hydrometeor fields may not be correctly
278 analyzed. The “hydrometeor classification” treatment of Gao and Stensrud (2012), and our use
279 of temperature-dependent background errors both try to introduce additional constraints into
280 the 3DVar system. To evaluate the true quality of analyses, the analyzed rainwater, snow and
281 hail mixing ratio fields, and their corresponding contributions to reflectivity in dBZ are shown
282 in Fig. 3. The contributions of rainwater (Z_r), snow (Z_s) or hail (Z_h) are defined as the
283 logarithmic equivalent reflectivity from the individual hydrometeors, e.g. $Z_r = \log_{10}(Z_{er})$.

284 In the truth simulation, non-zero rainwater (snow) mixing ratio mainly exists below
285 (above) the melting level (which is roughly at 4 km height), as shown in Fig. 3a and 3b. For
286 hail (Fig. 3c), there are significant hailstones below 5 km and the magnitude of low-level Z_h is
287 over 65 dBZ because hailstones can reach the ground before fully melt. While hail below the
288 5°C level is ignored by the hydrometeor classification scheme according to Eq. (7), it is
289 allowed when using the temperature-dependent background error profiles defined in Eq. (8).

290 In the forecast background (Fig. 3d, 3e and 3f), because the storm is much weaker than
291 the truth, the hydrometeor mixing ratios are also much less. In the background, little hail is
292 shown below the freezing level (Fig. 3f), while rainwater (snow) exists only below (above) the
293 freezing level (Fig. 3d and Fig. 3e).

294 (i) Analyzed hydrometeor fields from CTRL

295 With the assimilation of radar data (mainly of reflectivity data which have direct impact on
296 the hydrometeor fields in the first analysis while the effects of V_r data assimilation have to be
297 accumulated through 3DVar DA cycles), the analyzed hydrometeor distributions in CTRL
298 become narrower than in the background, and closer to the truth (Figs. 3g, 3h, 3i); the
299 magnitudes are also much increased. However, the partitioning of reflectivity observation

300 information across the hydrometeors is not quite right. In Fig. 3h, snow mixing ratio and
301 corresponding reflectivity below the melting level are much larger than those above the melting
302 level, which is obviously unphysical. The values at the lower levels are apparently large
303 because the coefficient of the equivalent reflectivity equation for wet snow (4.26×10^{11} in the
304 Eq. (5)) is much larger than that for dry snow (9.80×10^8 in the Eq.(4)). The larger is the
305 coefficient in the Z_e equation, the more sensitive Z_e is to the corresponding hydrometeor mixing
306 ratio, where the adjoint sensitivity is the gradient of Z_e with respect to the mixing ratio. Because
307 variational minimization tends to adjust most sensitive variables the most (e.g. using algorithms
308 having properties of steepest descent), the adjustment to snow mixing ratio below the freezing
309 level (where snow is expected to be wet) will be much larger than to snow above the freezing
310 level, leading to much more snow mixing ratio at the lower levels (Fig. 3h).

311 Since the same background error of 0.6 g kg^{-1} is used for all hydrometeor mixing ratios in
312 CTRL, the partition of the reflectivity observation information is almost exclusively controlled
313 by the sensitivity to (gradient with respect to) the hydrometeors in the reflectivity observation
314 operator. Below the melting level, because the reflectivity in the operator is most sensitive to
315 wet snow, the analyzed reflectivity component from snow is the largest in CTRL (Fig. 3h),
316 however such large values below the freezing level do not exist in the true simulation at all
317 (Fig. 3b). Above the melting level, some unphysical analysis increments of rainwater are also
318 found (Fig. 3g). In Fig. 3i, some hail analysis increments are found at the low levels but the
319 values in the truth simulation are much larger; this is because most of the adjustment had been
320 made to snow at the lower level, resulting in too small adjustments to hail mixing ratio.

321 (ii) Analyzed hydrometeor fields from 3DVarHC

322 In the “hydrometeor classification” version of reflectivity observation operator,

323 contributions of ice hydrometeors below the $5^{\circ}C$ level and rainwater above the $-5^{\circ}C$ level are
324 excluded, so that they cannot be directly produced by the reflectivity observation within the
325 3DVar analysis. Compared with the CTRL analysis below the melting level (Fig. 3g), rainwater
326 mixing ratio from 3DVarHC is much larger (Fig. 3j) because only rainwater is included in the
327 reflectivity operator at those levels. Similarly, the snow analysis increment at the upper levels
328 is increased than in CTRL because of the absence of competition from rainwater (Fig. 3k). Fig.
329 3k also shows that the large unphysical snow increments at the low levels in CTRL are greatly
330 suppressed due to the lack of direct link to snow in the reflectivity operator at those levels, but
331 snow below the melting level is still significantly overestimated. The unphysical rainwater
332 increment at the high levels in CTRL is correctly suppressed (Fig 3j).

333 However, if we compare the analyzed rainwater and snow fields in 3DVarHC with the truth
334 fields in more details, we still notice unphysical spreading of rainwater above the freezing level
335 (Fig. 3j), and unphysical spreading of snow below the freezing level, some of which even
336 reaches the ground (Fig. 3k). Another issue with “hydrometeor classification” is the exclusion
337 of hail below the $5^{\circ}C$ level, which results in small analysis increment at the low levels for hail
338 produced only by the vertical spreading via the background error correlation (Fig. 3l)..

