

# PODEn4DVar-based radar data assimilation scheme: formulation and preliminary results from real-data experiments with advanced research WRF (ARW)

By BIN ZHANG<sup>1,2</sup>, XIANGJUN TIAN<sup>2\*</sup>, JIANHUA SUN<sup>3</sup>, FENG CHEN<sup>4</sup>,  
YUANCHUN ZHANG<sup>3</sup>, LIFENG ZHANG<sup>1</sup> and SHENMING FU<sup>2</sup>, <sup>1</sup>College of Meteorology  
and Oceanography, PLA University of Science and Technology, Nanjing 211101, China; <sup>2</sup>ICCES, Institute of  
Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China; <sup>3</sup>LACS, Institute of Atmospheric  
Physics, Chinese Academy of Sciences, Beijing 100029, China; <sup>4</sup>Zhejiang Institute of Meteorological Sciences,  
Hangzhou, 310017, China

(Manuscript received 17 September 2014; in final form 21 January 2015)

## ABSTRACT

The Proper Orthogonal Decomposition (POD)-based ensemble four-dimensional variational (4DVar) assimilation method (referred to as PODEn4DVar) is a hybrid assimilation method that exploits the strengths of both the ensemble Kalman filter (EnKF) and the 4DVar assimilation method. Its feasibility and validity have been demonstrated using ideal models through observing system simulation experiments (OSSEs). In this study, we further utilise this approach to build a PODEn4DVar-based radar data assimilation scheme (PRAS). In a PRAS, radar observations including radial velocity and reflectivity, after some necessary data preprocessing, are assimilated directly to improve model initialisation. A group of single-observation-based OSSEs are first designed to generally evaluate the validity of PRAS. Subsequently, a group of comparison experiments are also carried out between PRAS and an LETKF-based radar assimilation scheme (LRAS), which shows that PRAS is able to produce results better than (at least as good as) LRAS. Thirdly, to evaluate the potential impact for PRAS in the operational context, a group of cycling assimilation experiments of radar data are performed, which demonstrates that PRAS can gradually improve the accuracy of analysis field by cycling assimilation. Finally, a heavy convective-rainfall case study was selected to investigate the performance of PRAS in assimilating real radar observations and the impacts of assimilating radar observations on numerical forecasts, with the Weather Research and Forecasting (WRF) model as our forecasting model. The results show that significant improvements in predicting heavy rainfall can be achieved due to the improved initial conditions for the convective system's dynamics and microphysics after assimilating the radar observations with PRAS. In summary, the results show that the PODEn4DVar is a promising method for atmospheric data assimilation.

*Keywords:* PODEn4Dvar, single-observation, radar data assimilation, WRF, heavy rainfall

## 1. Introduction

Recent studies have shown that assimilation of radar data can improve the short-term forecasts of convective systems by incorporating the initial conditions of mesoscale storm structures (Weygandt et al., 2002; Sun, 2005; Dawson and Xue, 2006; Hu et al., 2006; Zhao et al., 2006; Xiao et al., 2007, 2008; Pu et al., 2009; Li and John, 2010; Kawabata et al., 2011). In particular, both the four-dimensional

variational (4DVar) data assimilation method (e.g. Lewis and Derber, 1985; Courtier et al., 1994) and the ensemble Kalman filter (EnKF; e.g. Evensen, 1994, 2004; Houtekamer and Mitchell, 1998, 2001) show promising results in convective-scale radar data assimilation. The 4DVar promises to provide an initial condition that is consistent with the forecasting model and can make full use of multiple volume scans from radar. However, it is not an easy task to implement 4DVar radar data assimilation, largely because of the need for an adjoint model, whose coding, maintenance, and development are significantly labour intensive. Despite this, a few studies have looked at 4DVar-based analysis schemes for

\*Corresponding author.  
email: tianxj@mail.iap.ac.cn

radar data assimilation. Sun and Crook (1997, 1998) developed a 4DVar system called the variational Doppler radar analysis system (VDRAS) to assimilate radar reflectivity and radial velocity with an elastic non-hydrostatic model. Their results from both simulated and real-data experiments gave evidence of VDRAS improvement in retrieval of the cloud structure, especially for wind and moisture fields. Gao et al. (1998) also attempted to develop another 4DVar-based radar assimilation system using the Advanced Regional Prediction (ARPS) adjoint model. Kawabata et al. (2011) modified a cloud-resolving non-hydrostatic 4DVar data assimilation system to assimilate radar reflectivity data directly, which also forms a 4DVar radar data assimilation system. Very recently, Wang et al. (2013) and Sun and Wang (2013) assimilate radar data in the high-resolution Advanced Research WRF (ARW) for the improvement of short-term quantitative precipitation forecasting using the 4DVar data assimilation technique. However, if cold processes were not taken into account, 4DVar radar data assimilation will show strong limitations when applied to cold microphysics (Wu et al. 2000). As such, Caumont et al. (2010) further proposed an original ‘1D+3DVar method’ to assimilate radar reflectivities, which was used to mitigate difficulties in direct variational reflectivity assimilation. The method has been operationally implemented in the AROME Model at Météo-France (Wattrelot et al., 2014). On the other hand, the EnKF has been increasingly applied to the assimilation of simulated Doppler radar data for modelled convective storms (Snyder and Zhang, 2003; Zhang et al., 2004a; Tong and Xue, 2005) and real radar data (Dowell et al., 2011) because of its simple conceptual formulation and relative ease of implementation. It is well known that, by forecasting statistical characteristics, the EnKF can provide flow-dependent estimates of the background error covariance, which leads to many encouraging results from EnKF-based radar data assimilation. However, some impact studies using reflectivity assimilation methods have shown relatively small benefits in comparison with Doppler velocity data assimilation (Caya et al., 2005; Tong and Xue, 2005). Nevertheless, all studies report positive results with radar observations and improved forecast quality in a deterministic-forecasts framework (Aksoy et al., 2010).

Through a comprehensive comparison for storm-scale radar data assimilation between the 4DVar and the EnKF, Caya et al. (2005) concluded that, under a perfect-model assumption, the temporal smoothness constraint character made the 4DVar generate good, dynamically consistent analyses almost immediately. By contrast, it took longer for the EnKF to spin up, but ultimately the flow-dependent background error covariance utilised by the EnKF enabled it to perform well. Other studies (e.g. Hunt et al., 2007) also showed that each of the two data assimilation methods has its own strengths and weaknesses. In the 4DVar, the

physical model provides a temporal smoothness constraint and has the ability to simultaneously assimilate the observation data at multiple times, whereas only a static (not flow-dependent) background error covariance at the start of the assimilation window is adopted. By contrast, the EnKF lacks the temporal smoothness constraint of the 4DVar, because it is naturally designed to incorporate sequential information alone, but fortunately, its estimates of the background error covariance can model flow evolution.

Based on the characteristics of the 4DVar and the EnKF, major efforts have been devoted to improve atmospheric data assimilation by coupling 4DVar with EnKF with the goal of combining their strengths (e.g. Lorenc, 2003; Qiu et al., 2007; Liu et al., 2008; Tian et al., 2008, 2011; Zhang et al., 2009; Wang et al., 2010; Cheng et al., 2010). A hybrid method, referred to as PODEn4DVar, was proposed by Tian et al. (2008, 2011) based on the proper orthogonal decomposition (POD) and ensemble forecasting techniques. In the PODEn4DVar, the POD technique is adopted to transform the original ensemble coordinate system into an optimal one in the  $L^2$  norm (Ly and Tran, 2001), which contributes greatly to its enhanced assimilation performance. Its feasibility and effectiveness were demonstrated in an idealised model with simulated observations (Tian and Xie, 2012; Tian et al., 2011). It was found that the PODEn4DVar outperforms both the 4DVar and the EnKF under both perfect- and imperfect-model scenarios with lower computational costs than the EnKF (Tian et al., 2011). This method has also been successfully applied to land data assimilation (Tian et al., 2009, 2010) and Tan-Tracker joint data assimilation (Tian et al., 2014). Furthermore, as the first step to apply the PODEn4DVar in radar data assimilation, we have already built a three-dimensional (3D) case of PODEn4DVar-based radar assimilation system on the WRF model platform (Pan et al., 2012). The results show that assimilating real radar data with PODEn3DVar can effectively improve the description for the initial condition, thus achieving significant improvement in predicting precipitation.

In this study, we further report on the development and assessment of a new radar data assimilation scheme based on the PODEn4DVar approach [referred to as PODEn4DVar-based radar data assimilation scheme (PRAS)] and on an application of this scheme with the WRF model to assimilate real radar reflectivity and radial velocity data from single Doppler radar for a case of heavy rainfall. The rest of the paper is organised as follows. In Section 2, we describe our PODEn4DVar assimilation scheme, including a simple review of the PODEn4DVar, a description of the observation operator for the radial velocity, radar reflectivity, and localisation schemes. In Section 3, the performance of PRAS is comprehensively evaluated by four groups of numerical experiments including single-observation experiments,

comparisons between PRAS and LRAS [i.e. LETKF-based (Hunt et al., 2007) radar assimilation scheme], cycling assimilation, and real-data assimilation. Finally, a summary is given in Section 4.

## 2. The PODEn4DVar-based radar data assimilation scheme

### 2.1. Formulation of the PODEn4DVar

To provide a self-contained description of the radar data-assimilation scheme (referred to as PRAS) used here, we first present the formulation on the PODEn4DVar. Some equations of Tian et al. (2011) and Tian and Xie (2012) are repeated in this section.

The PODEn4DVar originates from the traditional 4DVar methodology. By minimising the following incremental format of the standard 4DVar cost function  $J(\mathbf{x}')$ , one can obtain an optimal increment ( $\mathbf{x}'_a$ ) of the initial condition (IC) at the initial time  $t_0$ :

$$J(\mathbf{x}') = \frac{1}{2}(\mathbf{x}')^T \mathbf{P}_b^{-1}(\mathbf{x}') + \frac{1}{2}[y'(\mathbf{x}') - y'_{obs}]^T \mathbf{R}^{-1}[y'(\mathbf{x}') - y'_{obs}] \quad (1)$$

in which  $\mathbf{x}' = \mathbf{x} - \mathbf{x}_b$  is the perturbation of the background field  $\mathbf{x}_b$  at  $t_0$ , and

$$\mathbf{y}'_{obs} = \begin{bmatrix} y'_{obs,1} \\ y'_{obs,2} \\ \vdots \\ y'_{obs,S} \end{bmatrix}, \quad (2)$$

$$\mathbf{y}' = y'(\mathbf{x}') = \begin{bmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_S \end{bmatrix}, \quad (3)$$

$$y'_k = y_k(\mathbf{x}_b + \mathbf{x}') - y_k(\mathbf{x}_b) \quad (4)$$

$$y'_{obs,k} = y_{obs,k} - y_k(\mathbf{x}_b), \quad (5)$$

$$y_k = H_k(M_{t_0 \rightarrow t_k}(\mathbf{x})), \quad (6)$$

and

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{R}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{R}_S \end{bmatrix}. \quad (7)$$

Here, the superscript  $T$  in (1) stands for transpose,  $b$  is the background value, index  $k$  denotes the observation time,  $S$  is the total number of observational times in the assimilation window,  $H_k$  is the observation operator, and matrices  $\mathbf{P}_b$  and  $\mathbf{R}_k$  are the background and observational error covariances, respectively.

