• Original Paper •

System of Multigrid Nonlinear Least-squares Four-dimensional Variational Data Assimilation for Numerical Weather Prediction (SNAP): System Formulation and Preliminary Evaluation

Hongqin ZHANG[1,2], Xiangjun TIAN[*1,2,3], Wei CHENG[4], and Lipeng JIANG[5]

[1]*International Center for Climate and Environment Sciences*, *Institute of Atmospheric Physics*,
Chinese Academy of Sciences, *Beijing* 100029, *China*
[2]*University of Chinese Academy of Sciences*, *Beijing* 100049, *China*
[3]*Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters*,
Nanjing University of Information Science and Technology, *Nanjing* 210044, *China*
[4]*Beijing Institute of Applied Meteorology*, *Beijing* 100029, *China*
[5]*National Meteorological Information Center*, *China Meteorological Administration*, *Beijing* 100081, *China*

ABSTRACT

A new forecasting system—the System of Multigrid Nonlinear Least-squares Four-dimensional Variational (NLS-4DVar) Data Assimilation for Numerical Weather Prediction (SNAP)—was established by building upon the multigrid NLS-4DVar data assimilation scheme, the operational Gridpoint Statistical Interpolation (GSI)−based data-processing and observation operators, and the widely used Weather Research and Forecasting numerical model. Drawing upon lessons learned from the superiority of the operational GSI analysis system, for its various observation operators and the ability to assimilate multiple-source observations, SNAP adopts GSI-based data-processing and observation operator modules to compute the observation innovations. The multigrid NLS-4DVar assimilation framework is used for the analysis, which can adequately correct errors from large to small scales and accelerate iteration solutions. The analysis variables are model state variables, rather than the control variables adopted in the conventional 4DVar system. Currently, we have achieved the assimilation of conventional observations, and we will continue to improve the assimilation of radar and satellite observations in the future. SNAP was evaluated by case evaluation experiments and one-week cycling assimilation experiments. In the case evaluation experiments, two six-hour time windows were established for assimilation experiments and precipitation forecasts were verified against hourly precipitation observations from more than 2400 national observation sites. This showed that SNAP can absorb observations and improve the initial field, thereby improving the precipitation forecast. In the one-week cycling assimilation experiments, six-hourly assimilation cycles were run in one week. SNAP produced slightly lower forecast RMSEs than the GSI 4DEnVar (Four-dimensional Ensemble Variational) as a whole and the threat scores of precipitation forecasts initialized from the analysis of SNAP were higher than those obtained from the analysis of GSI 4DEnVar.

Key words: data assimilation, numerical weather prediction, NLS-4DVar, multigrid, GSI

Article Highlights:

- The establishment of SNAP builds upon the multigrid NLS-4DVar assimilation scheme, the GSI-based observation operators, and the WRF model.
- The multigrid NLS-4DVar framework with fast localization corrects multiscale errors and accelerates iteration solutions.
- By assimilating conventional observations, the performance of SNAP is slightly better than that of GSI 4DEnVar.

---

* Corresponding author: Xiangjun TIAN
  Email: tianxj@mail.iap.ac.cn

1.   Introduction

The accuracy of the initial conditions largely determines the success or failure of numerical weather prediction (NWP). Data assimilation systems can provide accurate initial fields using optimization theory and methods to fully integrate the increasing amount of observations and numerical simulations, further improving NWP (Courtier et al., 1994; Evensen, 1994; Rabier et al., 2000). The four-dimensional variational data assimilation (4DVar) system, which is widely used in operational NWP centers (Lewis and Derber, 1985; Courtier et al., 1994; Rabier et al., 2000; Rosmond and Xu, 2006; Gauthier et al., 2007), has the following attractive features: (1) a numerical forecast model that, as a strong constraint, ensures a physically concordant analysis; (2) the capability to simultaneously assimilate multiple-time and multiple-source observations; and (3) the background error covariance is implicitly evolved by tangent linear and adjoint models over the assimilation window (Courtier et al., 1994; Lorenc, 2003a), beginning with a static one. In the process of minimizing the 4DVar cost function, the adjoint and tangent models are indispensable. However, coding, maintenance, and updating the adjoint/tangent models of the forecast model can be extremely difficult, especially when the forecast model is strongly nonlinear and the physical parameterization scheme includes discontinuities (Xu, 1996). The ensemble Kalman filter (EnKF) data assimilation system (Evensen, 1994, 2003; Houtekamer and Mitchell, 1998) has become increasingly popular due to its relatively simple concept and implementation, as well as the ensemble-estimated flow-dependent background error covariance (Houtekamer and Mitchell, 1998; Anderson, 2001; Houtekamer et al., 2005; Evensen, 2007). Notably, the Canadian Meteorological Center has operationally applied the EnKF-based ensemble forecasting system (Houtekamer and Mitchell, 2005). Nevertheless, it lacks the temporal smoothness constraint of the 4DVar system due to assimilating observations sequentially. Thus, although the 4DVar- and EnKF-based data assimilation systems have their own advantages and disadvantages, together they can be complementary. Great efforts have been made to advance data assimilation methods by coupling 4DVar and EnKF to exploit their strengths and offset their respective weaknesses (Hamill and Snyder, 2000; Lorenc, 2003b). The literature (i) contains many introductions to the development and application of hybrid 4DVar methods (Buehner et al., 2010a, b; Zhang and Zhang, 2012; Clayton et al., 2013; Kuhl et al., 2013; Lorenc, 2013), which solve the analysis increment under the 4DVar assimilation framework requiring adjoint and tangent linear models and partly introduce the ensemble-estimated flow-dependent background error covariance, and (ii) indicates 4DEnVar (Four-dimensional Ensemble Variational) makes the most of the linear assumption between the observation perturbations and the model perturbations to approximate a tangent linear operator and eliminate the dependence on the adjoint and tangent linear models. Therefore, the implementation of 4DEnVar is significantly

simplified (Qiu et al., 2007; Tian et al., 2008; Wang et al., 2010; Tian et al., 2011; Tian and Feng, 2015).

Nonlinear least-squares four-dimensional variational assimilation (NLS-4DVar; Tian and Feng, 2015; Tian et al., 2018) is a distinctive 4DEnVar method that transforms the cost function of 4DEnVar into a nonlinear least-squares problem. NLS-4DVar is solved by Gauss–Newton iteration, which is employed to handle non-quadratic, nonlinear forecast models and observation operators. Similarly, NLS-4DVar uses the ensemble-estimated flow-dependent background error covariance and no longer requires the tangent linear and adjoint models based on the assumption of a linear relationship between the model perturbations and the simulated observation perturbations. It is worth mentioning that Zhang and Tian (2018a) developed an ensemble expanding localization of NLS-4DVar based on an efficient local correlation matrix decomposition approach, which simplifies the complicated localization process and greatly improves the calculation efficiency and assimilation accuracy, granting the NLS-4DVar method great potential for operational application. In addition, it is well known that the multigrid technique is an effective iterative acceleration method for solving linear and nonlinear problems (Briggs et al., 2000) at different grid scales. Introducing the multigrid technique into data assimilation can correct errors at different grid scales (Xie et al., 2005, 2011; Li et al., 2008, 2010, 2013; Zhang and Tian, 2018b). At present, multigrid 3DVar is widely used, but the application of multigrid EnKF or multigrid 4DVar methods is rare. The former is mainly because multigrid EnKF requires ensembles at different resolutions, which incurs a high computational cost. The latter is mainly because the solving process of 4DVar is strongly dependent on adjoint and tangent linear models. Consequently, multigrid 4DVar naturally requires adjoint and tangent linear models at different grid scales with high computational cost and difficulty. Zhang and Tian (2018b) developed an effective multigrid NLS-4DVar that only needs to conduct ensemble simulations at the finest grid. Compared to the standard NLS-4DVar (Tian and Feng, 2015), the computational cost of multigrid NLS-4DVar decreases with higher assimilation accuracy.

The Gridpoint Statistical Interpolation (GSI) analysis system at the National Centers for Environmental Prediction (NCEP) is established in physical space, thus facilitating parallel computing and operational applications (Wu et al., 2002; Kleist et al., 2009), and originates from the operational Spectral Statistical Interpolation system. The GSI system has excellent observation operators and can simultaneously assimilate a variety of observations (including conventional, radar, and satellite observations; Benjamin et al., 2004; Skamarock and Klemp, 2008; Zhu et al., 2013; Pan et al., 2014, 2018; Benjamin et al., 2016). In terms of observation operators, GSI is one of the most advanced and mature analysis systems worldwide. Zhu et al. (2013) borrowed the data-processing and observation operators from GSI to establish a regional EnKF system. Currently, more than 20 conventional observations (including satellite retrievals) and satel-

lite radiance/brightness temperature observations from multiple satellites as well as others (containing global positioning system radio occultations and radar data, etc.) can be assimilated (Hu et al., 2018).

The purpose of this paper is to document the development and verification of the System of Multigrid Nonlinear Least-squares Four-dimensional Variational Data Assimilation for Numerical Weather Prediction (SNAP). The key components of SNAP are the multigrid NLS-4DVar assimilation scheme and the GSI-based data processing and observation operators. SNAP can assimilate all available observations in the operational GSI system. At present, conventional observations are assimilated, and the assimilation of radar and satellite observations are undergoing improvements. Furthermore, SNAP has been fully evaluated by a group of case evaluation experiments and another group of one-week cycling assimilation experiments by assimilating conventional observations.

