

## RESEARCH ARTICLE

# Assimilating polarimetric radar data with an ensemble Kalman filter: OSSEs with a tornadic supercell storm simulated with a two-moment microphysics scheme

Kefeng Zhu<sup>1</sup>  | Ming Xue<sup>1,2</sup>  | Kun Ouyang<sup>1</sup> | Youngsun Jung<sup>2</sup>

<sup>1</sup>Key Laboratory of Mesoscale Severe Weather/Ministry of Education and School of Atmospheric Sciences, Nanjing University, Nanjing, China

<sup>2</sup>Center for Analysis and Prediction of Storms and School of Meteorology, University of Oklahoma, Norman, Oklahoma

**Correspondence**

M. Xue, Center for Analysis and Prediction of Storms, University of Oklahoma, 120 David Boren Blvd, Norman, OK 73072, USA.  
Email: mxue@ou.edu

**Funding information**

National Natural Science Foundation of China, Grant/Award Number: 41730965; the National Key Research and Development Program of China, Grant/Award Number: 2018YFC1507604

**Abstract**

The impact of assimilating differential reflectivity  $Z_{DR}$  in addition to reflectivity ( $Z_H$ ) and radial velocity ( $V_r$ ) from a polarimetric radar on the analysis of a tornadic supercell storm using an ensemble Kalman filter (EnKF) is studied in an observing system simulation experiment (OSSE) framework assuming a perfect forecast model. A double-moment microphysics scheme is used to allow for proper simulation of polarimetric signatures. Root-mean-square errors of analysed state variables are calculated and the structure and intensity of analysed fields and derived quantities are examined. Compared to the baseline experiment assimilating radial velocity and reflectivity only, the assimilation of additional  $Z_{DR}$  further reduces the errors of all state variables. The analysed hydrometeor fields are improved in both pattern and intensity distributions. Polarimetric signatures including  $Z_{DR}$  and  $K_{DP}$  columns, and  $Z_{DR}$  arc in the supercell, are much better reproduced. Sensitivity experiments are performed that exclude the updating of hydrometeor number concentrations by  $Z_{DR}$  or of state variables not directly linked to  $Z_{DR}$  via observation operators. The results show that if number concentrations are not updated together with the mixing ratios, most of the benefit of assimilating  $Z_{DR}$  is lost. Among other state variables, the updating of water vapour mixing ratio  $q_v$  has the largest positive impact while the impact of updating vertical wind  $w$  comes in second. The updating of horizontal wind components or temperature has a much smaller but still noticeable impact. Reliable flow-dependent cross-covariances among the state variables and observation prior as derived from the forecast ensemble and used in EnKF are clearly very beneficial.

**KEYWORDS**

ensemble Kalman filter, polarimetric radar assimilation, two-moment microphysics

## 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 (Hu *et al.*, 2006; Stensrud *et al.*, 2009). Many studies have shown that radar data assimilation (DA) greatly reduces the spin-up time of a model and improves short-range precipitation forecasts (Xue *et al.*, 2003; Hu *et al.*, 2006; Hu and Xue, 2007; Xiao *et al.*, 2008; Dixon *et al.*, 2009; Zhu *et al.*, 2015).

To obtain additional information on precipitation MP, the entire US operational WSR-88D Doppler radar network was 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 their 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 (Cao *et al.*, 2013; Huang *et al.*, 2019; Zhang *et al.*, 2019), as can hydrometeor classification within storms (Ryzhkov and Zrnica, 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, that is, 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, that is, 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 analysed 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 analysed rainwater in the lower troposphere and snow in the mid- to upper troposphere for a mesoscale convective system. However, the impact of PRD assimilation was only assessed with respect to the analysis increments of rainwater and snow, and was limited to a portion of the analysed storm due to limited data coverage.<sup>1</sup>

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

More direct and quantitative use of PRD is desirable through direct assimilation. Direct DA methods compare simulated observations from the model state variables against observations, and make adjustments to the state variables to achieve optimal fit of the analysed state to observations and the prior guess of the state subjected to the weights related to their respective errors (Kalnay, 2002). Forward observation operators are needed to simulate PRD from the model state variables, and the forecast model should have a reasonable capability in simulating observed polarimetric signatures. Jung *et al.* (2008a) developed PRD observation operators based on calculations of electromagnetic wave scattering by hydrometeors then used power-law functions to fit backscattering amplitudes to obtain computationally more efficient operators. The contributions of wet snow and wet graupel/hail are also included. In Jung *et al.* (2010b), more accurate observation operators based on rigorous scattering calculations using

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

the T-matrix method (Vivekanandan *et al.*, 1991) are developed. Details of observation operators will be given in Section 2.1. In Jung *et al.* (2010b), they also compared the performance of single-moment (SM) and double-moment (DM). Their results showed that certain polarimetric signatures such as  $Z_{DR}$  arc,  $\rho_{HV}$  (cross-correlation coefficient) rings can only be correctly simulated by a DM scheme. Simulated PRD can have large uncertainties and can vary significantly with the use of an MP scheme, however (Putnam *et al.*, 2017a; 2017b).

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

Certain polarimetric signatures that depend on hydrometeor size sorting (Dawson *et al.*, 2014), such as the  $Z_{DR}$  arc in the supercell storms, can only be properly simulated using multi-moment MP schemes (Jung *et al.*, 2010b; Putnam *et al.*, 2014). In the only published study that directly assimilates real polarimetric observations using EnKF, Putnam *et al.* (2019) showed that the analysed  $Z_{DR}$  structures including the  $Z_{DR}$  arc in a supercell storm are improved with additional  $Z_{DR}$  assimilation. The study also showed that the analysed rain mean mass diameter is higher in the  $Z_{DR}$  arc region and the total rain number concentration is lower downshear in the forward flank, agreeing with observational estimations. Biases do exist in their EnKF analyses that require further investigations (Putnam *et al.*, 2019), however.

As far as we know, Putnam *et al.* (2019) is the only formally published study that examines the impact of directly assimilating additional PRD using EnKF combined with a multi-moment MP scheme. Many issues, including analysis biases, remain that require further studies as they pointed out. Being a real-data-based study, detailed verification of analysed state variables, especially those of MP, is difficult, because of the lack of truth. Errors in

the observational data can complicate the issues. To better understand the behaviours and impacts of assimilating additional PRD, observation system simulation experiments (OSSEs) can be very helpful. While Jung *et al.* (2008b; 2010b) examined the impacts of PRD data via OSSEs, their EnKF DA studies had limitations with the use of an SM MP scheme. For the above reasons, OSSEs are performed in this study with EnKF combined with a DM MP scheme and compatible observation operators, examining the impact of directly assimilating additional  $Z_{DR}$  data. Additional sensitivity experiments are performed to see the impacts of updating total number concentrations (the additional PSD moment associated with DM MP schemes) and updating state variables not directly linked to PRD via observation operators.

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

## 2 | EXPERIMENT CONFIGURATION AND SETTINGS

### 2.1 | The truth simulation and observation operators

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

For DM schemes, the shape parameter of three-parameter gamma distributions assumed of most hydrometeor PSDs is typically assumed constant (with zero being assumed most often). In this study, the shape parameter for rainwater in the Milbrandt–Yau DM scheme

is set to two while for other hydrometeors it is set to zero. Studies have found that most DM schemes tend to overestimate simulated reflectivity (e.g. Brown *et al.*, 2016; Putnam *et al.*, 2017b). One of the reasons, according to Brown *et al.* (2016), is that most schemes tend to produce a higher frequency of large raindrops than observed. Setting the rain shape parameter to two helps reduce the number of large raindrops and in turn reflectivity.

The observation operator for radial velocity is the same as that used in Jung *et al.* (2008a). However, there are some differences from OSSE experiments of Jung *et al.* (2008a) where the observation operators for radar reflectivity are calculated using a fitted approximation to T-matrix scattering amplitudes for rain and Rayleigh approximation for ice hydrometeors. This approximation may result in some error (Putnam *et al.*, 2019). In this study, more advanced observation operators using look-up tables calculated from the T-matrix method are used (Jung *et al.*, 2010b). In the following, the formula for radar reflectivity factors at the horizontal and vertical polarizations are based on the full T-matrix algorithm:

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

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

where

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

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

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

Here,  $\lambda$  is the wavelength of the radar and we assume a 10.7 cm wavelength S-band radar.  $K_w = 0.93$  is the dielectric factor for water.  $\bar{\phi}$  is the mean canting angle and  $\sigma$  is the standard deviation of the canting angle.  $\bar{\phi} = 0$  is assumed for all species.  $\sigma$  are  $0^\circ$ ,  $20^\circ$ ,  $60^\circ$  and  $60^\circ$  for rain, snow, graupel and hail, respectively.  $|\dots|$  represents the modulus of a complex number while  $\operatorname{Re}[\dots]$  represents the real part.

Superscript \* implies the conjugate. Subscript  $x$  can be rain ( $r$ ), rain–snow mixture ( $rs$ ), dry snow ( $ds$ ), rain–graupel mixture ( $rg$ ), dry graupel ( $dg$ ), rain–hail mixture ( $rh$ ) or dry hail ( $dh$ ).  $D$  is the diameter of a given hydrometeor and  $D_{\max}$  is the maximum diameter of each hydrometeor category. In this article, the

maximum diameters of rain drops, snow aggregates, graupels and hailstones are assumed to be 8, 30, 50 and 70 mm, respectively.  $n(D)$  is the number concentration of the hydrometeor at diameter  $D$ . To numerically integrate Equations 1 and 2, the integral ranges are partitioned into 100 bins. The backscattering amplitudes of each species with assumed drop size for polarizations along the major ( $f_a[\pi]$ ) and minor ( $f_b[\pi]$ ) axes are precomputed at the centre of each size bin and stored in look-up tables. For melting species including rain–snow, rain–graupel and rain–hail mixtures, the same tables are constructed at the uniform 5% water fraction interval. The fraction of water of each ice species 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 ice hydrometeors. More details on the PRD observation operators can be found in Jung *et al.* (2010b).

Once the radar reflectivity factors of all hydrometeor categories are calculated, the reflectivity in dBZ at horizontal and vertical are computed as follows:

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

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

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

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

## 2.2 | EnKF experiment settings and DA experiments

In this study, we use the ARPS EnKF package (Tong and Xue, 2005; Xue *et al.*, 2006) which uses the ensemble square-root filter algorithm (Whitaker and Hamill, 2002). The EnKF experiments employ 40 members in this study. With the DM Milbrandt–Yau MP scheme, the analysis variables include the three-dimensional wind components ( $u$ ,  $v$  and  $w$ ), pressure ( $p$ ), potential temperature ( $\theta$ ), water vapour mixing ratio ( $q_v$ ), as well as microphysical state variables including mixing ratios of cloud water ( $q_c$ ), rainwater ( $q_r$ ), cloud ice ( $q_i$ ), snow aggregate ( $q_s$ ), graupel ( $q_g$ ) and hail ( $q_h$ ), and their total number concentrations ( $Nq_c$ ,  $Nq_r$ ,  $Nq_i$ ,  $Nq_s$ ,  $Nq_g$  and  $Nq_h$ , respectively). Spin-up ensemble forecasts are run for 20 min, starting from initial ensemble states defined by the sounding profiles plus smoothed Gaussian random perturbations added in regions where observed reflectivity is larger than 10 dBZ. The mean standard deviations of added  $u$ ,  $v$ , and  $w$  perturbations are  $2 \text{ m}\cdot\text{s}^{-1}$  and that of  $\theta$  is 2 K. For water

vapour and hydrometeor mixing ratios, the mean standard deviations of added perturbations are  $0.0006 \text{ kg}\cdot\text{kg}^{-1}$ . Considering the large uncertainty of number concentrations, we did not add perturbation to those variables. After the 20 min spin-up forecasts, EnKF DA cycles are run over a 90 min period assimilating radar data every 5 min, corresponding to the model storm time period of 20 through to 110 min. Similar settings were in our earlier OSSE studies (Tong and Xue, 2005; Jung *et al.*, 2008a). The 90 min assimilation period is chosen mostly based on the life cycle of the storm. In the truth simulation, the main storm reaches its mature stage between 60 and 100 min. After that, the storm begins to weaken and move out of the simulation domain.