According to eqs. (3 and 4), the 4D ensemble of the observation perturbations (OPs)  $\mathbf{y}':\mathbf{y}'_1, \mathbf{y}'_2, \dots, \mathbf{y}'_N$  are constructed by using the non-linear observation operator  $H_k$ , the non-linear forecast model  $M_{t_0 \rightarrow t_k}$ , and the initial conditions  $\mathbf{x}':\mathbf{x}'_1, \mathbf{x}'_2, \dots, \mathbf{x}'_N$ . The POD of this OPs matrix  $\mathbf{y}'$  yields

$$\mathbf{y}'_r \mathbf{y}' = \mathbf{V} \Lambda^2 \mathbf{V}^T, \quad (8)$$

in which  $\mathbf{V}$  matrix is composed of orthogonal vectors and  $\Lambda^2$  is corresponding eigenvalues in POD. We define  $\mathbf{P}_y$  the POD-transformed OPs:

$$\mathbf{P}_y = \mathbf{y}' \mathbf{V}. \quad (9)$$

Assuming nearly linear relationship between the OPs and the model (or state) perturbations (MPs), we define  $\mathbf{P}_x$  the POD-transformed MPs:

$$\mathbf{P}_x = \mathbf{x}' \mathbf{V}. \quad (10)$$

The optimal solution  $\mathbf{x}'_a$  and its corresponding model equivalent  $\mathbf{y}'_a$  in observation space can be expressed by the linear combinations of the POD-transformed MPs and OPs, respectively, as follows:

$$\mathbf{x}'_a = \mathbf{P}_x \boldsymbol{\beta}, \quad (11)$$

and

$$\mathbf{y}'_a = \mathbf{P}_y \boldsymbol{\beta}. \quad (12)$$

Substituting (11–12) and the ensemble background covariance  $\mathbf{P}_b = \frac{\mathbf{P}_x \mathbf{P}_x^T}{N-1}$  into (1), the control variable becomes  $\boldsymbol{\beta}$  instead of  $\mathbf{x}'$ , so the control variable is expressed explicitly in the cost function:

$$J(\boldsymbol{\beta}) = \frac{1}{2}(N-1) \cdot \boldsymbol{\beta}^T \mathbf{P}_x^T (\mathbf{P}_x^T)^{-1} (\mathbf{P}_x)^{-1} \mathbf{P}_x \boldsymbol{\beta} + \frac{1}{2} \times (\mathbf{P}_y \boldsymbol{\beta} - \mathbf{y}'_{obs})^T \mathbf{R}^{-1} (\mathbf{P}_y \boldsymbol{\beta} - \mathbf{y}'_{obs}) \quad (13)$$

in which  $\mathbf{P}_x^T$  is the transpose of  $\mathbf{P}_x$ . Through simple calculations (see Tian et al., 2011 for more details), the solution for  $\mathbf{x}'_a$  is simplified into the following form:

$$\mathbf{x}'_a = \mathbf{P}_x \left[ (N-1)\mathbf{I} + \mathbf{P}_y^T \mathbf{R}^{-1} \mathbf{P}_y \right]^{-1} \mathbf{P}_y^T \mathbf{R}^{-1} \mathbf{y}'_{obs}. \quad (14)$$

One unique advantage embedded in the final 4D optimal solution (14) is that we can obtain the final analytical results at any time step  $t_k$  in the assimilation window if we replace  $\mathbf{x}'(t_0)$  in eq. (10) by  $\mathbf{x}'(t_k)$  and substitute  $\mathbf{x}_a(t_k)$ ,  $\mathbf{x}_b(t_k)$  for  $\mathbf{x}_a(t_0)$ ,  $\mathbf{x}_b(t_0)$  in eq. (14), respectively, which thus facilitates the implementation of data assimilation.

## 2.2. Observation operator for PRAS

The observation operator for radial velocity  $V_r$  is

$$V_r = u \frac{x - x_r}{r} + v \frac{y - y_r}{r} + (w - V_{Tm}) \frac{z - z_r}{r}, \quad (15)$$

in which  $(u, v, w)$  are the Cartesian velocity components from the model integration, and  $V_{Tm}$  is the mass-weighted terminal velocity of the precipitation, which is given by Sun and Crook (1997, 1998):

$$V_{Tm} = 5.40 a q_r^{0.125}. \quad (16)$$

The variable  $r$  is the distance between a model grid point  $(x, y, z)$  and the radar location  $(x_r, y_r, z_r)$ . The quantity  $a$  is a correction factor defined as

$$a = \left( \frac{P_0}{\bar{P}} \right)^{0.4}, \quad (17)$$

in which  $\bar{P}$  is the base state pressure, and  $P_0$  is the pressure at the ground.

The relationship between  $Z$  and  $q_r$  is derived analytically by assuming the Marshall–Palmer distribution of raindrop size (Sun and Crook, 1997); the observation operator for the reflectivity  $Z$  is

$$Z = 43.1 + 17.5 \log(\rho q_r), \quad (18)$$

in which the reflectivity is in units of dBZ,  $\rho$  is air density, and  $q_r$  is rainwater mixing ratio.

Consequently, the model-predicted radial velocity  $V_{r,k}$  and radar reflectivity  $Z_k$  at the time step  $t_k$ , which are used to compare with the observed vector  $\mathbf{y}_{obs,k} = \begin{pmatrix} V_{r,k}^{obs} \\ Z_k^{obs} \end{pmatrix}$  in the 4DVar cost function on the radar coordinates, can be calculated using the model outputs  $q_r$ ,  $(u, v, w)$ , and other model states from eqs. (15 and 18), respectively. Thus, eqs. (15 and 18) provide a link between the radar observation variables and the model states, which acts as the observation operator  $H_k$  in eq. (6) in PRAS.

For PRAS, the control variables are the three wind components ( $u$ ,  $v$ , and  $w$  winds), the water vapour mixing ratios ( $q_v$ ), and the rainwater mixing ratios ( $q_r$ ).

## 2.3. Localisation schemes

As an ensemble-based assimilation scheme, the localisation strategy was essential to ameliorate the spurious long-range correlations resulting from the limited number of ensemble members (Houtekamer and Mitchell, 1998). In PRAS, we use the following filter function as the horizontal correlation model (Gaspari and Cohn, 1999):

$$\rho_h = c_0(d_x/d_{x,0}) \cdot c_0(d_y/d_{y,0}), \quad (19)$$

and

$$c_0(r) = \begin{cases} -\frac{1}{4}r^5 + \frac{1}{2}r^4 + \frac{5}{8}r^3 - \frac{5}{3}r^2 + 1, & 0 \leq r \leq 1 \\ \frac{1}{12}r^5 - \frac{1}{2}r^4 + \frac{5}{8}r^3 + \frac{5}{3}r^2 - 5r + 4 - \frac{2}{3}r^{-1}, & 1 < r \leq 2 \\ 0, & 2 < r \end{cases}, \quad (20)$$

in which  $d_x$  and  $d_y$  are the zonal and meridional distances between the observation point and model grid point,  $d_{x,0}$  and  $d_{y,0}$  are zonal and meridional localisation Schur radii, respectively. Following Zhang et al. (2004b), vertical localisation is performed using the correlation function,

$$\rho_v(\Delta \log P) = \frac{1}{1 + 5 \times (\Delta \log P)^2} \quad (21)$$

in which  $\Delta \log P$  is the distance between two vertical levels in  $\log P$  space, and  $P$  is air pressure.

So, with the localisation scheme, the final PODen4DVar analysis solution  $\mathbf{x}_a$  in eq. (14) is formulated using the Schur product as follows:

$$\mathbf{x}_a = \mathbf{x}_b + (\rho_h \cdot \rho_v) \circ \{ \mathbf{P}_x \left[ (N-1)\mathbf{I} + \mathbf{P}_y^T \mathbf{R}^{-1} \mathbf{P}_y \right]^{-1} \mathbf{P}_y^T \mathbf{R}^{-1} \mathbf{y}'_{obs} \}, \quad (22)$$

in which the Schur product of two matrices having the same dimension is denoted by  $A = B \circ C$  and consists of the element-wise product such that  $a_{ij} = b_{ij} \cdot c_{ij}$ .

## 3. Evaluation experiments for PRAS

In this section, PRAS will be comprehensively assessed through a group of single-observation experiments, a group of comparison experiments between PRAS and LRAS, a group of cycling assimilation OSSEs and real-data assimilation experiments with the WRF model (Skamarock et al., 2008).

### 3.1. WRF configuration

Earlier versions of the WRF model have been widely used in atmospheric data assimilation (e.g. Barker et al., 2004; Xiao et al., 2007, 2008) and other applications. The WRF model version 3.3 in this study is used as the forecasting model in PRAS.

The experiments were conducted over a grid mesh of  $400 \times 400$  with a grid spacing of 6 km; 27 layers are present in the vertical direction. The main physical components of the WRF model used in our experiments include the Rapid Radiative Transfer Model (RRTM) long-wave radiation scheme, the Dudhia shortwave radiation scheme, the Yonsei University (YSU) PBL scheme, WRF single-moment six-classes microphysics (WSM6), and the Noah

LSM land scheme. Particularly, the cumulus parameterisation scheme is excluded from our experiments. The NCEP Final (FNL) Operational Model Global Tropospheric Analyses (<http://dss.ucar.edu/datasets/ds083.2/matrix.html>) were used as the first-guess field and boundary conditions in the experiments.

### 3.2. Single-observation experiments

In the single-observation experiments, following the experimental design by Wang et al. (2013), assimilation of a single radar reflectivity observation with PRAS is carried out to evaluate PRAS performance including in-depth analysis of background error covariance formed in PRAS.

*3.2.1. Experimental design.* To guarantee the linear assumption between the OPs and MPs and incorporate the radar observations as efficiently as possible, the assimilation window length is carefully chosen to be 1 hour according to Wang et al. (2012). A case with a convective system on 08 July 2010 is chosen to perform the single-observation experiments; correspondingly, the period from 0000 UTC 08 to 0100 UTC 08 July 2010 is chosen to be the assimilation window in PRAS. The same case and assimilation window are also chosen for the comparison experiments and the real-data assimilation experiments.

In the single-observation experiments, the analysis time is 0100 UTC 08 July 2010 (i.e. the end of the assimilation window), and correspondingly the single-observation is inserted at the analysis time. The background fields over the assimilation window are provided by the forecasting fields that are initialised using the FNL data from 1800 UTC 07 July 2010 with the WRF model (i.e. the background run introduced in the following), in which the first 6 hours are taken as the spin-up period. Figure 1 shows the background field on the eleventh model level (approximately 700 hPa) at the analysis time. Around the location

(30.60°N, 116.16°E), there exists a convection area (Fig. 1b) with the corresponding features of cyclonic convergence and southeast–northwest updraft (Fig. 1a). Then the single observation is placed on the grid point located at (30.60°N, 116.16°E) on the eleventh model level, marked by the red star in Figure 1. The simulated reflectivity for the background field is 47.15 dBZ at the single-observation location. Similar to the experimental design by Wang et al. (2013), two experiments are then conducted by assimilating two single-reflectivity observations of different magnitude. In the first experiment, the observed reflectivity (57.15 dBZ) is assumed to be larger than the value for the background field (hereafter SOE1, i.e. the background field underestimated the convective system). In contrast, we choose a smaller observed reflectivity (37.15 dBZ, also compared with the simulated value for the background field) in the second experiment (hereafter SOE2, i.e. the background field overestimated the convective system). Obviously, the two experiments are thus designed to investigate whether PRAS can effectively intensify (SOE1) or weaken (SOE2) the convection system fit to the observation by assimilating the single observed reflectivity.

For PRAS, to form the MPs  $\mathbf{x}'(x'_1, x'_2, \dots, x'_N)$  and OPs  $\mathbf{y}'(y'_1, y'_2, \dots, y'_N)$ , similar to Tian et al. (2014), in each assimilation cycle, the similar background run and sampling run are carried out. The background run is used to form the background fields, and the sampling run is used to form the ensemble samples by a 4D moving sampling strategy.