This paper is organized as follows. Section 2 provides an introduction to SNAP. Section 3 describes the case evaluation experiments for SNAP. The one-week cycling assimilation evaluation experiments are discussed in section 4. A summary and concluding remarks are presented in section 5.

2. SNAP

As noted previously, SNAP is constructed based on the multigrid NLS-4DVar method and the GSI-based data-processing and observation operator module, which uses the dynamic core of the Advanced Research version of the Weather Research and Forecasting model (WRF-ARW) and can assimilate multiple-source observations. Figure 1

presents a flowchart of SNAP. The system runs six-hourly assimilation cycles and the analysis time is at the beginning of the assimilation window. The WRF model is used for analysis–forecast cycles. SNAP has three grid scales, and at a certain grid scale the observation innovations $y_k - H_k(x_k)$ with multi-time ($k = 0, 1, 2, \ldots, S$ is the observation time) are calculated by the GSI-based data-processing and observation operator module, and the NLS-4DVar with an efficient localization scheme solves iteratively to obtain the analysis. Notably, the assimilation analysis of SNAP is in the model space and the analysis variables are the model prognostic variables. Currently, the analysis variables are the horizontal wind u/v, perturbation potential temperature T, perturbation pressure P, and water vapor mixing ratio q. The analysis variables can be added flexibly according to the specific assimilation problems. Additional details about the multigrid NLS-4DVar method are given below.

2.1. Multigrid NLS-4DVar method and covariance localization

The multigrid NLS-4DVar method described by Zhang and Tian (2018b) is used to obtain the analysis. The fundamental principle of multigrid NLS-4DVar is to sequentially minimize the 4DVar cost functions from the coarsest to the finest grid scales to obtain the analysis increment $x'_{(i)}$ ($i = 1, 2, \cdots, n$ is the grid scale) at the ith grid scale (outer iteration), which is solved iteratively with the NLS-4DVar method (inner iteration). Three grid scales are adopted by SNAP, i.e., $n = 3$. If the solution of NLS-4DVar is only at the finest grid scale ($n = 1$), we define it as "SNAP_S". It should be noted that multigrid NLS-4DVar can be recognized as the multi-scale iterative method for NLS-4DVar. In
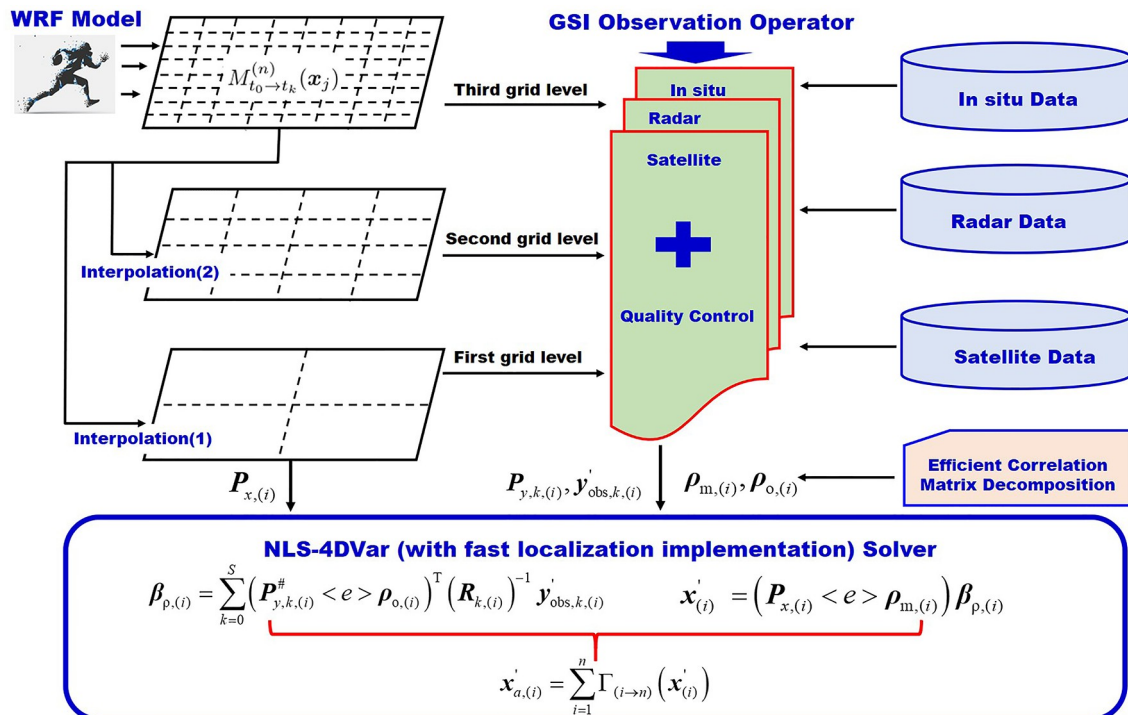

Fig. 1. Framework diagram of SNAP.

the multigrid NLS-4DVar scheme, the background is updated by the analysis of the previous grid level, which is the same as the usual iterative scheme adopted by traditional 4DVar. In such an iterative scheme, the initial value of the ith iterative step is updated by the analysis of the $(i-1)$th iterative step.

As an advanced 4DEnVar method, NLS-4DVar (Tian and Feng, 2015; Tian et al., 2018) assumes that the optimal analysis increment $x' = x - x_b$ ($x_b \in \mathbb{R}^{n_x}$ is the background, n_x is the dimension of the state variables) can be characterized as a linear combination of the initial ensemble perturbation P_x; that is, $x' = P_x\beta$, $P_x = (x'_1, x'_2, \cdots, x'_N)$, $\beta = (\beta_1, \beta_2, \cdots, \beta_N)^T$, $x'_j = x_j - x_b$, and x_j ($j = 1, 2, \cdots, N$) is the jth ensemble and the background error covariance B [$= P_x P_x^T/(N-1)$] is estimated by short-term forecast ensembles. By substituting the above assumptions and minimizing the incremental form of the 4DVar cost function by the Gauss−Newton iteration method (Dennis and Schnabel, 1996), Tian and Feng (2015) obtained

$$\beta^{(l)} = \beta^{(l-1)} + \sum_{k=0}^{S} \left(P_{y,k}^*\right)^T L'_k\left(P_x\beta^{(l-1)}\right) +$$
$$\sum_{k=0}^{S} \left(P_{y,k}^\#\right)^T \left[y'_{\text{obs},k} - L'_k\left(P_x\beta^{(l-1)}\right)\right] , \quad (1)$$

where:

$$\left(P_{y,k}^*\right)^T = -(N-1)\left[(N-1)I + \sum_{k=0}^{S} \left(P_{y,k}\right)^T \left(R_k\right)^{-1} P_{y,k}\right]^{-1}$$
$$\cdot \left[\sum_{k=0}^{S} \left(P_{y,k}\right)^T P_{y,k}\right]^{-1} \left(P_{y,k}\right)^T , \quad (2)$$

$$\left(P_{y,k}^\#\right)^T = \left[(N-1)I + \sum_{k=0}^{S} \left(P_{y,k}\right)^T \left(R_k\right)^{-1} P_{y,k}\right]^{-1} \left(P_{y,k}\right)^T \left(R_k\right)^{-1} . \quad (3)$$

Where $P_{y,k} = \left(y'_{1,k}, y'_{2,k}, \cdots, y'_{N,k}\right)$, and $y'_{j,k} = L'_k\left(x'_j\right)$. R_k is the observation error covariance matrix. $L'_k(x') = L'_k(x_b + x') - L_k(x_b)$, $L_k = H_k M_{t_0 \to t_k}$, $y'_{\text{obs},k} = y_{\text{obs},k} - L_k(x_b)$. H_k is the observation operator of GSI. $M_{t_0 \to t_k}(\cdot)$ is the nonlinear forecast model integration from t_0 to t_k. $y_{\text{obs},k} \in \mathbb{R}^{n_{y,k}}$ are the observations at t_k, $n_{y,k}(\sum_{k=0}^{S} n_{y,k} = n_y)$ is the dimension of $y_{\text{obs},k}$ and n_y is the total number of observations in the assimilation window. k is the observation time, $S + 1$ is the total number of observation times in the assimilation window. $l = 1, 2, \cdots, l_{\max}$ is the number of iterations and $l_{\max}$ is the maximum iteration number. The optimal analysis increment is:

$$x' = P_x\beta^{(l_{\max})} . \quad (4)$$

According to Zhang and Tian (2018b), the cost function can generally reach the minimization convergence standard after three iterations. Because the multigrid technique

can speed up the convergence and SNAP has three grid scales, the maximum iterations of each grid scale is $l_{\max} = 1$. In fact, the above formulas [Eqs. (1) and (4)] are the solution of NLS-4DVar without the localization scheme.