PRD from an assumed S-band radar with its centre located in the southwest corner ( $x = 2, y = 2$ ) are simulated from the truth simulation output, using the observation operators described in Section 2.1. Eleven elevations are assumed, based on the WSR-88D radar VCP-11 scan mode. Radar observation errors are assumed to be  $1 \text{ m}\cdot\text{s}^{-1}$ , 3 dBZ, 0.2 dB for radial velocity  $V_r$ , and  $Z_H$  and  $Z_{DR}$  in terms of standard deviation, respectively, and random errors of the corresponding magnitudes are added to the simulated PRD observations and assumed in the EnKF experiments.  $V_r$  observations are assimilated where observed  $Z_H > 10$  dBZ. For  $Z_{DR}$ , only values larger than 0.3 dB are used because smaller values tend to be very noisy. The covariance localization radii for radar observations are set to 4 km in the horizontal and 2 km in the vertical direction using the correlation function of Gaspari and Cohn (Gaspari and Cohn, 1999) for all state variables. The 4 km horizontal grid spacing spans four grid intervals in this study, which is consistent with most past studies in terms of grid intervals. For example, in Tong and Xue (2005) and Jung *et al.* (2008a), 6 to 8 km were suggested when a 2 km horizontal grid spacing was used. Sobash and Stensrud (2013) suggest 12 to 18 km horizontal radii when using a 3 km grid spacing. We have tested larger and smaller horizontal localization radii. The state analysis errors were found to be significantly larger when using a 6 km radius while the results using 3 km were slightly worse. To help maintain ensemble spread, multiplicative inflation (Anderson, 2001; Tong and Xue, 2005) is applied to all model state variables except for number concentrations, using an inflation coefficient of 1.2.

Table 1 lists all experiments presented in this article. Experiment VrZh assimilates  $V_r$  and  $Z_H$  data while experiment VrZhZdr assimilates additional  $Z_{DR}$  data. Both experiments update a full set of state variables in the model. Experiment VrZhZdr is considered a control experiment while VrZh is a reference for comparison purpose. Additional sensitivity experiments are conducted to help better understand how the assimilation of  $Z_{DR}$  improves the analysis. The first sensitivity experiment VrZhZdr\_NoNt,

excludes the updating of total number concentrations of the hydrometeors  $Nq_x$  by  $Z_{DR}$  observations compared to experiment VrZhZdr.  $Nq_x$  are still updated by  $V_r$  and  $Z_H$  though, just not by  $Z_{DR}$ .  $Nq_x$  arise from the use of a DM scheme and add complexity to the DA problem. The number concentrations of hydrometeors have very wide dynamic ranges, varying from 0 to as large as  $10^{12} \text{ m}^{-3}$ , implying that the relations between them and PRD observations can be very nonlinear. Updating both mixing ratios and total number concentrations at the same time may or may not be beneficial, especially when the correlations are unreliable or inconsistent with each other. VrZhZdr\_NoNt serves to examine the benefit, if any, of updating the total number concentrations using  $Z_{DR}$  observations.

Other sensitivity experiments serve to examine the impact of updating other state variables using  $Z_{DR}$ . Experiments VrZhZdr\_NoW, VrZhZdr\_NoUV, VrZhZdr\_NoQv and VrZhZdr\_NoPt exclude the updating of vertical velocity  $w$ , horizontal wind components  $u$  and  $v$ , water vapour mixing ratio  $q_v$ , and potential temperature  $\theta$ , respectively. In an intense tornadic supercell, a  $Z_{DR}$  column typically exists in the updraught region (Kumjian and Ryzhkov, 2008), indicating strong positive correlation between upward motion and  $Z_{DR}$ . Updraught regions are also associated with high moisture values. The largest theoretical benefit of EnKF method compared to 3D-Var and some of the other methods lies with the use of ensemble-derived correlations among all state variables, and hence among observation priors and state variables, which allows for the updating of state variables not directly observed (or involved in the observation operators). For such updating to be beneficial, the ensemble-derived correlations have to be sufficiently accurate and reliable. This second group of sensitivity experiments are designed to test the impacts of updating state variables that are not directly linked to  $Z_{DR}$  observations via the observation operators.

## 3 | RESULTS OF ENKF ANALYSES

### 3.1 | Evaluation of $Z_{DR}$ assimilation in the control experiment

Figure 1 shows the ensemble mean analysis and forecast RMSEs of model state variables during the assimilation cycles. Following Tong and Xue (2005) and many other studies, the RMSEs are calculated over grid points where the true reflectivity is greater than 10 dBZ, which roughly covers the precipitation regions. For most variables, VrZhZdr (red lines), which assimilates additional differential reflectivity, produces consistently better analyses and forecasts than VrZh (black lines), especially in

**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 $Z_{DR}$ DA does not update $w$
VrZhZdr_NoPt	$V_r, Z_H, Z_{DR}$	As VrZhZdr but $Z_{DR}$ DA does not update $\theta'$
VrZhZdr_NoQv	$V_r, Z_H, Z_{DR}$	As VrZhZdr but $Z_{DR}$ does not update $q_v$

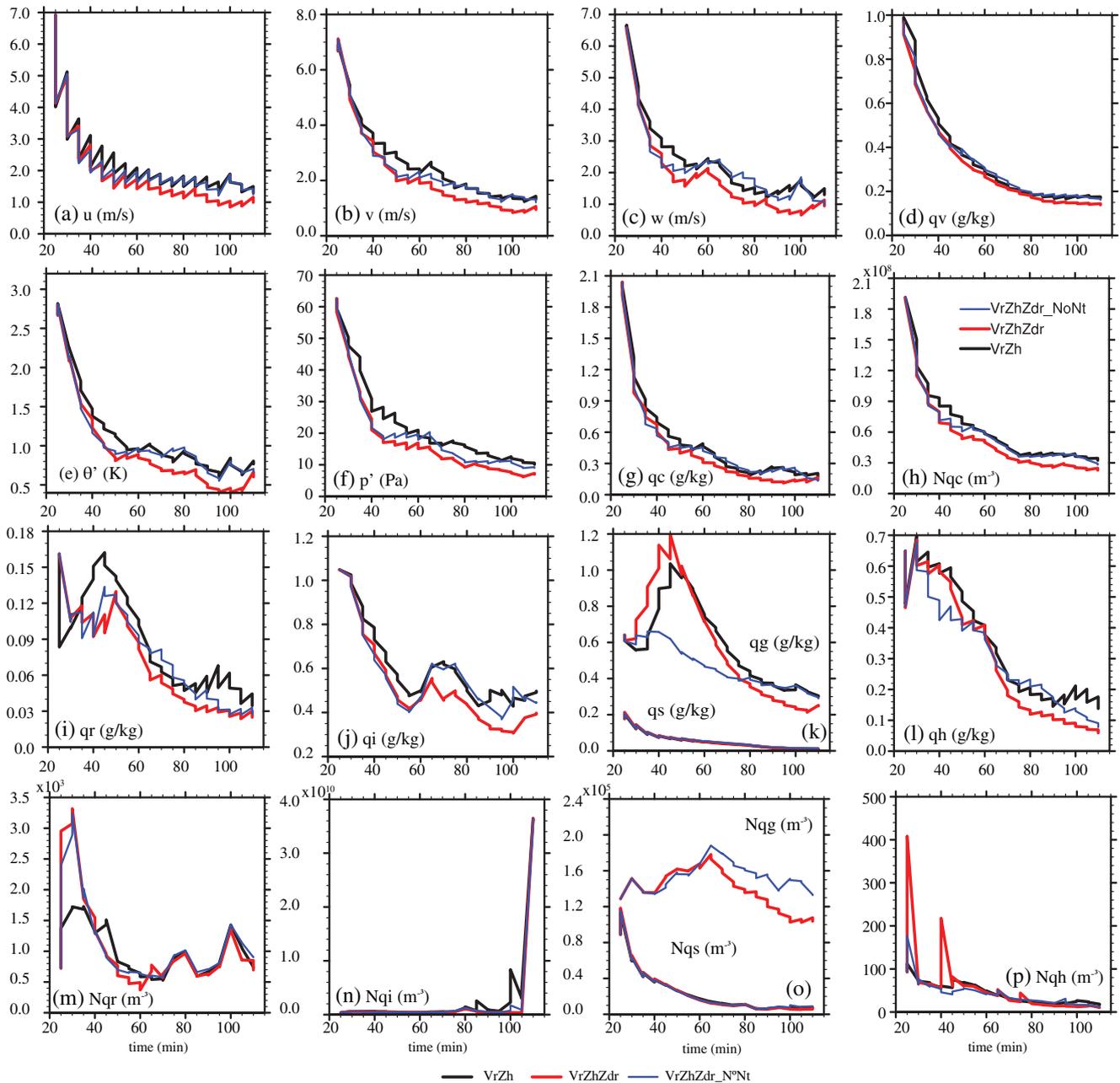
later cycles. Such results are quite similar to those of Jung *et al.* (2008b) which examined the impact of assimilating additional  $Z_{DR}$  data in OSSEs employing an SM MP scheme, except that RMSEs of most variables in the first few cycles are also reduced here. In Jung *et al.* (2008b), the assimilation of  $Z_{DR}$  does not show positive impact until later cycles. Additionally, we also examine the RMSEs of total number concentrations of hydrometeor variables, which were not predicted in Jung *et al.* (2008b). Here, for most number concentrations, the  $Z_{DR}$  assimilation shows neutral to positive impact. Among them, the number concentration for graupel,  $Nq_g$ , is improved most. As we will discuss later, it is probably benefiting from better analyses of liquid hydrometeor species, which in turn lead to more accurate analyses of ice hydrometeor species. In Figure 2, we show the RMSEs of the analyses and forecasts throughout the DA cycles in terms of radar observed variables, that is, the verifications in observation space. The results are consistent with the results in terms of the state variables, as shown in Figure 1; the assimilation of  $Z_{DR}$  data further reduces the differences between the analyses and forecasts from the observations in the observation space.

Figure 3 shows the vertical profiles of ensemble mean analysis and forecast RMSEs at 80 min, again averaged over grid points with observed  $Z_H$  exceeding 10 dBZ. At this time, the RMSEs of most variables have stabilized (Figure 1). It can be seen that the errors at most levels for most variables are reduced from the additional  $Z_{DR}$  assimilation. The largest improvements are mostly located where the errors are largest. As Jung *et al.* (2008b) pointed out, the direct improvements from  $Z_{DR}$  assimilation are mainly to those highly correlated variables such as  $q_v$  and  $q_r$  at the lower levels, where the  $Z_{DR}$  signatures are most prominent (given that large  $Z_{DR}$  is mostly associated with large raindrops). With more accurate analyses at the lower levels, the analysis fields at upper levels can also be improved through the dynamic interactions in the forecast

model. The weak and unreliable correlations between  $Z_{DR}$  and ice fields at the upper levels during the earlier cycles might be the reason for larger errors in  $q_h$  before 45 min (Figure 1k) while the errors become smaller in later cycles. Note that in Figures 1–3, the results of VrZhZdr\_NoNt are also included which will be discussed in Section 4.2 later.

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

The  $Z_{DR}$  structure near the hook echo region is similar to the classic supercell storm structure for both truth simulation and EnKF analyses (Figure 4d,e,f). Here, we only display the small hook area at 80 min because it shows a clear  $Z_{DR}$  column (Figure 6d). At 110 min, the  $Z_{DR}$  columns are not obvious (not shown). High  $Z_{DR}$  values are located at the leading edge of the high  $Z_H$  hook (black contours) (Figure 4d). Experiment VrZh shows generally similar patterns but the intensity is clearly underestimated for both  $Z_H$  and  $Z_{DR}$  (Figure 4b,e). With additional

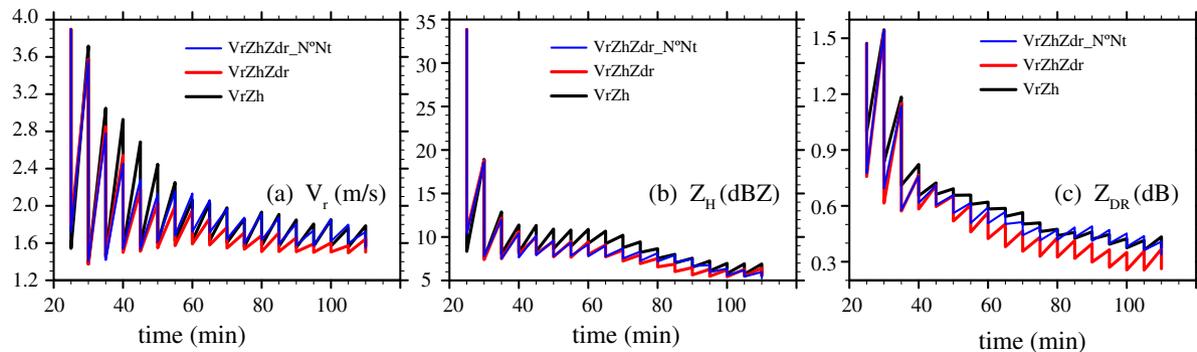


**FIGURE 1** The RMSEs of the ensemble-mean forecasts and analyses throughout the 5 min DA cycles, for experiments VrZh (black), VrZhZdr (red) and VrZhZdr\_NoNt (blue), averaged over points at which the observed reflectivity is greater than 10 dBZ. (a)  $u$ , (b)  $v$ , (c)  $w$ , (d) water vapour content  $q_v$ , (e) perturbation potential temperature  $\theta'$ , (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

$Z_{DR}$  assimilation, the shape of the  $Z_{DR}$  arc in experiment VrZhZdr looks closer to the truth than in experiment VrZh (Figure 4c). The intensity of  $Z_{DR}$  in the hook echo region is also much enhanced in VrZhZdr (Figure 4f). Moreover, the  $Z_H$  pattern has also been improved. The 35 dBZ  $Z_H$  contours in the southeast edge are much closer to the truth (Figure 4a–c), and the  $Z_H$  intensity in the hook region is greatly enhanced (Figure 4d–f). In all,  $Z_{DR}$  assimilation improves the polarimetric signatures of the

simulated storm, especially in the hook echo and forward flank regions.

