

1 **Assimilating Polarimetric Radar Data with an Ensemble Kalman Filter:**
2 **OSSEs with a Tornadoic Supercell Storm Simulated with a Two-Moment**
3 **Microphysics Scheme**

4
5 Kefeng Zhu¹, Ming Xue^{1,2}, Kun Ouyang¹ and Youngsun Jung²
6

7 ¹*Key Laboratory of Mesoscale Severe Weather/Ministry of Education and School of Atmospheric*
8 *Sciences, Nanjing University, Nanjing, 210093, China*
9

10 ²*Center for Analysis and Prediction of Storms and School of Meteorology, University of*
11 *Oklahoma 73072, United States*
12

13 May 2019

14 Submitted to Quarterly Journal of the Royal Meteorological Society

15 Revised November 2019
16

17 Corresponding author address:

18 Ming Xue

19 Center for Analysis and Prediction of Storms

20 University of Oklahoma

21 120 David Boren Blvd., Norman, OK 73072

22 mxue@ou.edu
23

24 **Abstract**
25

26 The impact of assimilating additional differential reflectivity Z_{DR} data from a polarimetric radar
27 on the analysis of a tornadoic supercell storm using an ensemble Kalman filter (EnKF) is studied in
28 an observing system simulation experiment (OSSE) framework assuming a perfect forecast model.
29 A double-moment microphysics scheme is used to allow for proper simulation of polarimetric
30 signatures. Root-mean square errors of analyzed state variables are calculated and the structure
31 and intensity of analyzed fields and derived quantities are examined. Compared to the baseline
32 experiment assimilating radial velocity and reflectivity only, the assimilation of additional Z_{DR}
33 further reduces the errors of all state variables. The analyzed hydrometeor fields are improved in
34 both pattern and intensity distributions. Polarimetric signatures including Z_{DR} and K_{DP} columns,
35 and Z_{DR} arc in the supercell are much better reproduced.

36 Sensitivity experiments are performed that exclude the updating of hydrometeor number
37 concentrations by Z_{DR} or of state variables not directly linked to Z_{DR} via observation operators. The
38 results show that if number concentrations are not updated together with the mixing ratios, most
39 of the benefit of assimilating Z_{DR} is lost. Among other state variables, the updating of water vapor
40 mixing ratio q_v has the largest positive impact while the impact of updating vertical wind w comes
41 in second. The updating of horizontal wind components or temperature has much smaller but still
42 noticeable impact. Reliable flow-dependent cross-covariances among the state variables and
43 observation prior as derived from the forecast ensemble and used in EnKF are clearly very
44 beneficial.

45 **Key points:**

- Z_{DR} assimilation improves the analyses for nearly all state variables of a simulated supercell storm.
- The updating of water value mixing ratio q_v has the largest impact followed by the updating of vertical velocity w .
- The hydrometeor total number concentrations should be updated together with the mixing ratios within EnKF.

1. Introduction

The forecast accuracy of high-resolution numerical weather prediction (NWP) models highly depends on the model initial state, especially for short-lived convective storms; the accuracy of initial microphysics (MP) state variables is key to successful short-range forecasting of precipitating systems (Sun et al. 2013). Radar is the only observational platform that can capture the internal structures of convective storms, at high spatial and temporal resolutions (Stensrud et al. 2009; Hu et al. 2006). Many studies have shown that radar data assimilation (DA) greatly reduces the spin up time of model and improves short-range precipitation forecasts (Dixon et al. 2009; Xiao et al. 2008; Hu and Xue 2007; Zhu et al. 2015; Xue et al. 2003; Hu et al. 2006).

To obtain additional information on precipitation MP, the entire U.S. operational WSR-88D Doppler radar network has been upgraded to dual polarization a few years ago (ROC 2013). In Europe, the number of operational dual-Doppler radars has grown steadily (Huuskonen et al. 2014). More countries such as China are in the process of upgrading its operational radars to dual polarization (Zhao et al. 2019). Compared to single-polarization radar, a dual-polarization radar measures hydrometeor particle scattering at both horizontal and vertical polarizations, and can thereby provide information on the shape and other characteristics of hydrometeor particles. From polarimetric radar data (PRD), rain drop/particle size distributions (PSDs) and related properties can be better retrieved (Huang et al. 2019; Zhang et al. 2019; Cao et al. 2013), as can be hydrometeor classification within storms (Ryzhkov and Zrnice 1998; Vivekanandan et al. 1999).

It is expected that the assimilation of PRD into NWP models would help improve the analysis (initialization) and prediction of precipitating systems. So far, studies on the assimilation of PRD are relatively few, however, and most of the studies assimilate PRD indirectly, i.e., retrieval of model state variables from data is performed first before assimilation. Wu et al (2000) assimilated rain and ice mixing ratios retrieved from Z_H and differential reflectivity Z_{DR} , assuming that only two hydrometeor categories, i.e., rain and ice, existed. In their study, the positive impact of assimilating PRD did not, however, last long in the forecast, and error associated with the very simple ice MP scheme used was suggested to be a reason.

Li and Mecikalski (2010) assimilated Z_H and Z_{DR} data based on warm-rain-only observation operators implemented within the Weather Research and Forecasting (WRF) three-dimensional variational (3DVAR) DA system. With the assimilation of both Z_H and Z_{DR} , in-storm structures were said to be better analyzed and short-range precipitation forecast was also improved. More recently, Li et al (2017) developed an observation operator for specific differential phase (K_{DP}) that includes an ice phase (snow) and found positive impact of assimilating extra K_{DP} data using WRF 3DVAR on analyzed rainwater in the lower troposphere and snow in the mid- to upper troposphere for a mesoscale convective system. The impact was limited to the examination of analysis increments of rainwater and snow at a single time though and the impact was limited to a portion of the analyzed storm due to limited data coverage¹.

¹ We note that their reflectivity operators including liquid and ice contained a significant error, so their results should be viewed with caution.

90 Some studies have attempted to assimilate information derived from polarimetric signatures
91 within convective storms. For example, in intense supercell storms, a column of high Z_{DR} or Z_{DR}
92 column is often found in the region of intense updraft, corresponding to large raindrops that can
93 be lofted above the freezing level in the form of supercooled liquid water (Kumjian and Ryzhkov
94 2008). In a proof-of-concept study, Carlin et al (2017), the moisture and temperature adjustments
95 within the ARPS cloud analysis system (Hu et al. 2006) were modified to be based on the detection
96 of Z_{DR} columns for two tornadic supercell storm cases. Both analyses and forecasts of the storms
97 were improved compared to the use of the original cloud analysis in both cases. While the
98 procedure appears to be effective for tornadic supercell storms, it will be hard to apply, however,
99 to weaker precipitating systems where Z_{DR} column is much less pronounced or absent. Such
100 methods also rely on empirical relations between PRD and model state variables.

101 More direct and quantitative use of PRD is desirable through direct assimilation. Direct DA
102 methods compare simulated observations from the model state variables against observations, and
103 make adjustments to the state variables to achieve optimal fit of the analyzed state to observations
104 and the prior guess of the state subjecting to the weights related to their respective errors (Kalnay
105 2002). Forward observation operators are needed to simulate PRD from the model state variables,
106 and the forecast model should have a reasonable capability in simulating observed polarimetric
107 signatures. Jung et al (2008a) developed PRD observation operators based on calculations of
108 electromagnetic wave scattering by hydrometeors then used power law functions to fit
109 backscattering amplitudes to obtain computationally more efficient operators. The contributions
110 of wet snow and wet graupel/hail are also included. In Jung et al. (2010b), more accurate
111 observation operators based on rigorous scattering calculations using the T-matrix method
112 (Vivekanandan et al. 1991) are developed. Details of observation operators will be given in section
113 2a. Jung et al. (2010b) compared the performances of single-moment (SM) and double-moment
114 (DM) microphysics in terms of the simulated polarimetric signatures. Their results showed that
115 certain polarimetric signatures such as Z_{DR} arc, ρ_{hv} (cross-correlation coefficient) rings can only
116 be correctly simulated by DM scheme. Simulated PRD can have large uncertainties and can vary
117 significantly with the use of MP scheme, however (Putnam et al. 2017b; Putnam et al. 2017a) .

118 The observation operators for PRD variables such as Z_{DR} are highly nonlinear. To variationally
119 assimilate PRD, linear tangent and adjoint of the observation operators are needed, and the high
120 nonlinearity often causes convergence problems with the variational minimization (Liu et al. 2019).
121 With the ensemble Kalman filter (EnKF) method that has been shown to work well with
122 complex MP schemes (Tong and Xue 2005), linear tangent or adjoint of the observation operators
123 is not needed. EnKF also has the ability to directly update state variables not directly involved in
124 the observation operators, through ensemble-estimated flow-dependent background error
125 covariances, even in the presence of complex mixed-phase microphysics (Tong and Xue 2005).
126 Jung et al (2008b) first assimilate PRD using EnKF with a SM MP schemes with positive impacts
127 achieved and Jung et al.(2010a) demonstrated the benefit of PRD in improving the estimation of
128 both microphysical state variables and PSD parameters associated with a SM MP scheme. Both of
129 these studies assimilated simulated PRD.

130 Certain polarimetric signatures that depend on hydrometeors size sorting (Dawson et al. 2014),
131 such as the Z_{DR} arc in the supercell storms, can only be properly simulated using multi-moment
132 MP schemes (Jung et al. 2010b; Putnam et al. 2014). In the only published study that directly
133 assimilates real polarimetric observations using EnKF, Putnam et al (2019) showed that the
134 analyzed Z_{DR} structures including the Z_{DR} arc in a supercell storm are improved with additional
135 Z_{DR} assimilation. The study also showed that the analyzed rain mean mass diameter is higher in

136 the Z_{DR} arc region and the total rain number concentration is lower downshear in the forward flank,
137 agreeing with observational estimations. Biases do exist in their EnKF analyses that require further
138 investigations (Putnam et al. 2019), however.

139 As far as we know, Putnam et al (2019) is the only formally published study that examines
140 the impact of directly assimilating additional PRD using EnKF combined with a multi-moment
141 MP scheme. Many issues, including analysis biases, remain that require further studies as they
142 pointed out. Being a real-data-based study, detailed verification of analyzed state variables,
143 especially those of MP, is difficult, because of the lack of truth. Errors in the observational data
144 can complicate the issues. To better understand the behaviors and impacts of assimilating
145 additional PRD, observation system simulation experiments (OSSEs) can be very helpful. While
146 Jung et al. (2008b) and Jung et al. (2010b) examined the impacts of PRD data via OSSEs, their
147 EnKF DA studies had limitations with the use of a SM MP scheme. For the above reasons, OSSEs
148 are performed in this study with EnKF combined with a DM MP scheme and compatible
149 observation operators, examining the impact of directly assimilating additional Z_{DR} data.
150 Additional sensitivity experiments are performed to see the impacts of updating total number
151 concentrations (the additional PSD moment associated with a DM MP schemes) and updating state
152 variable not directly linked to PRD via observation operators.

153 The rest of this paper is organized as follows. In section 2, the observation operators used in
154 this study together with configurations of the OSSE experiments are described. The results of
155 control and sensitivity experiments examining the impacts of PRD assimilation are presented and
156 discussed in section 3. Summary and conclusions are given in section 4.

157

158 **2. Experiment configuration and settings**

159 *a) The truth simulation and observation operators*

160 For the OSSEs, a truth simulation is produced using the Advanced Regional Prediction System
161 (ARPS, Xue et al. 2003) initialized from a sounding for the 1977 Del City, Oklahoma supercell
162 storm (Ray et al. 1981), as given in Xue et al (2001). A 4-K ellipsoidal thermal bubble with radii
163 of 10 km in the horizontal directions and 1.5 km in the vertical direction is used to initiate the
164 storm. Most of the configurations are inherited from Tong and Xue (2005) except for the MP
165 scheme used and the grid configuration. The SM Lin MP scheme is replaced by DM Milbrandt-
166 Yau MP scheme (Milbrandt and Yau 2005); as mentioned earlier, DM schemes can much better
167 reproduce Z_{DR} signatures (Jung et al. 2010b). The simulation domain has $105 \times 103 \times 53$ grid points
168 and the horizontal grid spacing is 1 km. The average vertical grid spacing is 400 m, and is stretched
169 from 50 m at the surface.

170 For DM schemes, the shape parameter of three-parameter gamma distributions assumed of
171 most hydrometeor PSDs is typically assumed constant (with zero being assumed most often). In
172 this study, the shape parameter for rainwater in the Milbrandt-Yau DM scheme is set to 2 while
173 for other hydrometers it is set to zero. Studies have found that most DM schemes tend to
174 overestimate simulated reflectivity (eg: Putnam et al. 2017b; Brown et al. 2016). One of the reasons,
175 according to Brown et al (2016), is that most schemes tend to produce a higher frequency of large
176 raindrops than observed. Setting the rain shape parameter to 2 helps reduce the number of large
177 raindrops and in turn reflectivity.

178 The observation operator for radial velocity is the same as that used in Jung et al (2008a).
179 However, there are some differences from OSSE experiments of Jung et al (2008a) where the
180 observation operators for radar reflectivity are calculated using a fitted approximation to T-matrix
181 scattering amplitudes for rain and Rayleigh approximation for ice hydrometeors. This

182 approximation may result in some error (Putnam et al. 2019). In this study, more advanced
 183 observation operators using look up tables calculated from T-matrix method are used (Jung et al.
 184 2010b). In the following are the formula for radar reflectivity factors at the horizontal and vertical
 185 polarizations based on the full T-matrix algorithm:

$$186 \quad Z_{H,x} = \frac{4\lambda^4}{\pi^4 |K_w|^2} \int_0^{D_{max,x}} [A |f_{a,x}(\pi)|^2 + B |f_{b,x}(\pi)|^2 + 2C \operatorname{Re}[f_{a,x}(\pi)f_{b,x}^*(\pi)]]n(D)dD, \quad (1)$$

$$187 \quad Z_{V,x} = \frac{4\lambda^4}{\pi^4 |K_w|^2} \int_0^{D_{max,x}} [B |f_{a,x}(\pi)|^2 + A |f_{b,x}(\pi)|^2 + 2C \operatorname{Re}[f_{a,x}(\pi)f_{b,x}^*(\pi)]]n(D)dD, \quad (2)$$

188 where

$$189 \quad A = \frac{1}{8} (3 + 4 \cos 2\bar{\phi}e^{-2\sigma^2} + \cos 4\bar{\phi}e^{-8\sigma^2}),$$

$$190 \quad B = \frac{1}{8} (3 - 4 \cos 2\bar{\phi}e^{-2\sigma^2} + \cos 4\bar{\phi}e^{-8\sigma^2}),$$

$$191 \quad C = \frac{1}{8} (1 - \cos 4\bar{\phi}e^{-8\sigma^2}).$$

192

193 Here, λ is the wavelength of the radar and we assume a 10.7 cm wavelength S-band radar. K_w
 194 $=0.93$ is the dielectric factor for water. $\bar{\phi}$ is the mean canting angle and σ is the standard deviation

195 of the canting angle. $\bar{\phi} = 0$ is assumed for all species. σ are 0° , 20° , 60° , and 60° for rain, snow,

196 graupel and hail, respectively. $|\dots|$ represents the modulus of complex number while $\operatorname{Re}[\]$

197 represents the real part. Superscript $*$ implies the conjugate.