However, due to the limited number of ensembles N, the ensemble-estimated B contains spurious correlations and further results in spurious analysis increments. In general, the localization process (Houtekamer and Mitchell, 1998; Hamill et al., 2001; Tian et al., 2018; Zhang and Tian, 2018a) can alleviate this problem. Tian and Feng (2015) considered spatial distance−based correlations between grid points and observation sites, which computes correlations repeatedly between the grid points and observation sites with higher calculation cost, especially at higher model resolutions and with a massive number of observations. Tian et al. (2018) proposed an equivalent fast localization scheme based on ensemble expanding localization, which is necessary to construct moderation functions to act on the model and observation space respectively. The analysis increment is as follows:

$$\beta_\rho^{(l)} = \beta_\rho^{(l-1)} + \sum_{k=0}^{S} \left(\rho_{\text{o},k} <e> P_{y,k}^*\right)^T L'_k\left(P_{x,\rho}\beta_\rho^{(l-1)}\right) +$$
$$\sum_{k=0}^{S} \left(\rho_{\text{o},k} <e> P_{y,k}^\#\right)^T \left[y'_{\text{obs},k} - L'_k\left(P_{x,\rho}\beta_\rho^{(l-1)}\right)\right] , \quad (5)$$

$$x' = P_{x,\rho}\beta_\rho^{(l_{\max})} , \quad (6)$$

where $P_{x,\rho} = (\rho_m <e> P_x)$, $\rho_m \in \mathbb{R}^{n_x \times r}$ and $\rho_{\text{o},k} \in \mathbb{R}^{n_{y,k} \times r}$ are the moderation functions generated by the efficient local correlation matrix decomposition approach (Zhang and Tian, 2018a). The subscripts "m" and "o" stand for the model and observation spaces, respectively. For the definition of the $(<e>)$ operator, $P_{x,\rho} = (\rho_m <e> P_x) = \left(\rho_{m,1}x'_1, \cdots, \rho_{m,1}x'_N, \rho_{m,2}x'_1, \cdots, \rho_{m,2}x'_N, \cdots, \rho_{m,r}x'_1, \cdots, \rho_{m,r}x'_N\right)$. It should be noted that the two localization schemes in Tian and Feng (2015) and Tian et al. (2018) are theoretically equivalent. However, the latter adopted the efficient local correlation matrix decomposition approach, which only uses a few truncated modes and does not need to repeatedly compute the correlations between the grid points and the observation sites. This greatly simplifies the complex localization process, especially when the model resolution and observations are increased (Zhang and Tian, 2018a; Tian et al., 2018).

2.2. *Initial ensemble generation and ensemble update*

The initial ensemble perturbations are generated by the random state variable method (Tian and Zhang, 2019, step2 b and c in section 2.2; Zhang, 2019, appendix) and then added to the background to obtain the initial ensembles. The random state variable includes the singular value decomposition and the random orthogonal matrix (Evensen, 2007). In the real assimilation system, the state variables are usually composed of multiple state variables, such as the u/v wind

components, perturbation potential temperature T, perturbation pressure P, and water vapor mixing ratio q. To reduce the calculation cost and programming difficulty, all state variables were determined individually. The cost of calculation and the difficulty of programming were reduced, and the differences in units and magnitude among variables were avoided, along with increasing ensemble spreads.

SNAP uses the covariance relaxation of Zhang et al. (2009) to inflate the background error covariance, which needs not only the prior perturbation, but also the posterior perturbation as follows:

$$P_{x,\mathrm{new}} = (1-\alpha)P_{x,\mathrm{a}} + \alpha P_x . \tag{7}$$

In this paper, α is equal to 0.8 (Zhang et al., 2009). The posterior ensemble perturbation matrix $P_{x,\mathrm{a}}$ (subscript a for analysis) is updated by the multiplication of P_x by a transform matrix T (Hunt et al., 2007; Tian and Xie, 2012):

$$T = [(N-1)P^*]^{1/2} , \tag{8}$$

$$P^* = \left[(N-1)I + \sum_{k=0}^{S}\left(P_{y,k}\right)^{\mathrm{T}}R_k^{-1}\left(P_{y,k}\right)\right]^{-1} . \tag{9}$$

2.3. *Verification techniques*

(1) Root-mean-square error (RMSE) and correlation coefficient (CC):

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(f_i - o_i)^2} , \tag{10}$$

$$\mathrm{CC} = \frac{\sum_{i=1}^{n}\left(f_i - \bar{f}\right)(o_i - \bar{o})}{\sqrt{\sum_{i=1}^{n}\left(f_i - \bar{f}\right)^2}\sqrt{\sum_{i=1}^{n}(o_i - \bar{o})^2}} , \tag{11}$$

where o_i is the observation, $\bar{o}$ is the mean of all observations, f_i is the forecast value, $\bar{f}$ is the mean of the forecast values, and n is the number of observation sites used for validation.

(2) Precipitation threat score (TS) and equitable threat score (ETS):

$$\mathrm{TS} = \frac{a}{a+b+c} , \tag{12}$$

$$\mathrm{ETS} = \frac{a-a_{\mathrm{r}}}{a+b+c-a_{\mathrm{r}}} , \tag{13}$$

$$a_{\mathrm{r}} = \frac{(a+b)(a+c)}{n} , \tag{14}$$

where a is the number of correct hits, b is the number of false alarms, c is the number of misses, d is the number of occasions that both forecast and observations are under a specific threshold, and $n = a + b + c + d$, as shown in Tables S1 and S2 in electronic supplementary material (ESM).

(3) The relative percentage improvement (RPI, unit: %) for the RMSE is computed as follows:

$$\mathrm{RPI}_{\mathrm{RMSE}} = 100\frac{\mathrm{RMSE}_{\mathrm{A}} - \mathrm{RMSE}_{\mathrm{B}}}{\mathrm{RMSE}_{\mathrm{A}}} . \tag{15}$$

If the $\mathrm{RPI}_{\mathrm{RMSE}}$ value is positive, this means that the experiment B has a smaller RMSE.

3. Case evaluation experiments for SNAP

First, a group of case evaluation experiments were designed to evaluate SNAP and SNAP_S, by assimilating conventional observations.

3.1. *Experimental setup*

From 0000 UTC 8 June 2010 to 0000 UTC 9 June 2010, heavy precipitation occurred in South China, at a concentrated precipitation range and high intensity. The rain band was zonally distributed from the southwest to the northeast. The 24-h accumulated precipitation exceeded 100 mm.

We used WRF-ARW version 3.7.1 as the numerical forecast model in the following numerical experiments. The domain covered the whole of China in the region (15.5°–43.5°N, 88.5°–131.5°E) with the central point of (30°N, 110°E). SNAP adopted three grid scales (coarsest, fine, and finest) to conduct the assimilation analysis and the model forecast was at the finest scale. The finest grid scale contained 120 × 100 (longitude × latitude) grid points in the horizontal direction, with 30-km grid spacing. The numbers of grid points in the coarsest and fine scales were 30 × 25 and 60 × 50, with horizontal resolutions of 120 km and 60 km, respectively. It is worth noting that the latitude and longitude ranges of the three grid scales were different, because the simulation domains of the three grid scales in these experiments were generated with the same center point (30°N, 110°E) and the map projection (Lambert), but the grid points and resolutions were different. In the vertical direction, we used 30 layers from $\eta = 0$ to $\eta = 1$. The top pressure of the model layer was 50 hPa. The main physical components of the WRF model included the rapid radiative transfer model for longwave radiation (Mlawer et al., 1997), the Dudhia shortwave radiation scheme (Dudhia, 1989), the Yonsei University planetary boundary layer scheme (Hong et al., 2006), the Purdue Lin explicit cloud microphysics parameterization (Lin et al., 1983; Rutledge and Hobbs, 1984; Chen and Sun, 2002), and the Noah land surface model land scheme (Chen and Dudhia, 2001). First-guess field and boundary conditions in the experiments were generated using NCEP final (FNL) operational global analysis data (http://rda.ucar.edu/datasets/ds083.2/).

Two window cycling assimilation experiments were designed. The length of each assimilation window was six hours ([−3, 3]). The first assimilation window (named W1)

was ranged from 2100 UTC 7 June 2010 to 0300 UTC 8 June 2010 and the second assimilation window (named W2) was ranged from 0300 UTC 8 June 2010 to 0900 UTC 8 June 2010. The analysis time was at the beginning of each assimilation window, at which time the optimal analysis is obtained by minimizing the cost function. In each assimilation window, observations were assimilated hourly; that is, each assimilation window contained seven observation bins, and the assimilated observations were reprocessed as hourly data batches, including data within ±3-h windows ((−3, −2.5], (−2.5, −1.5], (−1.5, −0.5], (−0.5, 0.5], (0.5, 1.5], (1.5, 2.5], (2.5, 3]). The background field of W1 was a 12-h model forecast initialized by NCEP/FNL data 12 h before the analysis time. The control (CNTL) was a 27-h model integration from the background field of W1 at the analysis time (2100 UTC 7 June 2010) initialized using NCEP/FNL operational global analysis data; 120 ensembles were used. The simulation observations and observation innovations were generated through the GSI-based data processing and observation operator module. After calibrating the performance sensitivity to localization radius experiments, the horizontal localization radius was 2100 km. The number of truncated modes used for the generation of ρ_m and ρ_o were $r_x = 9$ and $r_y = 7$ (Zhang and Tian, 2018a). The background field of W2 was obtained by a 6-h model integration initialized by the analysis field generated in W1 at its corresponding analysis time.

The conventional observations for assimilation were from the China Meteorological Administration (CMA) National Meteorological Information Center, and were used for China's first-generation global atmospheric reanalysis product (CRA-40), which consist of surface and upper-air observations. Surface observations were available from ships, drifting buoys, land stations, and airports. In-situ measurements of the upper air were available from radiosonde, pilot balloon, aircraft, and wind profile data. Liao et al. (2018) describe the integrated conventional data sources, the quality control process, evaluation procedure and rejected observations, etc. Figure 2 shows the horizontal distribution of the assimilated observations after the GSI-based data processing (including read-in of observations, data thinning, data time and localization check, and gross error check) and observation operator module. Different colors represent the observations assimilated at different times. For these evaluation experiments, we focused on precipitation verification, using the hourly precipitation observations from more than 2400 national observation sites.