The vertical cross-sections of analysed  $Z_H$ ,  $Z_{DR}$  and specific differential phase,  $K_{DP}$ , in the hook echo region through the low-level  $Z_H$  and  $Z_{DR}$  maximum centres at 80 min are shown in Figure 6. Here  $K_{DP}$  is not directly assimilated, but derived from analysed model state variables using the same equation as in Jung *et al.* (2010b). In general, both experiments VrZh and VrZhZdr produce



**FIGURE 2** Similar to Figure 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 [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

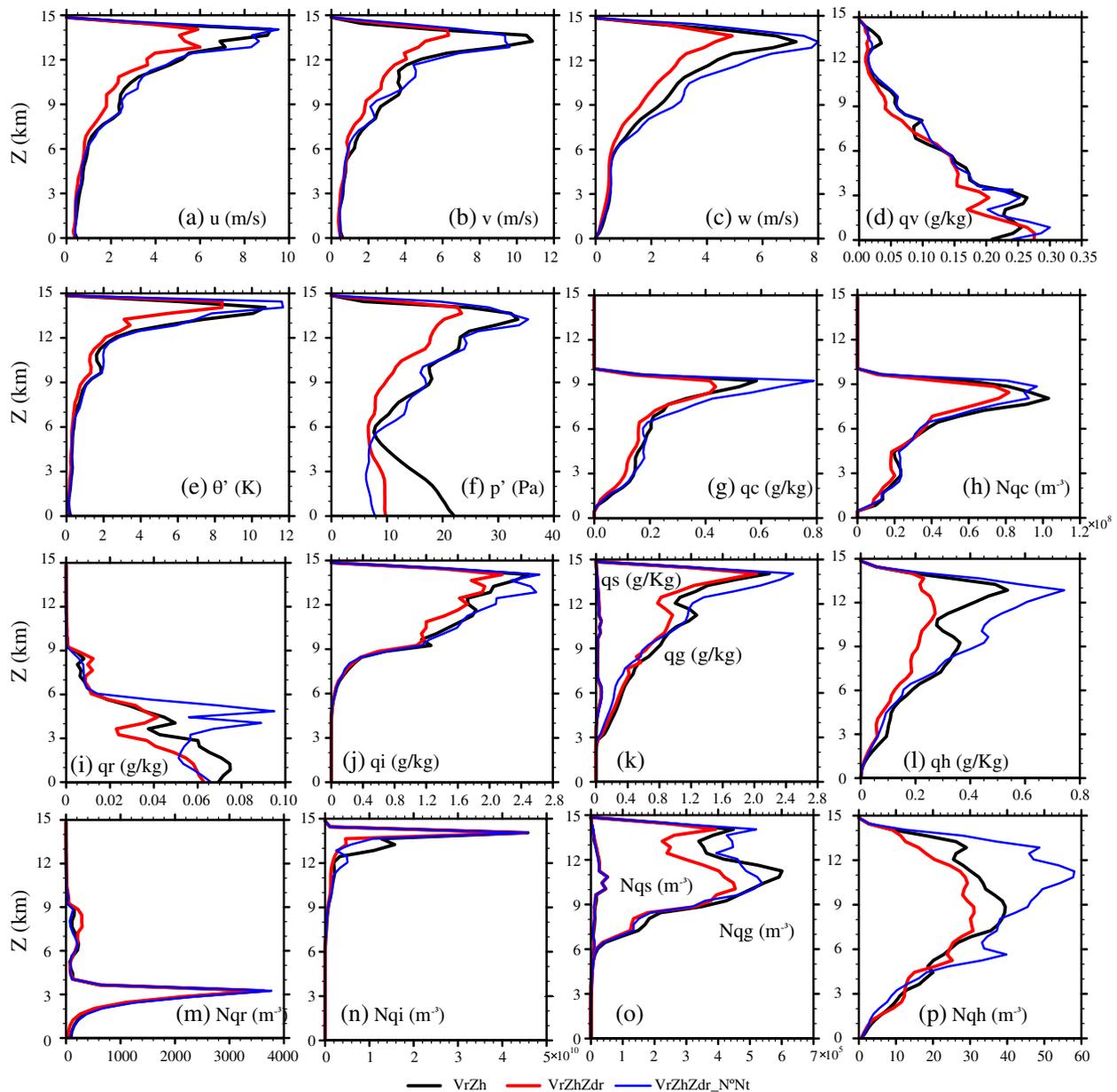
similar patterns of these fields. However, the intensities of  $Z_{DR}$  and  $K_{DP}$  are clearly underestimated in VrZh. For  $Z_H$ , the maximum values above 60 dBZ are right below the 0 °C contours in all cases (Figure 6a–c). The 45 dBZ  $Z_H$  contours (orange) extend up to above the –20 °C in the truth (Figure 6a) and in VrZhZdr (Figure 4c), but only to –10 °C line in VrZh (Figure 6b), indicating the analysed storm is less intense in VrZh. The improved vertical structure of  $Z_H$  indicates better analysis of the hydrometeor fields, which we will show more in Figure 7. The assimilation of  $Z_{DR}$  data also results in a more intense core updraught that is closer to the truth as indicated by the 10 m·s<sup>–1</sup>  $w$  contours in Figure 6a–c. With a stronger updraught, particles are more likely transported to high altitudes and also likely undergo more growth before falling to the ground. Associated with the updraught is a column of high  $Z_{DR}$  values that extend to the –10 °C level in the truth (Figure 6d) and in VrZhZdr (Figure 6f), while that in VrZh is clearly weaker (Figure 6e). Also, a column of high  $K_{DP}$  is also better reproduced in VrZhZdr (Figure 6i) than in VrZh (Figure 6h) compared to the truth (Figure 6g). High  $K_{DP}$  is mostly associated with high liquid water content, which is linked to intense updraught and heavy precipitation.

Figure 7 shows the analysed cloud water, hail and rain water mixing ratios from VrZh and VrZhZdr in the same vertical cross-sections as Figure 6, as compared to the truth. Since only  $Z_{DR}$  observations larger than 0.3 dB are assimilated, the direct impact from  $Z_{DR}$  is mostly limited in the lower levels (cf. Figure 6d). However, its benefit could be spread to the higher levels through spatial and cross-variable correlations, and through dynamic interactions within the forecast model. Figure 7 shows that the cloud ice field is better analysed all the way to the cloud top at ~9 km height in VrZhZdr (Figure 7c) and while that in VrZh is mostly limited to below the –20 °C level or about 6.5 km height; its maximum value is also too low (Figure 7b). For hail, VrZhZdr also much better reproduces the vertical distribution and intensity (Figure 7f)

than VrZh (Figure 7e); the latter severely underestimates hail at the higher levels. For rainwater, the analysis of VrZhZdr is also better, although the differences are smaller (Figure 7i,h). Overall, the assimilation of additional  $Z_{DR}$  produces analyses of the supercell storm whose intensity and structure are much closer to the truth, in terms of both observed parameters ( $Z_H$  and  $Z_{DR}$ ) and model state variables.

### 3.2 | The updating of hydrometeor number concentrations with $Z_{DR}$

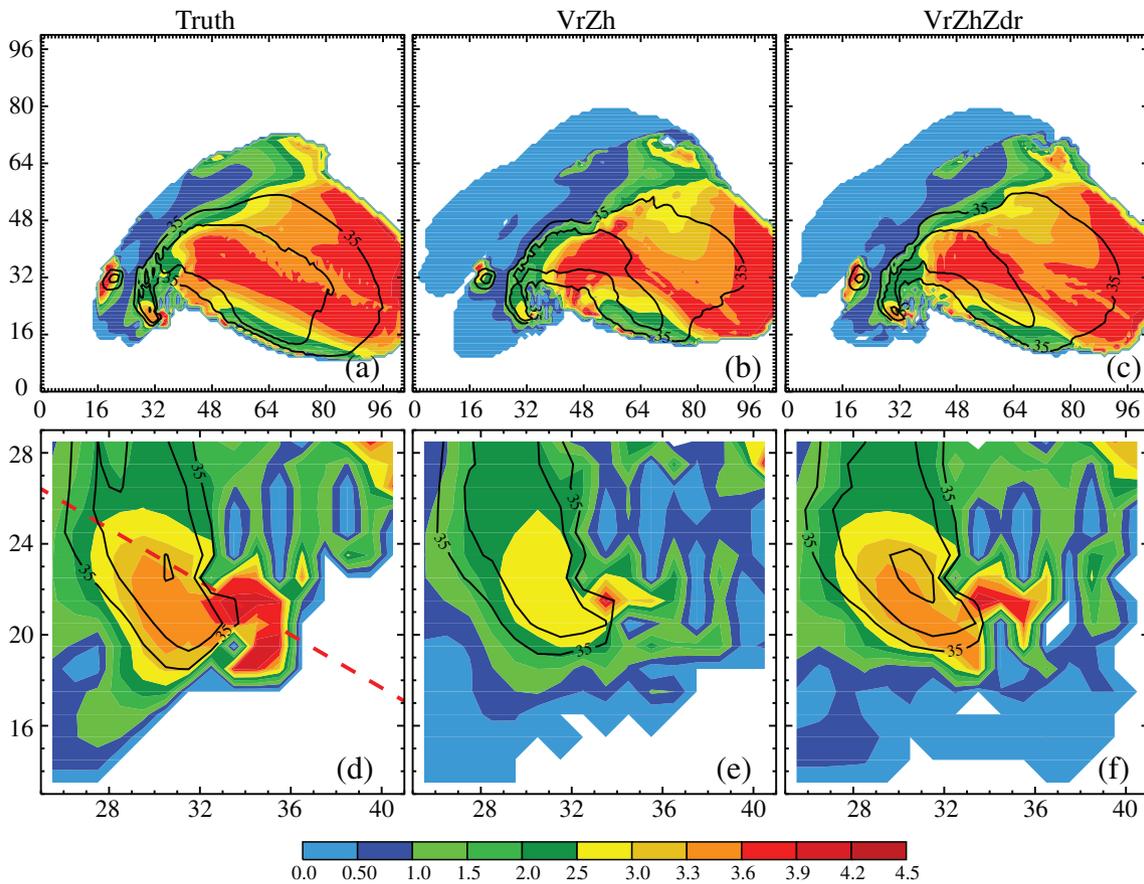
For the DM MP scheme, the hydrometeor number concentrations are part of the forecast variables which increase the degrees of freedom of the model state. As pointed out earlier, the values of number concentrations show a great range of variability. Additionally, for DM schemes,  $Z_{DR}$  depends mostly on the slope parameter of PSD which is a strong function of the third moment, the mass mixing ratio (Jung *et al.*, 2008b). It is not certain whether the updating of number concentrations by EnKF will improve the overall analysis. The RMSEs for most state variables and also for radar-observed variables of the experiment VrZhZdr\_NoNt that excludes the updating of number concentrations are shown in Figures 1–3. It can be seen that without updating  $Nq_x$ , the RMSE curves of VrZhZdr\_NoNt (blue lines) are closer to those of VrZh than VrZhZdr during the later DA cycles for most variables (Figure 1). Similar is true in terms of radar-observed variables  $V_r$ ,  $Z_H$  and  $Z_{DR}$  (Figure 2). For  $w$ ,  $q_r$  and  $q_h$ , the RMSEs of VrZhZdr\_NoNt even exceed those of VrZh in some of the cycles (Figure 1c,i,l). The deterioration of the analyses in VrZhZdr\_NoNt are clearer in the vertical profiles of RMSEs at 80 min (Figure 3). For  $w$  and most ice state variables, the RMSEs in VrZhZdr\_NoNt are larger than those of VrZh at the upper levels (Figure 3) while for  $q_r$  this happens at the mid-levels (Figure 3i). These results suggest that updating both mixing ratios and total number



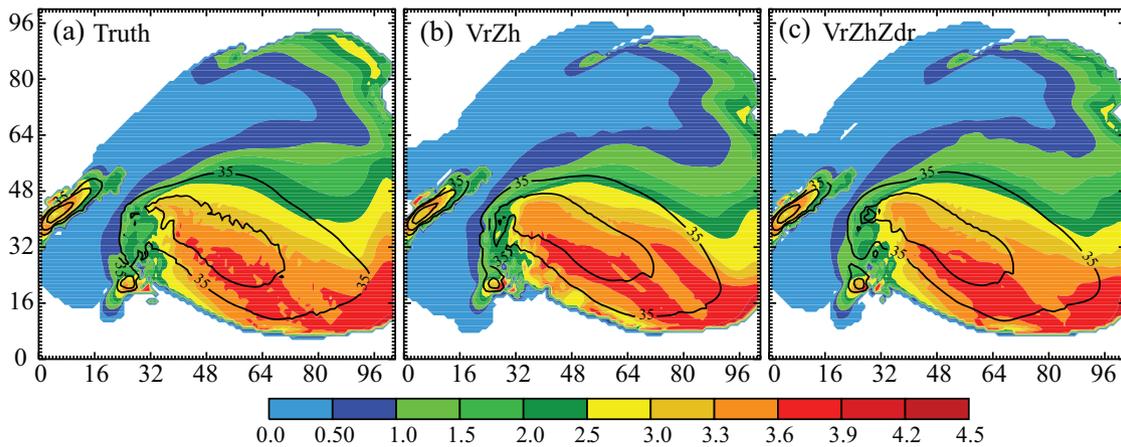
**FIGURE 3** The profiles of horizontally averaged ensemble mean analysis RMSEs averaged over points at which the truth reflectivity is greater than 10 dBZ at 80 min for variables (a)  $u$ , (b)  $v$ , (c)  $w$ , (d) water vapour content  $q_v$ , (e) perturbation potential temperature  $\theta'$  (f) 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 [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

concentrations of hydrometeor species associated with a DM MP scheme together when assimilating  $Z_{DR}$  is important; when only mixing ratios are updated, most of the benefit of assimilating  $Z_{DR}$  data is lost, and for some variables, those analyses may be even worse than not assimilating  $Z_{DR}$  data at all. This is presumably because serious imbalance or inconsistency is created between mixing ratios and corresponding number concentrations when only the former are updated.