198 Subscript x can be rain (r), rain-snow mixture (rs), dry snow (ds), rain-graupel mixture (rg), dry

199 graupel (dg), rain-hail mixture (rh) and dry hail (dh). D is the diameter of a given hydrometeor and

200 D_{max} is the maximum diameter of each hydrometeor category. In this paper, the maximum

201 diameters of rain drops, snow aggregates, graupels and hailstones are assumed to be 8, 30, 50 and

202 70 mm, respectively. $n(D)$ is the number concentration of the hydrometer at diameter D . To

203 numerically integrate Eqs.(1) and (2), the integral ranges are partitioned into 100 bins. The

204 backscattering amplitudes of each species with assumed drop size for polarizations along the major

205 ($f_a(\pi)$) and minor ($f_b(\pi)$) axes are precomputed at the center of each size bin and stored in look-up

206 tables. For melting species including rain-snow, rain-graupel and rain-hail mixtures, the same

207 tables are constructed at the uniform 5% water fraction interval. The fraction of water of each ice

208 specie is calculated as $f_{w,ix} = \frac{q_r}{q_r + q_{ix}}$. Here, q_r is the mixing ratio of rain while q_{ix} is one of the

209 ice hydrometeors. More details on the PRD observation operators can be found in Jung et al

210 (2010b).

211 Once the radar reflectivity factors of all hydrometeor categories are calculated. The

212 reflectivity in dBZ at horizontal and vertical are computed as follows:

$$213 \quad Z_H = 10 \log_{10} (Z_{h,r} + Z_{h,rs} + Z_{h,ds} + Z_{h,rg} + Z_{h,dg} + Z_{h,rh} + Z_{h,dh}) \quad (3)$$

$$214 \quad Z_V = 10 \log_{10} (Z_{v,r} + Z_{v,rs} + Z_{v,ds} + Z_{v,rg} + Z_{v,dg} + Z_{v,rh} + Z_{v,dh}) \quad (4)$$

215 The differential reflectivity Z_{DR} is calculated according to the following formula:

$$216 \quad Z_{DR} = Z_H - Z_V. \quad (5)$$

217

218 *b) EnKF experiment settings and DA experiments*

219 In this study, we use the ARPS EnKF package (Tong and Xue 2005; Xue et al. 2006) which
220 uses the ensemble square root filter algorithm (Whitaker and Hamill 2002). The EnKF experiments
221 employ 40 members in this study. With the DM Milbrandt-Yau MP scheme, the analysis variables
222 include the three-dimensional wind components (u , v , and w), pressure (p), potential temperature
223 (θ), water vapor mixing ratio (q_v), as well as microphysical state variables including mixing ratios
224 of cloud water (q_c), rainwater (q_r), cloud ice (q_i), snow aggregate (q_s), graupel (q_g) and hail (q_h),
225 and their total number concentrations (N_{q_c} , N_{q_r} , N_{q_i} , N_{q_s} , N_{q_g} and N_{q_h} , respectively). Spin-up
226 ensemble forecasts are run for 20 minutes, starting from initial ensemble states defined by the
227 sounding profiles plus smoothed Gaussian random perturbations added in regions where observed
228 reflectivity is larger than 10 dBZ. The mean standard deviations of added u , v , and w perturbations
229 are 2 m s^{-1} and that of θ is 2 K. For water vapor and hydrometeor mixing ratios, the mean standard
230 deviations of added perturbations are $0.0006 \text{ kg kg}^{-1}$. Consider the large uncertainty of number
231 concentrations, we did not add perturbation to those variables. After the 20-min long spin-up
232 forecasts, EnKF DA cycles are run over a 90-min period assimilating radar data every 5 minutes,
233 corresponding to the model storm time period of 20 through 110 minutes. Similar settings were in
234 our earlier OSSE studies (Jung et al. 2008a; Tong and Xue 2005). The 90-min assimilation period
235 is chosen mostly based on the life cycle of the storm. In the truth simulation, the main storm reaches
236 its mature stage between 60 to 100 min. After that, the storm begins to weaken and move out of
237 the simulation domain.

238 PRD from an assumed S-band radar with its center located in the southwest corner ($x=2, y=2$)
239 are simulated from the truth simulation output, using the observation operators described in section
240 2a. Eleven elevations are assumed, based on the WSR-88D radar VCP-11 scan mode. Radar
241 observation errors are assumed to be 1 m s^{-1} , 3 dBZ, 0.2 dB for radial velocity V_r , and Z_H and Z_{DR}
242 in terms of standard deviation, respectively, and random errors of the corresponding magnitudes
243 are added to the simulated PRD observations and assumed in the EnKF experiments. V_r
244 observations are assimilated where observed $Z_H > 10 \text{ dBZ}$. For Z_{DR} , only values larger than 0.3 dB
245 are used because smaller values tend to be very noisy. The covariance localization radii for radar
246 observations are set to 4 km in the horizontal and 2 km in the vertical direction using the correlation
247 function of Gaspari and Cohn (Gaspari and Cohn 1999) for all state variables. The 4 km horizontal
248 grid spacing spans 4 grid interval in this study, which is consistent with most past studies in terms
249 of grid intervals. For example, in Tong and Xue (2005) and Jung et al. (2008a), 6 to 8 km were
250 suggested when a 2 km horizontal grid spacing was used. Sobash and Stensrud (2013) suggest
251 suggested 12 to 18 km horizontal radii when using a 3 km grid spacing. We have tested larger and
252 smaller horizontal localization radii. The state analysis errors were found to be significantly larger
253 when using a 6 km radius while the results using 3 km were slightly worse. To help maintain
254 ensemble spread, multiplicative inflation (Anderson 2001; Tong and Xue 2005) is applied to all
255 model state variables except for number concentrations, using an inflation coefficient of 1.2.

256 Table 1 lists all experiments presented in this paper. Experiment VrZh assimilates V_r and Z_H
257 data while experiment VrZhZdr assimilates additional Z_{DR} data. Both experiments update a full
258 set of state variables in the model. Experiment VrZhZdr is considered a control experiment while
259 VrZh is a reference for comparison purpose. Additional sensitivity experiments are conducted to
260 help better understand how the assimilation of Z_{DR} improves the analysis. The first sensitivity
261 experiment VrZhZdr_NoNt, excludes the updating of total number concentrations of the
262 hydrometeors N_{q_x} by Z_{DR} observations compared to experiment VrZhZdr. N_{q_x} are still updated by

263 V_r and Z_H though, just not by Z_{DR} . Nq_x arise from the use of a DM scheme and adds additional
264 complexity to the DA problem. The number concentrations of hydrometers have very wide
265 dynamic ranges, varying from 0 to as large as 10^{12} m^{-3} , implying that the relations between them
266 and PRD observations can be very nonlinear. Updating both mixing ratios and total number
267 concentrations at the same time may or may not be beneficial, especially when the correlations are
268 unreliable or inconsistent with each other. VrZhZdr_NoNt serves to examine the benefit, if any,
269 of updating the total number concentrations using Z_{DR} observations.

270 Other sensitivity experiments serve to examine the impact of updating other state variables
271 using Z_{DR} . Experiments VrZhZdr_NoW, VrZhZdr_NoUV, VrZhZdr_NoQv and VrZhZdr_NoPt
272 excludes the updating of vertical velocity w , horizontal wind components u and v , water vapor
273 mixing ratio q_v and potential temperature θ , respectively. In an intense tornadic supercell, a Z_{DR}
274 column typically exists in the updraft region (Kumjian and Ryzhkov 2008), indicating strong
275 positive correlation between upward motion and Z_{DR} . Updraft regions are also associated with high
276 moisture values. The largest theoretical benefit of EnKF method compared to 3DVar and some of
277 the other methods lies with the use of ensemble-derived correlations among all state variables, and
278 hence among observation priors and state variables, which allows for the updating of state
279 variables not directly observed (or involved in the observation operators). For such updating to be
280 beneficial, the ensemble-derived correlations have to be sufficiently accurate and reliable. This
281 second group of sensitivity experiments are designed to test the impacts of updating state variables
282 that are not directly linked to Z_{DR} observations via the observation operators.

283

284 **3. Results of EnKF analyses**

285 *a) Evaluation of Z_{DR} assimilation in the control experiment*

286 Figure 1 shows the ensemble mean analysis and forecast RMSEs of model state variables
287 during the assimilation cycles. Following Tong and Xue (2005) and many other studies, the
288 RMSEs are calculated over grid points where the true reflectivity is greater than 10 dBZ, which
289 roughly covers the precipitation regions. For most variables, VrZhZdr (red lines), which
290 assimilates additional differential reflectivity, produces consistently better analyses and forecasts
291 than VrZh (black lines), especially in later cycles. Such results are quite similar to those of Jung
292 et al. (2008b) which examined the impact of assimilating additional Z_{DR} data in OSSEs employing
293 a SM MP scheme, except that RMSEs of most variables in the first few cycles are also reduced
294 here. In Jung et al (2008b), the assimilation of Z_{DR} does not show positive impact until later cycles.
295 Additionally, we also examine the RMSEs of total number concentrations of hydrometer variables,
296 which were not predicted in Jung et al (2008b). Here, for most number concentrations, the Z_{DR}
297 assimilation shows neutral to positive impact. Among them, the number concentration for graupel,
298 Nq_g , is improved most. As we will discuss later, it is probably benefiting from better analyses of
299 liquid hydrometeor species, which in turn lead to more accurate analyses of ice hydrometeor
300 species. In Fig. 2, we show the RMSEs of the analyses and forecasts throughout the DA cycles in
301 terms of radar observed variables, i.e., the verifications in observation space. The results are
302 consistent with the results in terms of the state variables, as shown in Fig. 1; the assimilation of
303 Z_{DR} data further reduces the differences between the analyses and forecasts from the observations
304 in the observation space.

305 Fig. 3 shows the vertical profiles of ensemble mean analysis and forecast RMSEs at 80 minutes,
306 again averaged over grid points with observed Z_H exceeds 10 dBZ. At this time, the RMSEs of
307 most variables have stabilized (Fig. 1). It can be seen that the errors at most levels for most
308 variables are reduced from the additional Z_{DR} assimilation. The largest improvements are mostly

309 located where the errors are largest. As Jung et al. (2008b) pointed out, the direct improvements
310 from Z_{DR} assimilation are mainly to those highly correlated variables such as q_v and q_r at the lower
311 levels, where the Z_{DR} signatures are most prominent (given that large Z_{DR} is mostly associated with
312 large raindrops). With more accurate analyses at the lower levels, the analysis fields at upper levels
313 can also be improved through the dynamic interactions in the forecast model. The weak and
314 unreliable correlations between Z_{DR} and ice fields at the upper levels during the earlier cycles
315 might be the reason for larger errors in q_h before 45 min (Fig. 1k) while the errors become smaller
316 in later cycles. Note that in Figs. 1-3, the results of VrZhZdr_NoNt are also included which will
317 be discussed in section 4b later.

318 In Fig. 4 and Fig. 5, we further examine the impact of Z_{DR} assimilation on the polarimetric
319 signatures of simulated storm. At 80 minutes (Fig. 4a), the Z_{DR} arc is not clearly seen in the truth
320 simulation. We can see a narrow high Z_{DR} band along the edge of 35 dBZ reflectivity. Between
321 this Z_{DR} band and main storm, there is a weak Z_{DR} area (green to light yellow) which is due to hail
322 falling and melting in this region. At 110 minutes (Fig. 4b), high Z_{DR} (red color) extends all the
323 way from the forward flank reflectivity core to the southern edge of forward flank precipitation
324 region; in fact, it extends beyond the 35 dBZ reflectivity contour, suggesting the existence of a
325 relatively small number of large rain drops there, giving rise to relatively high Z_{DR} values. Along
326 this edge, an arc of high Z_{DR} is often observed, due to hail stone and rain drop size sorting (Dawson
327 et al. 2014). For both analysis times, experiment VrZhZdr with additional Z_{DR} assimilation shows
328 better Z_{DR} structure than that of VrZh, especially for later analysis time. For experiment VrZh, the
329 pattern of high Z_{DR} area (red color) is not as good as experiment VrZhZdr when compared to truth
330 simulation at 110 minutes.

331 The Z_{DR} structure near hook echo region is similar to the classic supercell storm structure for
332 both truth simulation and EnKF analyses (Fig. 4d, e, f). Here, we only display small hook area at
333 80 minutes because it shows a clear Z_{DR} column (see Fig. 6d). At 110 minutes, the Z_{DR} columns
334 are not obvious (not shown). High Z_{DR} values are located at the leading edge of the high Z_H hook
335 (black contours) (Fig. 4d). Experiment VrZh shows generally similar patterns but the intensity is
336 clearly underestimated for both Z_H and Z_{DR} (Fig. 4b, e). With additional Z_{DR} assimilation, the shape
337 of Z_{DR} arc in experiment VrZhZdr looks closer to that of truth than in experiment VrZh (Fig. 4c).
338 The intensity of Z_{DR} in the hook echo region is also much enhanced in VrZhZdr (Fig. 4f). Moreover,
339 the Z_H pattern is also been improved. The 35 dBZ Z_H contours in the southeast edge are much
340 closer to those of truth (Fig. 4a, b, c), and the Z_H intensity in the hook region is greatly enhanced
341 (Fig. 4 d, e, f). In all, Z_{DR} assimilation improves the polarimetric signatures of the simulated storm,
342 especially in the hook echo and forward flank regions.