Considering the higher complexity and uncertainty for the convective system, differing slightly from Tian et al. (2014), two sampling runs are conducted to form more ensemble samples. More specifically, two 8-hour model integrations from 2000 UTC 07 July to 0400 UTC 08 July 2010 with two different initial conditions are conducted, outputting model results every 6 minutes (exactly consistent with the radar scanning frequency in China). The two different initial conditions at 2000 UTC 07 2010 July are the forecasting fields initiated from 1200 UTC 07 July and

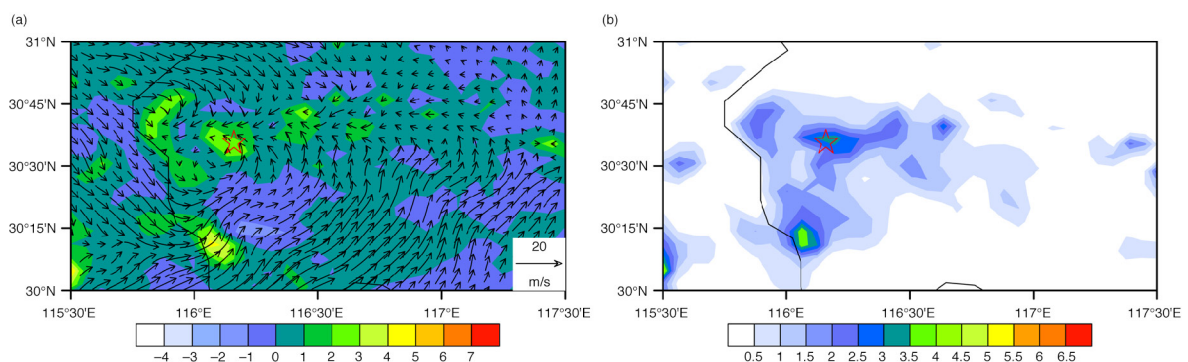


Fig. 1. The first guess field at 0010 UTC 08 July 2010. (a) Horizontal wind and vertical velocity (units: m/s) and (b) rainwater-mixing ratio (units: g/kg). The location of the single observation is marked by the red star.

0600 UTC 07 July 2010 with FNL data, respectively. Subsequently, similar to Wang et al. (2010), in the single-observation experiments, a 4D moving sampling strategy is adopted for the two 8-hour forecasts, as a result, 142 pairs of 4D model states samples are formed as follows:

$$\begin{aligned} & (x_{2000\text{UTC},07}, x_{2006\text{UTC},07}, \dots, x_{2100\text{UTC},07}), \\ & (x_{2006\text{UTC},07}, x_{2012\text{UTC},07}, \dots, x_{2106\text{UTC},07}), \\ & \dots \\ & (x_{0300\text{UTC},08}, x_{0306\text{UTC},08}, \dots, x_{0400\text{UTC},08}). \end{aligned}$$

which indicates there are 714D model states samples produced by one 8-hour forecast. For the single-observation experiments, the corresponding simulated-observation vectors are thus obtained by applying the observation operator to the 142 ensemble model samples [eq. (6)].

Additionally, the innovation vector [ $y'_{obs}$  in eq. (1)] is also easily prepared, the algorithm assumes the observation errors are additive, unbiased, and Gaussian, 1 dBZ are chosen as the standard deviation of reflectivity observations. Furthermore, a localisation radius with 16 grid points is adopted in PRAS (the same localisation radius is adopted in the following experiments). Consequently, PRAS is implemented to obtain the analysis states. It should be noted here, a group of sensitivity experiments conducted in advance demonstrate that PRAS performs best in case of all the orthogonal vectors with non-zero eigenvalues being retained to perform transformation for Eqs. (8 and 9). As a result, in all the following experiments, we used all the orthogonal vectors with non-zero eigenvalues for PRAS.

**3.2.2. Results.** The increments for the horizontal wind, vertical velocity, and rainwater mixing ratio on the eleventh model level from experiment SOE1 are shown in Fig. 2. Around the single observation, the horizontal wind increments display obvious cyclonic convergence (Fig. 2a), and the values of increment for the vertical velocity (Fig. 2a) and the rainwater mixing ratio (Fig. 2b) are entirely positive. The cross section of the increment of the rainwater mixing ratio along with that of the vertical velocity from SOE1 (Fig. 4a) also shows that the rainwater mixing ratio and updraft increase. These all demonstrate that the convective system is intensified, which indicates that PRAS can effectively assimilate radar observations. Second, as shown in Fig. 2, due to the localisation scheme adopted in this study, non-zero analysis increments are confined to the local patch with reduced magnitudes at increasing distance from the observation. As a result, the spurious increments resulting from the limited number of ensemble members won't appear far from the observation; meanwhile, the horizontal correlation model [eq. (20)] also alleviates the sharp discontinuity at the edge of the local patch (Xu et al., 2011). Finally, the increments all show an anisotropic distribution with a flow-dependent pattern, which displays a southeast–northwest orientation in agreement with that for the simulated convective system (Fig. 1), especially the horizontal wind increment. It displays cyclonic convergence and corresponds well to the horizontal wind for the background field, which demonstrates that the  $\mathbf{B}$  matrix formed in PRAS is flow dependent. This means that PRAS is able to spread observation information spatially in accordance with the flow-dependent error structure, which is important in analysis of rapidly evolving systems (Huang et al., 2009), such as convective systems, hurricanes, and cyclones.

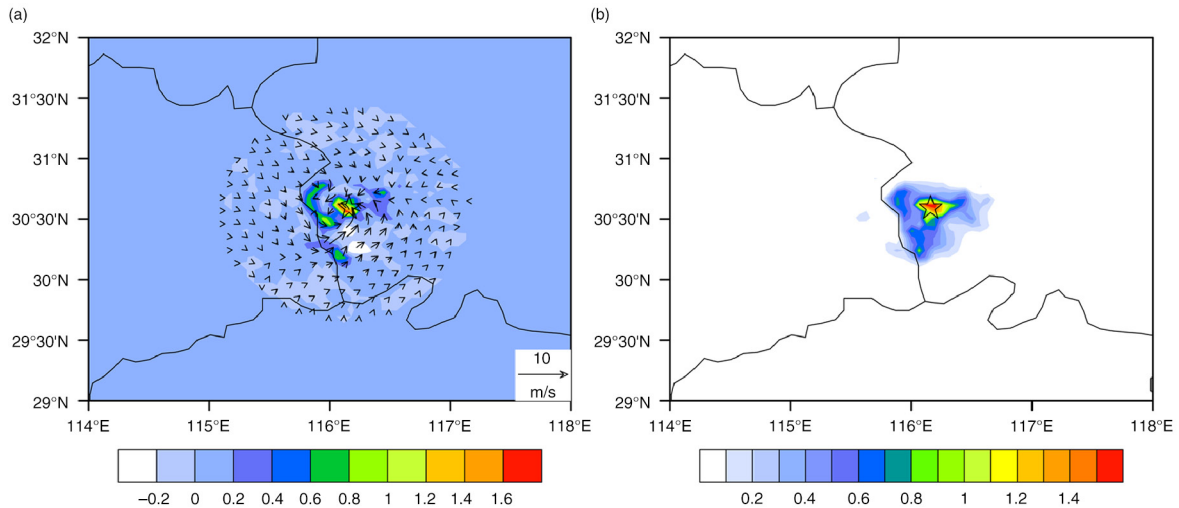


Fig. 2. The increments at 0100 UTC 08 July 2010 on the eleventh model level from experiment SOE1. (a) Wind vector and vertical velocity (shaded, units: m/s) and (b) rainwater-mixing ratio (units: g/kg). The location of the single observation is marked by the black star.



Figure 3 depicts the horizontal increments from experiment SOE2. As expected, the convective system in this experiment is weakened by the negative increment of the rainwater mixing ratio (Fig. 3b). The anticyclonic divergence and downdraft (Fig. 3a) also indicates that the convection is weakened; the cross section of the increment of the rainwater mixing ratio along with that of the vertical velocity from SOE2 (Fig. 4b) shows that the convection is weakened, as suggested by the decreases in the rainwater mixing ratio and updraft. The localisation scheme also confines the increments to a local patch with reduced magnitudes at increasing distances from the observation. The increments also display a flow-dependent structure.

In summary, the single-observation experiments demonstrate that PRAS is able to effectively assimilate the radar observation and spread the observation information spatially with a flow-dependent pattern in a local patch.

### 3.3. Comparison OSSEs between PRAS and LRAS

Next, PRAS performance will be further assessed by comparing with LRAS through the following OSSEs. OSSEs are considered the best benchmark tests to assess a data assimilation methodology or system, because it can provide both the ‘true’ state and the corresponding ‘observations’. Three experiments are designed including the control experiment (CTRL), the LRAS assimilation experiment (LRASAss), and PRAS assimilation experiment (PRASAss), which are introduced as follows.

*3.3.1. Experimental design.* For the comparison experiments, the analysis time is 0100 UTC 08 July 2010, and the LRAS is also accomplished at the same time step for

comparison with PRAS. The period from 0000 UTC 08 to 0100 UTC 08 July 2010 is chosen to be the assimilation window for PRASAss, and therefore three simulated-observation times in the assimilation window are inserted at 0000 UTC, 0030 UTC, and 0100 UTC 08 July 2010, respectively. The simulated observations are sampled at the site of the real observations (Wuhan station).

The CTRL run is initialised using the first guess field at 1800 UTC 07 July 2010 with FNL data and is integrated for 7 hours, in which the first 6 hours are taken as the spin-up period.

To conduct the LRASAss and PRASAss experiments, the assumed ‘true’ atmospheric state over the assimilation time window should first be provided. To create the ‘observations’ corresponding to the assumed ‘true’ atmospheric state, a random perturbation, which is derived from the background error covariance of the WRF 3DVar data assimilation system using an approach similar to the initial ensemble generating method in Houtekamer and Mitchell (2005), is added to the background field for CTRL at 2300 UTC 07 July 2010. Then, with the perturbed background field, an assumed ‘truth run’ is integrated for 2 hours. Thus, the simulated radar observations at 0000 UTC, 0030 UTC, and 0100 UTC 08 July 2010 are generated.

For PRASAss, to obtain the MPs and OPs, the same sampling runs are carried out as for the single-observation experiments. The 4D model state samples over the assimilation window can be formed by the 4D moving sampling strategy similar to that proposed by Wang et al. (2010); its corresponding 4D simulated-observation vector over the assimilation window is thus calculated by the observation operator. The 4D model state samples and simulated-observation vector samples are obtained using the two 8-hour forecasting by the 4D moving sampling strategy (Tian et al., 2014).