Figure 8 shows the analysed rainwater number concentrations  $Nq_r$  at 3 km height, and hail number concentrations  $Nq_h$  at  $z = 6$  km at 80 min from VrZhZdr and VrZhZdr\_NoNt, as compared to the truth. For the truth, highest  $Nq_r$  values are found in the southwest part of the supercell storm near the hook echo region and in the northwest part, corresponding to heavy rainfall in the rear flank and forward flank downdraught regions, respectively (Figure 8a). The patterns of analysed  $Nq_r$



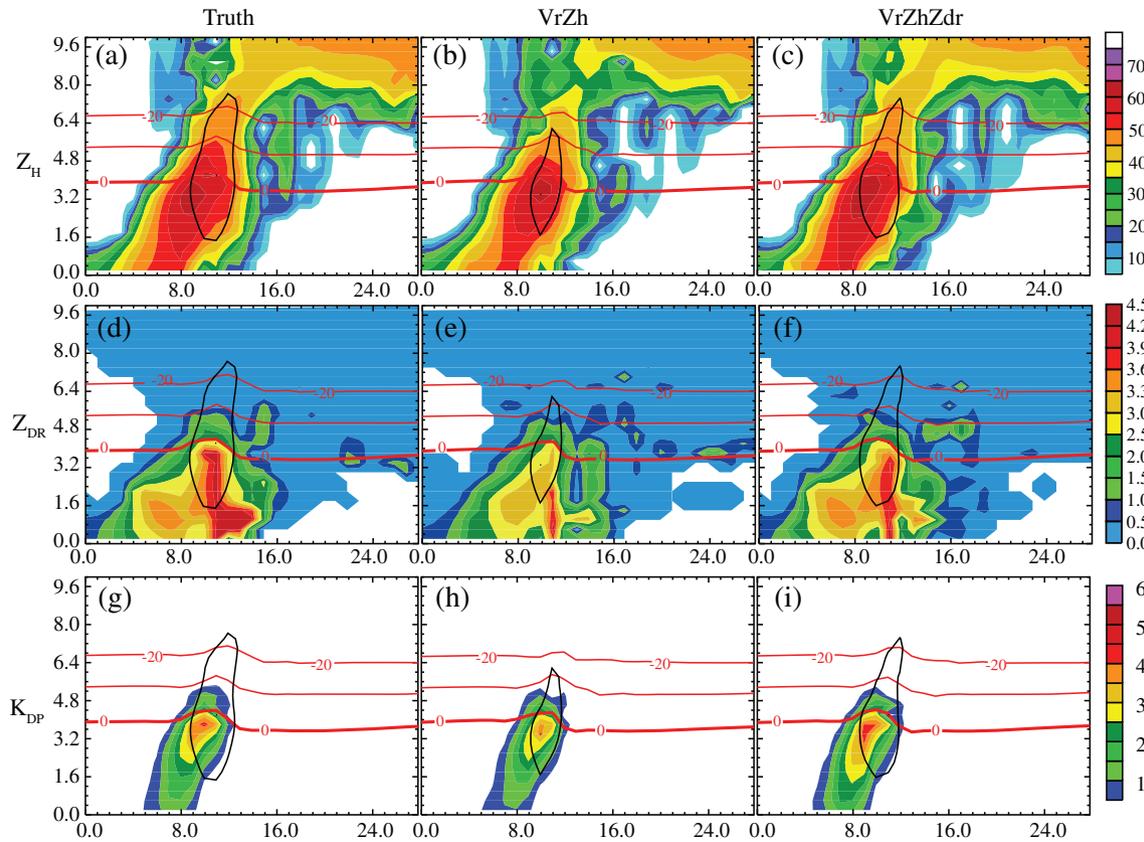
**FIGURE 4** Truth and analysed differential reflectivity  $Z_{DR}$  (colour shading) and reflectivity  $Z_H$  (black contours, starting from 35 dBZ at a 10 dBZ interval) at 500 m altitude at 80 min for (a) truth, (b) VrZh and (c) VrZhZdr. (d)–(f) are the zoomed-in plots for the hook echo area of the (a)–(c) plots, respectively



**FIGURE 5** Similar to Figure 4(a–c) but at 110 min

are similar (Figure 8b,c) although there is a larger area of overestimation in the forward flank region while the high values in the rear flank region are underestimated in VrZhZdr\_NoNt (Figure 8c). Both overestimation and underestimation are much less in VrZhZdr (Figure 8b). The hail number concentration  $N_{qh}$  for the truth exhibits

moderately high values in the southeastward-spreading forward flank and storm anvil regions at 6 km height (Figure 8d) while in the hook echo region, a ring of high  $N_{qh}$  is found around an  $N_{qh}$  hole, while the highest values are found on the west and southwest sides of the hole (Figure 8d). The hole should be associated



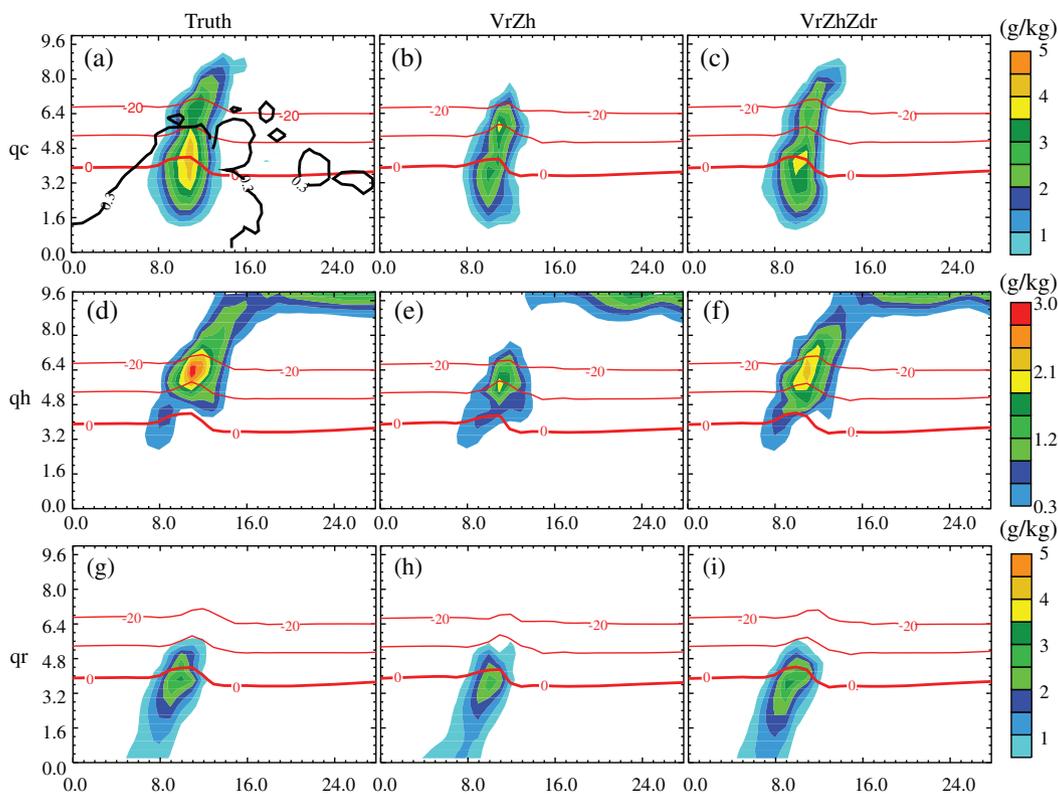
**FIGURE 6** (a–i) Vertical slices along the red dashed line in Figure 4, of reflectivity (shaded, upper row, unit: dBZ), differential reflectivity  $Z_{DR}$  (middle row, unit: dB) and specific differential phase  $K_{DP}$  (bottom row, unit:  $^{\circ}\cdot\text{km}^{-1}$ ), of truth (left column), experiment VrZh (central column) and VrZhZdr (right column). The black contours are for 10 m s<sup>-1</sup> vertical velocity  $w$  and the red contours are for temperature from 0 (bold) to  $-20$  at 10 °C intervals

with a bounded weak echo region typically found in intense supercell storms where hydrometeors are mostly absent, being swept away by the intense updraught. Within VrZhZdr\_NoNt, the “ring” structure is overestimated (Figure 8f) although the pattern of  $Nq_h$  in the forward flank region is a little better in VrZhZdr\_NoNt (Figure 8f) than in VrZhZdr (Figure 8e). Overall,  $Nq_r$  and  $Nq_h$  are better analysed in experiment VrZhZdr than in VrZhZdr\_NoNt.

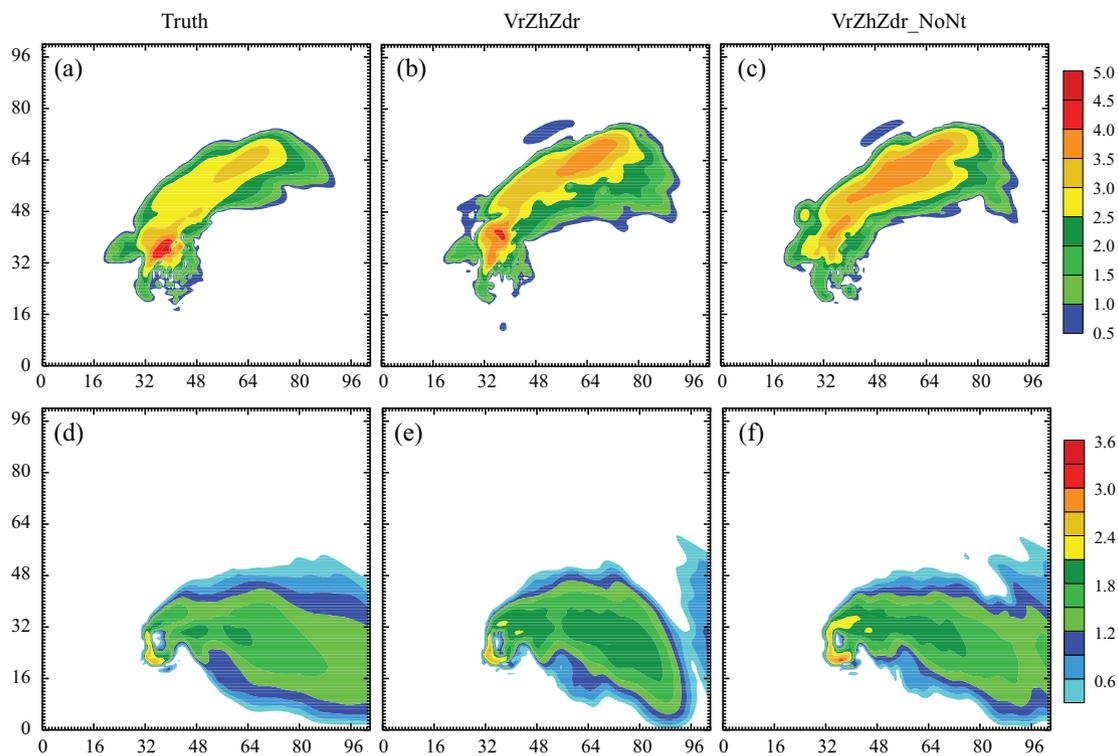
Figure 9 explain the possible reasons. Here, we calculate the correlation coefficients between the  $Z_{DR}$  observation prior and the hydrometeor state variables  $\rho(Z_{DR}, Nq_x)$  at 80 min from the forecast ensemble in a vertical cross-section passing through  $Z_{DR}$  prior which is located in the  $Z_{DR}$  column at  $(x, y, z) = (34, 22, 3.5)$  km. In general,  $Z_{DR}$  has clearly higher correlations to  $Nq_c$ ,  $Nq_r$  and  $Nq_h$  than to  $Nq_i$ ,  $Nq_s$  and  $Nq_g$ . This is because  $Z_{DR}$  is most sensitive to raindrop sizes and high  $Z_{DR}$  is found where there are a large number of large rain drops. Many large drops originate from the melting of falling hailstones (Dawson *et al.*, 2014). A column of high correlation is found for  $\rho(Z_{DR}, Nq_c)$ ,  $\rho(Z_{DR}, Nq_r)$  and  $\rho(Z_{DR}, Nq_h)$  near the main updraught. For  $Nq_i$ ,  $Nq_s$  and  $Nq_g$ , the correlations are weaker, and non-zero values are mostly found

above the freezing level (Figure 9). The coherent structures in the correlations between  $Z_{DR}$  and  $q_x$ , and between  $Z_{DR}$  and  $Nq_x$  suggest that the flow-dependent error covariances estimated and utilized within the EnKF should be physically reasonable, and hence the updating of  $Nq_x$  in addition to  $q_x$  can be beneficial.