343 The vertical cross sections of analyzed Z_H , Z_{DR} and specific differential phase, K_{DP} , in the hook
344 echo region through the low-level Z_H and Z_{DR} maximum centers at 80 minutes are shown in Fig.
345 6. In general, both experiments VrZh and VrZhZdr produce similar patterns of these fields.
346 However, the intensities of Z_{DR} and K_{DP} are clearly underestimated in VrZh. Here K_{DP} is not
347 directly assimilated, but derived from analyzed model state variables using the same equation as
348 in Jung et al (2010b). For Z_H , the maximum values above 60 dBZ are right below the 0 °C contours
349 in all cases (Fig. 6a, b, c). The 45 dBZ Z_H contours (orange) extend up to above the -20 °C in the
350 truth (see Fig. 6a) and in VrZhZdr (Fig. 4c), but only to -10 °C line in VrZh (see Fig. 6b), indicating
351 the analyzed storm is less intense in VrZh. The improved vertical structure of Z_H indicates better
352 analysis of the hydrometeor fields, which we will show more in Fig. 7. The assimilation of Z_{DR}
353 data also results in a more intense core updraft that is closer to the truth as indicated by the 10 m
354 s^{-1} w contours in Fig. 6a-c. With a stronger updraft, particles are more likely transported to high

355 altitudes and also likely undergo more growth before falling to the ground. Associated with the
356 updraft is a column of high Z_{DR} values that extend to the $-10\text{ }^{\circ}\text{C}$ level in the truth (Fig. 6d) and in
357 VrZhZdr (Fig. 6f), while that in VrZh is clearly weaker (Fig. 6e). Also, a column of high K_{DP} is
358 also better reproduced in VrZhZdr (Fig. 6i) than in VrZh (Fig. 6h) compared to the truth (Fig. 6g).
359 High K_{DP} is mostly associated with high liquid water content, which is linked to intense updraft
360 and heavy precipitation.

361 Fig. 7 shows the analyzed cloud water, hail and rain water mixing ratios from VrZh and
362 VrZhZdr in the same vertical cross sections as Fig. 6, as compared to the truth. Since only Z_{DR}
363 observations larger than 0.3 dB are assimilated, the direct impact from Z_{DR} are mostly limited in
364 the lower levels (c.f., Fig. 6d). However, its benefit could be spread to the higher levels through
365 spatial and cross-variable correlations, and through dynamic interactions within the forecast model.
366 Fig. 7 shows that the cloud ice field is better analyzed all the way to the cloud top at ~ 9 km height
367 in VrZhZdr (Fig. 7c) and while that in VrZh is mostly limited to below $-20\text{ }^{\circ}\text{C}$ level or about 6.5
368 km height; its maximum value is also too low (Fig. 7b). For hail, VrZhZdr also much better
369 reproduces the vertical distribution and intensity (Fig. 7f) than VrZh (Fig. 7e); the latter severely
370 underestimates hail at the higher levels. For rainwater, the analysis of VrZhZdr is also better,
371 although the differences are smaller (Fig. 7i, h). Overall, the assimilation of additional Z_{DR}
372 produces analyses of the supercell storm whose intensity and structure are much closer to the truth,
373 in terms of both observed parameters (Z_H and Z_{DR}) and model state variables.

374

375 *b) The updating of hydrometeor number concentrations with Z_{DR}*

376 For the DM MP scheme, the hydrometeor number concentrations are part of the forecast
377 variables which increase the degrees of freedom of the model state. As pointed out earlier, the
378 values of number concentrations show a great range of variability. Additionally, for DM schemes,
379 Z_{DR} depends mostly on the slope parameter of PSD which is a strong function of the third moment,
380 the mass mixing ratio (Jung et al. 2008b). It is not certain whether the updating of number
381 concentrations by EnKF will improve the overall analysis. The RMSEs for most state variables
382 and also for radar observed variables of the experiment VrZhZdr_NoNt that excludes the updating
383 of number concentrations are shown in Fig. 1 to Fig. 3. It can be seen that without updating Nq_x ,
384 the RMSE curves of VrZhZdr_NoNt (blue lines) are more close to those of VrZh than VrZhZdr
385 during the later DA cycles for most variables (Fig. 1). Similar is true in terms of radar observed
386 variables V_r , Z_H and Z_{DR} (Fig. 2). For w , q_r and q_h , the RMSEs of VrZhZdr_NoNt even exceed
387 those of VrZh in some of the cycles (Figs. 1c, 1i, 1l). The deterioration of the analyses in
388 VrZhZdr_NoNt are more clear in the vertical profiles of RMSEs at 80 minutes (Fig. 3). For w and
389 most ice state variables, the RMSEs in VrZhZdr_NoNt are larger than those of VrZh at the upper
390 levels (Fig. 3) while for q_r this happens at the mid-levels (Fig. 3i). These results suggest that
391 updating both mixing ratios and total number concentrations of hydrometeors species associated
392 with a DM MP scheme together when assimilating Z_{DR} is important; when only mixing ratios are
393 updated, most of the benefit of assimilating Z_{DR} data is lost, and for some variables, that analyses
394 may be even worse than not assimilating Z_{DR} data at all. This is presumably because serious
395 imbalance or inconsistency is created between mixing ratios and corresponding number
396 concentrations when only the former are updated.

397 Fig. 8 shows the analyzed rainwater number concentrations Nq_r at 3 km height, and hail
398 number concentrations Nq_h at $z = 6$ km from VrZhZdr and VrZhZdr_NoNt, as compared to the
399 truth. For the truth, highest Nq_r values are found in the southwest part of the supercell storm near
400 the hook echo region and in the northwest part, corresponding to heavy rainfall in the rear flank

401 and forward flank downdraft regions, respectively (Fig. 8a). The patterns of analyzed Nq_r are
 402 similar (Fig. 8b, c) although there is a larger area of over-estimation in the forward flank region
 403 while the high values in the rear flank region are under-estimated in VrZhZdr_NoNt (Fig. 8c).
 404 Both over-estimation and under-estimation are much less in VrZhZdr (Fig. 8b). The hail number
 405 concentration Nq_h for the truth exhibits moderately high values in the southeastward-spreading
 406 forward flank and storm anvil regions at the 6 km height (Fig. 8d) while in the hook echo region,
 407 a ring of high Nq_h is found around a Nq_h hole, while the highest values found on the west and
 408 southwest sides of the hole (Fig. 8d). The hole should be associated with bounded weak echo
 409 region typically found in intense supercell storms where hydrometeors are mostly absent being
 410 swept away by the intense updraft. Within VrZhZdr_NoNt, the ‘ring’ structure is over-estimated
 411 (Fig. 8f) although the pattern of Nq_h in the forward flank region is a little better in VrZhZdr_NoNt
 412 (Fig. 8f) than in VrZhZdr (Fig. 8e). Overall, Nq_r and Nq_h are better analyzed in experiment
 413 VrZhZdr than in VrZhZdr_NoNt.

414 Fig. 9 explain the possible reasons. Here, we calculate the correlation coefficients between the
 415 Z_{DR} observation prior and the hydrometeor state variables $\rho(Z_{DR}, Nq_x)$ from the forecast ensemble
 416 in a vertical cross section passing through Z_{DR} prior which is located in the Z_{DR} column at (x, y, z)
 417 $= (34, 22, 3.5)$ km. In general, Z_{DR} has clearly higher correlations to Nq_c , Nq_r and Nq_h than to Nq_i ,
 418 Nq_s and Nq_g . This is because Z_{DR} is most sensitive to raindrop sizes and high Z_{DR} is found where
 419 there are a large number of large rain drops. Many large drops originate from the melting of falling
 420 hail stones (Dawson et al. 2014). A column of high correlation is found for $\rho(Z_{DR}, Nq_c)$, $\rho(Z_{DR},$
 421 $Nq_r)$ and $\rho(Z_{DR}, Nq_h)$ near the main updraft. For Nq_i , Nq_s and Nq_g , the correlations are weaker,
 422 and non-zero values are mostly found above the freezing level (Fig. 9). The coherent structures in
 423 the correlations between Z_{DR} and q_x , and between Z_{DR} and Nq_x suggest that the flow-dependent
 424 error covariances estimated and utilized within the EnKF should be physically reasonable, and
 425 hence the updating of Nq_x in addition to q_x can be beneficial.

426 We also examine correlations between Z_{DR} at a 1.8 km height and hydrometeor state variables
 427 in a vertical cross section in the forward flank high Z_{DR} region (Fig. 10). The cloud water at this
 428 point is zero for all members. Therefore, the correlation is zero and is not shown. For other
 429 hydrometeor variables, similar to the point in the hook echo region, correlations $\rho(Z_{DR}, Nq_r)$ and
 430 $\rho(Z_{DR}, Nq_h)$ are clearly higher than $\rho(Z_{DR}, Nq_s)$ and $\rho(Z_{DR}, Nq_g)$. The patterns of correlation ρ
 431 (Z_{DR}, q_x) are also very similar to the corresponding $\rho(Z_{DR}, Nq_x)$ except that those of hail show
 432 opposite signs of correlation near the surface (Fig. 10e, j). The negative correlation between Z_{DR}
 433 at 1.8 km and q_h at the lower levels is consistent with the fact that hail stones tend to contribute
 434 little to Z_{DR} due to tumbling (which is the cause of Z_{DR} hole within supercell storms as a significant
 435 hail signature (Kumjian and Ryzhkov 2008)), while the positive correlation between Z_{DR} and Nq_h
 436 suggests that when a larger number of small hailstones exist, melting hailstones will cause less
 437 reduction to Z_{DR} . Given that large correlations between Z_{DR} and mixing ratio and between Z_{DR} and
 438 number concentration for rainwater and hail simultaneously, updating mixing ratios without
 439 updating corresponding number concentrations will create imbalances between different moments
 440 of the hydrometeor PSDs which in turn will negatively affect the analysis and forecast states.

441
 442 *c) The updating of other state variables with Z_{DR} assimilation*

443 Fig. 11 shows that analysis and forecast RMSEs of sensitivity experiments without updating
 444 certain state variables when assimilating Z_{DR} . RMSEs for experiment VrZhZdr are shown in black
 445 lines while those for VrZhZdr_NoPt, VrZhZdr_NoUV and VrZhZdr_NoW are shown in color.
 446 Among all potential temperature, water vapor, vertical and horizontal wind components, the

447 updating of water vapor q_v has the greatest impact. The RMSEs from VrZhZdr_NoQv (solid purple)
 448 are significantly larger for almost all forecast times and state variables and the differences are
 449 larger in later cycles. The updating of w has the second largest impact as the RMSEs of
 450 VrZhZdr_NoW (red) are noticeably larger for most variables especially during the intermediate
 451 cycles. The updating of horizontal wind components and potential temperature has less impact as
 452 the RMSEs of VrZhZdr_NoPt (solid green) and VrZhZdr_NoUV (blue) are rather close to those
 453 of VrZhZdr. These results are reasonable since water vapor is the primary fuel for intense
 454 convection while w provides the best measure of the intensity of convection. Given that VrZhZdr
 455 produces overall the lowest RMSEs, all state variables should be updated when assimilating Z_{DR} ,
 456 at least when no model error is present and the ensemble-estimated covariances are reasonably
 457 accurate.

458 The vertical RMSE profiles up to 5 km height at 80 minutes are shown in Fig. 12. Here, we
 459 focus on the low levels where Z_{DR} has largest impacts. Consistent with Fig. 11, experiment
 460 VrZhZdr_NoQv has the largest errors at essentially all vertical levels shown. Experiment
 461 VrZhZdr_NoW produces second largest RMSEs for most variables at most levels. The updating
 462 of potential temperature θ has the third largest impact (e.g., on q_v in Fig. 12d, on p' in Fig. 12f, and
 463 q_h in Fig. 12l), although for some variables not updating θ made little differences (e.g., for q_c in
 464 Fig. 12g and q_r in Fig. 12h). The updating of u and v has limited impact from lower to middle
 465 levels. In experiment VrZhZdr_NoUV, the analysis RMSEs are close to those of VrZhZdr below
 466 2 km, but larger above 2 km for variables including q_v (Fig. 12d), q_c (Fig. 12g) and q_r (Fig. 12h).
 467 This is better illustrated in Fig. 13 which show the correlations between the Z_{DR} and wind
 468 components, the mixing ratio q_v and potential temperature perturbation θ' . The Z_{DR} point is the
 469 same as the point in Fig. 9. For q_v and w , they show high and continuous correlation regions from
 470 the bottom to the top. For u , v and θ , the high correlation areas are clearly reduced and mostly
 471 located in the lower levels.

472 In Fig. 14, we further examine the impact of not updating certain variables on the dynamic
 473 structures of analyzed storm. Here, the vertical vorticity ζ at 2 km height in the main updraft region
 474 is shown, indicating low-level mesocyclone structure and intensity. Also plotted are the vertical
 475 velocity w and horizontal winds. The truth shows an ellipse shaped structure of ζ with its center
 476 located to south of w maximum (Fig. 14a). With all state variables updated in EnKF, experiment
 477 VrZhZdr obtains very similar structures of ζ and w with the horizontal winds flowing around the
 478 north side of the updraft core (Fig. 14b). Without updating horizontal winds in VrZhZdr_NoUV
 479 when assimilating Z_{DR} data, the overall structures of ζ and w and horizontal winds are not too
 480 different from those of VrZhZdr except that their intensities are somewhat underestimated (Fig.
 481 14c). The impact of not updating θ in VrZhZdr_NoPt by Z_{DR} is similar to not updating u and v
 482 (impact is relatively small), although the maximum w is slightly overestimated according to the w
 483 maximum values shown in the plots (Fig. 14f). Compared to u , v and θ , the impact of not updating
 484 w or q_v when assimilating Z_{DR} is much larger. Without updating w , the ζ pattern appears more
 485 circular and the updraft is more concentrated but its maximum is overestimated (Fig. 14d). Without
 486 updating q_v , the shapes of ζ structure and updraft region are still close to those of VrZhZdr and
 487 truth, but the maximum ζ is most overestimated among the sensitivity experiments, and w
 488 maximum is also overestimated (Fig. 14e), although slightly less so than in VrZhZdr_NoW.

489 The above results indicate the analyzed flow structures and intensity in the main updraft region
 490 are directly linked to the updating of w and q_v . This is more obvious in vertical cross sections. Fig.
 491 15 shows the vertical cross sections of ζ and w fields through maximum of three dimensional w in
 492 the y direction. In both truth and experiment VrZhZdr, a ζ maximum is found at ~ 1.6 km level

493 which corresponds to relatively strong vertical motion there, and the fields in VrZhZdr match those
494 of truth very closely. Except for experiment VrZhZdr_NoW, the general patterns of ζ and w in
495 other sensitivity experiments are similar except for under-estimation of the low-level vorticity
496 strength, especially in VrZhZdr_NoQv and VrZhZdr_NoPt. In VrZhZdr_NoW, mid-level ($z \sim 5$
497 km) w is over-estimated by nearly 50%, as is the column of high vertical vorticity (Fig. 15f). Fig.
498 14 and Fig. 15 provide more concrete ideas on the large impact of updating or not updating w and
499 q_v by Z_{DR} , results that are consistent with earlier findings based on RMSEs. This further confirms
500 that there are reliable, strong ensemble-derived correlation between w and Z_{DR} that enables
501 improved analysis of w and other fields by Z_{DR} observations.

502 Based on the above results, all model state variables should be updated when assimilating Z_{DR}
503 data with in the EnKF. This is at least true for perfect model OSSEs. In our OSSE framework for
504 a tornadic supercell storm, apart from the updating of hydrometeor state variables, the updating of
505 water vapor mixing ratio q_v has the largest impact on the overall analysis accuracy followed by
506 vertical wind w . However, no updating w leads to larger errors in the flows in the updraft region,
507 including updraft itself and vertical vorticity associated with it. The errors in q_v affect that storm
508 dynamics more indirectly through moist processes.