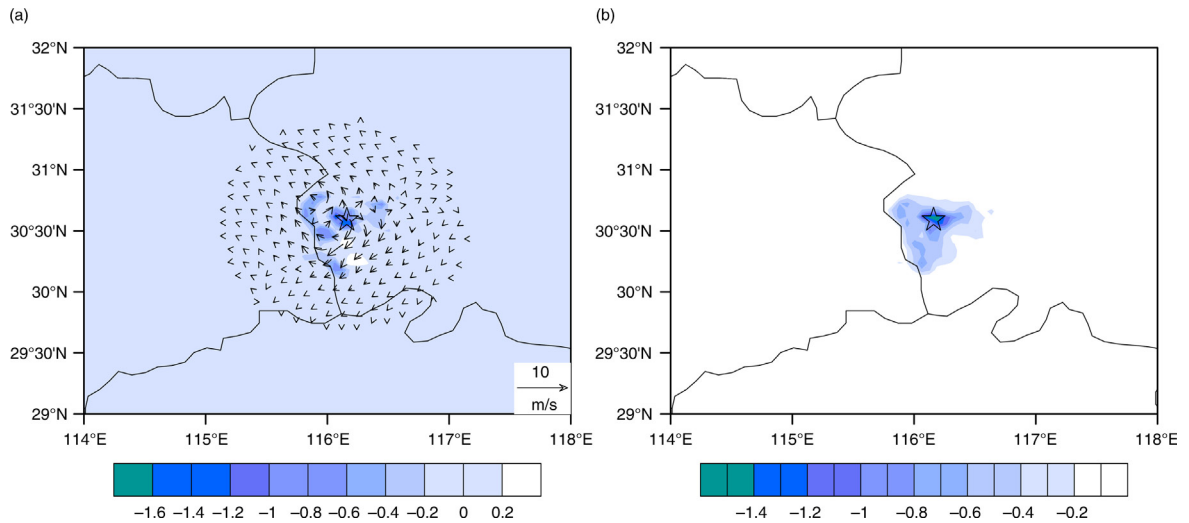


Fig. 3. As in Fig. 2, but for the increments from the experiment SOE2.

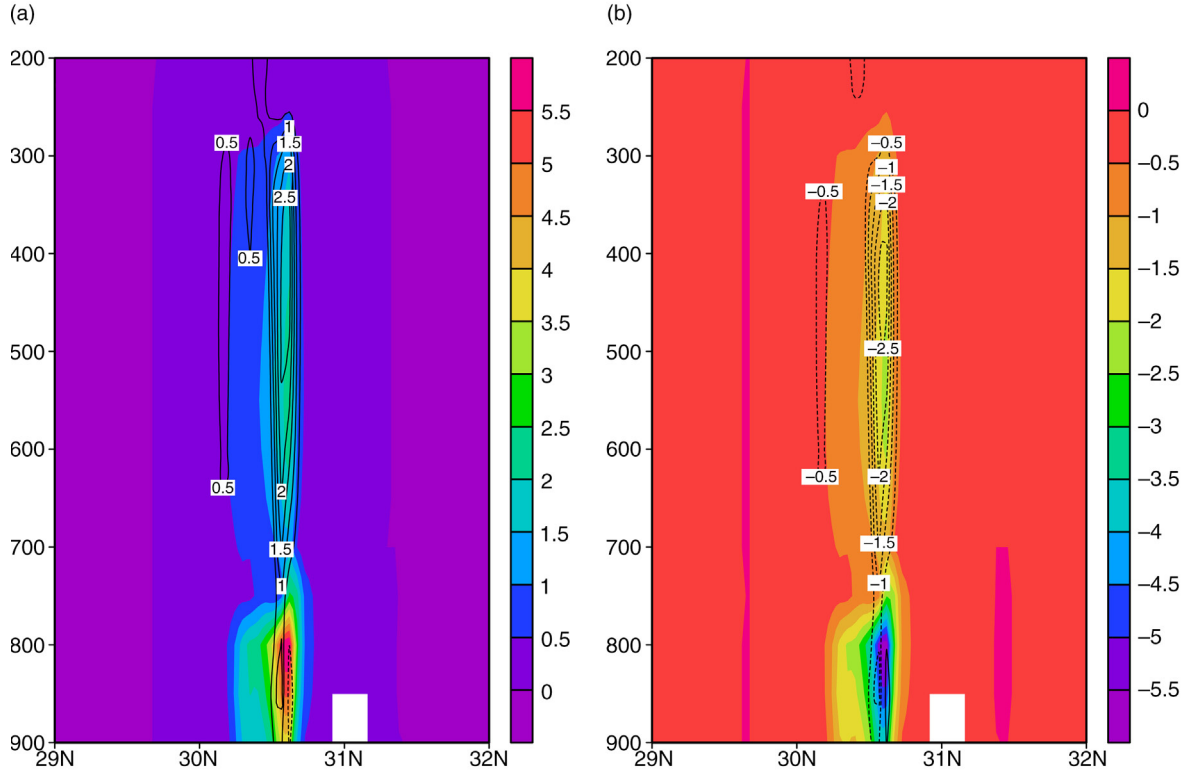


Fig. 4. The south–north vertical section of increment of rainwater-mixing ratio (colour; g/kg) and vertical velocity (contour; m/s) from (a) SOE1 and (b) SOE2. The vertical sections are along 30.60°N.

Additionally, the 4D innovation vector  $[\mathbf{y}'_{obs}]$  in eq. (1) over the assimilation window is also easily prepared for LRASAss and PRASAss. The algorithm assumes the observation errors are additive, unbiased, and Gaussian, 1 m/s and 1 dBZ are chosen as the standard deviation of radial velocity and reflectivity observations, respectively. Consequently, after implementing the two radar data assimilation schemes (LRAS and PRAS), the analysis field at 0100 UTC 08 July 2010 can be obtained for LRASAss and PRASAss.

**3.3.2. Results.** In Fig. 5, comparing with ‘true’ atmospheric state created by ‘truth run’, the vertical profiles of the root-mean-square errors (RMSEs) of some basic model variables are shown from the CTRL, LRASAss, and PRASAss results for the four basic model variables (i.e. u winds, v winds, w winds, and rainwater mixing ratio). The RMSEs from both LRASAss and PRASAss assimilation results are substantially smaller than those from CTRL (blue squares) at most model levels, especially for u winds and v winds, which means the two assimilation schemes can effectively assimilate radar observations to improve the accuracy of the atmospheric state. Furthermore, the RMSEs for PRASAss (green circles) are slightly smaller

than those from LRASAss (red triangles), which indicates that PRAS performance is superior to that of LRAS. In conclusion, the comparison experiments demonstrate that PRAS can effectively assimilate the radar observations and produce analysis results better than (or at least as good as) LRASAss.

### 3.4. Cycling assimilation OSSEs of PRAS

The following OSSEs for cycling assimilation are specially designed to study the feasibility of the PRAS to routinely assimilate radar data in operational context.

**3.4.1. Experimental design.** To conduct cycling assimilation OSSEs, the assumed ‘true’ atmospheric state should be provided. Similar with the above comparison experiments in Section 3.3, an assumed ‘truth run’ is first integrated for 4 hours with a perturbed background field at 2300 UTC 07 July 2010, thus, the simulated radar observations with an interval of 30 minutes from 0000 UTC to 0300 UTC 08 July 2010 are generated, which are located at the site of the real observations (Wuhan station). The length of assimilation window is also 1 hour, as per the experiments above.



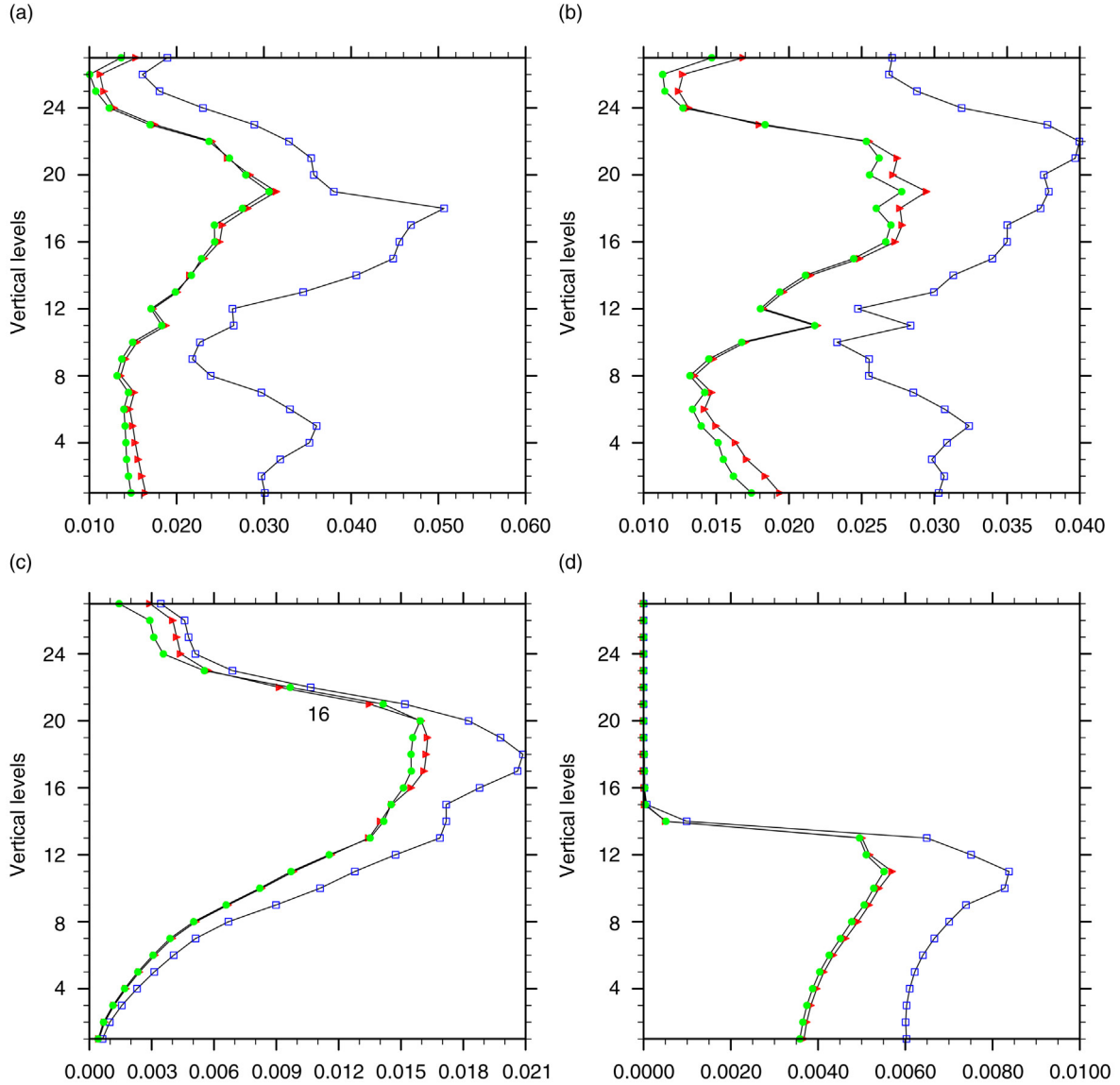


Fig. 5. The vertical profiles of RMSEs in (a)  $u$  winds (m/s), (b)  $v$  winds (m/s), (c)  $w$  winds (m/s) and (d) rainwater-mixing ratio (g/kg) from the CTRL, LRASAss, and PRASAss results. The results of CTRL, LRASAss and PRASAss analysis field are denoted by blue squares, red triangles and green circles, respectively.

Four experiments are then designed including the CTRL experiment, Cycle-1 experiment, Cycle-2 experiment, Cycle-3 experiment. The analysis time in all the four experiments is 0300 UTC 08 July 2010. The experiments are introduced simply as follows.

The CTRL experiment is initialised using the first guess field at 1800 UTC 07 July 2010 with FNL data and is integrated for 9 hours, in which the first 6 hours are taken as the spin-up period.

The Cycle-1 experiment assimilates the simulated radar data from 0000 UTC to 0100 UTC 08 July 2010, and then

integrated to 0300 UTC 08 July 2010, which means Cycle-1 just accomplishes one-cycle assimilation.

The Cycle-2 experiment assimilates the simulated radar data from 0000 UTC to 0100 UTC 08 July 2010 and from 0100 UTC to 0200 UTC 08 July 2010, and thus implements two-cycle assimilations.