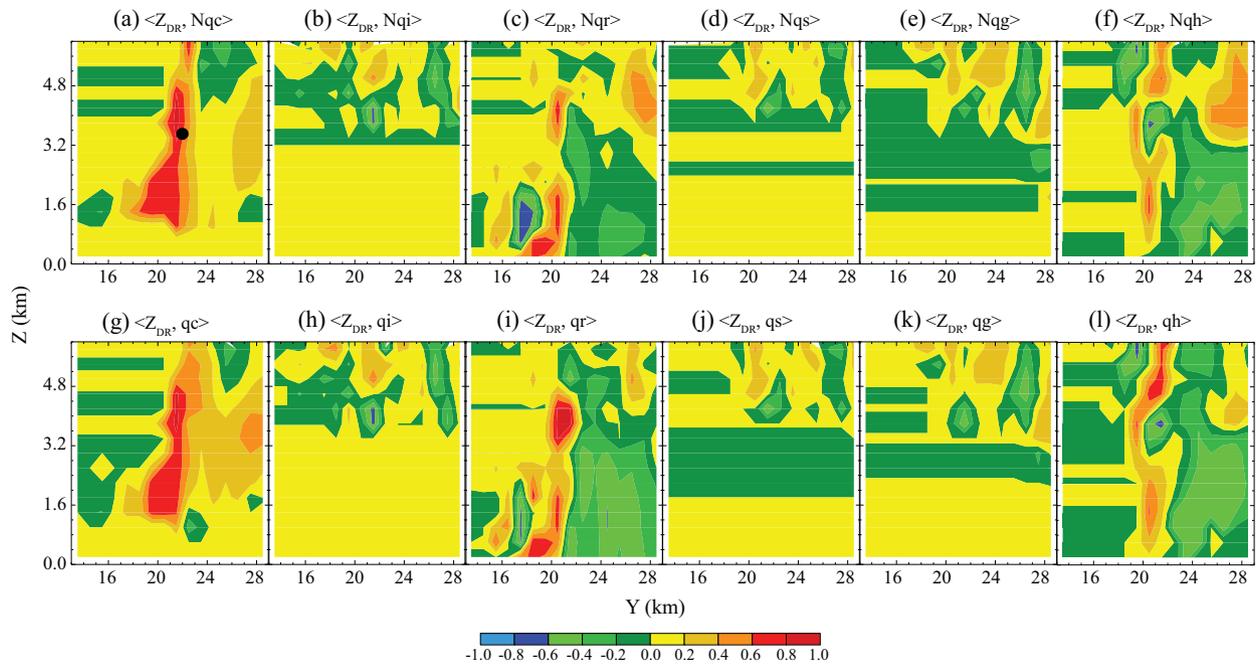
We also examine correlations between  $Z_{DR}$  at 1.8 km height and hydrometeor state variables at 80 min in a vertical cross-section in the forward flank high  $Z_{DR}$  region (Figure 10). The cloud water at this point is zero for all members. Therefore, the correlation is zero and is not shown. For other hydrometeor variables, similar to the point in the hook echo region, correlations  $\rho(Z_{DR}, Nq_r)$  and  $\rho(Z_{DR}, Nq_h)$  are clearly higher than  $\rho(Z_{DR}, Nq_s)$  and  $\rho(Z_{DR}, Nq_g)$ . The patterns of correlation  $\rho(Z_{DR}, q_x)$  are also very similar to the corresponding  $\rho(Z_{DR}, Nq_x)$  except that those of hail show opposite signs of correlation near the surface (Figure 10e,j). The negative correlation between  $Z_{DR}$  at 1.8 km and  $q_h$  at the lower levels is consistent with the fact that hailstones tend to contribute little to  $Z_{DR}$  due to tumbling (which is the cause of the  $Z_{DR}$  hole within supercell storms as a significant hail signature (Kumjian and Ryzhkov, 2008)), while the



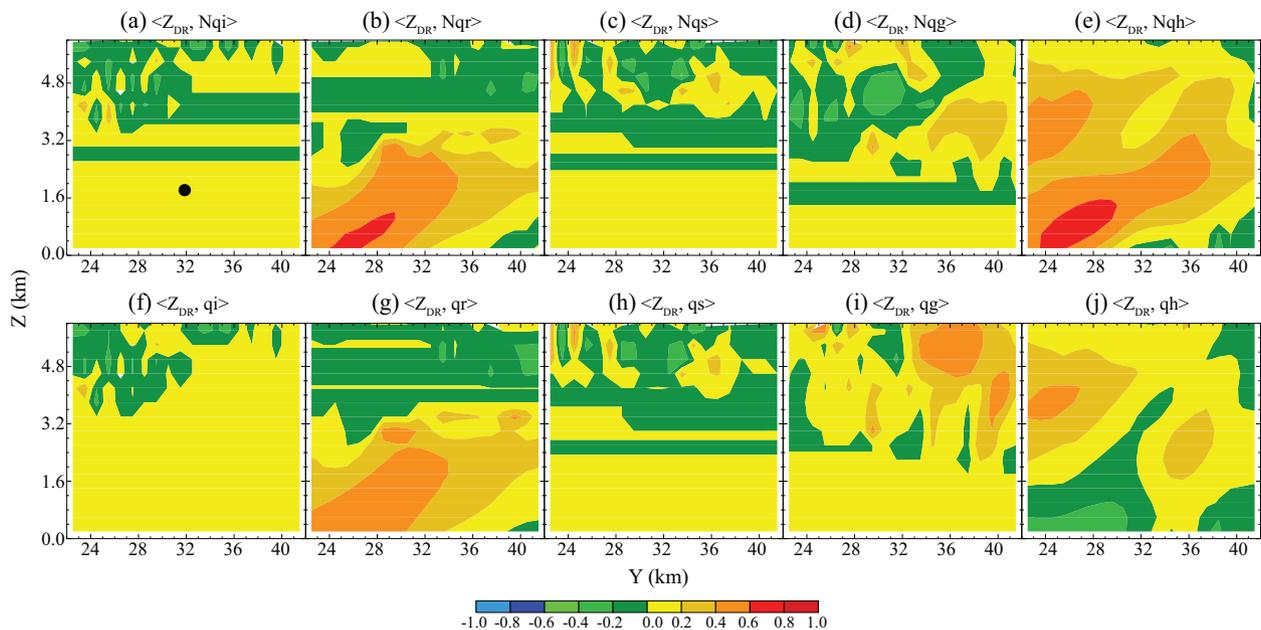
**FIGURE 7** Vertical cross-sections through the same slice at Figure 4, of the mixing ratios of cloud water  $q_c$  (shaded) for (a) truth, (b) VrZh and (c) VrZhZdr. (d–f) and (g–i) are similar to (a–c) but for hail and rain water. The black contour in (a) is for 0.3 dB  $Z_{DR}$  and the red contours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



**FIGURE 8** Logarithm of rainwater number concentrations ( $\log_{10}[Nqr]$ ) at 3 km height at 80 min 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



**FIGURE 9** (a–l) Correlation coefficients between  $Z_{DR}$  prior (at  $(x, y, z) = (34, 22, 3.5)$  km, black dot in (a)) and model hydrometeor state variables at 80 min. A  $y$ - $z$  vertical slice across  $x = 34$  km was plotted



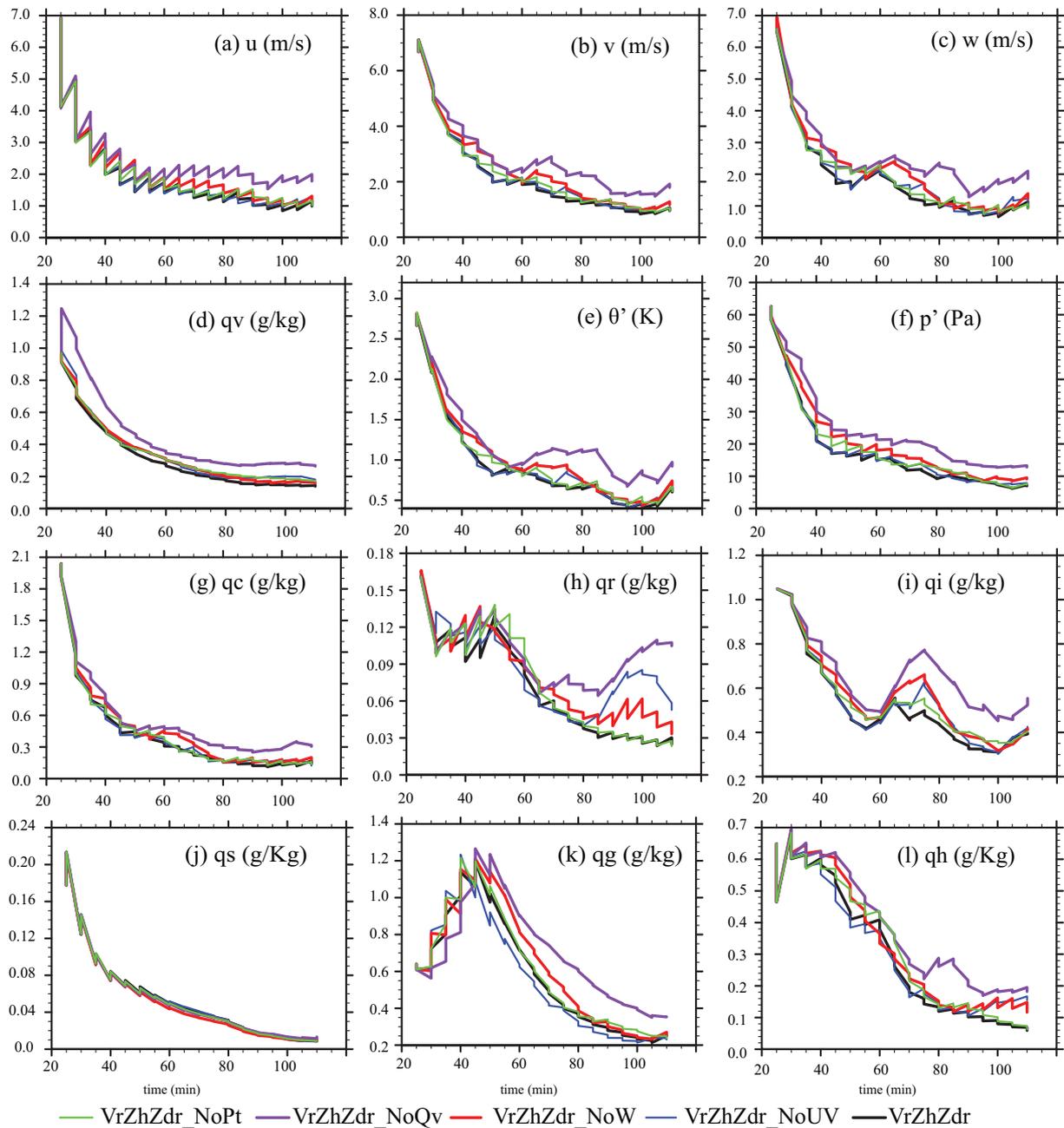
**FIGURE 10** (a–j) Correlation coefficients between  $Z_{DR}$  prior (at  $(x, y, z) = (64, 32, 1.8)$  km, black dot in (a)) and model hydrometeor state variables from control experiment at 80 min. A  $y$ - $z$  vertical slice across  $x = 64$  km was plotted

positive correlation between  $Z_{DR}$  and  $N_{qh}$  suggests that when a larger number of small hailstones exist, melting hailstones will cause less reduction to  $Z_{DR}$ . Given that large correlations between  $Z_{DR}$  and mixing ratio and between  $Z_{DR}$  and number concentration for rainwater and hail simultaneously, updating mixing ratios without updating corresponding number concentrations will create imbalances between different moments of the

hydrometeor PSDs which in turn will negatively affect the analysis and forecast states.