3.2. *Experimental results*

First, the experiments of the first assimilation window (W1) were used to comprehensively evaluate the correctness of SNAP. The precipitation forecast evaluation is described in section 3.2.1. The analysis of the improvements of the initial analysis field is discussed in section 3.2.2. The parameters of SNAP were determined by sensitivity experiments (section 3.2.3). Furthermore, the cycle assimilation performance of SNAP and the validity of the
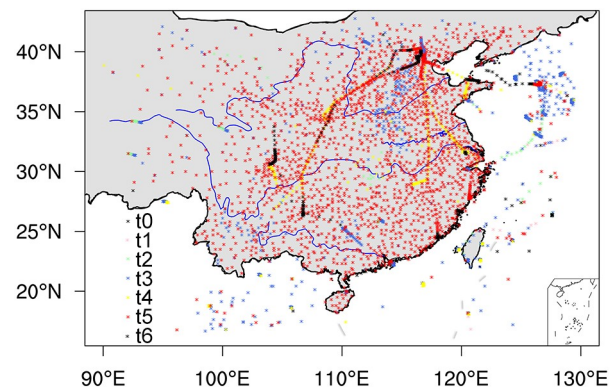

Fig. 2. Horizontal distribution of the assimilated observations in the first assimilation window. Different colors represent different observation times.

ensemble perturbation update scheme are illustrated through the experiments using the second assimilation window (section 3.2.4).

3.2.1. *Precipitation forecast skill*

Figure 3 shows the 24-h accumulated precipitation from 0000 UTC 8 June 2010 to 0000 UTC 9 June 2010, which initializes from 2100 UTC 7 June 2010. Figure 3a shows the precipitation observations (OBS) obtained from the hourly accumulated precipitation of more than 2400 national observation stations. Figures 3b–d show the precipitation forecasts obtained by model integration initialized from CTRL and assimilation analyses of SNAP_S and SNAP, respectively. The precipitation intensity predicted by the model (Figs. 3b–d) was greater than the cumulative precipitation observations (Fig. 3a). Heavy precipitation of up to 100 mm mainly occurred in Anhui, southeast Hubei, and central Hunan (Fig. 3a). At the same time, there were different degrees of precipitation in northern Jiangxi, eastern and western Guangxi, and western Guangdong. There was a false heavy precipitation center at the junction of Anhui and Hubei, where the precipitation reached 140 mm (Fig. 3b). There was obvious false precipitation in central Jiangxi, and the precipitation center in Hunan was to the southeast. At the same time, the precipitation forecast in western Guangxi was markedly stronger. However, there was little precipitation forecasting capability in western Guangdong. Figures 3c and d show the precipitation forecast simulated by models of different initial fields, generated through three iterations at a single grid scale (SNAP_S, Fig. 3c) and three grid scales with only one iteration at each grid scale (SNAP, Fig. 3d) after assimilating conventional observations. Compared to CTRL (Fig. 3b), the assimilation of conventional observations of SNAP reduced the false precipitation at the junction of central Jiangxi, Anhui, and Hubei, as well as to the west of Guangxi, but it could not forecast the precipitation in western Guangdong. Comparing Figs. 3c and d, the cumulative precipitation distribution of SNAP was closer to reality (Fig. 3a), especially in Anhui, northern Jiangxi, and other areas, which to some extent showed the importance of the multigrid assimilation framework.
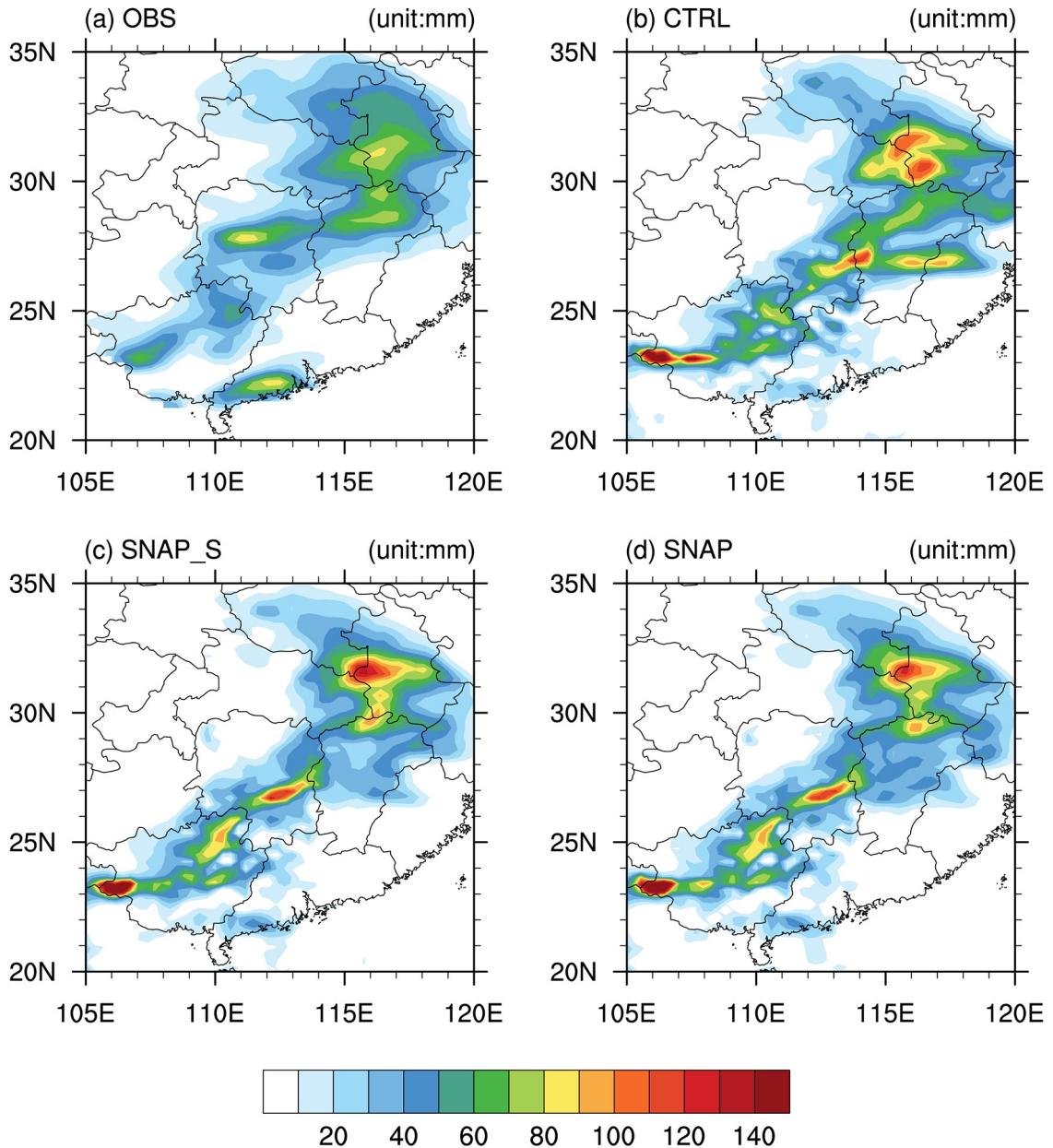
24-h Accumulated Percipitation

Fig. 3. The 24-h accumulated precipitation from 0000 UTC 8 June 2010 to 0000 UTC 9 June 2010 (units: mm): precipitation observations (a) OBS; and precipitation forecasts (b) CTRL; (c) SNAP_S; (d) SNAP.

Table 1 presents the results of the quantitative analysis of the RMSE and CC of precipitation forecasts and observations in the rainy region. The RMSE of CTRL and precipitation observations was 21.07174. After assimilating conventional observations, the RMSE of SNAP_S and precipitation observations was 18.87847. The RMSE of SNAP and precipitation observations (18.47027) was even smaller than that of SNAP_S, mainly due to using the multigrid NLS-4DVar assimilation framework to improve the assimilation accuracy. At the same time, the CC (passing the *t*-test at the 99% confidence level) between SNAP and precipitation observations (0.702526) was larger than those between

precipitation observations and CTRL (0.641511)/SNAP_S (0.686377), which further quantitatively indicated that the cumulative precipitation forecasts of SNAP were closer to the precipitation observations.

Figure 4 shows the 24-h cumulative precipitation TS values predicted from different initial fields. For the forecast of light rain, the scores were almost the same, although SNAP_S was slightly better. For moderate rain and heavy rain, SNAP_S and SNAP were better than CTRL, and SNAP was better than SNAP_S. For rainstorms, CTRL was better than SNAP_S and SNAP, and SNAP_S was better than SNAP. There are two possible reasons for these res-

Table 1. RMSE and CC values of 24-h cumulative precipitation forecasts with different initial fields and observations.