339 One way to reduce the unphysical vertical spreading of hydrometeor analysis increments
340 is to adjust the vertical background error decorrelation scale. This is done in experiment
341 3DVarHCS where the vertical scale is reduced from 1.1 km to 100 m. The 100 m is smaller
342 than the vertical grid spacing of 500 m in our experiment, therefore vertical spreading should
343 be virtually eliminated. This choice of a very small decorrelation scale is mainly for the purpose
344 of illustrating the effects of vertical increment spreading.

345 With the 100 m vertical decorrelation scale, the unphysical vertical spreading of analysis
346 increments is much reduced; for example, there is no wet snow below 2 km in 3DVarHCS

347 (Fig.3n). However, within regions of storms, there should be vertical error correlations where
348 updrafts and downdrafts are deep and strong (see, e.g., Tong and Xue 2005). Kong et al. (2018)
349 found via OSSEs that the optimal vertical decorrelation scale for 3DVar for this case should be
350 about 1 km. The vertical decorrelation scale of 100 meters in this experiment is obviously too
351 short to obtain optimal analyses. Moreover, the issue of significantly overestimating snow in
352 the melting layer still exists in 3DVarHCS (Fig. 3n) because of very large sensitivity to wet
353 snow in the melting layer; the hydrometeor classification treatment dose not eliminate this issue.

354 *(iii) Analyzed hydrometeor fields from 3DVarTD*

355 In the temperature-dependent background error profiles, the background error of snow
356 below the $5^{\circ}C$ level and of rainwater above the $-5^{\circ}C$ level are set to zero, which means that
357 the analysis system will very much trust the background. This is obviously based on physical
358 conditions. When error statistics are obtained from model predicted fields, similar temperature
359 dependences are expected, although the actual statistical distribution is left for future studies.
360 With this setting of 3DVarTD, non-physical analysis increments of snow (rainwater) below
361 (above) the melting level are no longer produced (Fig 3q and Fig. 3p). Because the physical
362 constraint is considered via background error covariance profiles, proper vertical spreading of
363 hydrometeor analysis increments can be taken into account via vertical variations of the error
364 magnitudes.

365 In CTRL, 3DVarHC and 3DVarHCS, wet snow is significantly over-estimated below the
366 melting level (Fig. 3h, 3k and 3n) because the cost function sensitivity to snow is large where
367 the snow is expected to be wet. In contrast, the background error variance of snow is much
368 smaller below the melting level in 3DVarTD (see Fig. 1) so that the analysis increment of wet
369 snow is much smaller (Fig. 3q), alleviating the over-estimation problem.

370 For hail, significant magnitudes of analysis increments are found below the melting level
371 (Fig. 3r), much closer to the truth than in other experiments; this is attributed to the fact that
372 hail and its error variance is considered in the background error profile of hail (Fig. 1).

373 Overall, the spatial patterns and magnitudes of the analyzed rainwater, snow and hail
374 mixing ratios and corresponding reflectivity fields are much better reproduced by the single-
375 time 3DVar analysis using our proposed background error profiles than those using a
376 homogeneous error covariance. While the analyses using the “hydrometeor classification”
377 observation operator are improved over CTRL, some issues are still shown. Given that only a
378 single time analysis is performed, the analyzed hydrometeor fields are generally weaker than
379 the truth, which is expected.

380 *b. Results of cycled analyses*

381 The previous section presents the results from the first data assimilation cycle. What we
382 also want to know is the quality of the analyses and forecasts through the rest of the DA cycles
383 when the storm structures are more accurately established. For this purpose, we examine the
384 errors of the analysis and forecast hydrometeor fields in all DA cycles, and the fields at the end
385 of the cycles.

386 The RMSEs of the wind components and hydrometeor mixing ratios are calculated in
387 regions where reflectivity exceeds 10 dBZ in the true simulation. As shown in Fig. 4, the
388 RMSEs of all experiments generally decrease with the forecast-analysis cycle. In addition, the
389 RMSEs from the three experiments are very close except for the snow mixing ratio. The
390 RMSEs of snow mixing ratio from 3DVarTD are lower than those of CTRL and 3DVarHC.
391 Moreover, the analysis RMSEs of CTRL and 3DVarHC at the first analysis time are actually
392 increased by the data assimilation, mainly because unphysical snow increments are created at
393 the lower levels as discussed earlier and shown in Figs. 3h and 3k.

394 To further examine the differences in the analyzed hydrometeor fields from CTRL,
395 3DVarHC and 3DVarTD after one hour of DA cycles, the rainwater, snow, and hail mixing
396 ratios and the corresponding reflectivity components at the last cycle are shown in Fig. 5 in a
397 vertical cross section through the maximum vertical velocity of the truth simulation. Like the
398 first time analyses discussed in section 4a, the unphysical snow mixing ratio at the low levels
399 and the unphysical rainwater mixing ratio at the high levels are still shown in the analyses of
400 CTRL and 3DVarHC after one hour of DA cycles.