### 3.3 | The updating of other state variables with $Z_{DR}$ assimilation

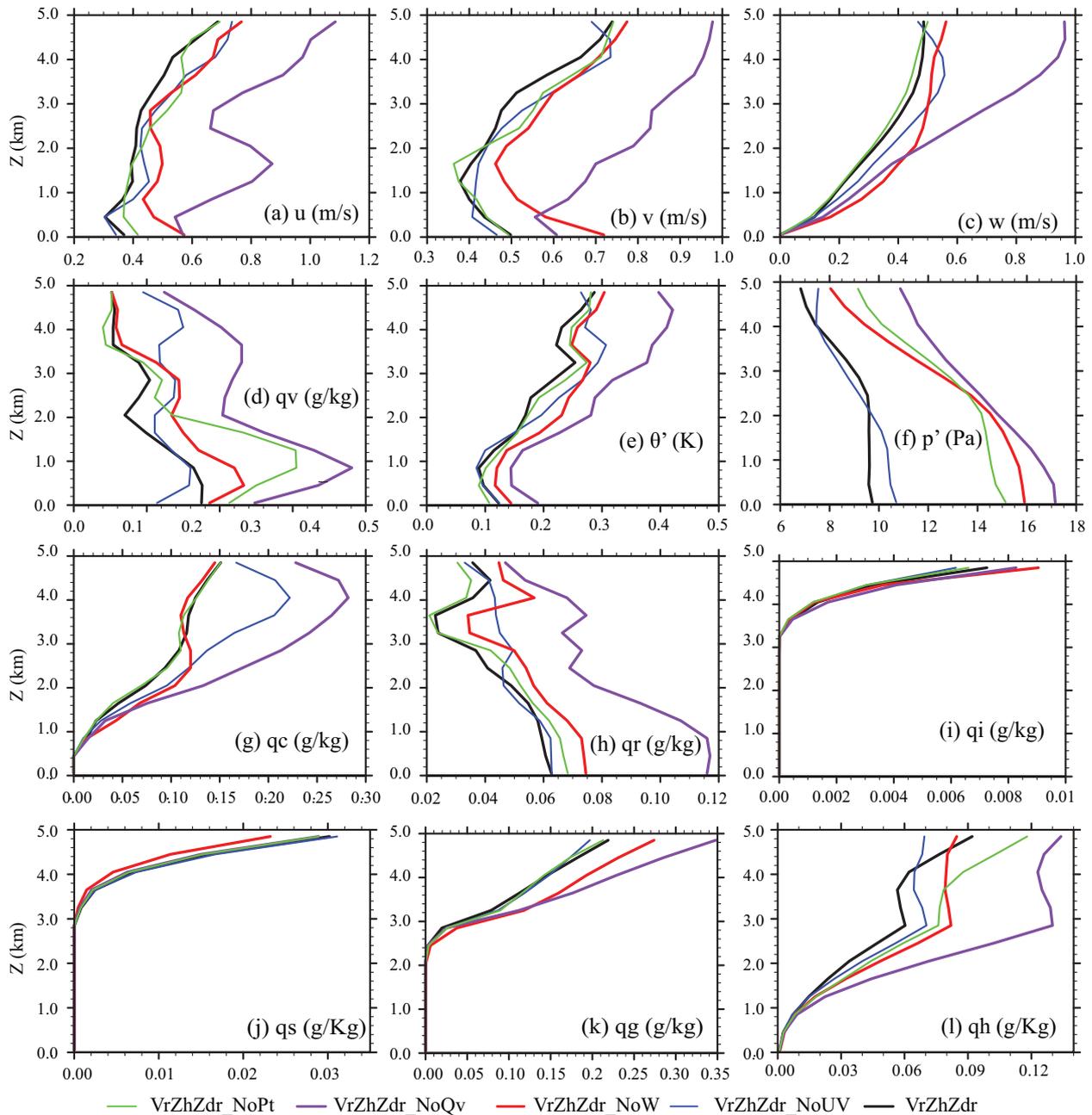
Figure 11 shows that analysis and forecast RMSEs of sensitivity experiments without updating certain state



**FIGURE 11** Forecast and analysis RMSEs of experiments VrZhZdr\_NoPt (green), VrZhZdr\_NoQv (purple), VrZhZdr\_NoW (red), VrZhZdr\_NoUV (blue) and VrZhZdr (black) for variables (a)  $u$ , (b)  $v$ , (c)  $w$ , (d) water vapour mixing ratio  $q_v$ , (e) perturbation potential temperature  $\theta'$ , (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  $q_h$

variables when assimilating  $Z_{DR}$ . RMSEs for experiment VrZhZdr are shown in black lines while those for VrZhZdr\_NoPt, VrZhZdr\_NoUV and VrZhZdr\_NoW are shown in colour. Among all potential temperature, water vapour, vertical and horizontal wind components, the updating of water vapour  $q_v$  has the greatest impact. The RMSEs from VrZhZdr\_NoQv (solid purple) are significantly larger for almost all forecast times and state variables and the differences are larger in later cycles. The updating of  $w$  has the second largest impact as

the RMSEs of VrZhZdr\_NoW (red) are noticeably larger for most variables especially during the intermediate cycles. The updating of horizontal wind components and potential temperature has less impact as the RMSEs of VrZhZdr\_NoPt (solid green) and VrZhZdr\_NoUV (blue) are rather close to those of VrZhZdr. These results are reasonable since water vapour is the primary fuel for intense convection while  $w$  provides the best measure of the intensity of convection. Given that VrZhZdr produces overall the lowest RMSEs, all state variables should be updated

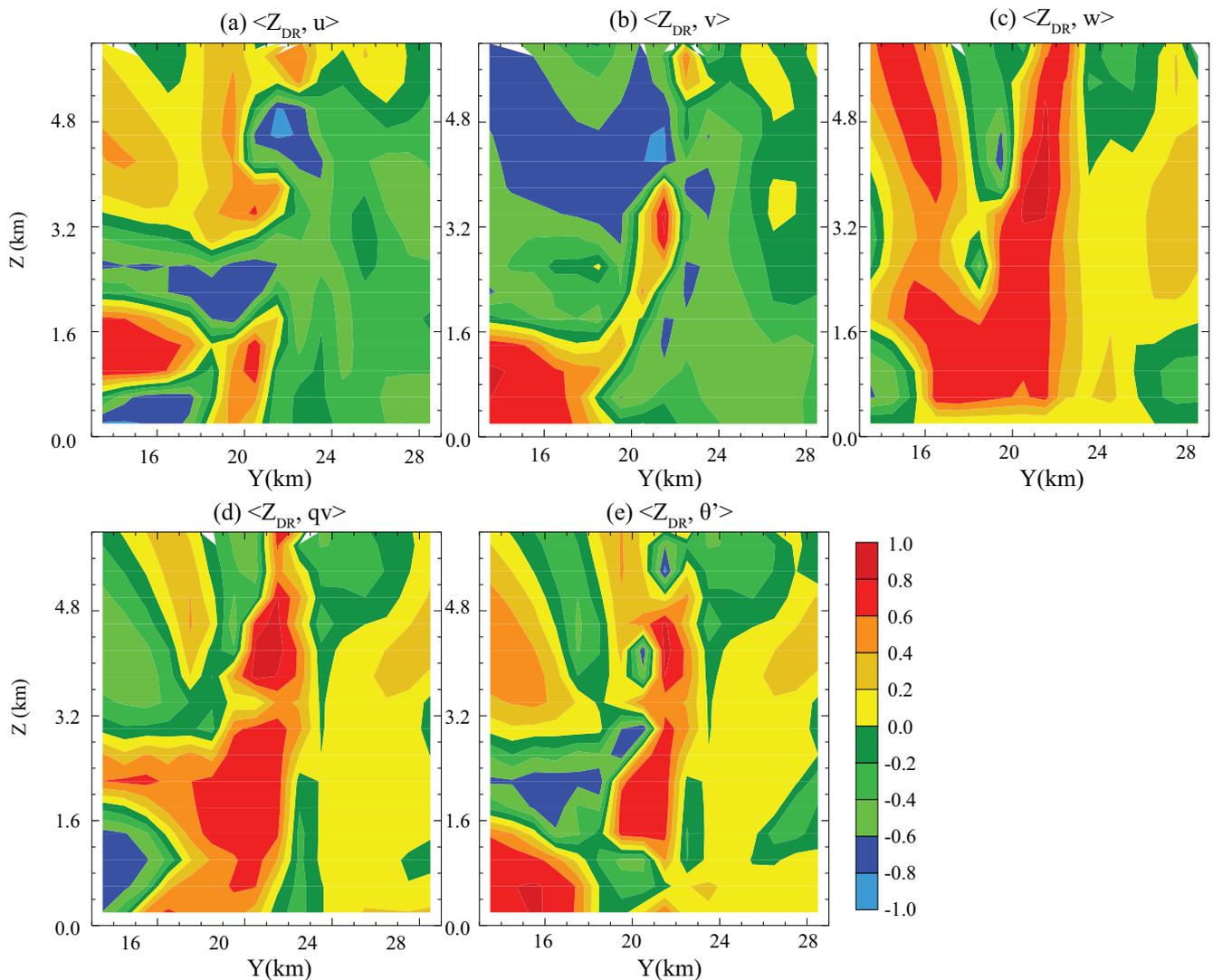


**FIGURE 12** (a–l) Similar to Figure 11 but for vertical profiles of analysis RMSEs for different variables at 80 min

when assimilating  $Z_{DR}$ , at least when no model error is present and the ensemble-estimated covariances are reasonably accurate.

The vertical RMSE profiles up to 5 km height at 80 min are shown in Figure 12. Here, we focus on the low levels where  $Z_{DR}$  has largest impacts. Consistent with Figure 11, experiment VrZhZdr\_NoQv has the largest errors at essentially all vertical levels shown. Experiment VrZhZdr\_NoW produces the second largest RMSEs for most variables at most levels. The updating of potential temperature  $\theta$  has the third largest impact (e.g. on  $q_v$  in Figure 12d, on  $p'$  in

Figure 12f, and  $q_h$  in Figure 12l), although for some variables not updating  $\theta$  made little difference (e.g. for  $q_c$  in Figure 12g and  $q_r$  in Figure 12h). The updating of  $u$  and  $v$  has limited impact from lower to middle levels. In experiment VrZhZdr\_NoUV, the analysis RMSEs are close to those of VrZhZdr below 2 km, but larger above 2 km for variables including  $q_v$  (Figure 12d),  $q_c$  (Figure 12g) and  $q_r$  (Figure 12h). This is better illustrated in Figure 13 which shows the correlations between the  $Z_{DR}$  and wind components, the mixing ratio  $q_v$  and potential temperature perturbation  $\theta'$ . The  $Z_{DR}$  point is the same as the point in