The Cycle-3 experiment assimilates the simulated radar data from 0000 UTC to 0100 UTC 08 July 2010, from 0100 UTC to 0200 UTC 08 July 2010 and from 0200 UTC to 0300 UTC 08 July 2010, and conducts three-cycle assimilations as a result.

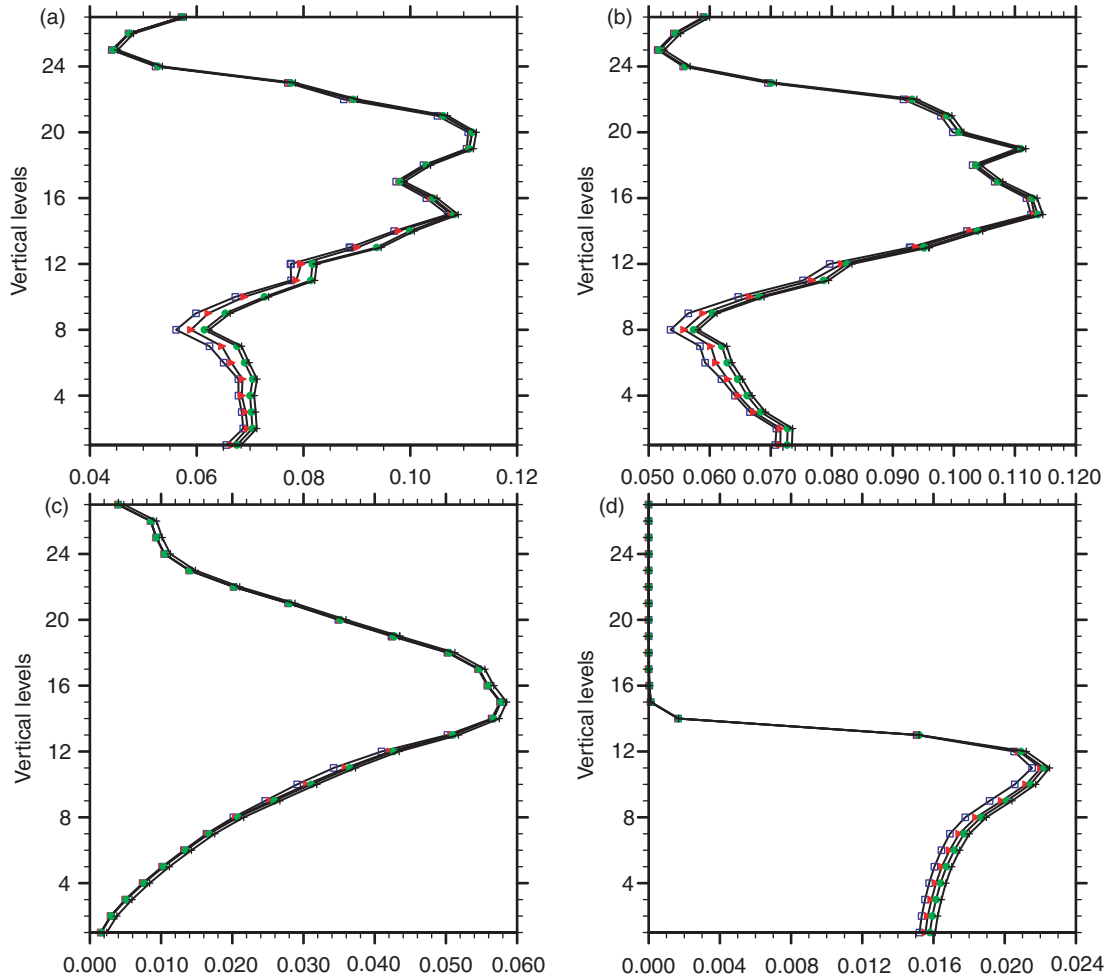


Fig. 6. The vertical profiles of RMSEs in (a)  $u$  winds (m/s), (b)  $v$  winds (m/s), (c)  $w$  winds (m/s) and (d) rainwater-mixing ratio (g/kg) from the CTRL, Cycle-1, Cycle-2 and Cycle-3 results. The results of CTRL, Cycle-1, Cycle-2 and Cycle-3 analysis field are denoted by black plus, red triangles, green circles, and blue squares, respectively.

**3.4.2. Results.** In Fig. 6, the vertical profiles of the RMSEs of some basic model variables are shown from the CTRL, Cycle-1, Cycle-2 and Cycle-3 results for the four basic model variables (i.e.  $u$  winds,  $v$  winds,  $w$  winds, and rainwater mixing ratio). The RMSEs from Cycle-1 (green circles), Cycle-2 (red triangles) and Cycle-3 (blue squares) are substantially smaller than those from CTRL (black plus) at most model levels, which illustrates that the PRAS can effectively assimilate radar observations to improve the accuracy of the atmospheric state. Furthermore, the RMSEs for Cycle-3 are smallest in all experiments, and the RMSEs for Cycle-2 are smaller than those for Cycle-1, which demonstrates PRAS can gradually improve the accuracy of the atmospheric state by increasing cycling assimilation times.

In conclusion, the cycling assimilation experiments further demonstrate that PRAS can effectively assimilate the radar

observations and gradually improve the accuracy of the atmospheric state by increasing cycling assimilation times. In the following, we will continue to apply PRAS to assimilate real radar observations to exploit the potential of PRAS in real numerical weather prediction (NWP).

### 3.5. Real-data assimilation of PRAS

**3.5.1. Rainfall event.** To illustrate PRAS performance in real-data assimilation, a case was selected that occurred on July 08 2010 with abrupt heavy rainfall in central China. The observed 24-hour accumulated precipitation (Fig. 8a) indicates that the heaviest precipitation occurred in the eastern Hubei Province, where the centre intensity reached 287 mm. The heavy rainfall also occurred at the borders between Hubei and Hunan Provinces and between Anhui and Jiangxi Provinces. In most areas, the amount of

rainfall exceeded 90 mm. The area shown in Fig. 8a is set as the verification area in this study.

*3.5.2. Description of the experiments.* To evaluate the effect of assimilating real radar observations with PRAS on the accuracy of precipitation forecasting, two experiments, including the simulated experiment (Exp-Sim) without radar data assimilation and the assimilated data experiment (PRASAss) with radar data assimilation by PRAS, are conducted.

The analysis time is 0000 UTC 08 July 2010, and the first guess field at the analysis time is obtained from the 6-hour forecasting field initiated from 1800 UTC 07 July 2010, in which the first 6 hours are taken as the spin-up period. Then, Exp-Sim is carried out by 24-hour integration from 0000 UTC 08 to 0000 UTC 09 July 2010 without radar observation assimilation.

For PRASAss, to obtain the MPs and OPs for PRAS, the same two sampling runs as those in the comparison OSSEs are also carried out. Consequently, PRAS is implemented to obtain the analysis state at 0000 UTC 08 July 2010. Finally, with such an analysis field, PRASAss is completed by a 24-hour WRF model integration from 0000 UTC 08 to 0000 UTC 09 July 2010.

*3.5.3. Radar data and preprocessing.* Radar data obtained through the China Meteorological Administration (CMA) Wuhan radar are used in this study. The CMA Wuhan radar is located in Hubei Province (114.38°E, 30.52°N) at an altitude of 135.7 m. The radar has nine elevation scans with elevation angles of 0.5°, 1.5°, 2.4°, 3.4°, 4.3°, 6.0°, 9.9°, 14.6°, and 19.5°.

The raw data of Wuhan radar have substantially high resolution, which is typically updated every 6 minutes or so with a 250-m resolution for radial velocity and 1-km resolution for reflectivity. To utilise the radar data effectively in PRAS, several data preprocessing procedures, including data de-noising (Zhang and Wang, 2006; Jiang et al., 2009), erasing folded velocity and clutter, data thinning, and observation error estimation (Tong and Xue, 2005), are performed; specifically, the reflectivity value is set to 5 dBZ wherever the raw reflectivity data is no greater than a threshold of 5 dBZ or where the reflectivity is flagged as missing in the raw data (Lopez and Bauer, 2007; Aksoy et al., 2009, 2010), then, the radar data processed are converted to the model grid space for PRAS.

*3.5.4. Analysis of sensitivity to the initial ensemble size.* The initial ensemble size and the selection of the initial ensemble members are important issues in using ensembles

of discrete forecasting to approximate the evolution of an initial probability distribution in a model (Anderson, 1996), which directly influence the result of PRAS. If the ensemble size is large enough to capture the initial probability distribution, the ensemble size should be selected to be as small as possible to alleviate the computational costs. To study the sensitivity of PRAS to the initial ensemble size, the size of the initial ensemble is set as 100, 120 and 142 in this paper. Figure 7 shows the increments for the wind and rainwater-mixing ratio on the eleventh model level. In general, the increment distribution for the three different ensemble sizes is similar, except that in the northern area, the increments for the ensemble size of 100 are a little different from those for the sizes of 120 and 142. However, the increments for the ensemble size of 120 are almost the same as those for the ensemble size of 142, which indicates that the ensemble size of 120 is enough to represent the initial probability distribution; hence, it is reasonable to take 120 as the initial ensemble size in the following study.

*3.5.5. Comparisons of precipitation predictions with and without radar assimilation.* In this section, the predicted precipitation for Exp-Sim and PRASAss are compared to assess the performance of PRAS in real-data assimilation.

With the 24-hour accumulated precipitation from 0000 UTC 08 to 0000 UTC 09 July 2010 for the observed data, the predicted Exp-Sim and PRASAss are shown in Fig. 8. Comparing these with the observation (Fig. 8a), Exp-Sim (Fig. 8b) almost failed to predict the heavy rainfall that occurred in eastern Hubei Province, which is the maximum precipitation area in this rainfall event. Furthermore, in the southeastern Hubei Province, Exp-Sim predicted considerable spurious heavy rainfall, with the intensity for the maximum precipitation centre exceeding 240 mm. However, after assimilating the radar observations with PRAS, PRASAss predicted the heavy rainfall in the eastern Hubei Province well (Fig. 8c), which is very close to the observed precipitation, except that the location is slightly southeasterly. At the same time, PRASAss eliminated the spurious heavy rainfall predicted by Exp-Sim in southeastern Hubei Province. PRASAss had some ability to predict the heavy rainfall at the border between Hubei and Hunan Provinces (Fig. 8a), but Exp-Sim nearly failed to predict it. The heavy rainfall at the border between Anhui and Jiangxi Provinces (Fig. 8a) was not well predicted by either experiment, but PRASAss eliminated the spurious heavy rainfall predicted by Exp-Sim in northeastern Jiangxi Province.