	CTRL	SNAP_S	SNAP
RMSE	21.07174	18.87847	18.47027
CC	0.641511	0.686377	0.702526
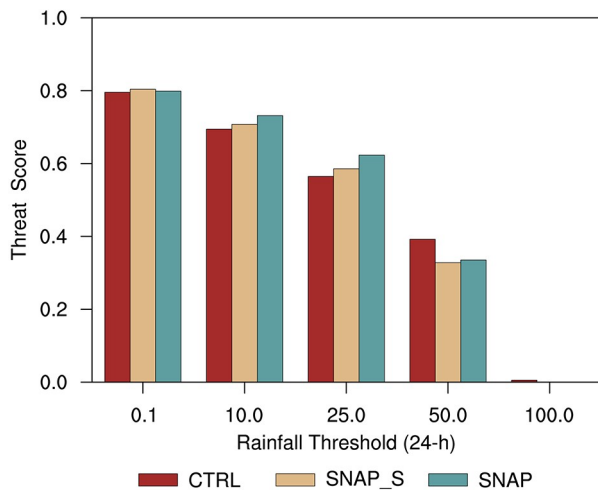
Fig. 4. The TS of 24-h cumulative precipitation classifications from 0000 UTC 8 June 2010 to 0000 UTC 9 June 2010.

ults; one is the relatively coarse resolution of the experiments, and the other is that only conventional observations were assimilated, and they were sparse, representing a large scale. For torrential rainfall, the score was almost 0. It was not surprising that the assimilation of conventional observations led to a higher TS than CTRL. These results further demonstrate that the multigrid NLS-4DVar assimilation framework can further improve the initial field and precipitation forecast.

Table 2 presents a comparison of the CPU times required for SNAP and SNAP_S to solve the optimal analysis field. The numerical experiments were conducted on the TH-1A system of the National Supercomputer Center in Tianjin, with 600 CPUs (50 nodes × 12 cores) and 5 TB of memory, and all assimilation calculations were serial on a single node single core. The total CPU time required by SNAP_S to obtain the optimal analysis field was 129.52 s, with 31 584 observations used for each iteration. The total CPU time of SNAP was 102.22 s (Table 2). Therefore, SNAP was more efficient than SNAP_S. This was because the observation operators of each grid scale were different in the multigrid assimilation framework and the longitudes and latitudes of the three grid scales were different (the finest grid scale covers the maximum domain, the fine grid scale comes next, and the coarsest grid scale is the minimum coverage). Consequently, the number of observations assimilated at each grid scale had a certain difference. In this experiment, 28 120, 30 338, and 31 584 assimilated observations were included, respectively, from the coarsest to the finest grid. In summary, using the multigrid assimilation framework SNAP can improve the assimilation accuracy

Table 2. CPU times required for SNAP and SNAP_S to solve the optimal analysis field, in which $l_i (i = 1, 2, 3)$ represents the number of iterations of SNAP_S and $L_i (i = 1, 2, 3)$ represents the ith grid scale of SNAP.

	CPU time (s)	
	SNAP_S	SNAP
l_1/L_1	43.04	27.27
l_2/L_2	42.98	32.02
l_3/L_3	43.30	42.93
Total CPU time	129.32	102.22

using fewer observations and less computational cost (Table 2), while revising multiscale errors (Figs. 3 and 4, Table 1).

3.2.2. *Optimal initial analysis field*

The improvement of precipitation forecasts is attributed to the assimilation of conventional observations by SNAP and SNAP_S to obtain the optimal initial field. Thus, from the perspective of the initial field increment, the reasons for the improvement of precipitation forecasts were analyzed. Figure 5 shows the analysis increment of water vapor mixing ratio (SNAP-CTRL and SNAP_S-CTRL) at the analysis time of the 12th layer of the model (850 hPa). The analysis increments of water vapor mixing ratio in central Jiangxi, north of central Hunan, and northeast Anhui were negative, and they were revised to different degrees by SNAP and SNAP_S (Figs. 5a and b), which was consistent with the decreases in false precipitation in central Jiangxi, Anhui, Hubei, and central Hunan of SNAP_S and SNAP compared to CTRL (Figs. 3b−d). Figure 6 shows the vertical distribution of the water vapor mixing ratio analysis increment (SNAP-CTRL and SNAP_S-CTRL) along 28°N. In the vertical direction, the analysis increments of the water vapor mixing ratio within the region 110−118°E of SNAP_S and SNAP were negative, especially below 400 hPa (Fig. 6). This is consistent with SNAP_S and SNAP weakening false precipitation in central Jiangxi and central Hunan (Fig. 3). By comparing the precipitation forecasts in Fig. 3 and the analyses increment fields in Figs. 5 and 6, it could be seen that central Jiangxi, Hunan, and the junction of Anhui and Hubei in the central region of the precipitation forecast and analysis increment field had a good corresponding relationship, and the SNAP_S and SNAP assimilation systems could absorb well the observation information, improving the structure of the initial field, thereby improving the precipitation forecast.

3.2.3. *Sensitivity to the system parameters*

Next, we conducted sensitivity experiments for the horizontal localization radius and the number of truncated modes selected in SNAP, based on the RMSE, CC and TS of 24-h precipitation forecasts and the CPU time required for assimilation calculations [Fig. S1, and Tables S3 in Electronic Supplementary Material (ESM) and Table 3]. When the localization radius was 2100 km, the RMSE was smal-
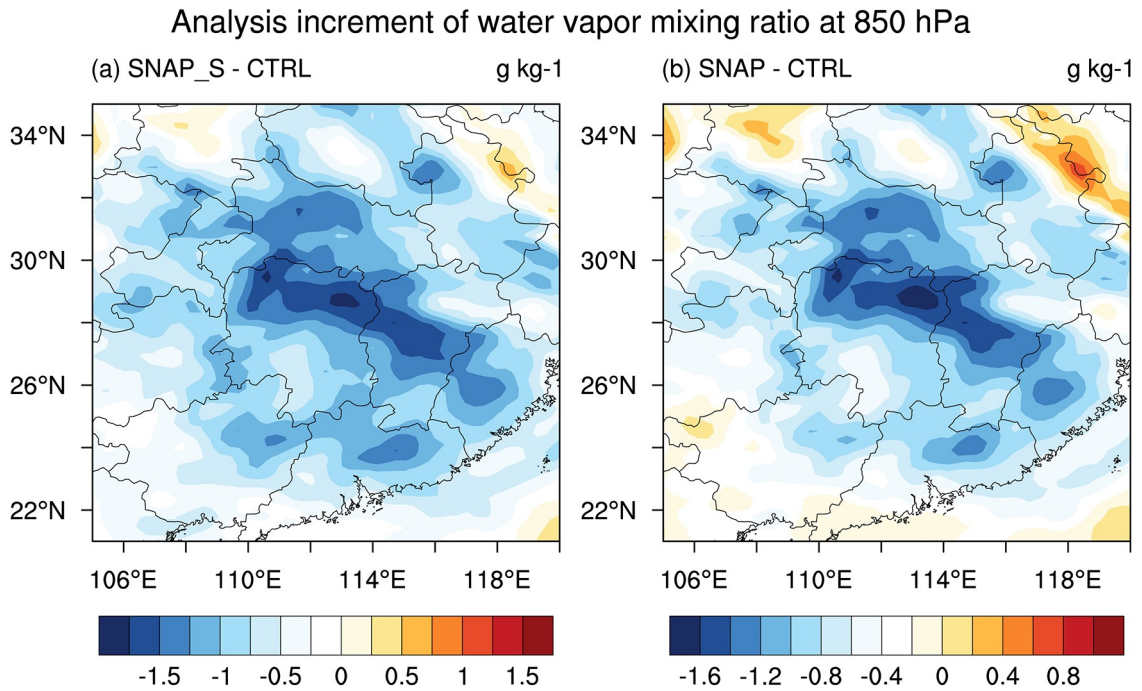
Analysis increment of water vapor mixing ratio at 850 hPa

Fig. 5. Analysis increment of the water vapor mixing ratio (units: g kg^{-1}) at the 12th layer of the model (850 hPa): (a) SNAP_S-CTRL; (b) SNAP-CTRL.
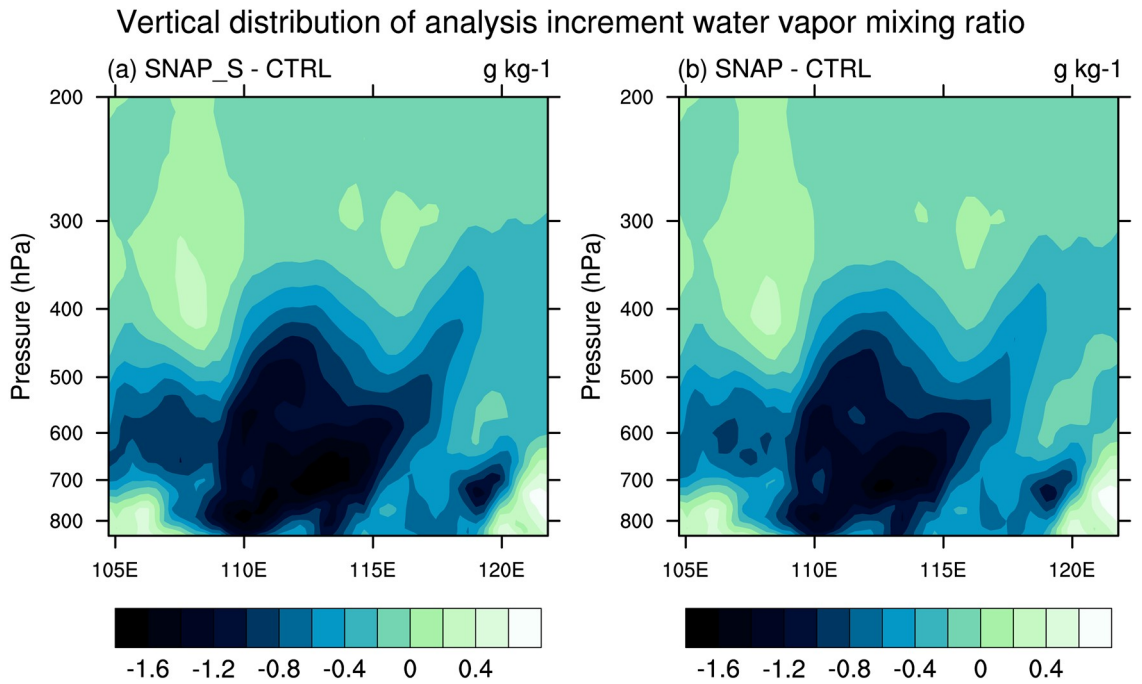
Vertical distribution of analysis increment water vapor mixing ratio