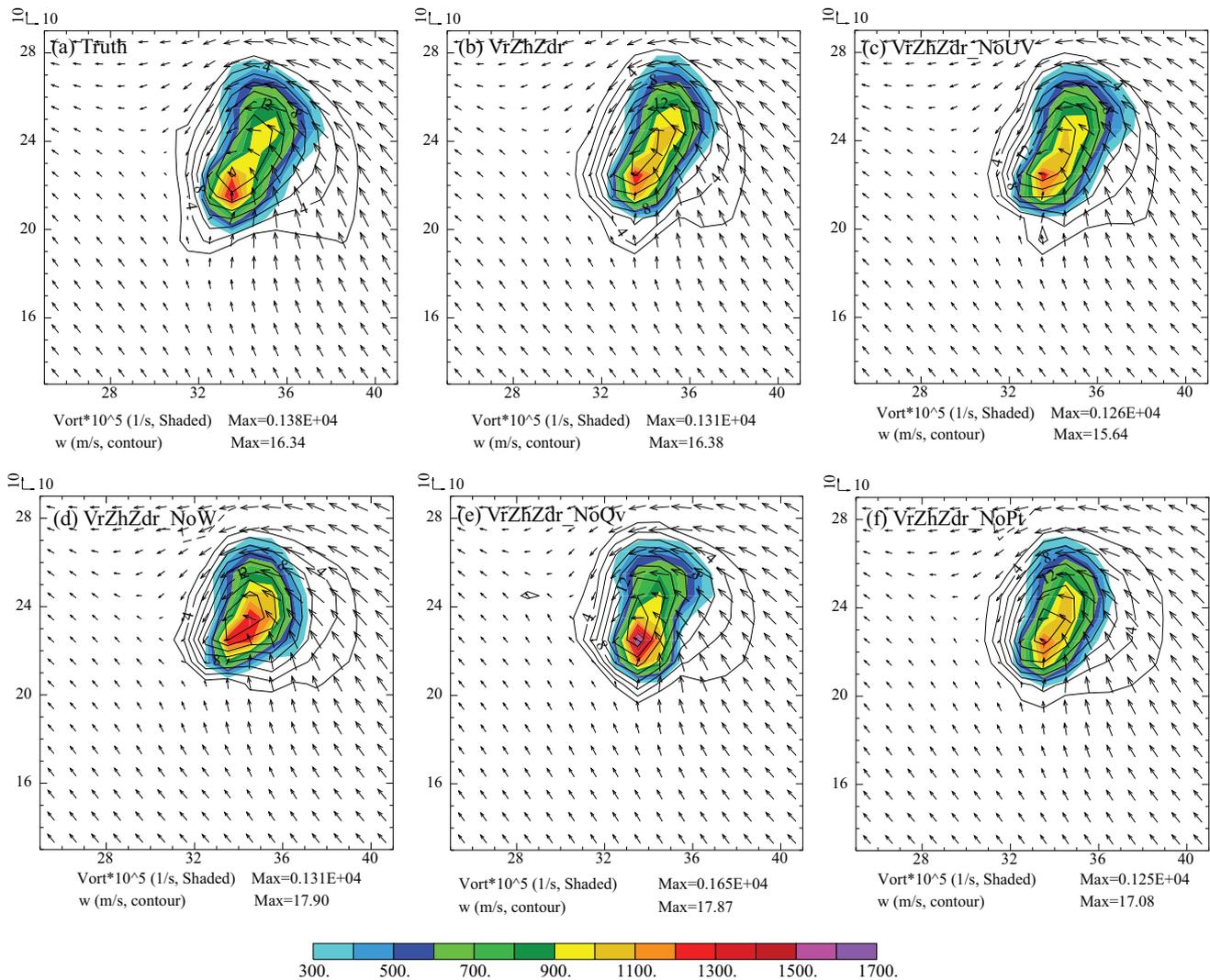


**FIGURE 13** Similar to Figure 9 but for correlations between  $Z_{DR}$  and (a)  $u$ , (b)  $v$ , (c)  $w$ , (d)  $q_v$  and (e)  $\theta'$ , respectively

Figure 9. For  $q_v$  and  $w$ , they show high and continuous correlation regions from the bottom to the top. For  $u$ ,  $v$  and  $\theta$ , the high correlation areas are clearly reduced and mostly located in the lower levels.

In Figure 14, we further examine the impact of not updating certain variables on the dynamic structures of the analysed storm. Here, the vertical vorticity  $\zeta$  at 2 km height in the main updraught region is shown, indicating low-level mesocyclone structure and intensity. Also plotted are the vertical velocity  $w$  and horizontal winds. The truth shows an ellipse-shaped structure of  $\zeta$  with its centre located to the south of the  $w$  maximum (Figure 14a). With all state variables updated in EnKF, experiment VrZhZdr obtains very similar structures of  $\zeta$  and  $w$  with the horizontal winds flowing around the north side of the updraught core (Figure 14b). Without updating horizontal winds in VrZhZdr\_NoUV when assimilating  $Z_{DR}$  data, the overall

structures of  $\zeta$  and  $w$  and horizontal winds are not too different from those of VrZhZdr except that their intensities are somewhat underestimated (Figure 14c). The impact of not updating  $\theta$  in VrZhZdr\_NoPt by  $Z_{DR}$  is similar to not updating  $u$  and  $v$  (impact is relatively small), although the maximum  $w$  is slightly overestimated according to the  $w$  maximum values shown in the plots (Figure 14f). Compared to  $u$ ,  $v$  and  $\theta$ , the impact of not updating  $w$  or  $q_v$  when assimilating  $Z_{DR}$  is much larger. Without updating  $w$ , the  $\zeta$  pattern appears more circular and the updraught is more concentrated but its maximum is overestimated (Figure 14d). Without updating  $q_v$ , the shapes of  $\zeta$  structure and updraught region are still close to those of VrZhZdr and truth, but the maximum  $\zeta$  is most overestimated among the sensitivity experiments, and the  $w$  maximum is also overestimated (Figure 14e), although slightly less so than in VrZhZdr\_NoW.

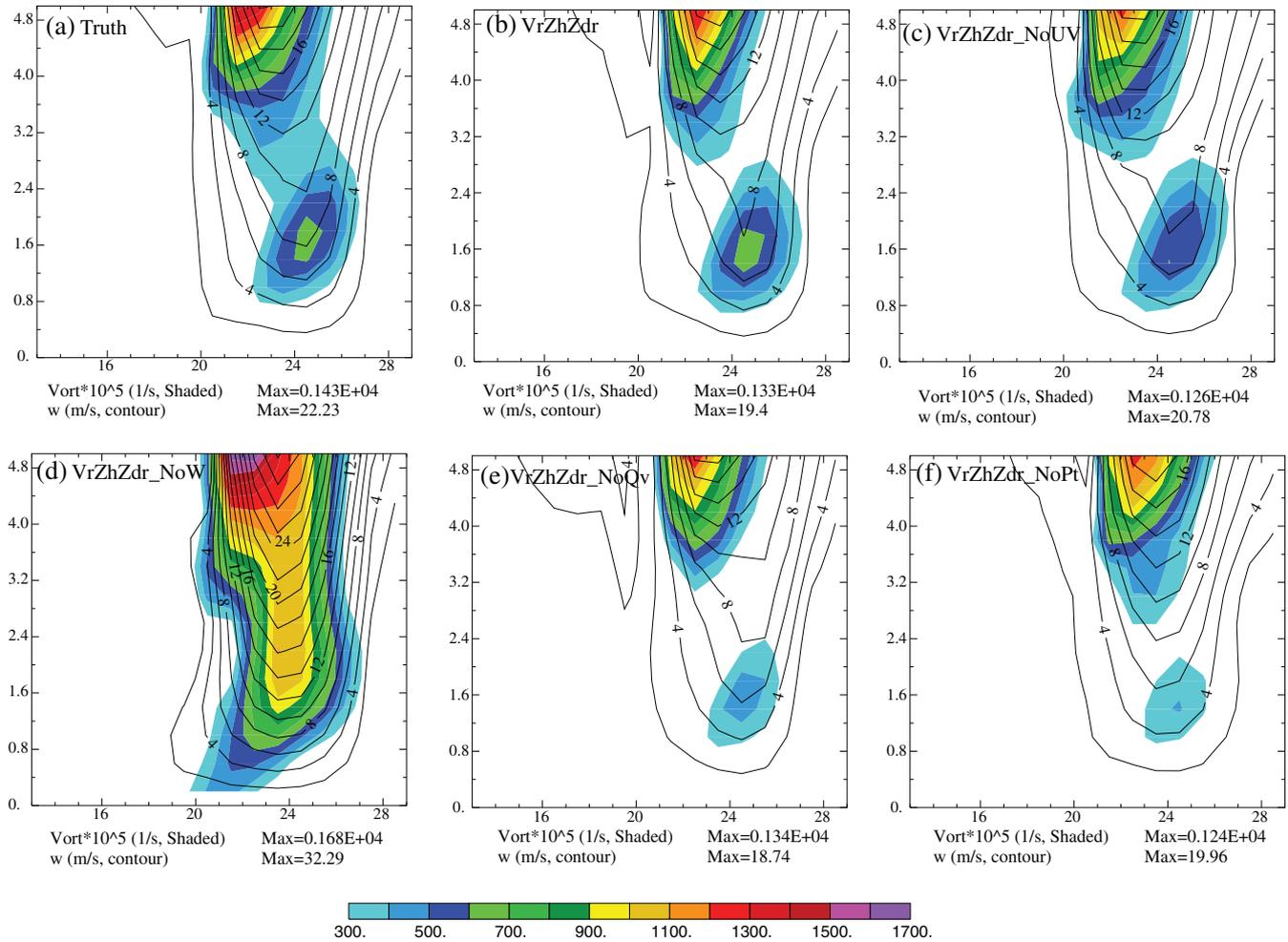


**FIGURE 14** Vertical vorticity (shaded), velocity (contours) and horizontal wind vectors at 2 km height for (a) truth, (b) VrZhZdr, (c) VrZhZdr\_NoUV, (d) VrZhZdr\_NoW, (e) VrZhZdr\_NoQv and (f) VrZhZdr\_NoPt

The above results indicate the analysed flow structures and intensity in the main updraught region are directly linked to the updating of  $w$  and  $q_v$ . This is more obvious in vertical cross-sections. Figure 15 shows the vertical cross-sections of  $\zeta$  and  $w$  fields through the maximum of three-dimensional  $w$  in the  $y$  direction. In both truth and experiment VrZhZdr, a  $\zeta$  maximum is found at  $\sim 1.6$  km level which corresponds to relatively strong vertical motion there, and the fields in VrZhZdr match the truth very closely. Except for experiment VrZhZdr\_NoW, the general patterns of  $\zeta$  and  $w$  in other sensitivity experiments are similar except for underestimation of the low-level vorticity strength, especially in VrZhZdr\_NoQv and VrZhZdr\_NoPt. In VrZhZdr\_NoW, mid-level ( $z \sim 5$  km)  $w$  is overestimated by nearly 50%, as is the column of high vertical vorticity (Figure 15f). Figures 14 and 15 provide more concrete ideas on the large impact of updating or not

updating  $w$  and  $q_v$  by  $Z_{DR}$ , results that are consistent with earlier findings based on RMSEs. This further confirms that there are reliable, strong ensemble-derived correlation between  $w$  and  $Z_{DR}$  that enables improved analysis of  $w$  and other fields by  $Z_{DR}$  observations.

Based on the above results, all model state variables should be updated when assimilating  $Z_{DR}$  data with in the EnKF. This is at least true for perfect model OSSEs. In our OSSE framework for a tornadic supercell storm, apart from the updating of hydrometeor state variables, the updating of water vapour mixing ratio  $q_v$  has the largest impact on the overall analysis accuracy followed by vertical wind  $w$ . Not updating  $w$  leads to large errors in the flows in the updraught region, including the updraught itself and vertical vorticity associated with it. The errors in  $q_v$  affect the storm dynamics more indirectly through moist processes.



**FIGURE 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 Figure 14). (a) Truth, (b) VrZhZdr, (c) VrZhZdr\_NoUV, (d) VrZhZdr\_NoW, (e) VrZhZdr\_NoQv and (f) VrZhZdr\_NoPt

## 4 | CONCLUSIONS AND DISCUSSIONS

In this study, the impact of assimilating differential reflectivity  $Z_{DR}$  data within an EnKF framework is investigated using observing system simulation experiments with simulated data for a tornadic supercell storm. The Milbrandt and Yau (2005) double-moment microphysics scheme is used in both truth simulation and for EnKF DA; with this double-moment scheme, previous studies have shown the reasonable ability to simulate most important polarimetric radar signatures found in supercell storms. Radar observations are simulated using a polarimetric radar data simulator developed by Jung *et al.* (2010b), in which the T-matrix method is used to calculate the hydrometeor scattering magnitudes for particles of particular sizes. Observation errors of realistic magnitudes are added to the simulated observations, and the same error variances are specified in the EnKF DA. The observation

operators from the simulator are also used in the EnKF DA, which is run over a 90 min period assimilating radar data every 5 min spanning the developing and mature stages of the supercell. Experiments are conducted with and without assimilating  $Z_{DR}$  data in addition to reflectivity at horizontal polarization  $Z_H$  and radial velocity  $V_r$  to examine the impact of  $Z_{DR}$  assimilation. Results show that the assimilation of  $Z_{DR}$  reduces the RMSEs for almost all model state variables at almost all analysis times. The polarimetric signatures of tornadic storm including the  $Z_{DR}$  and  $K_{DP}$  columns and  $Z_{DR}$  arc are all improved. Analyses show that the structures and intensities of hydrometeor fields at both lower and upper levels are improved, even though the strongest  $Z_{DR}$  signatures are mostly found at the lower levels due to the concentration of large raindrops there.

Additional sensitivity experiments are conducted to understand the benefit and impact of updating different state variables when assimilating  $Z_{DR}$ . The first sensitivity

experiment excludes the updating of the total number concentrations of all hydrometeors, which arise from the use of a double-moment microphysics scheme. Although the number concentrations have very large dynamic ranges, their relations with  $Z_{DR}$  are highly nonlinear, and the ensemble-derived error correlations with  $Z_{DR}$  may or may not be reliable enough to produce improved analyses, the results show that updating number concentrations together with the mixing ratios are very beneficial. The number concentrations have high correlations with  $Z_{DR}$  at the lower levels that are comparable to those of mixing ratios. If the number concentrations are not updated by  $Z_{DR}$  observations, most of the benefit of assimilating  $Z_{DR}$  data is lost, and in fact, for vertical velocity, rainwater and hail mixing ratios, the analysis RMSEs are larger in intermediate DA cycles than those in the experiment not assimilating  $Z_{DR}$  data at all. Clearly, updating both mixing ratios and total number concentrations of hydrometeors leads to much more physically consistent analyses.