509 **4. Conclusions and discussions**

511 In this study, the impact of assimilating differential reflectivity Z_{DR} data within an EnKF
512 framework is investigated using observing system simulation experiments with simulated data for
513 a tornadic supercell storm. The Milbrandt and Yau (2005) double-moment microphysics scheme
514 is used in both truth simulation and for EnKF DA; with this double-moment scheme, previous
515 studies have shown the reasonable ability to simulate most important polarimetric radar signatures
516 found in supercell storms. Radar observations are simulated using a polarimetric radar data
517 simulator developed by Jung (2010b), in which T-matrix method is used to calculate the
518 hydrometeor scattering magnitudes for particles of particular sizes. Observation errors of realistic
519 magnitudes are added to the simulated observations, and the same error variances are specified in
520 the EnKF DA. The observation operators from the simulator are also used in the EnKF DA, which
521 is run over a 90 minute period assimilating radar data every 5 minutes spanning the developing
522 and mature stages of the supercell. Experiments are conducted with and without assimilating Z_{DR}
523 data in addition to reflectivity at horizontal polarization Z_H and radial velocity V_r to examine the
524 impact of Z_{DR} assimilation. Results show that the assimilation of Z_{DR} reduces the RMSEs for
525 almost all model state variables at almost all analysis times. The polarimetric signatures of tornadic
526 storm including the Z_{DR} and K_{DP} columns and Z_{DR} arc are all improved. Analyses show that the
527 structures and intensities of hydrometeor fields at both lower and upper levels are improved, even
528 though the strongest Z_{DR} signatures are mostly found at the lower levels due to the concentration
529 of large raindrops there.

530 Additional sensitivity experiments are conducted to understand the benefit and impact of
531 updating different state variables when assimilating Z_{DR} . The first sensitivity experiment excludes
532 the updating of the total number concentrations of all hydrometeors, which arise from the use of a
533 double-moment microphysics scheme. Although the number concentrations have very large
534 dynamic ranges, and their relations with Z_{DR} are highly nonlinear, and the ensemble-derived error
535 correlations with Z_{DR} may or may not be reliable enough to produce improved analyses, the results
536 show that updating number concentrations together with the mixing ratios are very beneficial. The
537 number concentrations have high correlations with Z_{DR} at the lower levels that are comparable to
538 those of mixing ratios. If the number concentrations are not updated by Z_{DR} observations, most of

539 the benefit of assimilating Z_{DR} data is lost, and in fact, for vertical velocity, rainwater and hail
540 mixing ratios, the analysis RMSEs are larger in intermediate DA cycles than those in the
541 experiment not assimilating Z_{DR} data at all. Clearly, updating both mixing ratios and total number
542 concentrations of hydrometeors leads to much more physically consistent analyses.

543 In other sensitivity experiments, the updating of horizontal wind components, vertical velocity,
544 water vapor or potential temperature by Z_{DR} data is excluded, respectively. This allows us to
545 examine the impact and importance of updating these state variables, which are not directly or are
546 only weakly linked to Z_{DR} via the observation operators. Among these state variables, the updating
547 of water vapor mixing ratio q_v has the largest impact, which is followed by the updating of vertical
548 wind w . The updating of horizontal wind components or potential temperature has much smaller
549 though still noticeable impact. Further analysis shows that the updating of q_v or w has significant
550 effects on the intensity and structures of vertical vorticity and vertical velocity in the main updraft
551 region, and significant under-estimation and over-estimation are seen, respectively, in the vertical
552 cross section through the main updraft when q_v or w is not updated. Clearly the updating of w has
553 more direct effect on the storm intensity than the updating of q_v , but the effect of the latter via
554 moist processes is apparently very significant. Overall, updating all model state variables when
555 assimilating Z_{DR} data produces the best results, and the RMSEs of analyzed state variables are
556 consistently lower than those of experiment without assimilating Z_{DR} data.

557 Finally, we point out that the results presented in this paper are limited to OSSE tests with a
558 single supercell storm, and no model error is included. When model error is present, as is with all
559 real data cases, the conclusions may be somewhat different. In addition, other polarimetric
560 measurements, including specific differential phase K_{DP} and co-polar correlation coefficient ρ_{hv} ,
561 also contain valuable information on the hydrometeors and their PSDs. The assimilation of these
562 parameters and their impact on analyzed storm and subsequent forecasts were not considered in
563 this study, or in the real data study of Putnam et al. (2019); they require further research and
564 investigations. The impact of PRD assimilation on other types of precipitation systems also
565 requires study.

566 *Acknowledgements:* This work was primarily supported by the National Key Research and
567 Development Program of China under grant 2018YFC1507604 and the National Natural Science
568 Foundation of China (Grant No. 41730965 and 41405100).

569

570 **References**

- 571 Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea.*
572 *Rev.*, **129**, 2884-2903.
- 573 Brown, B. R., M. M. Bell, and A. J. Frambach, 2016: Validation of simulated hurricane drop size
574 distributions using polarimetric radar. *Geophys Res Lett*, **43**, 910-917.
- 575 Cao, Q., G. Zhang, and M. Xue, 2013: A Variational Approach for Retrieving Raindrop Size
576 Distribution from Polarimetric Radar Measurements in the Presence of Attenuation. *Journal*
577 *of Applied Meteorology Climatol.*, **52**, 169-185.
- 578 Carlin, J. T., J. Gao, J. C. Snyder, and A. V. Ryzhkov, 2017: Assimilation of ZDR Columns for
579 Improving the Spinup and Forecast of Convective Storms in Storm-Scale Models: Proof-of-
580 Concept Experiments. *Mon Weather Rev*, **145**, 5033-5057.
- 581 Dawson, D. T., E. R. Mansell, Y. Jung, L. J. Wicker, M. R. Kumjian, and M. Xue, 2014: Low-
582 Level ZDR Signatures in Supercell Forward Flanks: The Role of Size Sorting and Melting of
583 Hail. *J. Atmos. Sci.*, **71**, 276-299.
- 584 Dixon, M., Z. H. Li, H. Lean, N. Roberts, and S. Ballard, 2009: Impact of Data Assimilation on
585 Forecasting Convection over the United Kingdom Using a High-Resolution Version of the
586 Met Office Unified Model. *Mon Weather Rev*, **137**, 1562-1584.
- 587 Gaspari, G. and S. E. Cohn, 1999: Construction of correlation functions in two and three
588 dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723-757.
- 589 Hu, M. and M. Xue, 2007: Impact of configurations of rapid intermittent assimilation of WSR-
590 88D radar data for the 8 May 2003 Oklahoma City tornadic thunderstorm case. *Mon. Wea.*
591 *Rev.*, **135**, 507-525.
- 592 Hu, M., M. Xue, and K. Brewster, 2006: 3DVAR and cloud analysis with WSR-88D level-II data
593 for the prediction of Fort Worth tornadic thunderstorms. Part I: Cloud analysis and its impact.
594 *Mon. Wea. Rev.*, **134**, 675-698.
- 595 Huang, H., G. Zhang, K. Zhao, S. Liu, L. Wen, G. Chen, and Z. Yang, 2019: Uncertainty in
596 Retrieving Raindrop Size Distribution from Polarimetric Radar Measurements. *J. Atmos.*
597 *Oceanic Technol.*, **36**, 585-605.
- 598 Huuskonen, A., E. Saltkoff, and I. Holleman, 2014: The Operational Weather Radar Network in
599 Europe. *Bull. Amer. Meteor. Soc.*, **95**, 897-907.
- 600 Jung, Y., G. Zhang, and M. Xue, 2008a: Assimilation of simulated polarimetric radar data for a
601 convective storm using the ensemble Kalman filter. Part I: Observation operators for
602 reflectivity and polarimetric variables. *Mon. Wea. Rev.*, **136**, 2228-2245.
- 603 Jung, Y., M. Xue, and G. Zhang, 2010a: Simultaneous estimation of microphysical parameters and
604 the atmospheric state using simulated polarimetric radar data and ensemble Kalman filter in
605 the presence of observation operator error. *Mon Weather Rev*, **138**, 539-562.
- 606 Jung, Y., M. Xue, and G. Zhang, 2010b: Simulations of polarimetric radar signatures of a supercell
607 storm using a two-moment bulk microphysics scheme. *J. Appl. Meteor. Climatol.*, **49**, 146-
608 163.
- 609 Jung, Y., M. Xue, G. Zhang, and J. Straka, 2008b: Assimilation of simulated polarimetric radar
610 data for a convective storm using the ensemble Kalman filter. Part II: Impact of polarimetric
611 data on storm analysis. *Mon. Wea. Rev.*, **136**, 2246-2260.
- 612 Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation, and Predictability*. Cambridge
613 University Press, 341 pp.
- 614 Kumjian, M. R. and A. V. Ryzhkov, 2008: Polarimetric signatures in supercell thunderstorms. *J.*
615 *Appl. Meteor. Climatol.*, **47**, 1940-1961.

616 Li, X. and J. R. Mecikalski, 2010: Assimilation of the dual - polarization Doppler radar data for a
617 convective storm with a warm - rain radar forward operator. *J. Geophys. Res.*, **115**, D16208.

618 Li, X., J. R. Mecikalski, and D. Posselt, 2017: An Ice-Phase Microphysics Forward Model and
619 Preliminary Results of Polarimetric Radar Data Assimilation. *Mon Weather Rev*, **145**, 683-
620 708.

621 Liu, C., M. Xue, and R. Kong, 2019: Direct variational assimilation of radar reflectivity and radial
622 velocity data: Issues with nonlinear reflectivity operator and solutions. *Mon Weather Rev*,
623 Accepted.

624 Milbrandt, J. A. and M. K. Yau, 2005: A multi-moment bulk microphysics parameterization. Part
625 I: Analysis of the role of the spectral shape parameter. *J. Atmos. Sci.*, **62**, 3051-3064.

626 Putnam, B. J., M. Xue, Y. Jung, N. Snook, and G. Zhang, 2014: The Analysis and Prediction of
627 Microphysical States and Polarimetric Radar Variables in a Mesoscale Convective System
628 Using Double-Moment Microphysics, Multinetwork Radar Data, and the Ensemble Kalman
629 Filter. *Mon Weather Rev*, **142**, 141-162.

630 Putnam, B. J., M. Xue, Y. Jung, N. Snook, and G. Zhang, 2017a: Ensemble Probabilistic Prediction
631 of a Mesoscale Convective System and Associated Polarimetric Radar Variables Using
632 Single-Moment and Double-Moment Microphysics Schemes and EnKF Radar Data
633 Assimilation. *Mon Weather Rev*, **145**, 2257-2279.

634 Putnam, B. J., M. Xue, Y. Jung, G. Zhang, and F. Kong, 2017b: Simulation of Polarimetric Radar
635 Variables from 2013 CAPS Spring Experiment Storm-Scale Ensemble Forecasts and
636 Evaluation of Microphysics Schemes. *Mon Weather Rev*, **145**, 49-73.

637 Putnam, B. J., M. Xue, Y. Jung, N. Snook, and G. Zhang, 2019: Ensemble Kalman Filter
638 Assimilation of Polarimetric Radar Observations for the 20 May 2013 Oklahoma Tornadoic
639 Supercell Case. *Mon Weather Rev*, **147**, 2511-2533.

640 Ray, P. S., B. Johnson, K. W. Johnson, J. S. Bradberry, J. J. Stephens, K. K. Wagner, R. B.
641 Wilhelmson, and J. B. Klemp, 1981: The morphology of severe tornadic storms on 20 May
642 1977. *J. Atmos. Sci.*, **38**, 1643-1663.

643 ROC, 2013: WSR-88D dual polarization deployment progress. *NOAA/RadarOperations Center*,
644 6 pp. [Available online at
645 <http://www.roc.noaa.gov/WSR88D/PublicDocs/DualPol/DPstatus.pdf>.]

646 Ryzhkov, A. V. and D. S. Zrnic, 1998: Discrimination between Rain and Snow with a Polarimetric
647 Radar. *Journal of Applied Meteorology*, **37**, 1228-1240.

648 Sobash, R. A. and D. J. Stensrud, 2013: The Impact of Covariance Localization for Radar Data on
649 EnKF Analyses of a Developing MCS: Observing System Simulation Experiments. *Mon*
650 *Weather Rev*, **141**, 3691-3709.

651 Stensrud, D. J., M. Xue, L. J. Wicker, K. E. Kelleher, M. P. Foster, J. T. Schaefer, R. S. Schneider,
652 S. G. Benjamin, S. S. Weygandt, J. T. Ferree, and J. P. Tuell, 2009: Convective-Scale Warn-
653 on-Forecast System. *Bull. Amer. Meteor. Soc.*, **90**, 1487-1500.

654 Sun, J., M. Xue, J. W. Wilson, I. Zawadzki, S. P. Ballard, J. Onvlee-Hooimeyer, P. Joe, D. M.
655 Barker, P.-W. Li, B. Golding, M. Xu, and J. Pinto, 2013: Use of NWP for Nowcasting
656 Convective Precipitation: Recent Progress and Challenges. *Bull. Amer. Meteor. Soc.*, **95**, 409-
657 426.

658 Tong, M. and M. Xue, 2005: Ensemble Kalman filter assimilation of Doppler radar data with a
659 compressible nonhydrostatic model: OSS Experiments. *Mon. Wea. Rev.*, **133**, 1789-1807.

660 Vivekanandan, J., W. M. Adams, and V. N. Bringi, 1991: Rigorous Approach to Polarimetric
661 Radar Modeling of Hydrometeor Orientation Distributions. *Journal of Applied Meteorology*,
662 **30**, 1053-1063.

663 Vivekanandan, J., S. M. Ellis, R. Oye, D. S. Zrnic, A. V. Ryzhkov, and J. Straka, 1999: Cloud
664 microphysics retrieval using S-band dual-polarization radar measurements. *B. Amer. Meteor.*
665 *Soc.*, **80**, 381-388.

666 Whitaker, J. S. and T. M. Hamill, 2002: Ensemble data assimilation without perturbed
667 observations. *Mon. Wea. Rev.*, **130**, 1913-1924.

668 Wu, B., J. Verlinde, and J. Sun, 2000: Dynamical and microphysical retrievals from Doppler radar
669 observations of a deep convective cloud. *J. Atmos. Sci.*, **57**, 262-283.

670 Xiao, Q., J. Sun, W.-C. Lee, Y.-H. Kuo, D. M. Barker, E. Lim, D.-J. Won, M.-S. Lee, W.-J. Lee,
671 J.-Y. Cho, D.-K. Lee, and H.-S. Lee, 2008: Doppler Radar Data Assimilation in KMA's
672 Operational Forecasting. *Bull. Amer. Meteor. Soc.*, **89**, 39-43.