401 Figure 5a shows that the rainwater in the truth simulation can reach above 7 km because
402 of the strong updraft in the storm. The rainwater analyses in both CTRL and 3DVarTD are
403 much weaker than that of truth simulation while 3DVarHC has better rainwater intensity
404 analysis since snow and hail are excluded in the reflectivity operator at the lower levels (Fig.
405 5d, g and j). Because the background error of rainwater above the freezing level is set to zero,
406 no rainwater analysis increment above the freezing level can be directly produced in 3DVarTD.
407 Further tuning to the rainwater error profile can be made, by allowing some rainwater error
408 above the freezing level. The presence of super-cooled liquid water does depend on the updraft
409 intensity, which could be accounted for by flow-dependent background error covariance from
410 ensemble forecasts in a hybrid system developed recently for ARPS 3DVar (Kong et al. 2018).
411 The snow analysis at the high levels of CTRL is clearly underestimated (Fig 5e) while the
412 unphysical wet snow at the low levels of 3DVarHC (Fig 5h) becomes more prominent than that
413 of the first analysis. For 3DVarTD, the snow analysis at the high levels (Fig. 5k) is slightly
414 underestimated but no unphysical wet snow at the low levels is produced. For hail analysis in
415 3DVarHC (Fig. 5i), there are spurious hail below the over-hanging reflectivity anvil at the low
416 levels, which is apparently due to inappropriate analysis increment spreading associated with
417 the constant background error variance and vertical spatial correlation. It is noted that the

418 reflectivity core at the low levels in 3DVarTD is still under-estimated even after DA cycling. We
419 have found that this is because the gradient of the cost function is much smaller (larger) where
420 background reflectivity is large (small) so that the hydrometeors near reflectivity core are not
421 sufficiently adjusted by observations, especially when clear-air reflectivity is assimilated. This
422 issue is related to the nonlinearity of reflectivity operator, and separate treatments to alleviate this
423 problem will be the topic of a separate paper.

424 Overall, the analyzed hydrometeor fields at the end of the one-hour radar DA cycles
425 remain the most accurate and physical in the experiment that uses our proposed temperature-
426 dependent background error profiles for the hydrometeors. Issues found in the first 3DVar
427 analysis with the “hydrometeor classification” treatment and with the constant background
428 errors remain at the end of the one-hour DA cycles.

429 **5. Summary and conclusions**

430 To effectively and directly assimilate radar data, especially radar reflectivity data, using
431 ARPS 3DVar, temperature-dependent background error profiles are proposed for the
432 hydrometeor state variables associated with liquid- and ice-phase microphysics to better
433 partition radar-observed precipitation information among hydrometeors. The proposed method
434 is compared via OSSEs with another approach proposed by Gao and Stensrud (2012) that
435 modifies the reflectivity observation operator so that the involvement of the hydrometeor state
436 variables in the operator is temperature dependent, a treatment that they refer to as
437 “hydrometeor classification”.

438 Simulated radial velocity and reflectivity observations are assimilated every 5 minutes
439 over a 60-minute period for a simulated classic supercell storm. Three DA experiments are
440 performed using 3DVar with invariant hydrometeor background error profiles (CTRL), with
441 an additional “hydrometeor classification” treatment in the reflectivity observation operator

442 (3DVarHC), and with the new temperature-dependent background error profiles and the
443 original reflectivity operator (3DVarTD). The analyzed fields after the first 3DVar analysis and
444 at the end of one hour DA cycles are examined, together with the RMSEs of model state
445 variables through the DA cycles. Major conclusions are summarized as follows:

446 1) All experiments are able to produce total reflectivity analyses that fit the truth
447 reflectivity reasonably well. This is not surprising since reflectivity is observed and directly
448 assimilated by the 3DVar. In contrast, significant differences are found with the analyzed
449 hydrometeor fields.

450 2) When using constant background errors for the hydrometeors in CTRL, unphysical
451 hydrometeor analyses are produced, including the presence of snow at the ground level for this
452 summer storm. Wet snow below the freezing level is seriously overestimated due to its larger
453 backscattering coefficient than rainwater and hail in the reflectivity observation operator. Dry
454 snow at the high levels is seriously underestimated due to its smaller backscattering coefficient
455 and the analysis of rainwater at the high levels.

456 3) In 3DVarHC, by employing temperature dependence of the reflectivity operator,
457 unphysical analysis increment of rainwater above -5°C level and of snow below 5°C level
458 cannot be directly produced from the reflectivity observation. However, because of the
459 homogeneous vertical background error correlation used, the analysis increment of the
460 rainwater is unphysically spread to the high levels and the analysis increment of snow is
461 unphysically spread to the low levels. In addition, the exclusion of hail in the observation
462 operator below 5°C level leads to underestimated hail at the low levels.

463 4) When using the newly introduced temperature-dependent background error profiles,
464 unphysical hydrometeor increments produced from reflectivity observations and unphysical
465 vertical spreading when using homogenous background errors are much reduced. The analysis

466 of snow mixing ratio is improved most over both CTRL and 3DVarHC.

467 Further improvement may be possible when the background error profiles of the
468 hydrometeors are obtained statistically, from for example convective scale model forecasts
469 using a procedure similar to the NMC method. The error statistics should be calculated using
470 data from precipitation regions only, and the statistics should be stratified according to
471 temperature, and may need to consider other factors such as precipitation cycles. This can be a
472 topic for future study. We also note that even with the procedure proposed in this study, the
473 3DVar system used in this study for hydrometeor analysis is univariate for the hydrometeors
474 and does not fully utilize model equation constraints. Significant improvement is expected
475 when ensemble-derived flow-dependent multi-variate ensemble covariances are introduced
476 into the variational framework, to formulate a hybrid ensemble-variational system. However,
477 even in such a system, a proper and best possible treatment of the static background error
478 covariance is still desirable. In fact, the proposed temperature-dependent error profiles are used
479 in a recent 3DVar, En3DVar (ensemble 3DVar) and EnKF inter-comparison study reported in
480 Kong et al. (2018).