Fig. 6. Vertical distribution of the analysis increment of the water vapor mixing ratio (units: g kg^{-1}) along 28°N: (a) SNAP_S-CTRL; (b) SNAP-CTRL.

ler and the CC was larger (Fig. S1). At the same time, the TS showed that for light rain, moderate rain, and heavy rain, the local precipitation forecast with a radius of 2100 km had an absolute advantage. In fact, when the localization radius is about 1000 km, the distribution and intensity of precipitation has been significantly improved (not shown). However, by comparing all the indexes of assimilation accuracy, we think that 2100 km is the best localization radius in these experiments. For rainstorms and torrential rainfall, the TS of CTRL was higher. For the selected number of optimal truncation modes, this section focuses on the assimilation accuracy and calculation efficiency, as shown in Table 3. When the cumulative variance was greater than 90%, the assimilation accuracy was significantly improved (Table 3). In terms of statistical error, when the cumulative variance was 95%, there was a smaller RMSE and a larger CC (passing

Table 3. RMSE and CC values of 24-h precipitation observations and forecasts, and TSs of 24-h cumulative precipitation classifications with different cumulative variances (truncated modes) and CTRL. CPU times required for SNAP to solve the optimal analysis with different cumulative variances are also shown.

	CTRL	Cumulative variances		
		90% ($r_x = 7$, $r_y = 6$)	95% ($r_x = 9$, $r_y = 7$)	99% ($r_x = 11$, $r_y = 9$)
RMSE	21.07174	19.06320	18.87847	19.09410
CC	0.641511	0.681729	0.686377	0.682230
Threshold=0.1 (mm)	0.7957	0.8030	0.8039	0.8019
Threshold=10.0 (mm)	0.6945	0.7114	0.7076	0.7095
Threshold=25.0 (mm)	0.5647	0.5843	0.5856	0.5963
Threshold=50.0 (mm)	0.3922	0.3315	0.3279	0.3405
Threshold=100.0 (mm)	0.05	0.0	0.0	0.0
Time (s)	–	38.99	43.04	51.09

the *t*-test at the 99% confidence level). For precipitation, the TS of a precipitation forecast with a cumulative variance of 99% was better than that for 95%, except for light rain. However, with the increase of the truncated mode number (cumulative variance), the CPU time for the assimilation calculation also increased. Therefore, considering the assimilation accuracy and calculation efficiency, we chose the optimal truncated mode number with a cumulative variance of 95%; namely, $r_x = 9$ and $r_y = 7$.

3.2.4. *Results of the cycle system*

The two-window cyclic assimilation experiments used to evaluate SNAP were designed by continuous assimilation of observations (total of 12 hours), and the optimal analysis field was obtained at the start of the second window (0300 UTC 9 June 2010). To verify the cyclic assimilation performance of SNAP, the 12-h cumulative precipitation results were selected for evaluation in this section. Figure 7 shows the 12-h cumulative precipitation distribution from 0300 to 1500 on 9 June 2010. SNAP and SNAP_S improved the precipitation forecast through the assimilation of conventional observations, which was closer to observations (Figs. 7a, c and d). Table 4 shows the RMSE and CC of 12-h cumulative precipitation forecasts of different initial fields (CTRL, SNAP_S, and SNAP). The RMSE was lower for SNAP_S and SNAP than for CTRL, while CC was greater than CTRL. In addition, SNAP was better than SNAP_S. Furthermore, the TS of the precipitation forecast was better than that of CTRL except for torrential rainfall (Fig. 8). SNAP_S and SNAP were almost the same for the precipitation forecast of light rainfall, moderate rainfall, and rainstorms. SNAP was better than SNAP_S at forecasting heavy rainfall. The above results all demonstrate the cyclic assimilation capacity of SNAP and the effectiveness of the ensemble perturbation updating scheme used for the second assimilation window.

4. One-week cycling data assimilation experiments

One-week cycling data assimilation experiments were

designed to further evaluate SNAP compared to GSI 4DEn-Var by assimilating conventional observations. Two experiments using the multigrid NLS-4DVar (called SNAP) and GSI 4DEnVar (called GSI) methods were conducted to obtain the analysis, respectively. The generation and update schemes of ensembles adopted in SNAP were also examined.

4.1. *Experimental setup*

The one-week (16−23 July 2016) evaluation experiments were designed with continuous six-hourly assimilation cycles throughout this period, which started at 0300 UTC 16 July 2016 and ended at 0300 UTC 23 July 2016. This period included an extreme rainstorm in North China (35°−43°N, 113°−122°E) that occurred between 18 and 21 July 2016 (Fig. 9). There were two heavy rainfall centers in North China. The first one was located in the Taihang Mountains and occurred as a consequence of convective precipitation. The second one was located in south-central Beijing and occurred as a consequence of stratiform precipitation. A 30-h forecast was generated, initialed from the six-hourly cycled multigrid NLS-4DVar assimilation analyses. Parallel experiments using the GSI 4DEnVar system were also run to enable comparison with the GSI 4DEnVar scheme. It should be noted that there were many differences between SNAP and GSI including the analysis time and variables. The SNAP/GSI analysis time was at the beginning/middle of the assimilation window, respectively. Therefore, a 27-h forecast was generated, initialed from the six-hourly cycled GSI 4DEnVar assimilation analyses assimilating the same observations as in SNAP. The SNAP analysis variables were model variables, and the GSI analysis variables were control variables. The horizontal localization radius of SNAP was 300 km. Both the experiments used seven time levels of each assimilation window (as in section 3.1). Sixty ensemble members were employed. Only the ensemble-based background error covariance was used in both SNAP and GSI. The total number of iterations solved by GSI was 100, including 2 outer loops and 50 inner loops. First-guess field and boundary conditions were generated using ECMWF ERA-Interim global analysis data (https://apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc/),
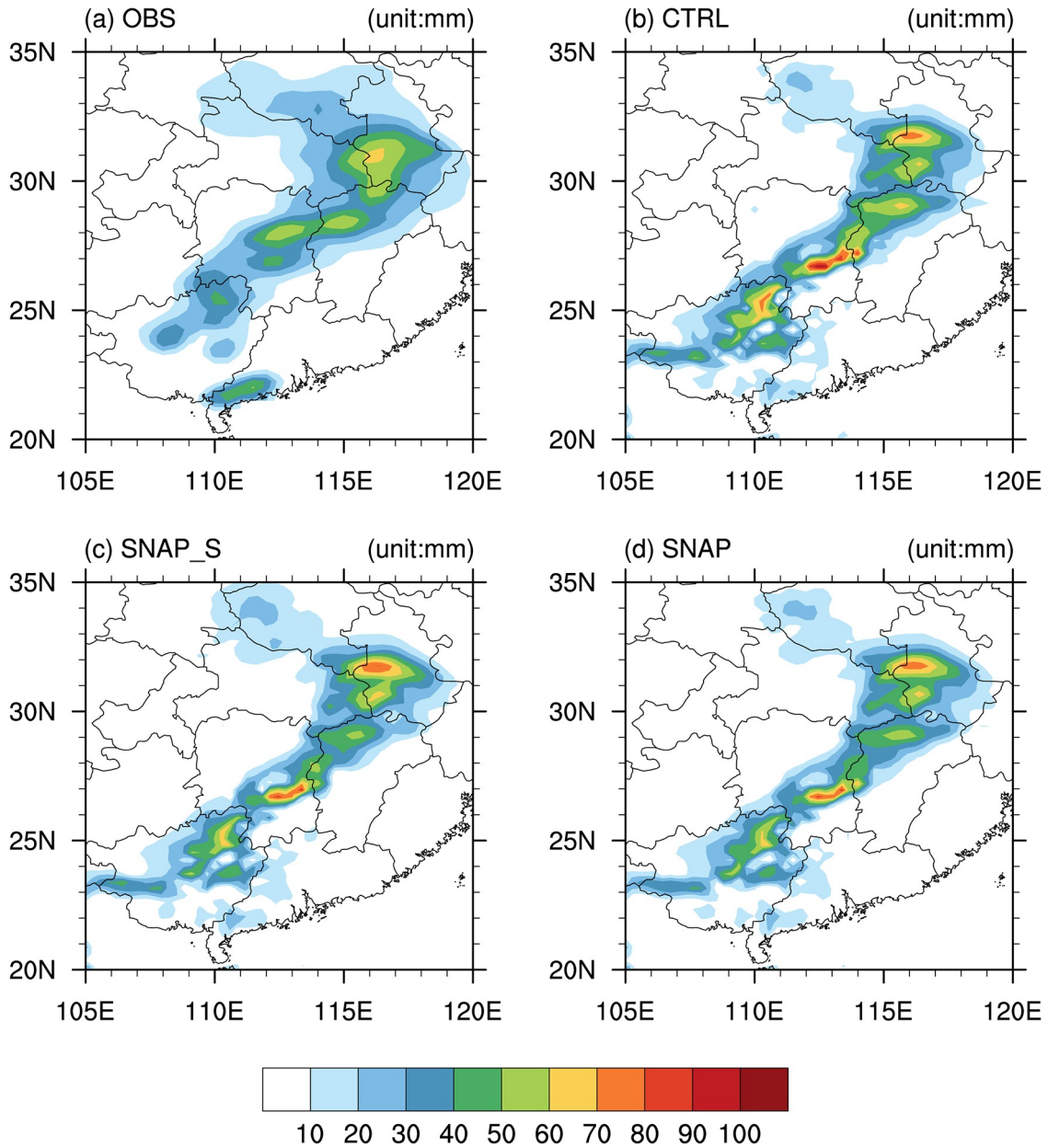
12-h Accumulated Percipitation