673 Xue, M., M. Tong, and K. K. Droegemeier, 2006: An OSSE framework based on the ensemble
674 square-root Kalman filter for evaluating impact of data from radar networks on thunderstorm
675 analysis and forecast. *J. Atmos. Oceanic Technol.*, **23**, 46-66.

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

679 Xue, M., K. K. Droegemeier, V. Wong, A. Shapiro, K. Brewster, F. Carr, D. Weber, Y. Liu, and
680 D. Wang, 2001: The Advanced Regional Prediction System (ARPS) - A multiscale
681 nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and
682 applications. *Meteor. Atmos. Phys.*, **76**, 143-165.

683 Zhang, G., V. N. Mahale, B. J. Putnam, Y. Qi, Q. Cao, A. D. Byrd, P. Bukovcic, D. S. Zrnic, J.
684 Gao, M. Xue, Y. Jung, H. D. Reeves, P. L. Heinselman, A. Ryzhkov, R. D. Palmer, P. Zhang,
685 M. Weber, G. M. Mcfarquhar, B. Moore, Y. Zhang, J. Zhang, J. Vivekanandan, Y. Al-Rashid,
686 R. L. Ice, D. S. Berkowitz, C. Tong, C. Fulton, and R. J. Doviak, 2019: Current Status and
687 Future Challenges of Weather Radar Polarimetry: Bridging the Gap between Radar
688 Meteorology/Hydrology/Engineering and Numerical Weather Prediction. *Adv. Atmos. Sci.*,
689 **36**, 571-588.

690 Zhao, K., H. Huang, M. Wang, W.-C. Lee, G. Chen, L. Wen, J. Wen, G. Zhang, M. Xue, Z. Yang,
691 L. Liu, C. Wu, Z. Hu, and S. Chen, 2019: Recent Progress in Dual-Polarization Radar
692 Research and Applications in China. *Adv. Atmos. Sci.*, **36**, 961-974.

693 Zhu, K., Y. Yang, and M. Xue, 2015: Percentile-based neighborhood precipitation verification and
694 its application to a landfalling tropical storm case with radar data assimilation. *Adv. Atmos.*
695 *Sci.*, **32**, 1449-1459.

696

697 **List of Figures**

698 Fig. 1 The RMSEs of the ensemble-mean forecasts and analyses throughout the 5-min DA cycles,
699 for experiments VrZh (black), VrZhZdr (red) and VrZhZdr_NoNt (blue), averaged over points
700 at which the observed reflectivity is greater than 10 dBZ. (a) u , (b) v , (c) w , (d) water vapor content
701 qv , (e) perturbation potential temperature θ' ,

702 Fig. 3 The profiles of horizontally averaged ensemble mean analysis RMSEs averaged over points
703 at which the truth reflectivity is greater than 10 dBZ at 80 minutes for variables (a) u , (b) v , (c) w ,
704 (d) water vapor content qv , (e) perturbation potential temperature θ' , (f) perturbation pressure p' ,
705 (g) mixing ratio of cloud water qc and (h) total number concentrations of cloud water, mixing
706 ratios of (i) rain qr , (j) cloud ice qi , (k) snow qs and graupel qg , (l) hail qh , and (m)-(p) their
707 corresponding total number concentrations.

708 Fig. 4 Truth and analyzed differential reflectivity ZDR (color shading) and reflectivity ZH (black
709 contours, starting from 35 dBZ at a 10 dBZ interval) at a 500 m altitude at 80 minutes for (a) Truth,
710 (b) VrZh and (c) VrZhZdr.

711 Fig. 5 Similar as Fig. 4 (a)-(c) but at 110 minutes.

712 Fig. 6 Vertical slices along the red dashed line in Fig. 4, of reflectivity (shaded, upper row, unit:
713 dBZ), differential reflectivity ZDR (middle row, unit: dB) and specific differential phase KDP
714 (bottom row, unit: $^{\circ}\text{km}^{-1}$), of truth (right column), experiment VrZh (central column) and VrZhZdr
715 (right column). The black contours are for 10 m s⁻¹ vertical velocity w and the red contours are
716 for temperature from 0 (bold) to -20 at 10 $^{\circ}\text{C}$ intervals.

717 Fig. 7 Vertical cross sections through the same slice at Fig. 4, of the mixing ratios of cloud water
718 qc (shaded) for (a) Truth, (b) VrZh and (c) VrZhZdr. (d)-(f) and (g)-(i) are similar to (a)-(c) but for
719 rain water and hail.

720 Fig. 8 Logarithm of rainwater number concentrations ($\log_{10}(Nqr)$) at 3 km height of (a) Truth, (b)
721 VrZhZdr and (c) VrZhZdr_NoNt

722 Fig. 9. Correlation coefficients between ZDR prior (at $(x, y, z) = (34, 22, 3.5)$ km, black dot in (a))
723 and model hydrometeor state variables.

724 Fig. 10. Correlation coefficients between ZDR prior (at $(x, y, z) = (64, 32, 1.8)$ km, black dot in (a))
725 and model hydrometeor state variables from control experiment at 80 minutes.

726 Fig. 11. Forecast and analysis RMSEs of experiments VrZhZdr_NoPt (green), VrZhZdr_NoQv
727 (purple), VrZhZdr_NoW (red), VrZhZdr_NoW (blue) and VrZhZdr (black) for variables (a) u , (b)
728 v , (c) w , (d) water vapor mixing ratio qv , (e) perturbation potential temperature θ' , (f) perturbation
729 pressure p' , mixing ratios of (g) cloud water qc , (h) rain qr , (i) cloud ice qi , (j) snow qs , (k) graupel
730 qg , (l) hail qh .

731 Fig. 12. Similar to Fig. 11 but for vertical profiles of analysis RMSEs for different variables at 80
732 minutes.

733 Fig. 13. Similar as Fig. 9 but for correlations between ZDR and (a) u , (b) v , (c) w , (d) qv and (e)
734 θ' , respectively.

735 Fig. 14. Vertical vorticity (shaded), velocity (contours) and horizontal wind vectors at 2 km height
736 for (a) Truth, (b) VrZhZdr, (c) VrZhZdr_NoUV, (d) VrZhZdr_NoW

737 Fig. 15. Vertical south-north slice of vertical vorticity (shaded) and velocity (contours) through the
738 maximum w ($x = 35.5$ km for VrZhZdr_NoW and $x = 36.5$ km for truth and other experiments in
739 Fig. 14) (a) Truth, (b) VrZhZdr, (c) VrZhZdr_NoUV, (d) VrZhZdr_NoW

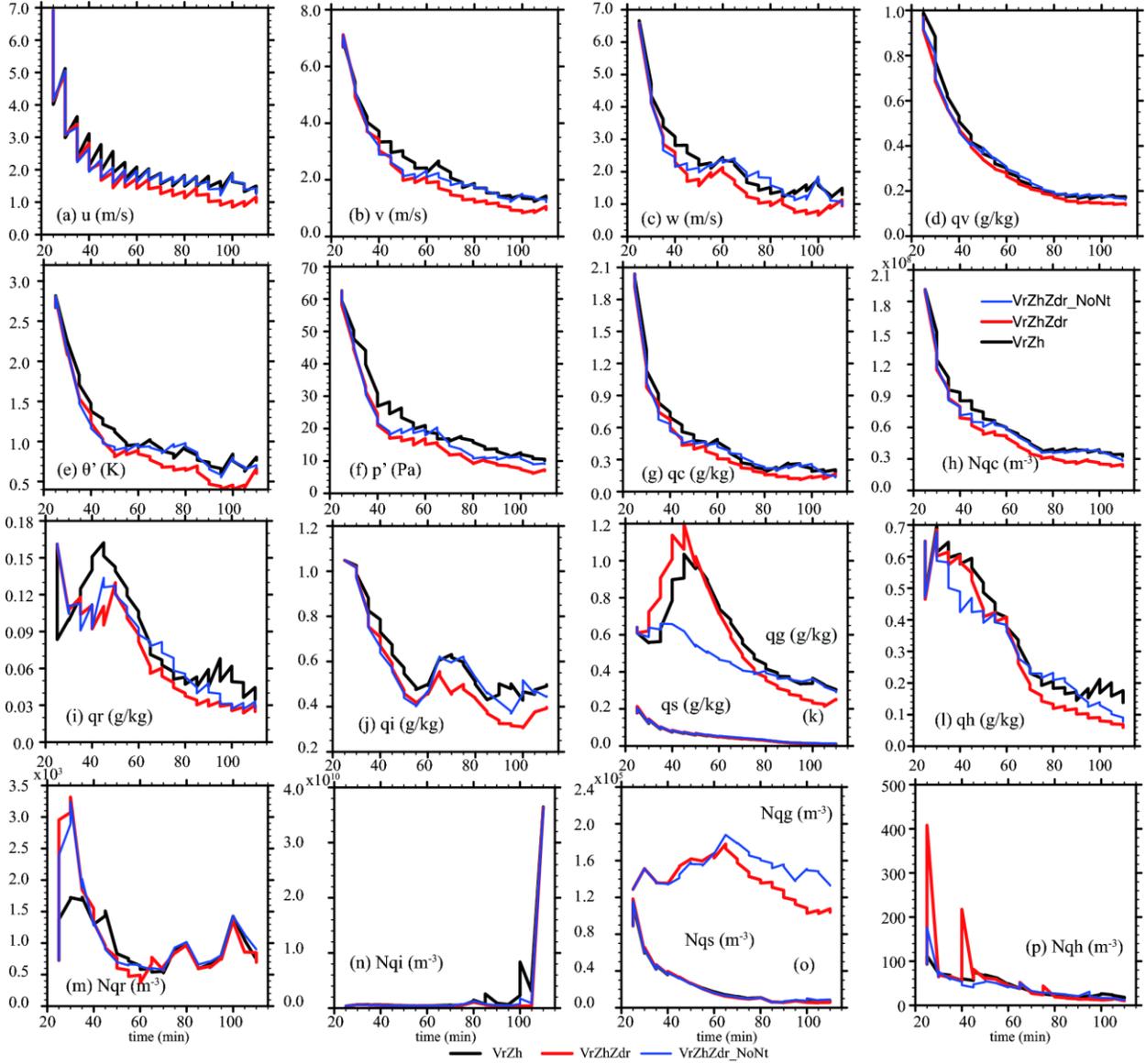


Fig. 1 The RMSEs of the ensemble-mean forecasts and analyses throughout the 5-min DA cycles, for experiments VrZh (black), VrZhZdr (red) and VrZhZdr_NoNt (blue), averaged over points at which the observed reflectivity is greater than 10 dBZ. (a) u , (b) v , (c) w , (d) water vapor content q_v , (e) perturbation potential temperature θ' , (f) perturbation pressure p' , (g) mixing ratio of cloud water q_c , (h) total number concentrations of cloud water, mixing ratios of (i) rain q_r , (j) cloud ice q_i , (k) snow q_s and graupel q_g , (l) hail q_h , and (m)-(p) their corresponding number concentrations.

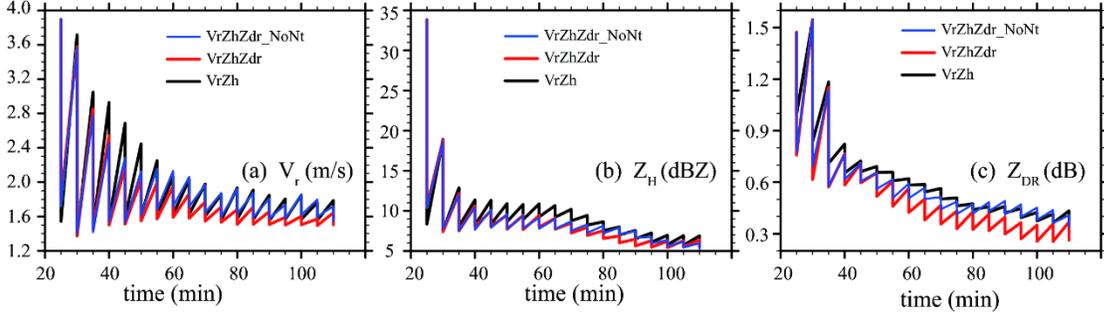


Fig. 2 Similar to Fig. 1 but for radar variables (a) V_r , (b) Z_H and (c) Z_{DR} . Here, the ensemble-mean forecasts and analyses are converted into the observation space using the observation operators, and the RMSEs of the simulated radar variables are calculated against the corresponding observations.

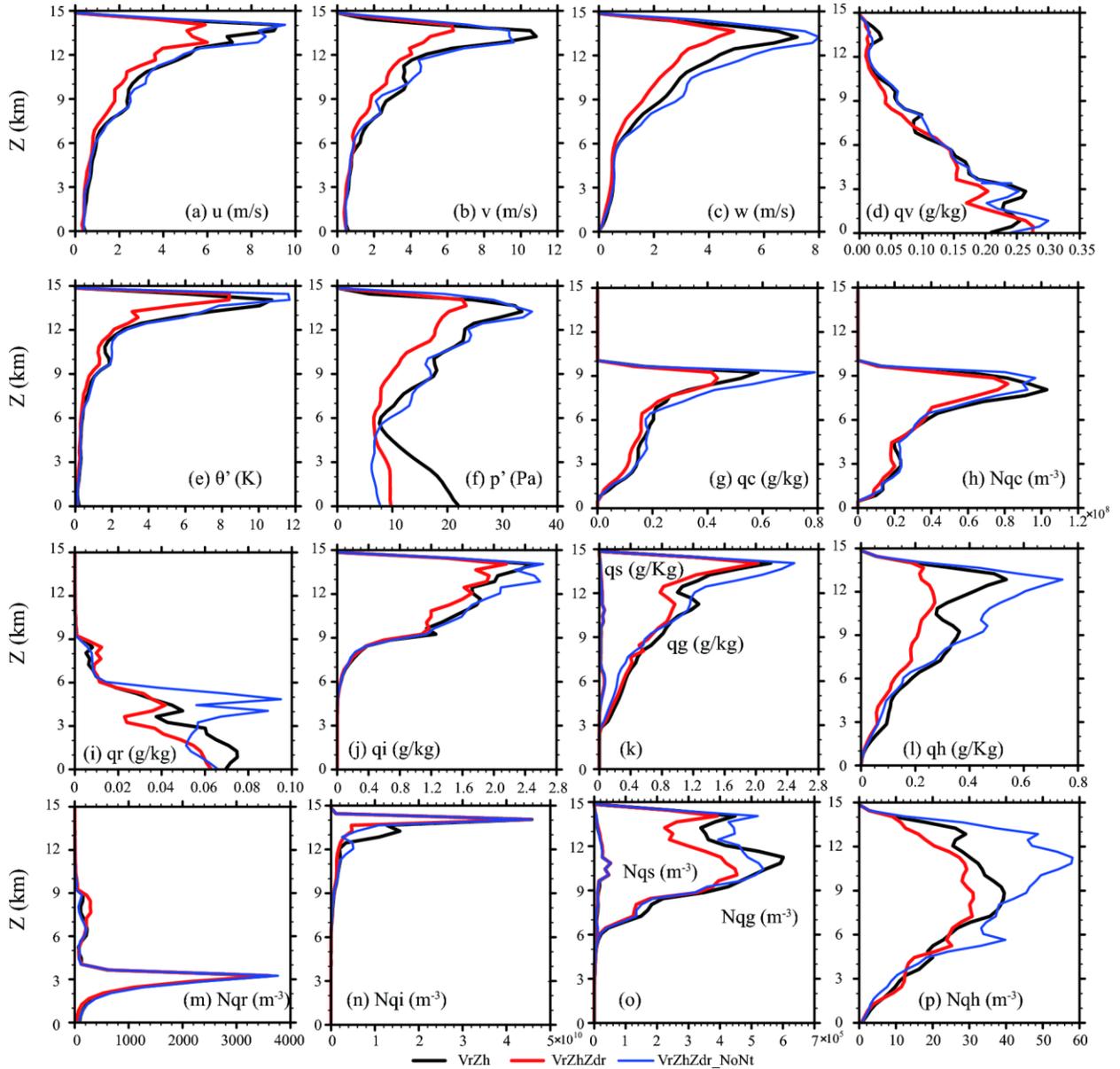


Fig. 3 The profiles of horizontally averaged ensemble mean analysis RMSEs averaged over points at which the truth reflectivity is greater than 10 dBZ at 80 minutes for variables (a) u , (b) v , (c) w , (d) water vapor content q_v , (e) perturbation potential temperature θ' , (f) perturbation pressure p' , (g) mixing ratio of cloud water q_c and (h) total number concentrations of cloud water, mixing ratios of (i) rain q_r , (j) cloud ice q_i , (k) snow q_s and graupel q_g , (l) hail q_h , and (m)-(p) their corresponding total number concentrations. Error profiles for experiments VrZh, VrZhZdr and VrZhZdr_NoNt are shown in black, red and blue, respectively. The RMSEs were calculated at 1 km intervals in the vertical.