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

Finally, we point out that the results presented in this article are limited to OSSE tests with a single supercell storm, and no model error is included. When model error is present, as with all real data cases, the conclusions may be somewhat different. In addition, other polarimetric measurements, including specific differential phase  $K_{DP}$  and co-polar correlation coefficient  $\rho_{hv}$ , also contain valuable information on the hydrometeors and their PSDs.

The assimilation of these parameters and their impact on analysed storm and subsequent forecasts were not considered in this study, or in the real data study of Putnam *et al.* (2019); they require further research and investigations. The impact of PRD assimilation on other types of precipitation systems also requires study.

## ACKNOWLEDGEMENTS

This work was primarily supported by the National Key Research and Development Program of China under grant 2018YFC1507604 and the National Natural Science Foundation of China (Grant Nos. 41730965, 41975124 and 41405100). This work was also supported by the Open Research Program of the State Key Laboratory of Severe Weather.

## ORCID

Kefeng Zhu  <https://orcid.org/0000-0003-3334-8743>

Ming Xue  <https://orcid.org/0000-0003-1976-3238>

## REFERENCES

- Anderson, J.L. (2001) An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Review*, 129, 2884–2903.
- Brown, B.R., Bell, M.M. and Frambach, A.J. (2016) Validation of simulated hurricane drop size distributions using polarimetric radar. *Geophysical Research Letters*, 43, 910–917.
- Cao, Q., Zhang, G. and Xue, M. (2013) A Variational approach for retrieving raindrop size distribution from polarimetric radar measurements in the presence of attenuation. *Journal of Applied Meteorology and Climatology*, 52, 169–185.
- Carlin, J.T., Gao, J., Snyder, J.C. and Ryzhkov, A.V. (2017) Assimilation of ZDR columns for improving the spinup and forecast of convective storms in storm-scale models: proof-of-concept experiments. *Monthly Weather Review*, 145, 5033–5057.
- Dawson, D.T., Mansell, E.R., Jung, Y., Wicker, L.J., Kumjian, M.R. and Xue, M. (2014) Low-level ZDR signatures in supercell forward flanks: the role of size sorting and melting of hail. *Journal of Atmospheric Sciences*, 71, 276–299.
- Dixon, M., Li, Z.H., Lean, H., Roberts, N. and Ballard, S. (2009) Impact of data assimilation on forecasting convection over the United Kingdom using a high-resolution version of the Met Office Unified Model. *Monthly Weather Review*, 137, 1562–1584.
- Gaspari, G. and Cohn, S.E. (1999) Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125, 723–757.
- Hu, M. and Xue, M. (2007) Impact of configurations of rapid intermittent assimilation of WSR-88D radar data for the 8 May 2003 Oklahoma City tornadic thunderstorm case. *Monthly Weather Review*, 135, 507–525.
- Hu, M., Xue, M. and Brewster, K. (2006) 3DVAR and cloud analysis with WSR-88D level-II data for the prediction of Fort Worth tornadic thunderstorms. Part I: Cloud analysis and its impact. *Monthly Weather Review*, 134, 675–698.
- Huang, H., Zhang, G., Zhao, K., Liu, S., Wen, L., Chen, G. and Yang, Z. (2019) Uncertainty in retrieving raindrop size distribution from

- polarimetric radar measurements. *Journal of Atmospheric and Oceanic Technology*, 36, 585–605.
- Huuskonen, A., Saltkoff, E. and Holleman, I. (2014) The operational weather radar network in Europe. *Bulletin of the American Meteorological Society*, 95, 897–907.
- Jung, Y., Xue, M. and Zhang, G. (2010a) Simultaneous estimation of microphysical parameters and the atmospheric state using simulated polarimetric radar data and ensemble Kalman filter in the presence of observation operator error. *Monthly Weather Review*, 138, 539–562.
- Jung, Y., Xue, M. and Zhang, G. (2010b) Simulations of polarimetric radar signatures of a supercell storm using a two-moment bulk microphysics scheme. *Journal of Applied Meteorology and Climatology*, 49, 146–163.
- Jung, Y., Xue, M., Zhang, G. and Straka, J. (2008b) Assimilation of simulated polarimetric radar data for a convective storm using ensemble Kalman filter. Part II: Impact of polarimetric data on storm analysis. *Monthly Weather Review*, 136, 2246–2260.
- Jung, Y., Zhang, G. and Xue, M. (2008a) Assimilation of simulated polarimetric radar data for a convective storm using ensemble Kalman filter. Part I: Observation operators for reflectivity and polarimetric variables. *Monthly Weather Review*, 136, 2228–2245.
- Kalnay, E. (2002) *Atmospheric Modeling, Data Assimilation, and Predictability*. New York: Cambridge University Press, 341 p.
- Kumjian, M.R. and Ryzhkov, A.V. (2008) Polarimetric signatures in supercell thunderstorms. *Journal of Applied Meteorology and Climatology*, 47, 1940–1961.
- Li, X. and Mecikalski, J.R. (2010) Assimilation of the dual-polarization Doppler radar data for a convective storm with a warm-rain radar forward operator. *Journal of Geophysical Research*, 115, D16208. <https://doi.org/10.1029/2009JD013666>.
- Li, X., Mecikalski, J.R. and Posselt, D. (2017) An ice-phase microphysics forward model and preliminary results of polarimetric radar data assimilation. *Monthly Weather Review*, 145, 683–708.
- Liu, C., Xue, M. and Kong, R. (2019) *Direct variational assimilation of radar reflectivity and radial velocity data: issues with nonlinear reflectivity operator and solutions*. *Monthly Weather Review*, 137, 17–29. <https://doi.org/10.1175/MWR-D-19-0149.1>.
- Milbrandt, J.A. and Yau, M.K. (2005) A multimoment bulk microphysics parameterization. Part I: Analysis of the role of the spectral shape parameter. *Journal of the Atmospheric Sciences*, 62, 3051–3064.
- Putnam, B.J., Xue, M., Jung, Y., Snook, N. and Zhang, G. (2014) The analysis and prediction of microphysical states and polarimetric radar variables in a mesoscale convective system using double-moment microphysics, multinet network radar data, and the ensemble Kalman filter. *Monthly Weather Review*, 142, 141–162.
- Putnam, B.J., Xue, M., Jung, Y., Snook, N. and Zhang, G. (2017a) Ensemble probabilistic prediction of a mesoscale convective system and associated polarimetric radar variables using single-moment and double-moment microphysics schemes and EnKF radar data assimilation. *Monthly Weather Review*, 145, 2257–2279.
- Putnam, B.J., Xue, M., Jung, Y., Snook, N. and Zhang, G. (2019) Ensemble Kalman filter assimilation of polarimetric radar observations for the 20 May 2013 Oklahoma tornadic supercell case. *Monthly Weather Review*, 147, 2511–2533.
- Putnam, B.J., Xue, M., Jung, Y., Zhang, G. and Kong, F. (2017b) Simulation of polarimetric radar variables from 2013 CAPS spring experiment storm-scale ensemble forecasts and evaluation of microphysics schemes. *Monthly Weather Review*, 145, 49–73.
- Ray, P.S., Johnson, B., Johnson, K.W., Bradberry, J.S., Stephens, J.J., Wagner, K.K., Wilhelmson, R.B. and Klemp, J.B. (1981) The morphology of severe tornadic storms on 20 May 1977. *Journal of the Atmospheric Sciences*, 38, 1643–1663.
- ROC. (2013) *WSR-88D dual polarization deployment progress*. NOAA/RadarOperations Center, 6 pp. Available at: <http://www.roc.noaa.gov/WSR88D/PublicDocs/DualPol/DPstatus.pdf> [Accessed 24 June 2013].
- Ryzhkov, A.V. and Zrnich, D.S. (1998) Discrimination between rain and snow with a polarimetric radar. *Journal of Applied Meteorology*, 37, 1228–1240.
- Sobash, R.A. and Stensrud, D.J. (2013) The impact of covariance localization for radar data on EnKF analyses of a developing MCS: observing system simulation experiments. *Monthly Weather Review*, 141, 3691–3709.
- Stensrud, D.J., Xue, M., Wicker, L.J., Kelleher, K.E., Foster, M.P., Schaefer, J.T., Schneider, R.S., Benjamin, S.G., Weygandt, S.S., Ferree, J.T. and Tuell, J.P. (2009) Convective-scale warn-on-forecast system. *Bulletin of the American Meteorological Society*, 90, 1487–1500.
- Sun, J., Xue, M., Wilson, J.W., Zawadzki, I., Ballard, S.P., Onville-Hoomeyer, J., Joe, P., Barker, D.M., Li, P.-W., Golding, B., Xu, M. and Pinto, J. (2013) Use of NWP for nowcasting convective precipitation: recent progress and challenges. *Bulletin of the American Meteorological Society*, 95, 409–426.
- Tong, M. and Xue, M. (2005) Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. *Monthly Weather Review*, 133, 1789–1807.
- Vivekanandan, J., Adams, W.M. and Bringi, V.N. (1991) Rigorous approach to polarimetric radar modeling of hydrometeor orientation distributions. *Journal of Applied Meteorology*, 30, 1053–1063.
- Vivekanandan, J., Ellis, S.M., Oye, R., Zrnich, D.S., Ryzhkov, A.V. and Straka, J. (1999) Cloud microphysics retrieval using S-band dual-polarization radar measurements. *Bulletin of the American Meteorological Society*, 80, 381–388.
- Whitaker, J.S. and Hamill, T.M. (2002) Ensemble data assimilation without perturbed observations. *Monthly Weather Review*, 130, 1913–1924.
- Wu, B., Verlinde, J. and Sun, J. (2000) Dynamical and microphysical retrievals from Doppler radar observations of a deep convective cloud. *Journal of the Atmospheric Sciences*, 57, 262–283.
- Xiao, Q., Sun, J., Lee, W.-C., Kuo, Y.-H., Barker, D.M., Lim, E., Won, D.-J., Lee, M.-S., Lee, W.-J., Cho, J.-Y., Lee, D.-K. and Lee, H.-S. (2008) Doppler radar data assimilation in KMA's operational forecasting. *Bulletin of the American Meteorological Society*, 89, 39–43.
- Xue, M., Droegemeier, K.K., Wong, V., Shapiro, A., Brewster, K., Carr, F., Weber, D., Liu, Y. and Wang, D. (2001) The advanced regional prediction system (ARPS) – a multiscale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteorology and Atmospheric Physics*, 76, 143–165.
- Xue, M., Tong, M. and Droegemeier, K.K. (2006) An OSSE framework based on the ensemble square-root Kalman filter for evaluating impact of data from radar networks on thunderstorm analysis

- and forecast. *Journal of Atmospheric and Oceanic Technology*, 23, 46–66.
- Xue, M., Wang, D., Gao, J., Brewster, K. and Drogemeier, K.K. (2003) The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation. *Meteorology and Atmospheric Physics*, 82, 139–170.
- Zhang, G., Mahale, V.N., Putnam, B.J., Qi, Y., Cao, Q., Byrd, A.D., Bukovcic, P., Zrnic, D.S., Gao, J., Xue, M., Jung, Y., Reeves, H.D., Heinselman, P.L., Ryzhkov, A., Palmer, R.D., Zhang, P., Weber, M., McFarquhar, G.M., Moore, B., Zhang, Y., Zhang, J., Vivekanandan, J., Al-Rashid, Y., Ice, R.L., Berkowitz, D.S., Tong, C., Fulton, C. and Doviak, R.J. (2019) Current status and future challenges of weather radar polarimetry: bridging the gap between radar meteorology/hydrology/engineering and numerical weather prediction. *Advances in Atmospheric Sciences*, 36, 571–588.
- Zhao, K., Huang, H., Wang, M., Lee, W.-C., Chen, G., Wen, L., Wen, J., Zhang, G., Xue, M., Yang, Z., Liu, L., Wu, C., Hu, Z. and Chen, S. (2019) Recent progress in dual-polarization radar research and applications in China. *Advances in Atmospheric Sciences*, 36, 961–974.
- Zhu, K., Yang, Y. and Xue, M. (2015) Percentile-based neighborhood precipitation verification and its application to a landfalling tropical storm case with radar data assimilation. *Advances in Atmospheric Sciences*, 32, 1449–1459.

**How to cite this article:** Zhu K, Xue M, Ouyang K, Jung Y. Assimilating polarimetric radar data with an ensemble Kalman filter: OSSEs with a tornadic supercell storm simulated with a two-moment microphysics scheme. *QJR Meteorol Soc.* 2020;146:1880–1900. <https://doi.org/10.1002/qj.3772>