Fig. 7. The 12-h accumulated precipitation forecast from 0300 UTC 9 June 2010 to 1500 UTC 9 June 2010 (unit: mm): precipitation observations (a) OBS; and precipitation forecast (b) CTRL; (c) SNAP_S; (d) SNAP.

Table 4. RMSE and CC values of 12-h cumulative precipitation observations and forecasts for different initial fields.

	CTRL	SNAP_S	SNAP
RMSE	8.831729	8.529635	8.470109
CC	0.7460064	0.7666475	0.7688572

which were available every 6 h and updated every day. The model settings were the same as in section 3.1.

The conventional observational data for assimilation in these evaluation experiments were GDAS PrepBUFR data, including the surface observations (land-reporting stations, ships, buoys, etc.) and the upper-air observations (radio-sondes, aircrafts, wind profilers, etc.). The observations were treated in accordance with the time levels of the background and ensembles. The forecasts were verified against the conventional observations from the CMA National Meteorological Information Center used for China's first-generation global atmospheric reanalysis product (CRA-40) after the GSI-based data-processing and observation operator module, which converts relative humidity to humidity. Precipitation verification used hourly precipitation observations from 2380 national observation stations.

4.2. *Experimental results*

Figure 10 shows the one-week and domain-averaged

RMSEs at different forecast hours out of the assimilation window, verified against all the conventional observations for the *u/v* wind components, temperature, and humidity. It can be seen that the averaged RMSEs of SNAP were slightly lower than those of GSI at most forecast hours, especially for *u* wind and temperature. And for *v* wind and humidity,

the forecast improvements were evident. This may be due to the multigrid NLS-4DVar being able to correct multiscale errors to improve the initial field and forecast.

Figures 11 and 12 show the vertical profiles of 6-h and 24-h forecast averaged RMSEs, verified against all conven-
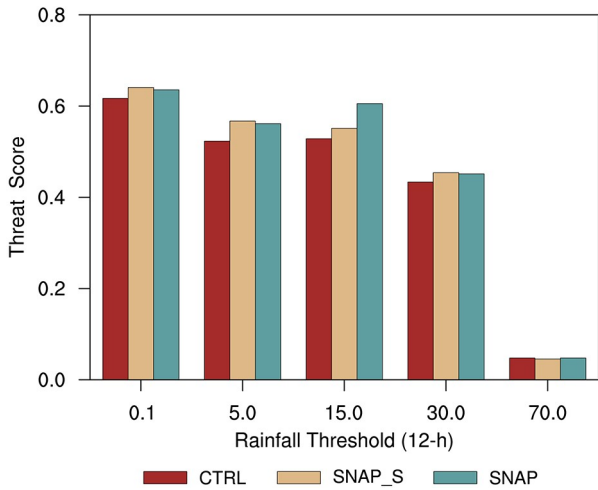
Fig. 8. The TS of 12-h cumulative precipitation classifications from 0300 UTC 9 June 2010 to 1500 UTC 9 June 2010.
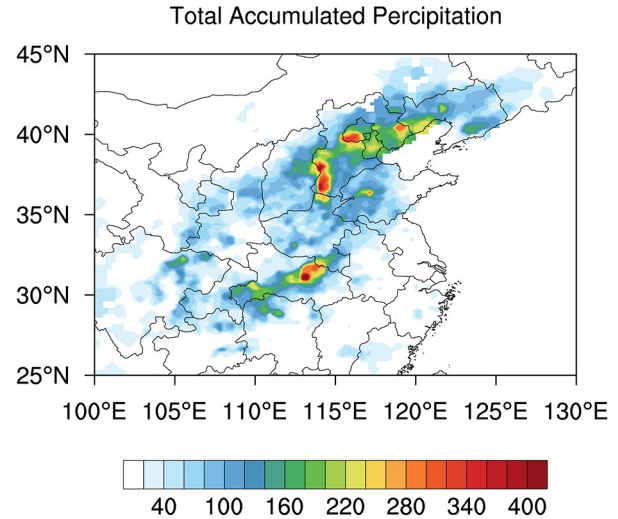
Fig. 9. The accumulated precipitation observations from 0000 UTC 18 July 2016 to 0000 UTC 22 July 2016 (unit: mm).
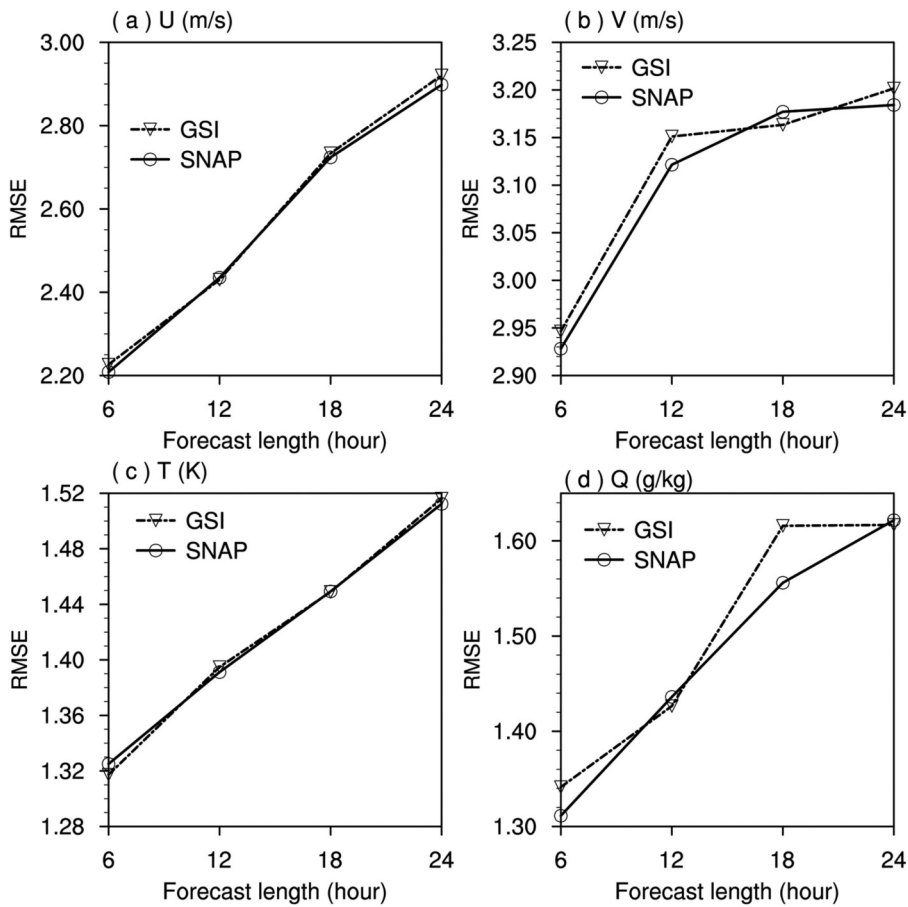