To objectively and quantitatively verify the precipitation forecast, a Structure, Amplitude, and Location (SAL) quantitative assessment method proposed by Wernli et al. (2008) is adopted. The precipitation forecast is verified

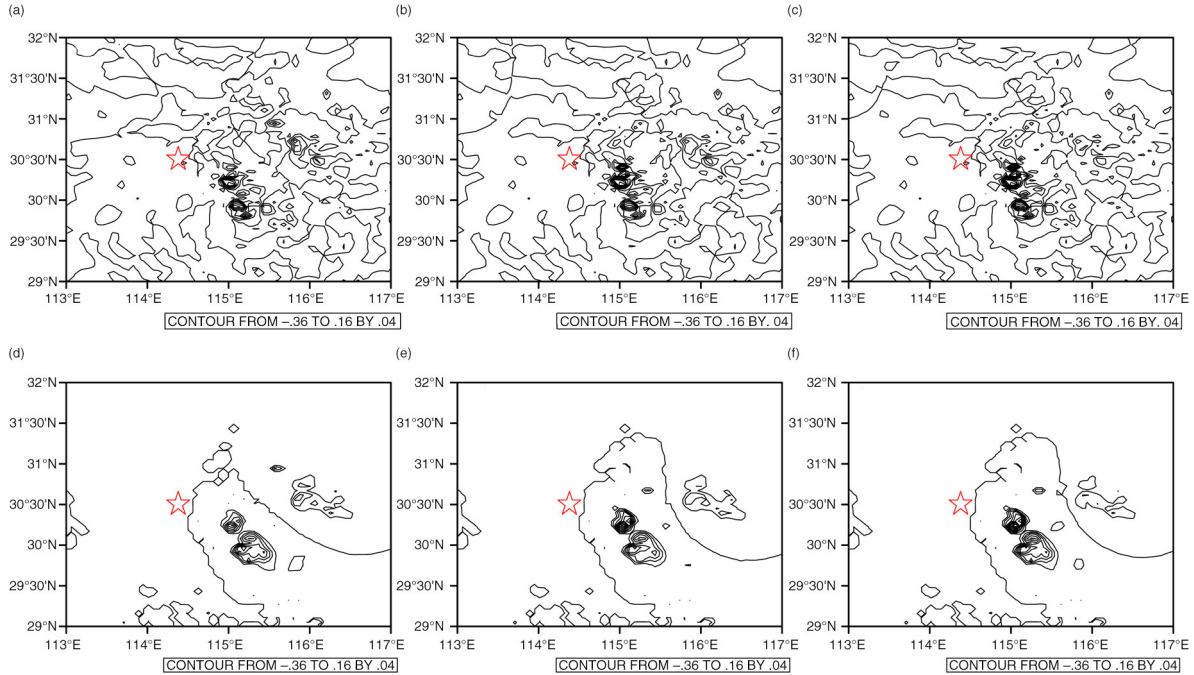


Fig. 7. The increments for vertical velocity and rainwater-mixing ratio on the eleventh model level: (a, b and c) vertical velocity (m/s) for the sizes 100, 120 and 142, respectively; (d, e, and f) rainwater-mixing ratio (g/kg) for the sizes 100, 120 and 142, respectively. The red star represents the location of the Doppler radar in Wuhan.

according to the structure, the amplitude, and the location. The smaller the absolute values of  $S$ ,  $A$ , and  $L$  are, the better is the forecast. The SAL verification results for Exp-Sim and PRASAss are shown in Fig. 9.

The  $L$  represents the effect of the precipitation location on the forecast and is the most important parameter in SAL. Obviously, the value of  $L$  for PRASAss is smaller than that for Exp-Sim, which indicates that the ability to forecast the location of the rainband can be improved through assimilating the radar observations with PRAS. The negative value of  $A$  indicates that the predicted precipitation for Exp-Sim and PRASAss both are weaker than the observed precipitation; the absolute value of  $A$  for PRASAss is slightly larger than that for Exp-Sim, which means the intensity of the precipitation predicted by Exp-Sim is stronger than that predicted by PRASAss. We may speculate that this is due to the fact that more spurious precipitation patterns are predicted by Exp-Sim. The positive value of  $S$  for Exp-Sim indicates that the scale of the predicted precipitation is greater than that of the observed precipitation, and the negative value of  $S$  for PRASAss indicates that the scale of the predicted precipitation is smaller than that of the observed precipitation. The predicted precipitation for PRASAss mainly concentrates at the border between Hubei and Anhui provinces.

Obviously, the predicted precipitation for PRASAss is closer to the observation than is that for Exp-Sim, indi-

cating that PRAS can assimilate real radar data to improve precipitation forecasting.

*3.5.6. Effects of radar data assimilation on model initialisation.* Doppler radar observations with rich information about the 3D structures of mesoscale convection are a major data source for initial conditions in high-resolution NWP of severe weather events. The improvements in precipitation forecasting discussed above are the result of the change in the initial condition with PRAS. Therefore, the impact of radar assimilation with PRAS on the initial field will be examined in what follows.

The analysis increments of both the dynamical and microphysical structures at 0000 UTC 08 July 2010 are illustrated in Figs. 10 and 11. In the horizontal direction (Fig. 10), in the heavy rainfall area in eastern Hubei Province, on the eighth model level ( $\sigma=0.85$ , approximately 850 hPa), the intensified northerly wind in the north and the intensified southerly wind in the south strengthen wind convergence. Meanwhile, the vertical velocity and the rainwater-mixing ratio are also intensified in the vertical direction (Fig. 11). The increments show that the northerly winds north of  $30.20^{\circ}\text{N}$  and the southerly winds south of  $30.20^{\circ}\text{N}$  are significantly enhanced below 750 hPa (Fig. 11a), leading to intensification of the low-level convergence. Figure 10b shows that the rainwater-mixing ratio is significantly

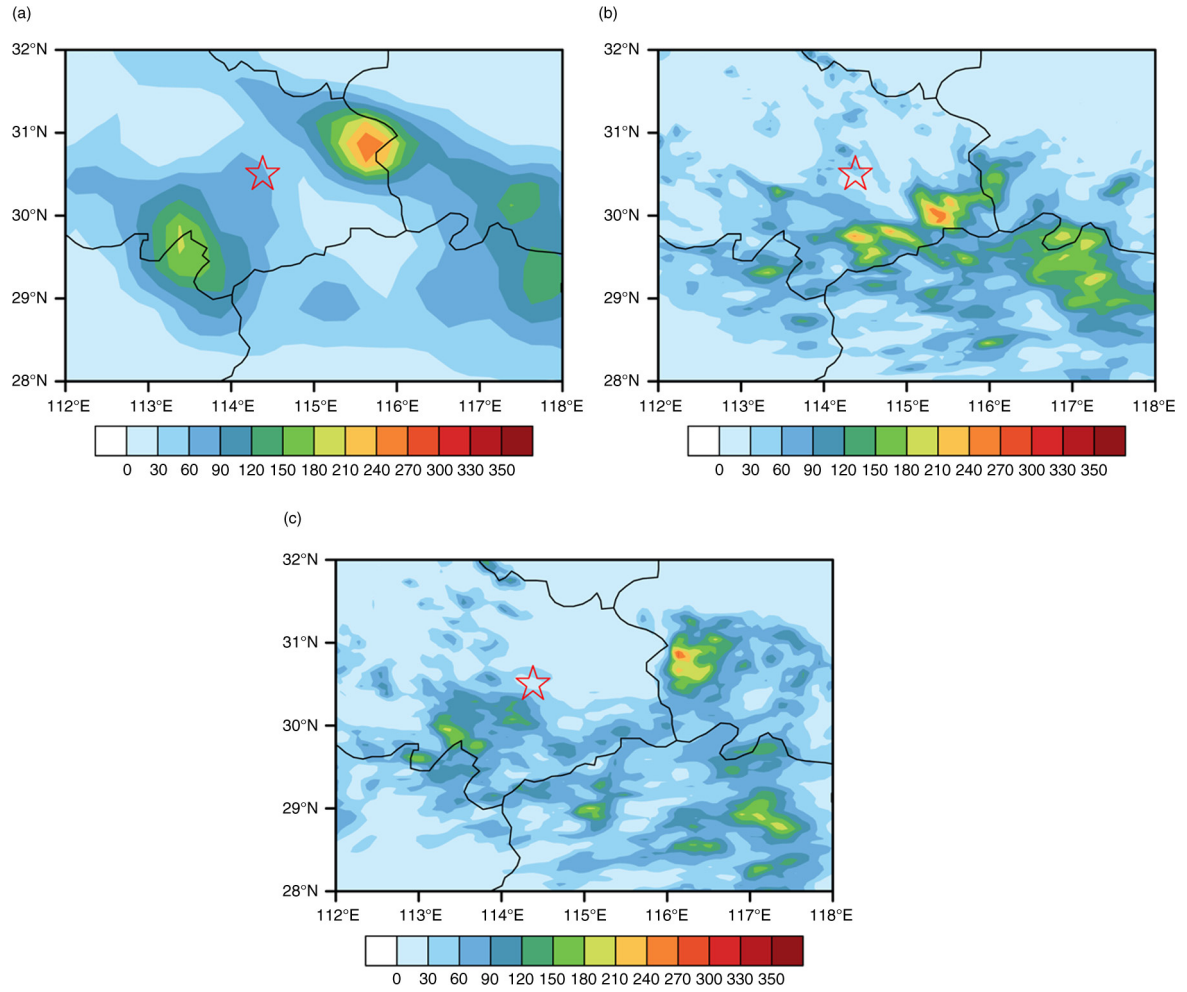


Fig. 8. The 24-hour accumulated precipitation for (a) the observed (b) Exp-Sim and (c) PRASAss from 0000 UTC 08 to 0000 UTC 09 July 2010 (units: mm); the red star represents the location of the Doppler radar in Wuhan.

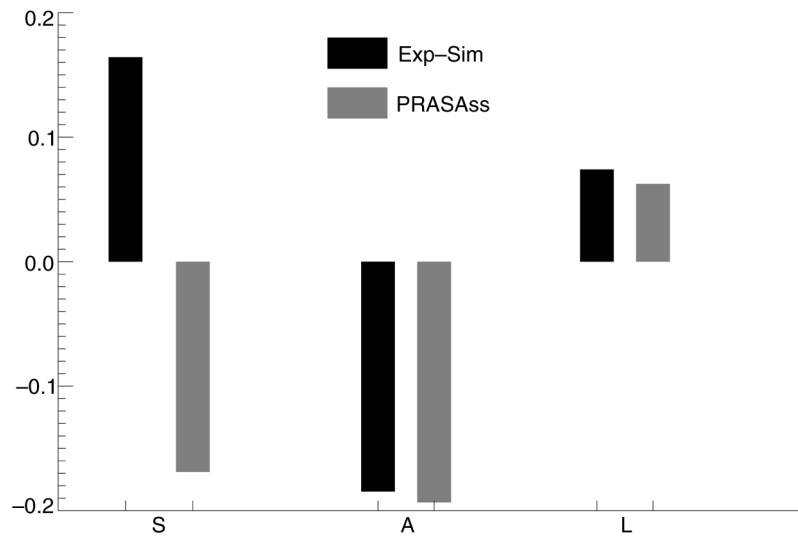


Fig. 9. The SAL verification of 24-hour forecast precipitation for Exp-Sim and PRASAss.