481 *Acknowledgments* This research was primarily supported by the NOAA Warn-on-Forecast
482 (WoF) grant NA160AR4320115. The second author was also supported by NSF grants AGS
483 0941491 and AGS-1046171. Computational resources of the Oklahoma Supercomputing
484 Center for Research and Education (OSCER), and the NSF Xsede Supercomputing Centers
485 were used.

486 **REFERENCE**

487 **Uncategorized References**

- 488 Aksoy, A., D. C. Dowell, and C. Snyder, 2009: A Multicase Comparative Assessment of the
489 Ensemble Kalman Filter for Assimilation of Radar Observations. Part I: Storm-Scale
490 Analyses. *Mon Weather Rev*, **137**, 1805-1824.
- 491 ———, 2010: A Multicase Comparative Assessment of the Ensemble Kalman Filter for
492 Assimilation of Radar Observations. Part II: Short-Range Ensemble Forecasts. *Mon*
493 *Weather Rev*, **138**, 1273-1292.
- 494 Albers, S. C., J. A. McGinley, D. L. Birkenheuer, and J. R. Smart, 1996: The local analysis
495 and prediction system (LAPS): Analyses of clouds, precipitation, and temperature.
496 *Weather Forecast*, **11**, 273-287.
- 497 Gao, J., and D. J. Stensrud, 2012: Assimilation of reflectivity data in a convective-scale,
498 cycled 3DVAR framework with hydrometeor classification. *J. Atmos. Sci.*, **69**, 1054-
499 1065.
- 500 Gao, J., M. Xue, K. Brewster, and K. K. Droegemeier, 2004: A three-dimensional variational
501 data analysis method with recursive filter for Doppler radars. *J. Atmos. Oceanic Tech.*,
502 **21**, 457-469.
- 503 Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter - 3D variational
504 analysis scheme. *Mon. Wea. Rev.*, **128**, 2905-2919.
- 505 Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman
506 filter technique. *Mon. Wea. Rev.*, **126**, 796-811.
- 507 Houtekamer, P. L., H. L. Mitchell, and X. Deng, 2009: Model error representation in an
508 operational ensemble Kalman filter. *Mon. Wea. Rev.*, **137**, 2126-2143.
- 509 Hu, M., M. Xue, and K. Brewster, 2006: 3DVAR and cloud analysis with WSR-88D level-II

510 data for the prediction of Fort Worth tornadic thunderstorms. Part I: Cloud analysis
511 and its impact. *Mon. Wea. Rev.*, **134**, 675-698.

512 Jung, Y. S., G. F. Zhang, and M. Xue, 2008a: Assimilation of simulated polarimetric radar
513 data for a convective storm using the ensemble Kalman filter. Part I: Observation
514 operators for reflectivity and polarimetric variables. *Mon Weather Rev*, **136**, 2228-
515 2245.

516 Jung, Y. S., M. Xue, G. F. Zhang, and J. M. Straka, 2008b: Assimilation of simulated
517 polarimetric radar data for a convective storm using the ensemble Kalman filter. Part
518 II: Impact of polarimetric data on storm analysis. *Mon Weather Rev*, **136**, 2246-2260.

519 Kain, J. S., and Coauthors, 2010: Assessing advances in the assimilation of radar data within
520 a collaborative forecasting-research environment. *Wea. Forecasting*, **25**, 1510-1521.

521 Kong, R., M. Xue, and C. Liu, 2018: Development of a hybrid en3DVar data assimilation
522 system and comparisons with 3DVar and EnKF for radar data assimilation with
523 observing system simulation experiments. *Mon. Wea. Rev.*, **146**: 175-198.

524 Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a
525 cloud model. *J. Climat. Appl. Meteor.*, **22**, 1065-1092.

526 Parrish, D. F., and J. C. Derber, 1992: The National-Meteorological-Centers Spectral
527 Statistical-Interpolation Analysis System. *Mon Weather Rev*, **120**, 1747-1763.

528 Smith, P. L., Jr., C. G. Myers, and H. D. Orville, 1975: Radar reflectivity factor calculations
529 in numerical cloud models using bulk parameterization of precipitation processes. *J.*
530 *Appl. Meteor.*, **14**, 1156-1165.

531 Sun, J., and N. A. Crook, 1997: Dynamical and microphysical retrieval from Doppler radar
532 observations using a cloud model and its adjoint. Part I: Model development and
533 simulated data experiments. *J. Atmos. Sci.*, **54**, 1642-1661.

534 Sun, J., and Coauthors, 2014: Use of NWP for nowcasting convective precipitation: Recent
535 progress and challenges. *Bulletin of the American Meteorological Society*, **95**, 409-
536 426.

537 Tong, M. J., and M. Xue, 2005: Ensemble Kalman filter assimilation of Doppler radar data
538 with a compressible nonhydrostatic model: OSS experiments. *Mon Weather Rev*, **133**,
539 1789-1807.

540 Wang, H., J. Sun, S. Fan, and X.-Y. Huang, 2013a: Indirect assimilation of radar reflectivity
541 with WRF 3D-Var and its impact on prediction of four summertime convective
542 events. *J. Appl. Meteor. Climatol.*, **52**, 889–902.

543 Wang, H., J. Sun, X. Zhang, X.-Y. Huang, and T. Auligné, 2013b: Radar data assimilation
544 with WRF 4D-Var: Part I. System development and preliminary testing. *Mon. Wea*
545 *Rev.*, **141**, 2224–2244.