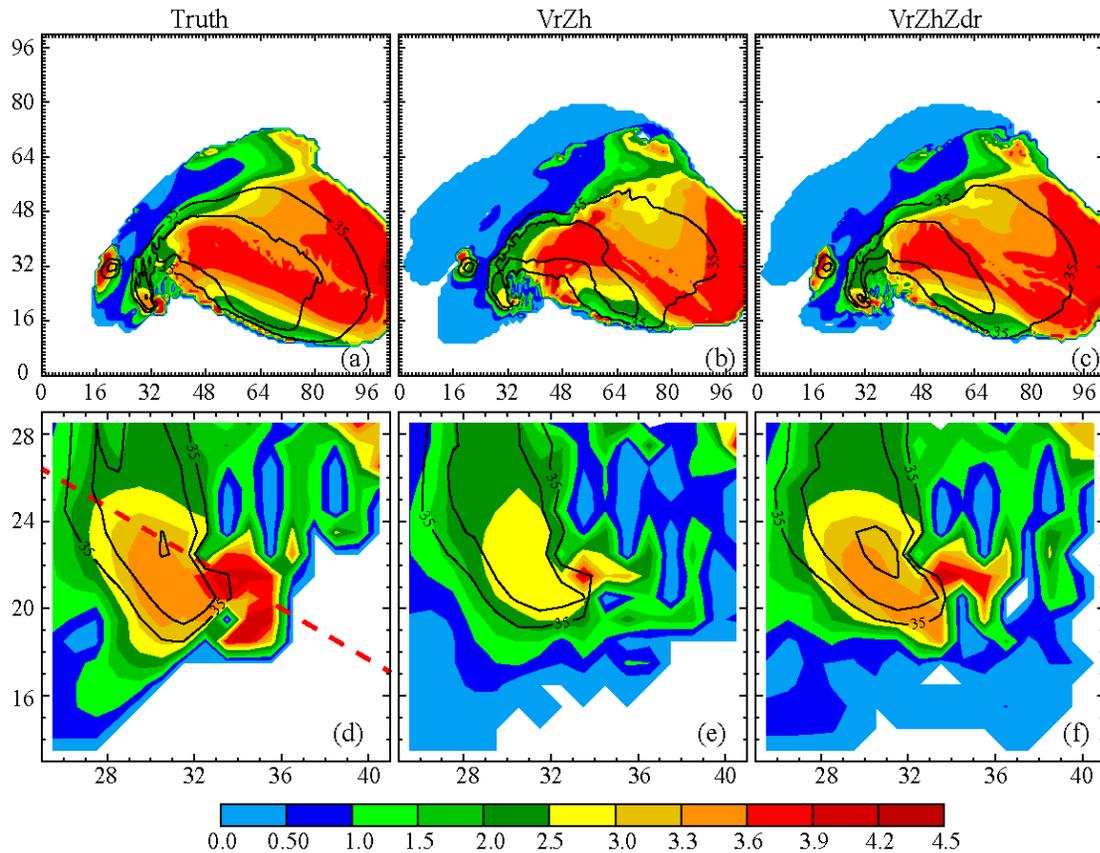


Fig. 4 Truth and analyzed differential reflectivity Z_{DR} (color shading) and reflectivity Z_H (black contours, starting from 35 dBZ at a 10 dBZ interval) at a 500 m altitude at 80 minutes for (a) Truth, (b) VrZh and (c) VrZhZdr. (d)-(f) are the zoomed-in plots for the hook echo area of the (a)-(c) plots, respectively.

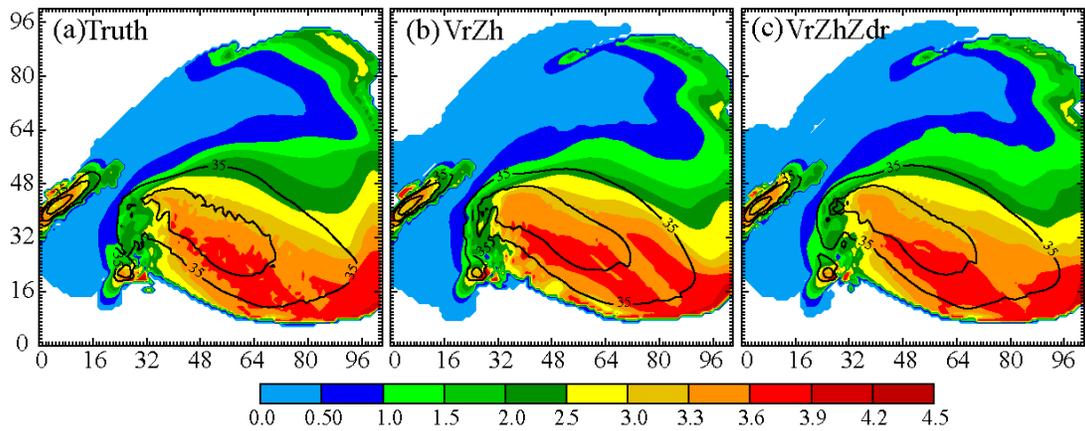


Fig. 5 Similar as Fig. 4 (a)-(c) but at 110 minutes.

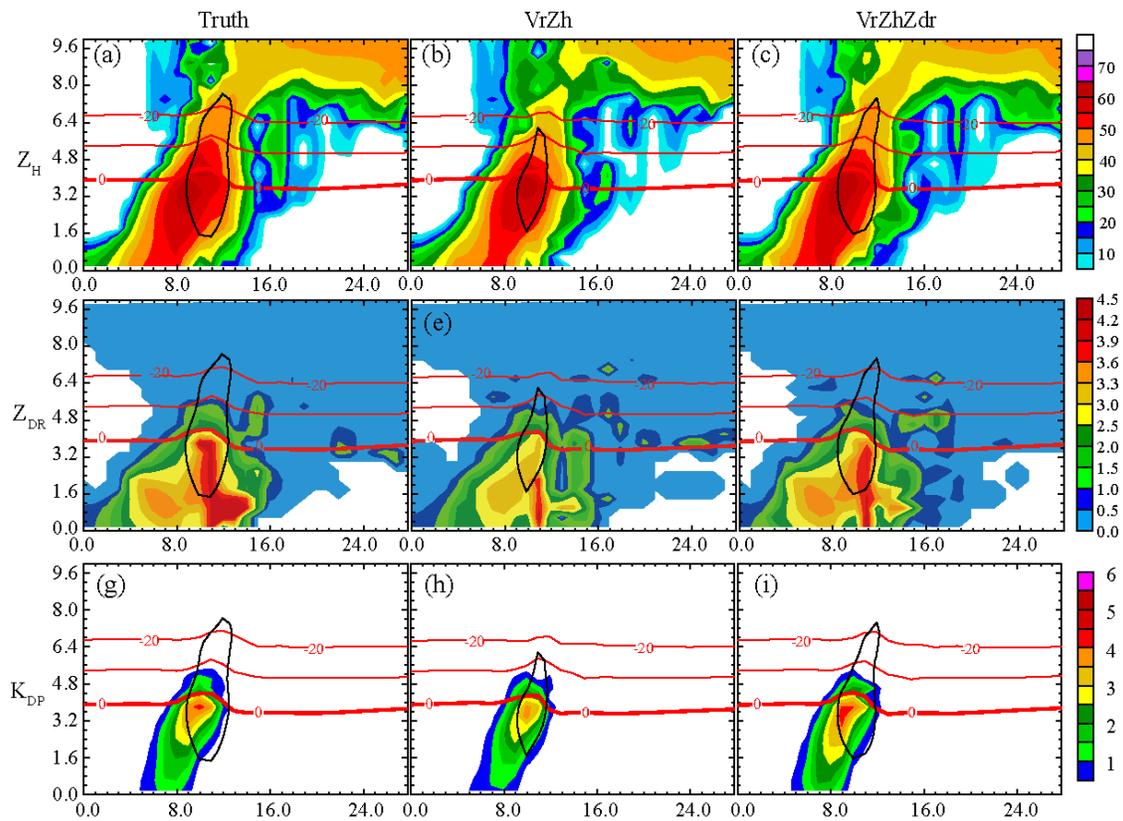


Fig. 6 Vertical slices along the red dashed line in Fig. 4, of reflectivity (shaded, upper row, unit: dBZ), differential reflectivity Z_{DR} (middle row, unit: dB) and specific differential phase K_{DP} (bottom row, unit: $^{\circ}\text{km}^{-1}$), of truth (right column), experiment VrZh (central column) and VrZhZdr

(right column). The black contours are for 10 m s^{-1} vertical velocity w and the red contours are for temperature from 0 (bold) to -20 at 10°C intervals.

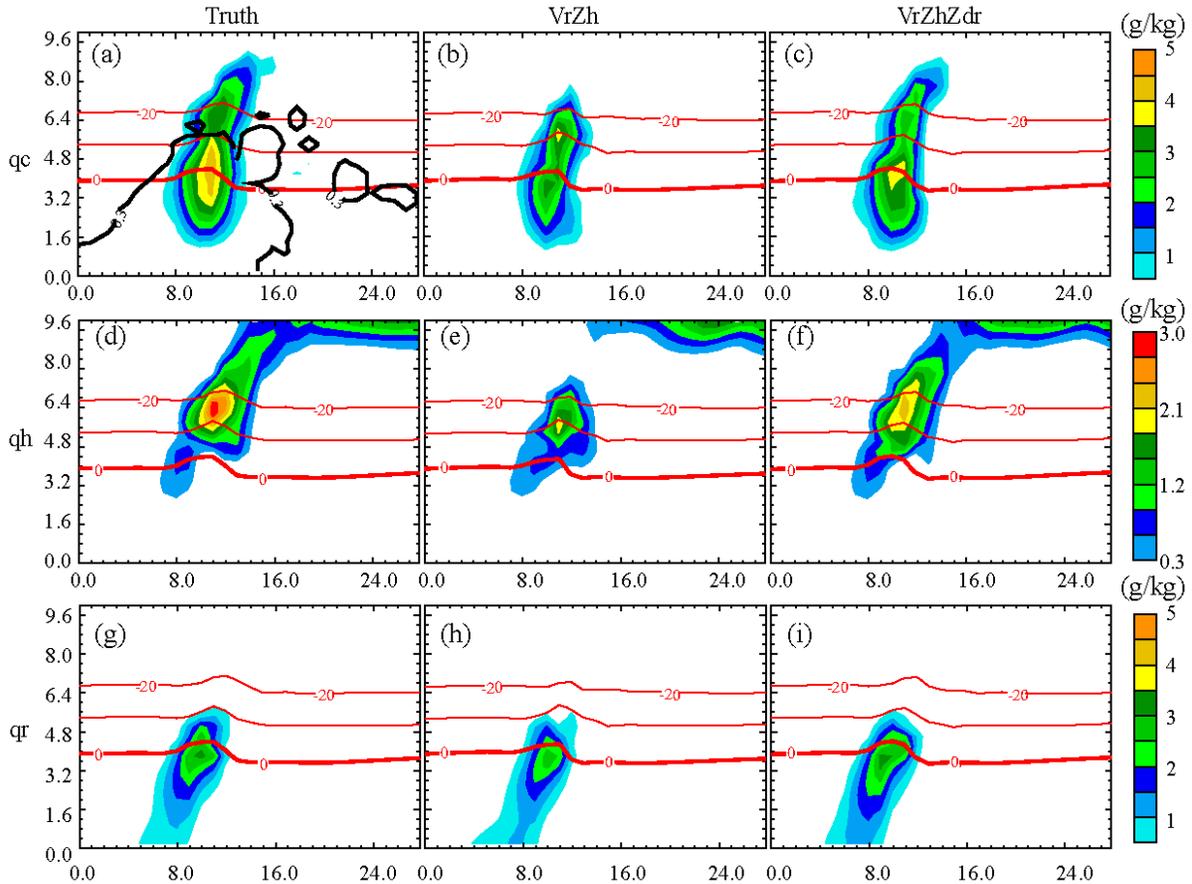


Fig. 7 Vertical cross sections through the same slice at Fig. 4, of the mixing ratios of cloud water q_c (shaded) for (a) Truth, (b) VrZh and (c) VrZhZdr. (d)-(f) and (g)-(i) are similar to (a)-(c) but for rain water and hail. The black contour in (a) is for $0.3 \text{ dB } Z_{DR}$ and the red contours are for temperature from 0 (bold) to -20 at 10°C intervals.