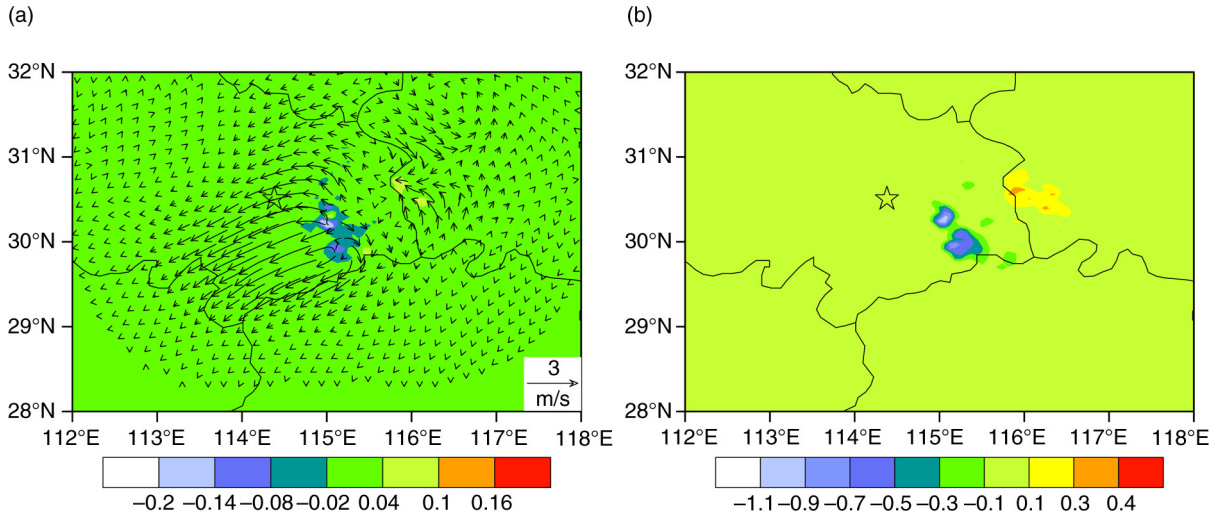


Fig. 10. The increments for wind field and rainwater-mixing ratio on the eighth model level. (a) Horizontal wind (vector, units: m/s) and vertical velocity (shaded, units: m/s); (b) rainwater-mixing ratio (units: g/kg). The black star represents the location of the Doppler radar in Wuhan.

enhanced, especially in the low-level atmosphere, and the vertical velocity is also enhanced from 900 to 250 hPa, which increase the precipitation intensity; this explains why the heavy rainfall in eastern Hubei Province was improved by PRASAss. At the same time, in southeastern Hubei

Province (Fig. 10), the horizontal wind increment intensifies wind divergence, and the vertical velocity and rainwater-mixing ratio are also weakened in the vertical direction (not shown), which eliminated the spurious heavy rainfall predicted by Exp-Sim (Fig. 8b). At the border

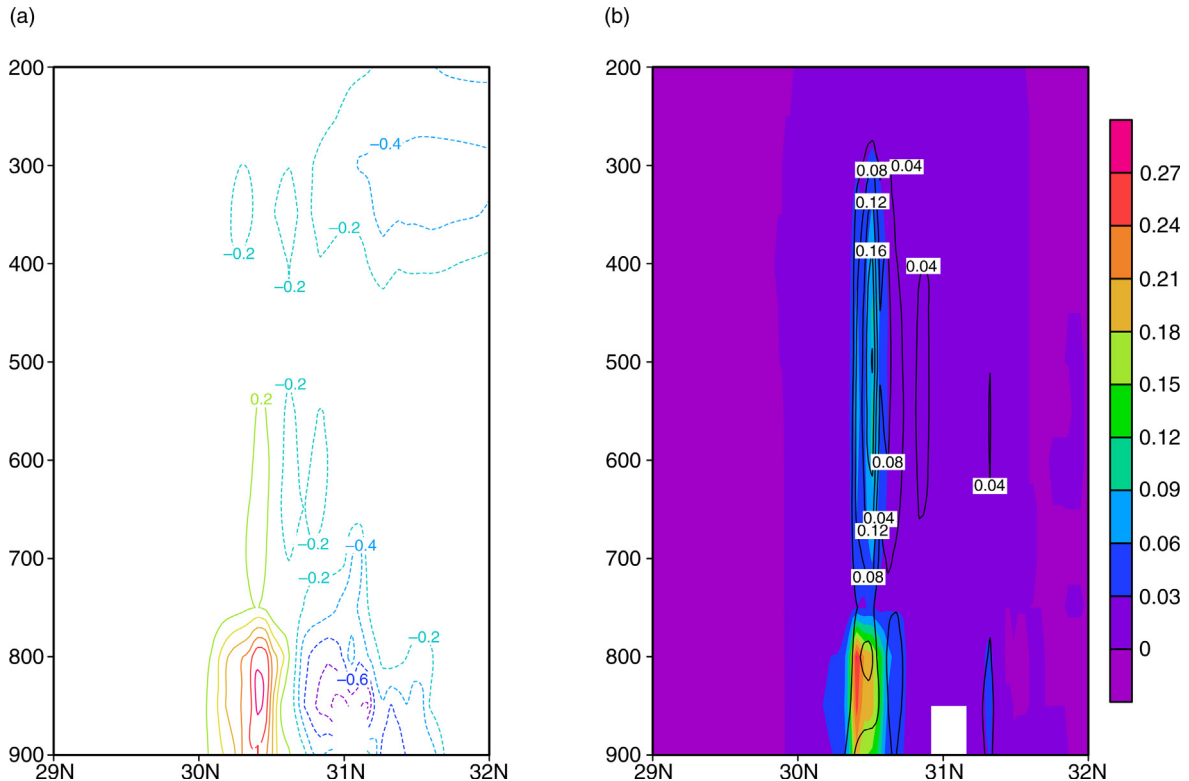


Fig. 11. The south-north vertical section of increment along 116.2°E. (a)  $V$  wind (m/s). (b) rainwater-mixing ratio (colour; g/kg) and vertical velocity (contour; m/s).



between Hubei and Hunan provinces, the predicted precipitation is also strengthened by the wind convergence resulting from the horizontal wind increment (Fig. 10). However, at the border between Anhui and Jiangxi provinces, the horizontal wind, vertical velocity, and rainwater-mixing ratio all change little, so the predicted precipitation is also hardly improved. Meanwhile, the difference between assimilated observed reflectivity and corresponding background reflectivity at 2.4° scan surface is shown in Fig. 12. Obviously, the background field overestimates the convective intensity in the south of 30.50°N. Logically, the convective intensity would be weakened if radar data could be assimilated successfully. In Figs. 10 and 11, the corresponding analysis increments in the south of 30.50°N indicates that the convective intensity is weakened, which further demonstrates the PRAS can efficiently incorporate radar data information.

In summary, assimilating radar observations with PRAS can properly add mesoscale information on the initial field to improve the accuracy of heavy rainfall forecasting.

#### 4. Summary and concluding remarks

The standard 4DVar technique is increasingly used for synoptic and global-scale atmospheric data assimilation at operational NWP centres around the world due to the two following attractive features: (1) the physical model can provide a temporal smoothness constraint, and (2) it has the ability to simultaneously assimilate the observation data at multiple times in an assimilation window. Unfortunately, the need for an adjoint model in 4DVar makes it considerably difficult to implement 4DVar-based radar data assimilation for improving model initialisation. Furthermore, the adoption of a static background error

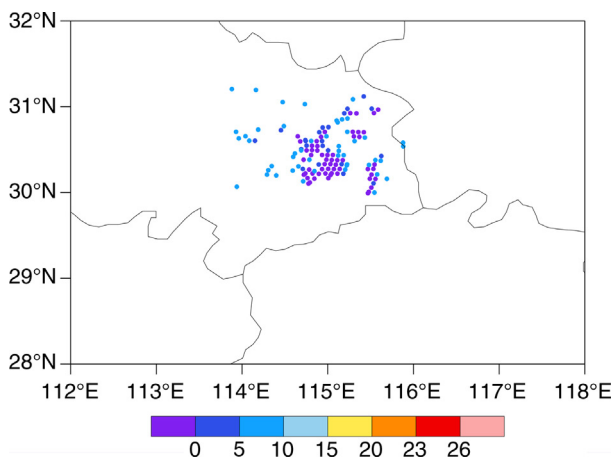


Fig. 12. The difference between assimilated observed reflectivity and corresponding background reflectivity. All fields are computed at 2.4° scan surface.

covariance in the standard 4DVar potentially degrades its assimilation performance. To address this issue, a hybrid method, namely the POD-based ensemble 4-D variational assimilation (POD<sub>EN</sub>4DVar) method, was proposed by Tian et al. (2008, 2011) based on ensemble forecasting and the POD techniques. Its robust performance and significant strengths (compared with the usual EnKF, 4DVar, and several other popular ensemble-based assimilation methods) have been demonstrated and highlighted by idealised models through observing system simulation experiments (OSSEs) (Tian et al., 2008, 2011; Tian and Xie, 2012). In this study, we further utilise this approach to build a radar data assimilation scheme (PRAS) for exploiting its wider applications.

To assess the ability of PRAS to assimilate radar observations, the single-observation experiments, the OSSEs, and the real-data assimilation experiments were carried out. The results from the single-observation experiments show that PRAS is effective in assimilating radar observations and that the background error covariance formed in PRAS is flow dependent, demonstrating the validity of PRAS in assimilating radar observations. In the Comparison OSSEs, the performance of PRAS was compared with of the LETKF, and the results demonstrate that PRAS can assimilate radar observations at least as good LETKF can. In the cycling assimilation OSSEs, the results demonstrate the PRAS can effectively assimilate the radar observations and gradually improve the accuracy of the atmospheric state by increasing cycling assimilation times. In the real-data assimilation experiments, a heavy rainfall case over central China was selected to assess PRAS; after the radar observations from Wuhan station were assimilated by PRAS, the initial dynamic and micro-physical fields were improved so as to improve the accuracy of heavy rain forecasting. These encouraging results presented here show that PRAS provides a promising tool for radar data assimilation.

As the radar observations in this study came from single radar, which only covers a limited area, the multiple radar observations are expected to be assimilated in PRAS; at the same time, the other observations, such as satellite observations that cover a larger space, are also expected to be assimilated in PRAS in the future.

#### 5. Authors' contributions

The work presented here was carried out in collaboration between all authors. Xiangjun Tian designed the framework of the POD<sub>EN</sub>4DVar radar data assimilation system (PRAS). Bin Zhang, Xiangjun Tian and Feng Chen co-coded the PRAS. Bin Zhang and Yangchun Zhang co-worked on associated data collection and their processing. Bin Zhang, Xiangjun Tian, Jianhua Sun, Lifeng Zhang and Shenming Fu co-designed the experiments, discussed analyses,

interpretation and presentation. Bin Zhang drafted the manuscript. All authors have contributed to, seen and approved the manuscript.

## 6. Acknowledgements

We thank Dr. Aiguo Dai at NCAR for his helpful comments and suggestions. This work was partially supported by the National High Technology Research and Development Program of China (Grant No. 2013AA122002, Grant No. 2012AA091801), the Knowledge Innovation Program of the Chinese Academy of Sciences (Grant No. KZCX2-EW-QN207), the National Natural Science Foundation of China (Grant Nos. 41375063). Two anonymous reviewers are thanked for helpful comments and suggestions that helped to reformulate the PRAS.