546 Xiao, Q., Y.-H. Kuo, J. Sun, W.-C. Lee, D. M. Barker, and E. Lim, 2007: An approach of
547 radar reflectivity data assimilation and its assessment with the inland QPF of typhoon
548 Rusa (2002) at landfall. *J. Appl. Meteor. Climatology* **46**, 14-22.

549 Xu, Q., 1996: Generalized adjoint for physical processes with parameterized discontinuities -
550 Part I: Basic issues and heuristic examples. *J. Atmos. Sci.*, **53**, 1123-1142.

551 Xue, M., M. J. Tong, and K. K. Droegemeier, 2006: An OSSE framework based on the
552 ensemble square root Kalman filter for evaluating the impact of data from radar
553 networks on thunderstorm analysis and forecasting. *J Atmos Ocean Tech*, **23**, 46-66.

554 Xue, M., D.-H. Wang, J.-D. Gao, K. Brewster, and K. K. Droegemeier, 2003: The Advanced
555 Regional Prediction System (ARPS), storm-scale numerical weather prediction and
556 data assimilation. *Meteor. Atmos. Phy.*, **82**, 139-170.

557 Xue, M., and Coauthors, 2001: The Advanced Regional Prediction System (ARPS) - A multi-

558 scale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model
559 physics and applications. *Meteor. Atmos. Phys.*, **76**, 143-165.
560
561

562

563

564 Table 1: The parameters in the background error profiles for rainwater, snow and hail mixing
565 ratios.

566

	q_r	q_s	q_h
T_h (°C)	-5	-30	-30
T_l (°C)	5	5	5
E_h (k kg⁻¹)	0	1.2	0.6
E_l (k kg⁻¹)	0.8	0	0.3

567

568

569 **List of figures**

570 FIG. 1. The temperature-dependent background error profiles of the mixing ratio of rainwater
571 (black), snow (blue) and hail (red).

572 FIG. 2. The reflectivity (shading) and wind vectors in a vertical cross section through the
573 maximum vertical velocity of simulated truth storm at the time of first analysis (25 min), for
574 (a) the truth simulation, (b) the background, (c) the CTRL, (d) 3DVarHC and (e) 3DVarTD
575 experiments.

576 FIG. 3. The rainwater (1st column), snow (2nd column) and hail (3rd column) mixing ratios
577 (contours starting at 0.5 g kg⁻¹ and at intervals of 0.5 g kg⁻¹) and corresponding reflectivity
578 (shading) in a vertical cross section through the maximum vertical velocity of simulated truth
579 storm at the first analysis time, from truth simulation (1st row), background (2nd row), CTRL
580 (3rd row), 3DVarHC (4th row), 3DVarHCS (5th row), and 3DVarTD (6th row).

581 FIG. 4. The RMSEs of analyses and forecasts during the DA cycles for CTRL (black),
582 3DVarHC (blue), and 3DVarTD (red), for different state variables as labelled.

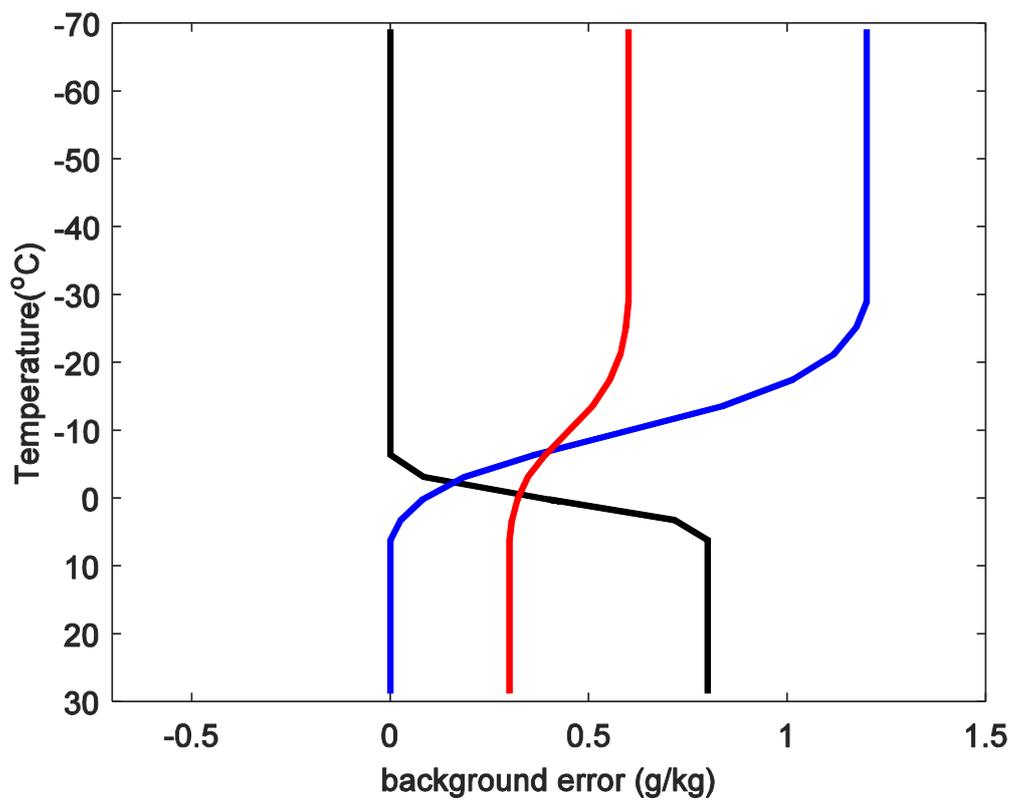
583 FIG. 5. The rainwater (1st column), snow (2nd column) and hail (3rd column) mixing ratios
584 (contours starting at 0.5 g kg⁻¹ and at intervals of 0.5 g kg⁻¹) and the corresponding
585 reflectivity (shading) in a vertical cross section through the maximum vertical velocity of the
586 simulated truth storm at the end of DA cycles for truth simulation (1st row), CTRL (2nd row),
587 3DVarHC (3rd row), and 3DVarTD (4th row).