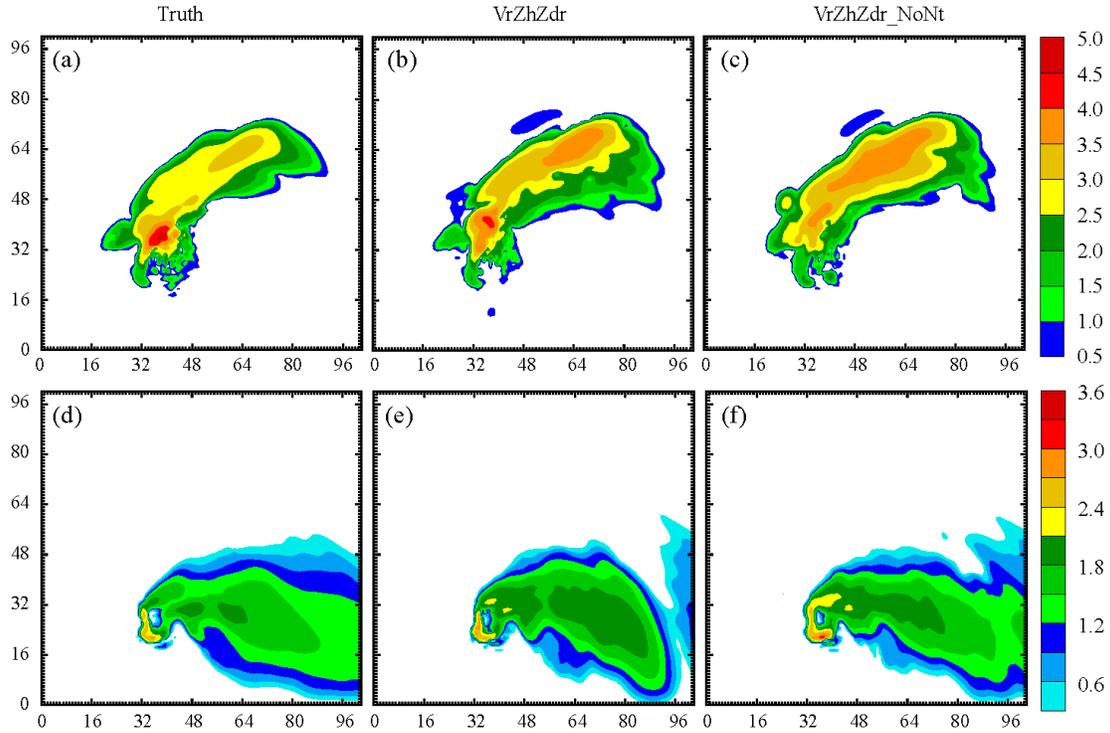


Fig. 8 Logarithm of rainwater number concentrations ($\log_{10}(Nq_r)$) at 3 km height of (a) Truth, (b) VrZhZdr and (c) VrZhZdr_NoNt. (d)-(f) are similar to (a)-(c) but for hail number concentrations at 6 km height.

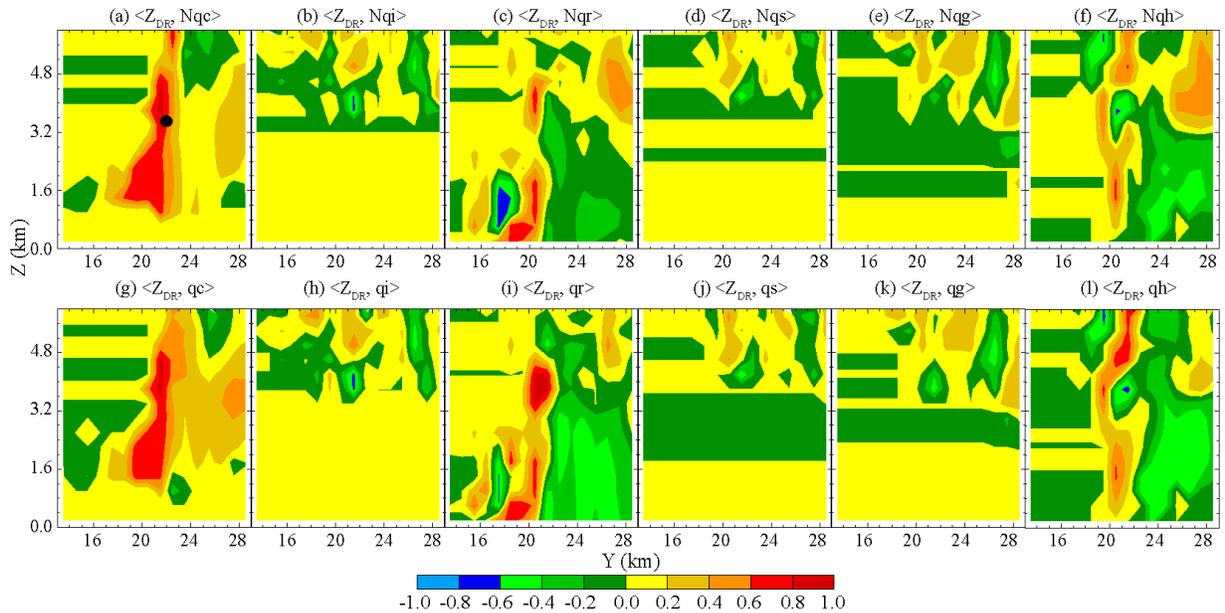


Fig. 9. Correlation coefficients between Z_{DR} prior (at $(x, y, z) = (34, 22, 3.5)$ km, black dot in (a)) and model hydrometeor state variables. Y-Z vertical slice across $x=34$ km was plotted.

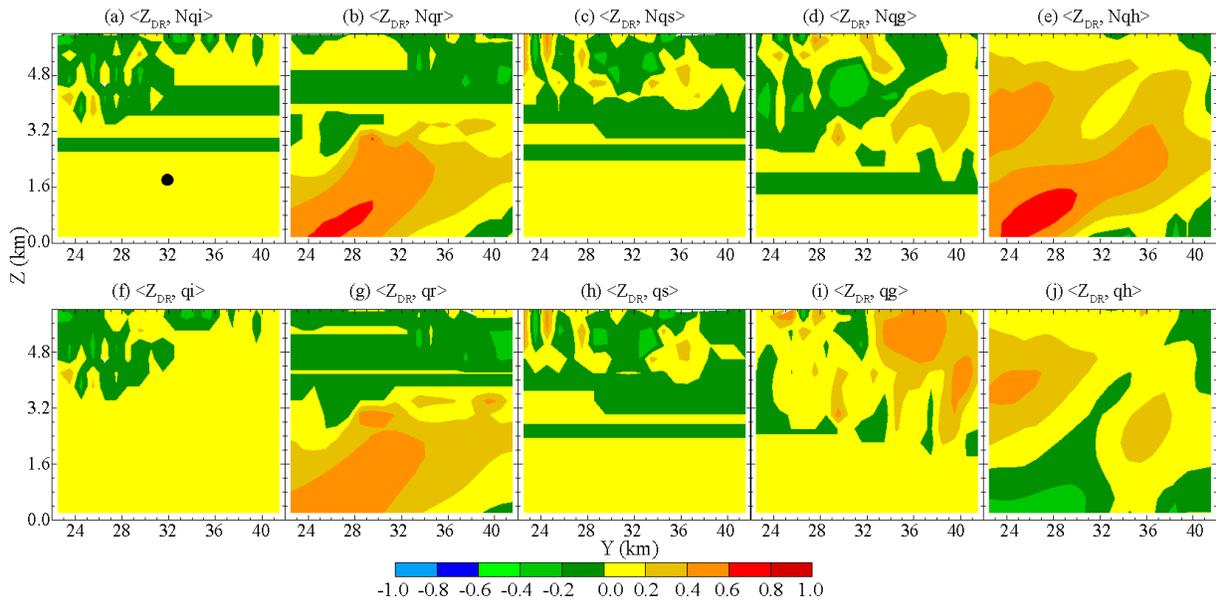


Fig. 10. Correlation coefficients between Z_{DR} prior (at $(x, y, z) = (64, 32, 1.8)$ km, black dot in (a)) and model hydrometeor state variables from control experiment at 80 minutes. Y-Z vertical slice across $x=64$ km was plotted.

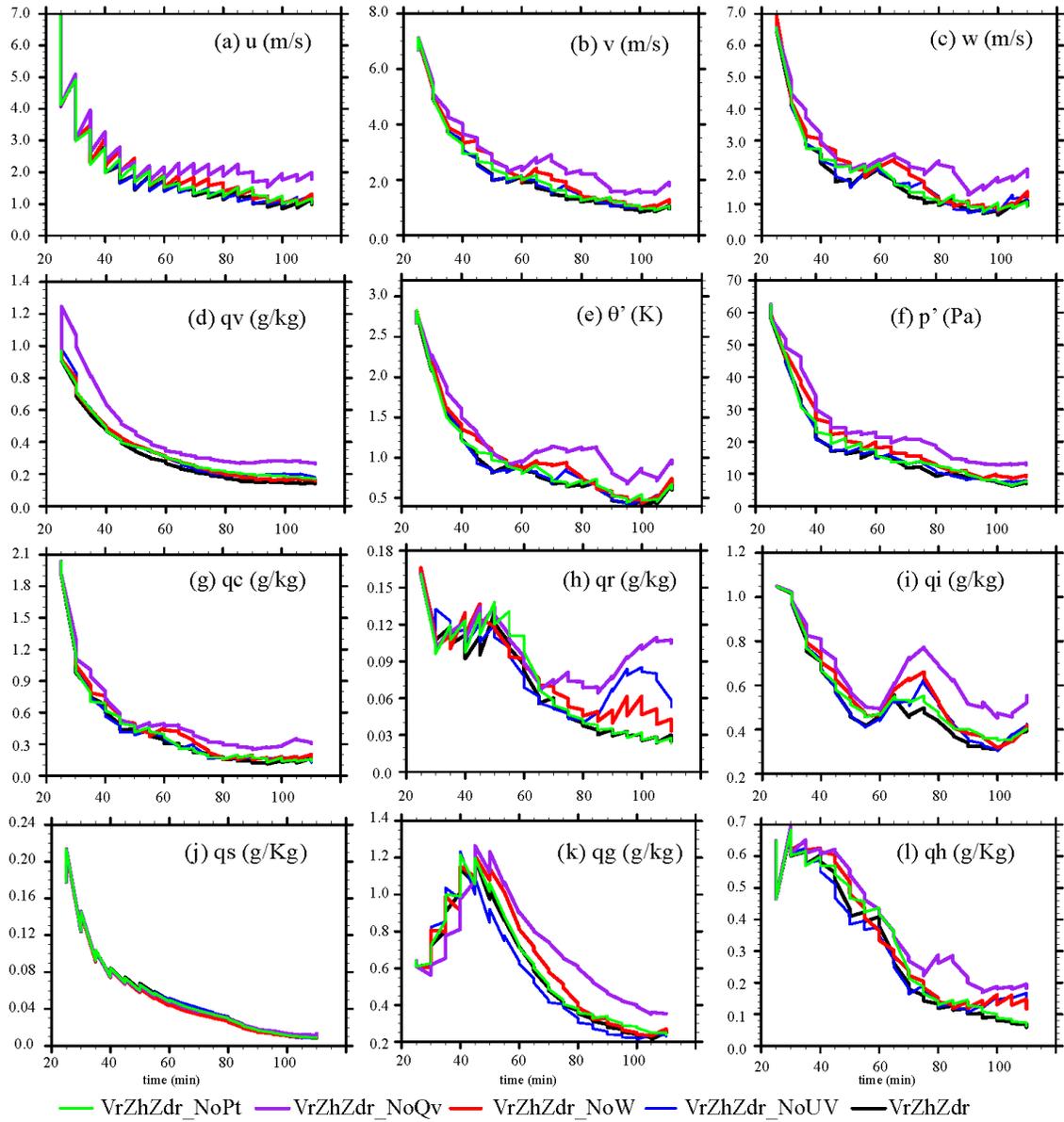


Fig. 11. Forecast and analysis RMSEs of experiments VrZhZdr_NoPt (green), VrZhZdr_NoQv (purple), VrZhZdr_NoW (red), VrZhZdr_NoW (blue) and VrZhZdr (black) for variables (a) u , (b) v , (c) w , (d) water vapor mixing ratio q_v , (e) perturbation potential temperature θ' , (f) perturbation pressure p' , mixing ratios of (g) cloud water q_c , (h) rain q_r , (i) cloud ice q_i , (j) snow q_s , (k) graupel q_g , (l) hail q_h .

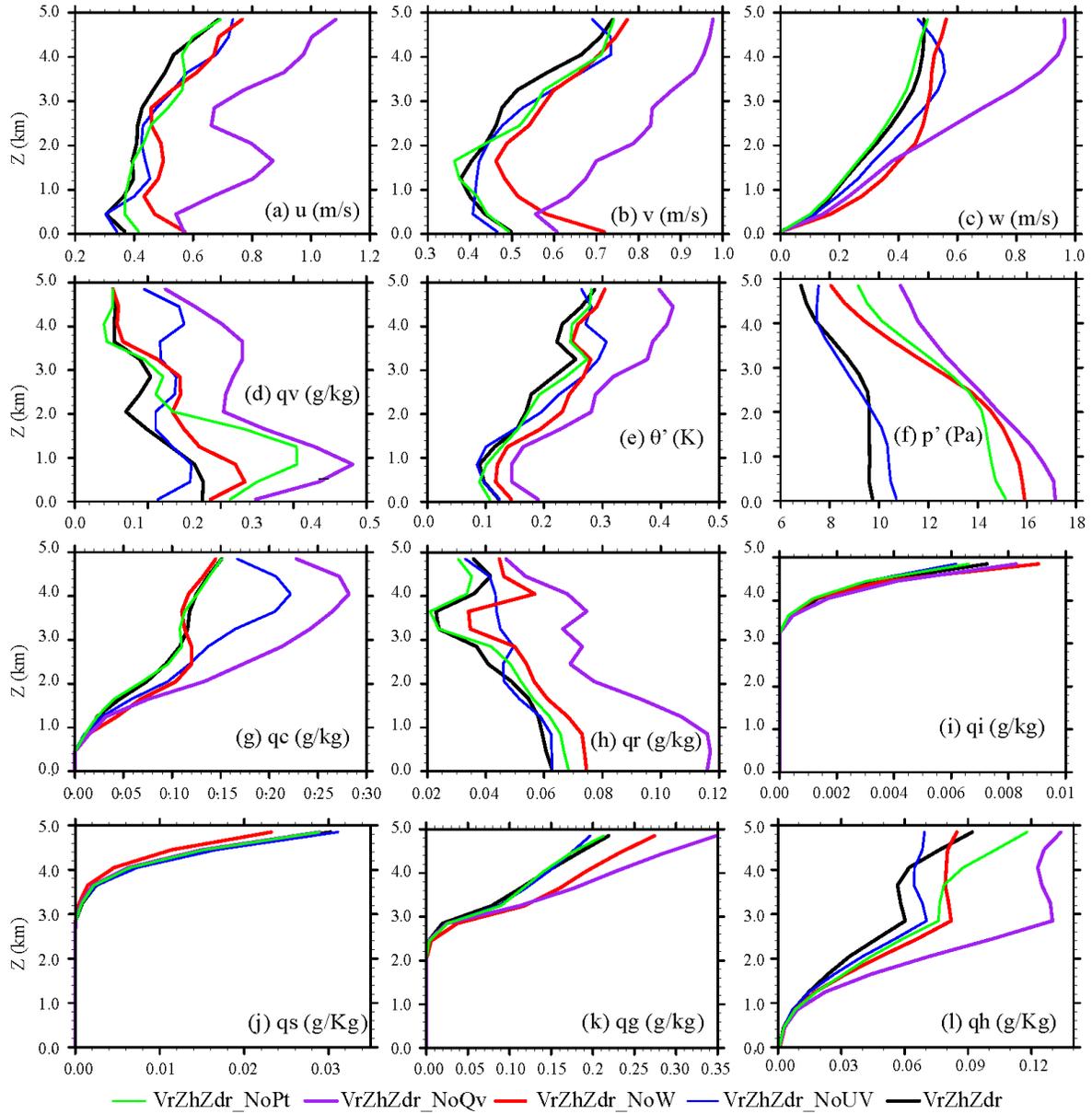


Fig. 12. Similar to Fig. 11 but for vertical profiles of analysis RMSEs for different variables at 80 minutes.