## References

- Anderson, J. L. 1996. Selection of initial conditions for ensemble forecasts in a simple perfect model framework. *J. Atmos. Sci.* **53**(1), 22–36. DOI: 10.1175/15200469(1996)053<0022:SOICFE>2.0.CO;2.
- Aksoy, A., Dowell, D. C. and Snyder, C. 2009. A multicasel comparative assessment of the ensemble Kalman Filter for assimilation of radar observations. Part I: Storm-scale analyses. *Mon. Weather Rev.* **137**, 1805–1824.
- Aksoy, A., Dowell, D. C. and Snyder, C. 2010. A multicasel comparative assessment of the ensemble Kalman Filter for assimilation of radar observations. Part II: Short-range ensemble forecasts. *Mon. Weather Rev.* **138**, 1273–1292.
- Barker, D. M., Huang, W., Guo, Y.-R., Bourgeois, A. and Xiao, X. N. 2004. A three-dimensional variational data assimilation system for MM5: implementation and initial results. *Mon. Weather Rev.* **132**, 897–914.
- Caumont, O., Ducrocq, V., Wattrelot, É., Jaubert, G. and Pradier-Vabre, S. 2010. 1D+ 3DVar assimilation of radar reflectivity data: a proof of concept. *Tellus A.* **62**, 173–187.
- Caya, A., Sun, J. and Snyder, C. 2005. A comparison between the 4DVar and the ensemble Kalman filter techniques for radar data assimilation. *Mon. Weather Rev.* **133**, 3801–3094.
- Cheng, H., Jardak, M., Alexe, M. and Sandu, A. 2010. A hybrid approach to estimating error covariances in variational data assimilation. *Tellus A.* **62**, 288–297.
- Courtier, P., Thepaut, J. N. and Hollingsworth, A. 1994. A strategy for operational implementation of 4DVar using an incremental approach. *Q. J. Roy. Meteorol. Soc.* **120**, 1367–1387.
- Dawson, D. T. and Xue, M. 2006. Numerical forecasts of the 15–16 June 2002 Southern Plains mesoscale convective system: impact of mesoscale data and cloud analysis. *Mon. Weather Rev.* **134**, 1607–1629.
- Dowell, D. C., Wicker, L. J. and Snyder, C. 2011. Ensemble Kalman Filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: influences of reflectivity observations on storm-scale analyses. *Mon. Weather Rev.* **139**, 272–294. DOI: 10.1175/2010MWR3438.1.
- Evensen, G. 1994. Sequential data assimilation with a nonlinear quasigeostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **99**(C5), 10143–10162.
- Evensen, G. 2004. Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dynam.* **54**, 539–560. DOI: 10.1007/s10236-004-0099-2.
- Gao, J., Xue, M., Wang, Z. and Droegemeier, K. K. 1998. The initial condition and explicit prediction of convection using ARPS adjoint and other retrievals methods with WSR-88D data. Preprints, In: *12th Conference on Numerical Weather Prediction*, Phoenix, AZ, American Meteorological Society, pp. 176–178.
- Gaspari, G. and Cohn, S. E. 1999. Construction of correlation functions in two and three dimensions. *Q. J. Roy. Meteorol. Soc.* **125**, 723–757.
- Houtekamer, P. L. and Mitchell, H. L. 1998. Data assimilation using an ensemble Kalman Filter technique. *Mon. Weather Rev.* **126**, 796–811.
- Houtekamer, P. L. and Mitchell, H. L. 2001. A sequential ensemble Kalman Filter for atmospheric data assimilation. *Mon. Weather Rev.* **129**, 123–137.
- Houtekamer, P. L. and Mitchell, H. L. 2005. Ensemble Kalman Filtering. *Q. J. Roy. Meteorol. Soc.* **131**, 3269–3289.
- Hu, M., Xue, M. and Brewster, K. 2006. 3DVAR and cloud analysis with WSR-88D Level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms: Part I. Cloud analysis and its impact. *Mon. Weather Rev.* **134**, 675–698.
- Huang, X., Xiao, Q. and Barker, D. M. 2009. Four-dimensional variational data assimilation for WRF: formulation and preliminary results. *Mon. Weather Rev.* **137**, 299–314.
- Hunt, B. R., Kostelich, E. J., Ott, E. and Szunyogh, I. 2007. Efficient data assimilation spatiotemporal chaos: a local ensemble transform Kalman Filter. *Physica D.* **230**, 112–126.
- Jiang, Y., Liu, L. P. and Zhuang, W. 2009. Statistical characteristics of clutter and improvements of ground clutter identification technique with Doppler weather radar. *J. Appl. Meteorol. Sci.* **20**(2), 203–213. (In Chinese).
- Kawabata, T., Kuroda, T., Seko, H. and Saito, K. 2011. A cloud-resolving 4DVAR assimilation experiment for a local heavy rainfall event in the Tokyo metropolitan area. *Mon. Weather Rev.* **139**, 1911–1930.
- Lewis, J. M. and Derber, J. C. 1985. The use of the adjoint equation to solve a variational adjustment problem with advective constraints. *Tellus A.* **37**, 309–322.
- Li, X. and John, R. M. 2010. Assimilation of the dual-polarization Doppler radar for a convective storm with a warm-rain radar forward operator. *J. Geophys. Res.* **115**, D1628. DOI: 10.1029/2009JD013666.
- Liu, C., Xiao, Q. and Wang, B. 2008. An ensemble-based four-dimensional variational data assimilation scheme. Part I: Technique formulation and preliminary test. *Mon. Weather Rev.* **136**, 3363–3373.
- Lopez, P. and Bauer, P. 2007. “1D+4DVAR” assimilation of NCEP stage-IV radar and gauge hourly precipitation data at ECMWF. *Mon. Weather Rev.* **135**, 2506–2524.
- Lorenc, A. 2003. The potential of the ensemble Kalman Filter for NWP: a comparison with 4Dvar. *Q. J. Roy. Meteorol. Soc.* **129**, 3183–3203.

- Ly, H. V. and Tran, H. T. 2001. Modeling and control of physical processes using proper orthogonal decomposition. *Math. Comput. Model.* **33**, 223–236.
- Pan, X., Tian, X., Li, X., Xie, Z., Shao, A. and co-authors. 2012. Assimilating Doppler radar radial velocity and reflectivity observations in the weather research and forecasting model by a proper orthogonal decomposition based ensemble three-dimensional variational assimilation method. *J. Geophys. Res.* **117**, D17113. DOI: 10.1029/2012JD017684.
- Pu, Z., Li, X. and Sun, J. 2009. Impact of airborne Doppler radar data assimilation on the numerical simulation of intensity change of Hurricane Dennis (2005) near a landfall. *J. Atmos. Sci.* **66**, 3351–3365.
- Qiu, C., Shao, A., Xu, Q. and Wei, L. 2007. Fitting model fields to observations by using singular value decomposition: an ensemble-based 4DVar approach. *J. Geophys. Res.* **112**, D11105. DOI: 10.1029/2006JD007994.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M. and co-authors. 2008. *A Description of the Advanced Research WRF Version 3*. NCAR Tech. Note TN-475 STR, 113 pp.
- Snyder, C. and Zhang, F. 2003. Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Weather Rev.* **131**, 1663–1677.
- Sun, J. 2005. Initialization and numerical forecasting of a supercell storm observed during STEPS. *Mon. Weather Rev.* **133**, 793–813.
- Sun, J. and Crook, N. A. 1997. Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint: Part I. Model development and simulated data experiments. *J. Atmos. Sci.* **54**, 1642–1661.
- Sun, J. and Crook, N. A. 1998. Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint: Part II. Retrieval experiments of an observed Florida convective storm. *J. Atmos. Sci.* **55**, 835–852.
- Sun, J. and Wang, H. 2013. Radar data assimilation with WRF 4D-Var. Part II: Comparison with 3D-Var for a squall line over the U.S. Great Plains. *Mon. Weather Rev.* **141**, 2245–2264.
- Tian, X. and Xie, Z. 2012. Implementations of a square-root ensemble analysis and a hybrid localization into the POD-based ensemble 4DVar. *Tellus A.* **64**, 18375. DOI: 10.3402/tellusa.v64i0.18375.
- Tian, X., Xie, Z. and Dai, A. 2008. An ensemble-based explicit four-dimensional variational assimilation method. *J. Geophys. Res.* **113**, D21124. DOI: 10.1029/2008JD010358.
- Tian, X., Xie, Z., Dai, A., Jia, B. and Shi, C. 2010. A microwave land data assimilation system: scheme and preliminary evaluation over China. *J. Geophys. Res.* **115**, D21113. DOI: 10.1029/2010JD014370.
- Tian, X., Xie, Z., Dai, A., Shi, C., Jia, B. and co-authors. 2009. A dual-pass variational data assimilation framework for estimating soil moisture profiles from AMSR-E microwave brightness temperature. *J. Geophys. Res.* **114**, D16102. DOI: 10.1029/2008JD011600.
- Tian, X., Xie, Z. and Liu, Y. 2014. A joint data assimilation system (Tan-Tracker) to simultaneously estimate surface CO<sub>2</sub> fluxes and 3-D atmospheric concentrations from observations. *Atmos. Chem. Phys.* **14**, 13281–13293.
- Tian, X., Xie, Z. and Sun, Q. 2011. A POD-based ensemble four-dimensional variational assimilation method. *Tellus A.* **63**, 805–816.
- Tong, M. and Xue, M. 2005. Ensemble Kalman Filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. *Mon. Weather Rev.* **133**, 1789–1807.
- Wang, B., Liu, J., Wang, S., Cheng, W., Liu, J. and co-authors. 2010. An economical approach to four-dimensional variational data assimilation. *Adv. Atmos. Sci.* **27**(4), 715–727. DOI: 10.1007/s00376-009-9122-3.
- Wang, H., Auligné, T. and Morrison, H. 2012. Impact of microphysics scheme complexity on the propagation of initial perturbations. *Mon. Weather Rev.* **140**, 2287–2296.
- Wang, H., Sun, J., Zhang, X., Huang, X. and Auligné, T. 2013. Radar data assimilation with WRF 4D-Var. Part I: System development and preliminary testing. *Mon. Weather Rev.* **141**, 2224–2244.
- Wattrelot, E., Caumont, O. and Mahfouf, J. F. 2014. Operational implementation of the 1D+3D-Var assimilation of radar reflectivity data in the AROME model. *Mon. Weather Rev.* **142**, 1852–1873.
- Wernli, H., Paulat, M., Hagen, M. and Frei, C. 2008. SAL – a novel quality measure for the verification of quantitative precipitation forecasts. *Mon. Weather Rev.* **136**, 4470–4487.
- Weygandt, S. S., Shapiro, A. and Droegemeier, K. K. 2002. Retrieval of model initial fields from single-Doppler observations of a supercell thunderstorm: Part II. Thermodynamic retrieval and numerical prediction. *Mon. Weather Rev.* **130**, 454–476.
- Wu, B., Verlinde, J. and Sun, J. 2000. Dynamical and microphysical retrievals from Doppler radar observations of a deep convective cloud. *J. Atmos. Sci.* **57**, 262–283.
- Xiao, Q., Sun, J., Lee, W. C., Kuo, Y.-H. and Barker, D. M. 2008. Doppler radar data assimilation in KMA's operational forecasting. *Bull. Am. Meteorol. Soc.* **89**, 39–43.
- Xiao, Q., Sun, J., Lee, W. C., Lim, E., Barker, D. M. and co-authors. 2007. An approach of radar reflectivity data assimilation and its assessment with the inland QPF of Typhoon Rusa (2002) at landfall. *J. Clim. Appl. Meteorol.* **46**, 14–22.
- Xu, D., Shao, A. and Qiu, C. 2012. Doppler radar data assimilation with a local SVD-En3DVar method. *Acta Meteorol. Sinica.* **6**, 717–733.
- Zhang, F., Snyder, C. and Sun, J. 2004a. Impacts of initial estimate and observations on the convective-scale data assimilation with an ensemble Kalman Filter. *Mon. Weather Rev.* **132**, 1238–1253.
- Zhang, F. Q., Zhang, M. and Hansen, J. A. 2009. Coupling ensemble Kalman Filter with four dimensional variational data assimilation. *Adv. Atmos. Sci.* **26**(1), 1–8. DOI: 10.1007/s00376-009-0001-8.
- Zhang, H., Xue, J., Zhuang, S., Zhu, G. and Zhu, Z. 2004b. GRAPeS 3D-Var data assimilation system ideal experiments. *Acta Meteorol. Sinica.* **62**, 31–41.
- Zhang, J. and Wang, S. X. 2006. An automated 2D multipass Doppler radar velocity dealiasing scheme. *J. Atmos. Ocean. Technol.* **23**, 1239–1248.
- Zhao, Q., Cook, J., Xu, Q. and Harasti, P. R. 2006. Using radar wind observations to improve mesoscale numerical weather prediction. *Weather Forecast.* **21**, 502–522.