588

589

590

591



592

593 FIG. 1. The temperature-dependent background error profiles of the mixing ratio of rainwater
594 (black), snow (blue) and hail (red).

595

596

597

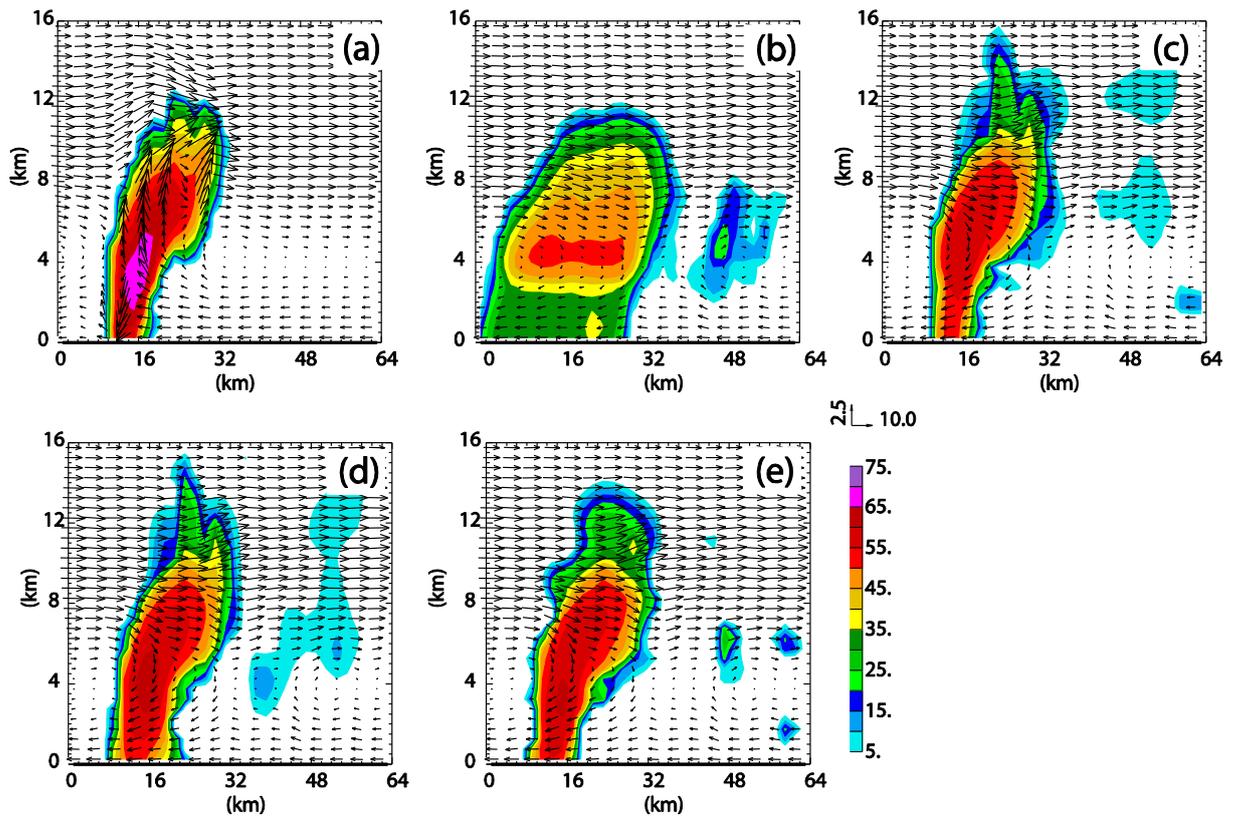
598

599

600

601

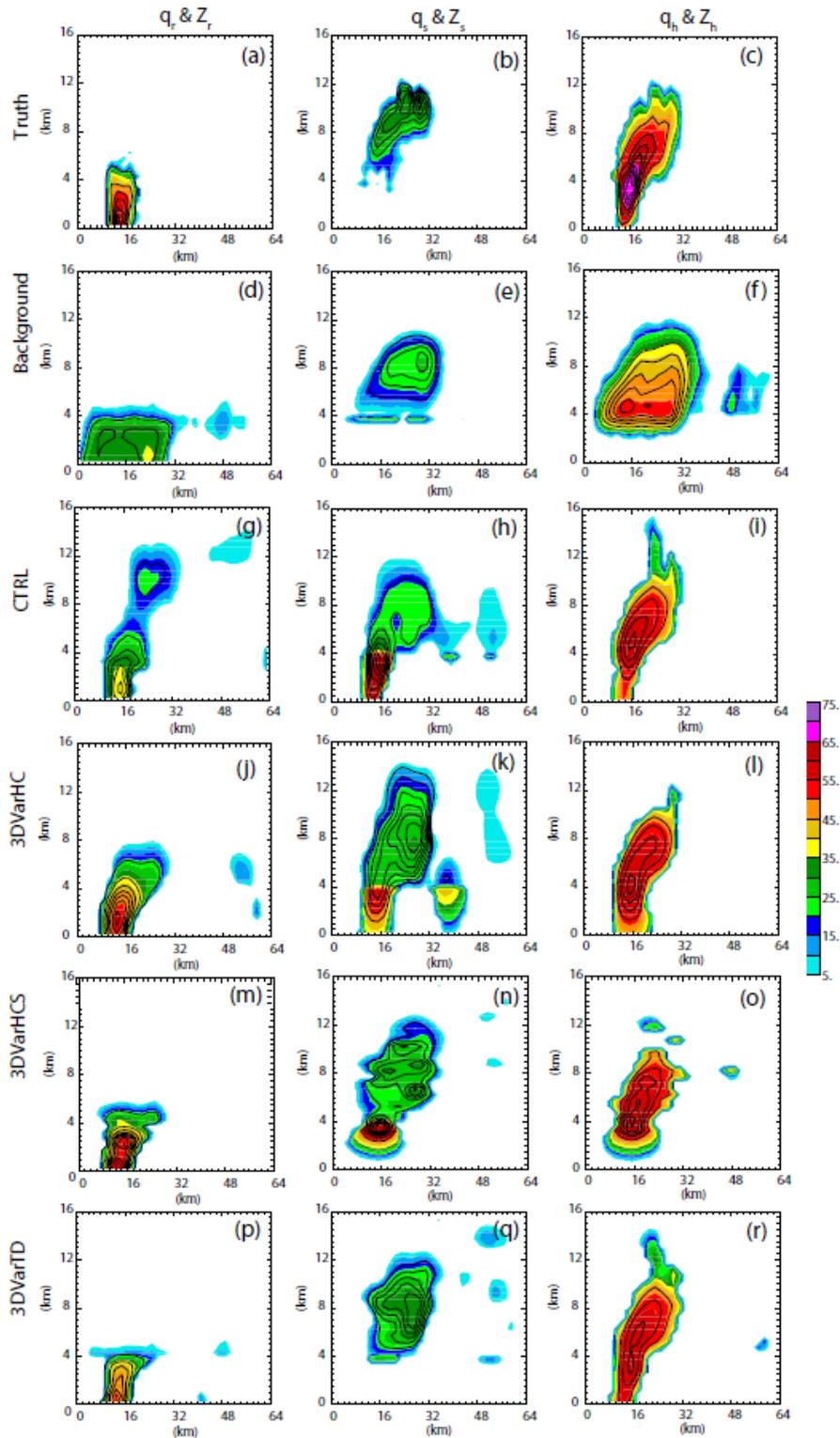
602



603

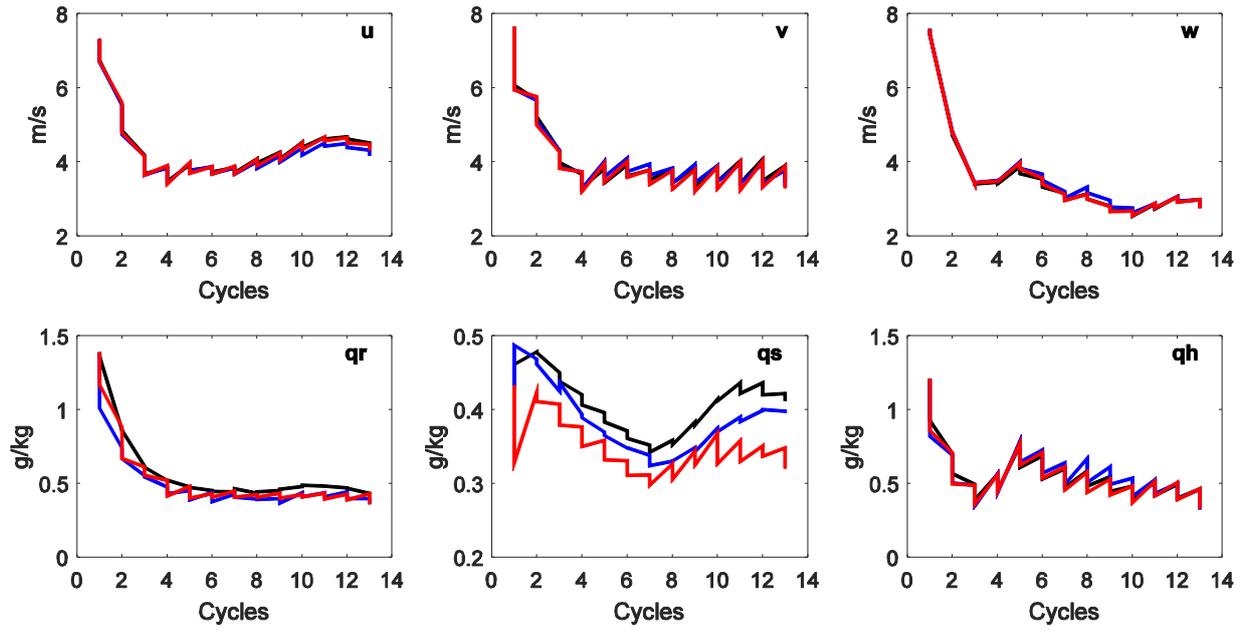
604 FIG. 2. The reflectivity (shading) and wind vectors in a vertical cross section through the
605 maximum vertical velocity of simulated truth storm at the time of first analysis (25 min), for
606 (a) the truth simulation, (b) the background, (c) the CTRL, (d) 3DVarHC and (e) 3DVarTD
607 experiments.