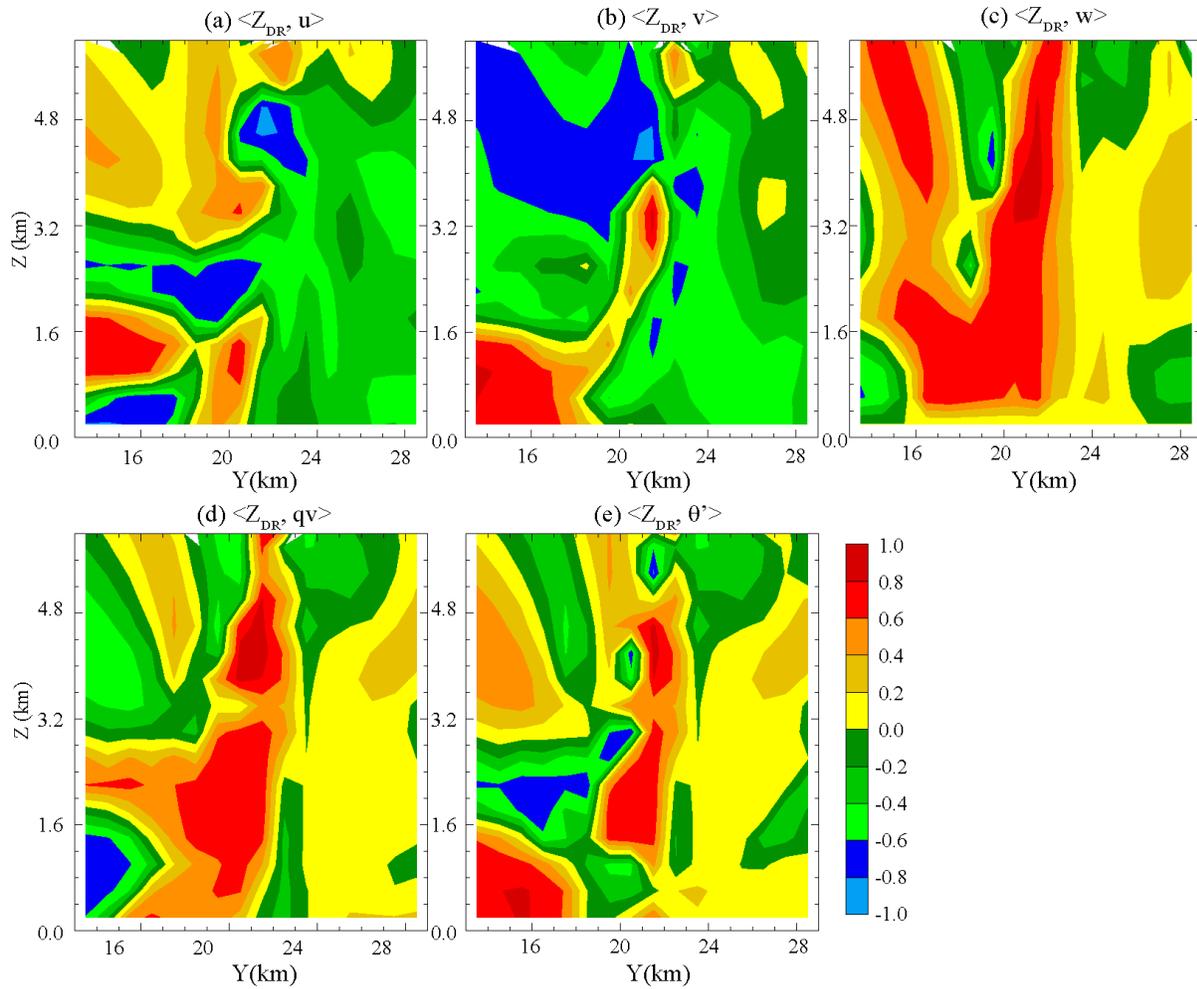


Fig. 13. Similar as Fig. 9 but for correlations between Z_{DR} and (a) u , (b) v , (c) w , (d) q_v and (e) θ' , respectively.

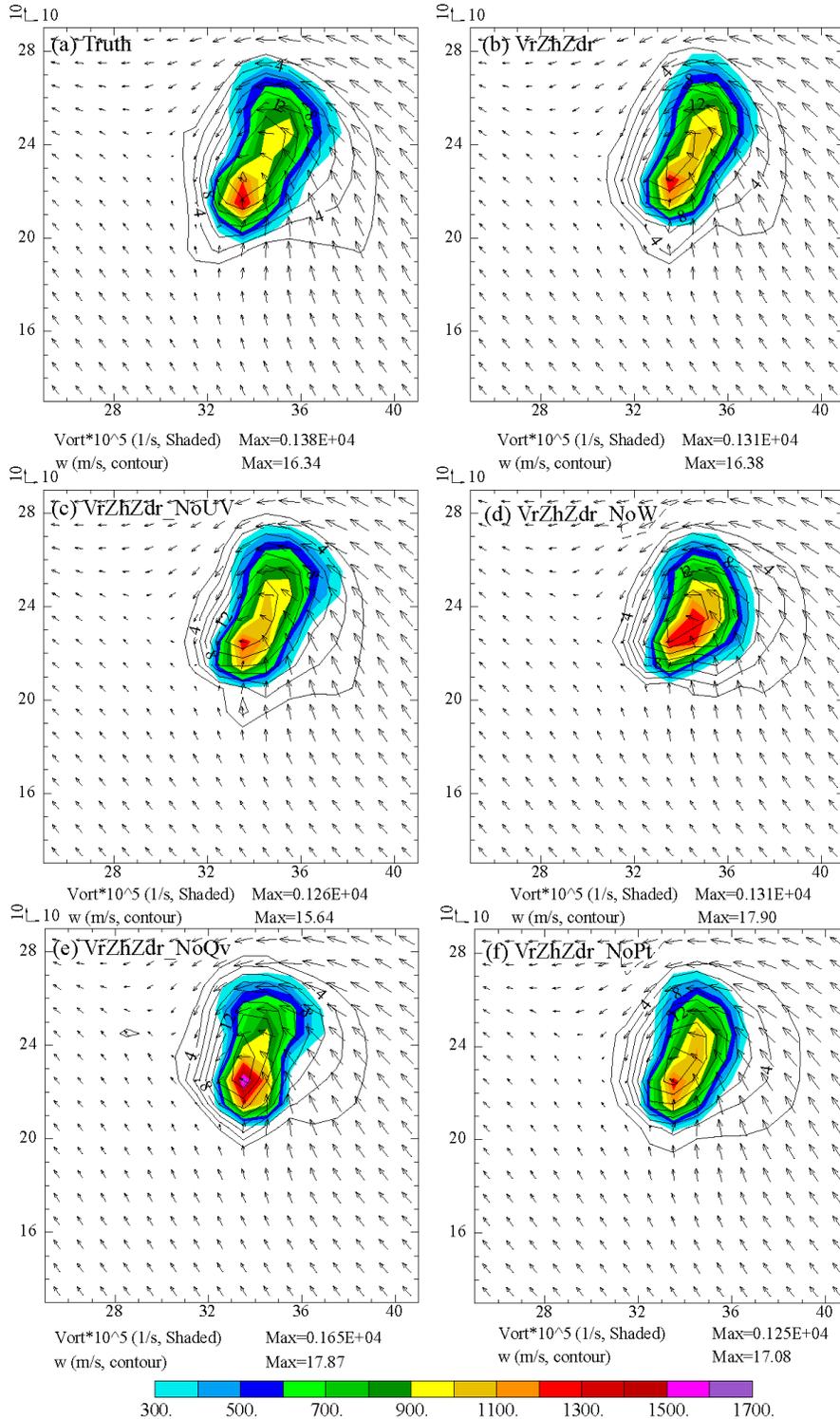


Fig. 14. Vertical vorticity (shaded), velocity (contours) and horizontal wind vectors at 2 km height for (a) Truth, (b) VrZhZdr, (c) VrZhZdr_NoUV, (d) VrZhZdr_NoW, (e) VrZhZdr_NoQv and (f) VrZhZdr_NoPt.

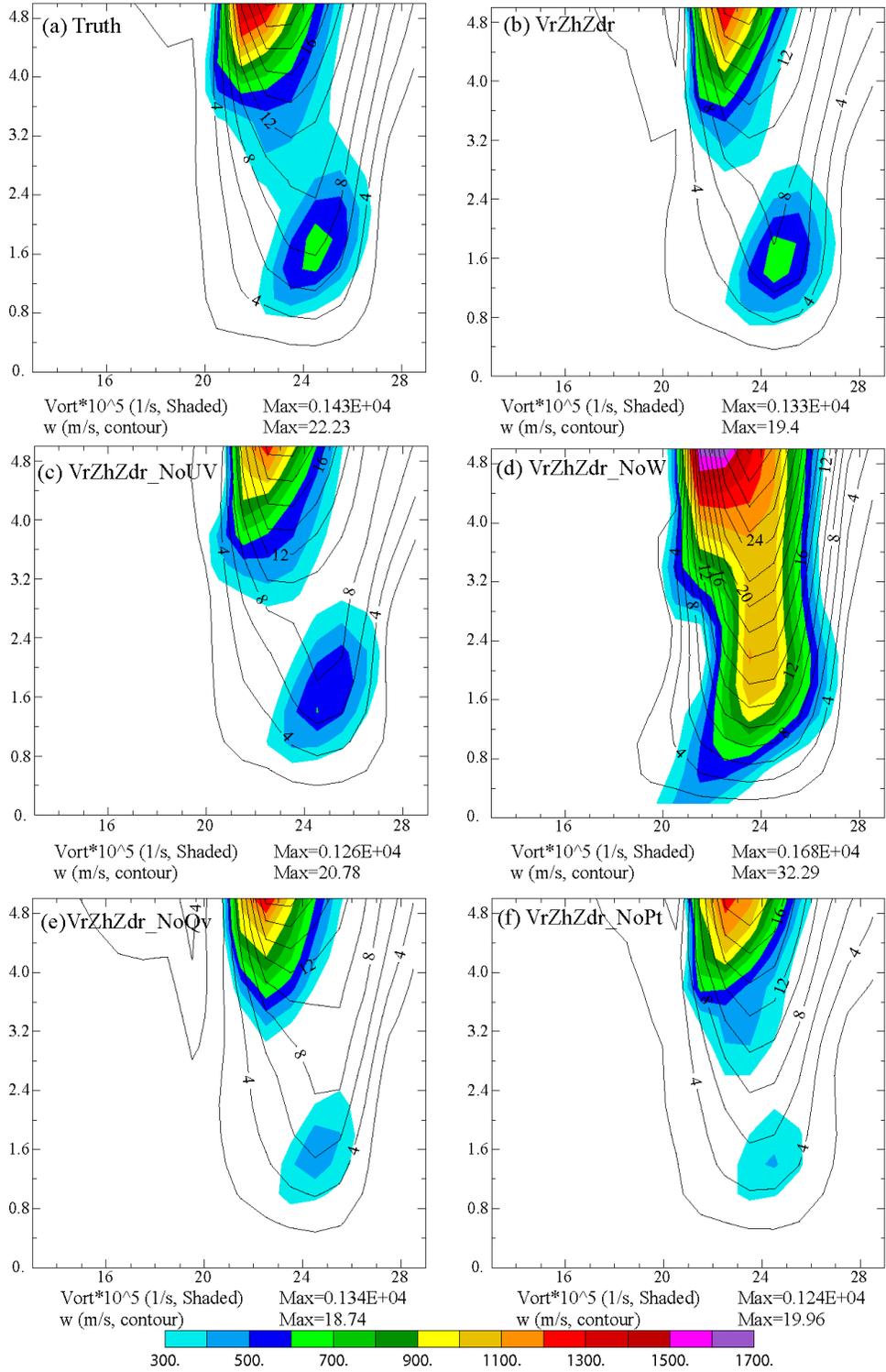


Fig. 15. Vertical south-north slice of vertical vorticity (shaded) and velocity (contours) through the maximum w ($x = 35.5$ km for VrZhZdr_NoW and $x = 36.5$ km for truth and other experiments in Fig. 14) (a) Truth, (b) VrZhZdr, (c) VrZhZdr_NoUV, (d) VrZhZdr_NoW, (e) VrZhZdr_NoQv and (f) VrZhZdr_NoPt.

Table 1. List of experiments

Experiment name	Assimilated data	State variables updated by EnKF
VrZh	V_r, Z_H	$u, v, w, p, \theta, q_v, q_x$, and Nq_x ($x=c, r, i, s, g$ or h)
VrZhZdr	V_r, Z_H, Z_{DR}	As VrZh but with additional Z_{DR} DA
VrZhZdr_NoNt	V_r, Z_H, Z_{DR}	As VrZhZdr but Z_{DR} DA does not update Nq_x . Note that Nq_x are still updated by V_r and Z_H . Similarly for the following experiments.
VrZhZdr_NoUV	V_r, Z_H, Z_{DR}	As VrZhZdr but Z_{DR} DA does not update u and v
VrZhZdr_NoW	V_r, Z_H, Z_{DR}	As VrZhZdr but for Z_{DR} DA does not update w
VrZhZdr_NoPt	V_r, Z_H, Z_{DR}	As VrZhZdr but Z_{DR} DA does not update θ'
VrZhZdr_NoQv	V_r, Z_H, Z_{DR}	As VrZhZdr but Z_{DR} does not update q_v