608

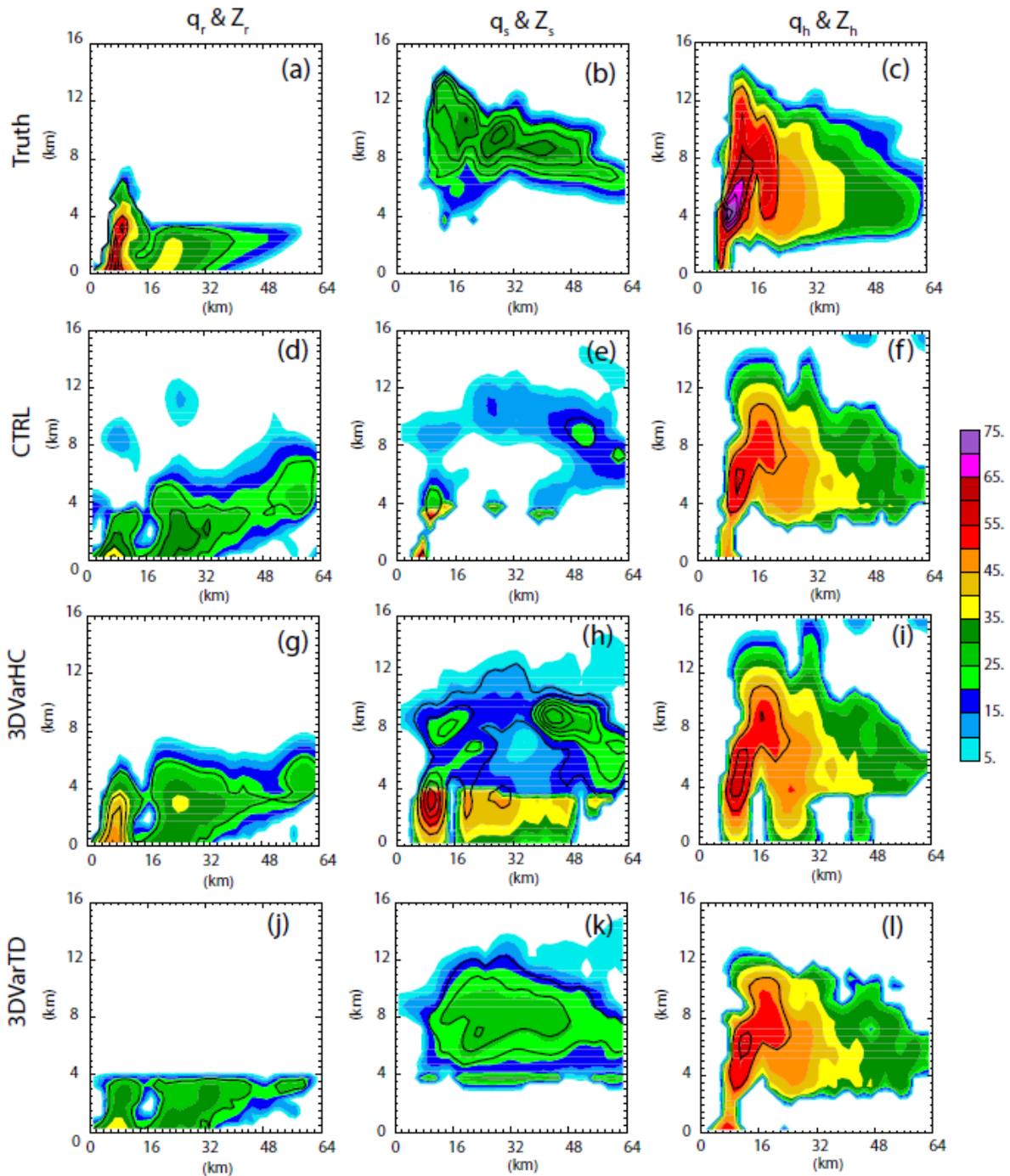


609

610 FIG. 3. The rainwater (1st column), snow (2nd column) and hail (3rd column) mixing ratios
 611 (contours starting at 0.5 g kg^{-1} and at intervals of 0.5 g kg^{-1}) and corresponding reflectivity
 612 (shading) in a vertical cross section through the maximum vertical velocity of simulated truth
 613 storm at the first analysis time, from truth simulation (1st row), background (2nd row), CTRL
 614 (3rd row), 3DVarHC (4th row), 3DVarHCS (5th row), and 3DVarTD (6th row).



616 FIG. 4. The RMSEs of analyses and forecasts during the DA cycles for CTRL (black),
 617 3DVarHC (blue), and 3DVarTD (red), for different state variables as labelled.
 618



619

620

621

622

623

624

625

626

FIG. 5. The rainwater (1st column), snow (2nd column) and hail (3rd column) mixing ratios (contours starting at 0.5 g kg^{-1} and at intervals of 0.5 g kg^{-1}) and the corresponding reflectivity (shading) in a vertical cross section through the maximum vertical velocity of the simulated truth storm at the end of DA cycles for truth simulation (1st row), CTRL (2nd row), 3DVarHC (3rd row), and 3DVarTD (4th row).