Fig. 10. The one-week and domain-averaged RMSEs at different forecast hours out of the assimilation window from SNAP and GSI, verified against all conventional observations for the (a) *u* and (b) *v* wind components, (c) temperature, and (d) humidity. The horizontal axis shows the forecast hour.
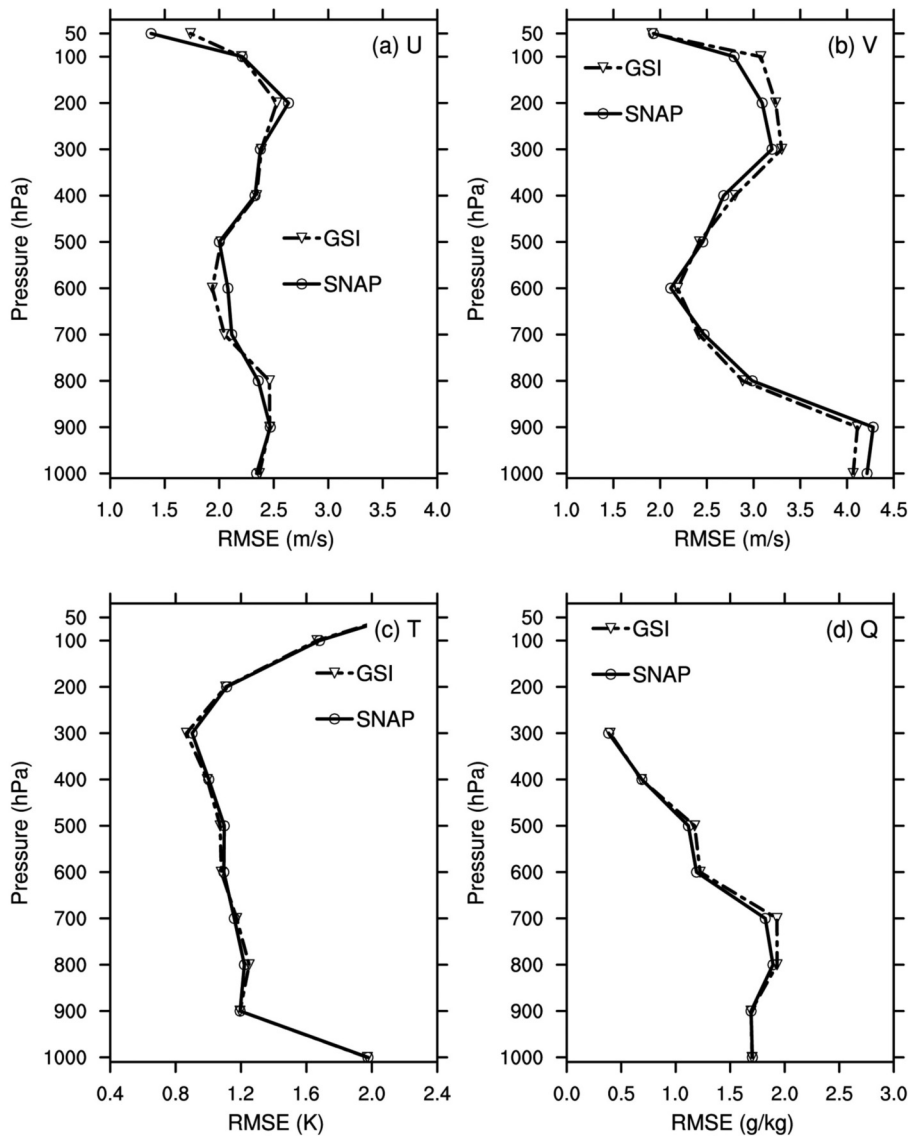
Fig. 11. Vertical profiles of the 6-h averaged RMSEs of the SNAP and GSI forecasts' fit to conventional observations for the (a) *u* and (b) *v* wind components, (c) temperature and (d) humidity for the testing period.

tional observations for the *u*/*v* wind components, temperature, and humidity. It should be noted that the statistical values of 1000 hPa include the results for cases with pressure greater than 1000 hPa. As can be seen from Fig. 11, for the *u*/*v* wind and temperature, SNAP and GSI had their own advantages in different pressure levels. According to Figs. 10 and 11, the 6-h forecast averaged RMSEs of SNAP were slightly lower than those of GSI for *u*/*v* wind. The 6-h forecast averaged RMSEs of SNAP were also lower than those of GSI for the humidity variable at most pressure layers, suggesting a better analysis of all layer structures (Fig. 11 and Table 5). Except for the higher RMSEs at the upper level for *u* wind, and at lower middle levels (700 hPa) for humidity, the performance of SNAP was superior to that of GSI throughout the 24-h forecast (Fig. 12 and Table 5). For temperature, SNAP was better at the upper pressure levels, suggesting that the SNAP forecast was generally a better fit to

the observations than the GSI forecast.

To quantify the improvement of SNAP over GSI, the RPI for RMSE was computed (Table 5). It can be seen from Table 5 that SNAP produced slightly lower forecast RMSEs than GSI 4DEnVar as a whole in the prediction verification of *u*/*v* wind, temperature and humidity (Figs. 10–12). The 6-h forecast averaged RMSE of *u* wind was improved by 20.81% at 50 hPa, which represents the largest improvement among all variables and the 24-h RPI at above 400 hPa was positive, which means that SNAP has a good forecast of *u* wind in the middle and lower layers. For the humidity, except for the 700- and 400-hPa levels, the values of the 24-h RPI were positive. For the *v* wind and temperature, SNAP and GSI have their own advantages in different pressure layers.

Figure 13 shows the 12-h accumulated precipitation for a case of extreme precipitation from 1800 UTC 19 to 0600

Table 5.   The RPI of the RMSE for 6-h and 24-h forecasts over all forecast cycles throughout the experimental period.

| Pressure (hPa)	RPI$_{RMSE}$							
	u (%)		v (%)		T (%)		q (%)	
	6-h	24-h	6-h	24-h	6-h	24-h	6-h	24-h
50	20.81	−6.26	−0.85	0.07	−0.82	−5.44	–	–
100	−0.29	−0.82	9.29	−1.24	−0.9	0.37	–	–
200	−4.26	−0.23	4.46	2.89	−0.52	2.91	–	–
300	0.37	2.61	3.13	2.81	−3.99	3.29	3.36	0.35
400	0.56	−1.06	4.29	−4.51	−0.94	5.33	−0.72	−1.38
500	0.75	1.41	−1.447	0.96	−2.35	2.34	4.92	2.75
600	−7.35	1.51	3.31	−0.06	−1.42	0.54	2.67	4.34
700	−3.27	2.82	−2.27	1.93	1.24	−1.38	5.56	−11.95
800	4.16	2.91	−3.68	2.85	2.32	−1.19	1.93	0.01
900	−0.19	1.47	−4.09	0.68	−0.04	−0.54	−0.01	0.56
1000	1.06	0.96	−3.64	−1.04	−0.17	−1.28	−0.15	1.41
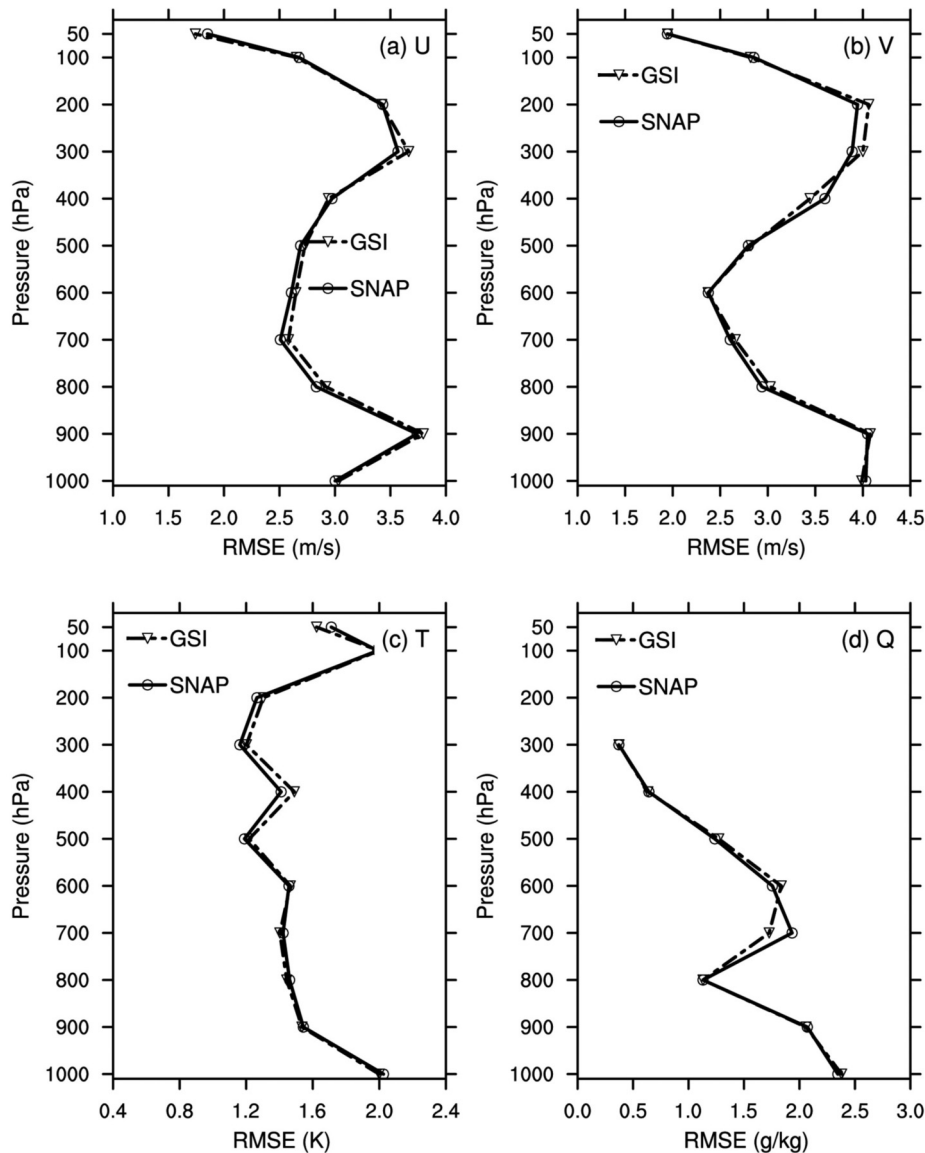
Fig. 12. As in Fig. 11 but for the 24-h forecast averaged RMSEs.
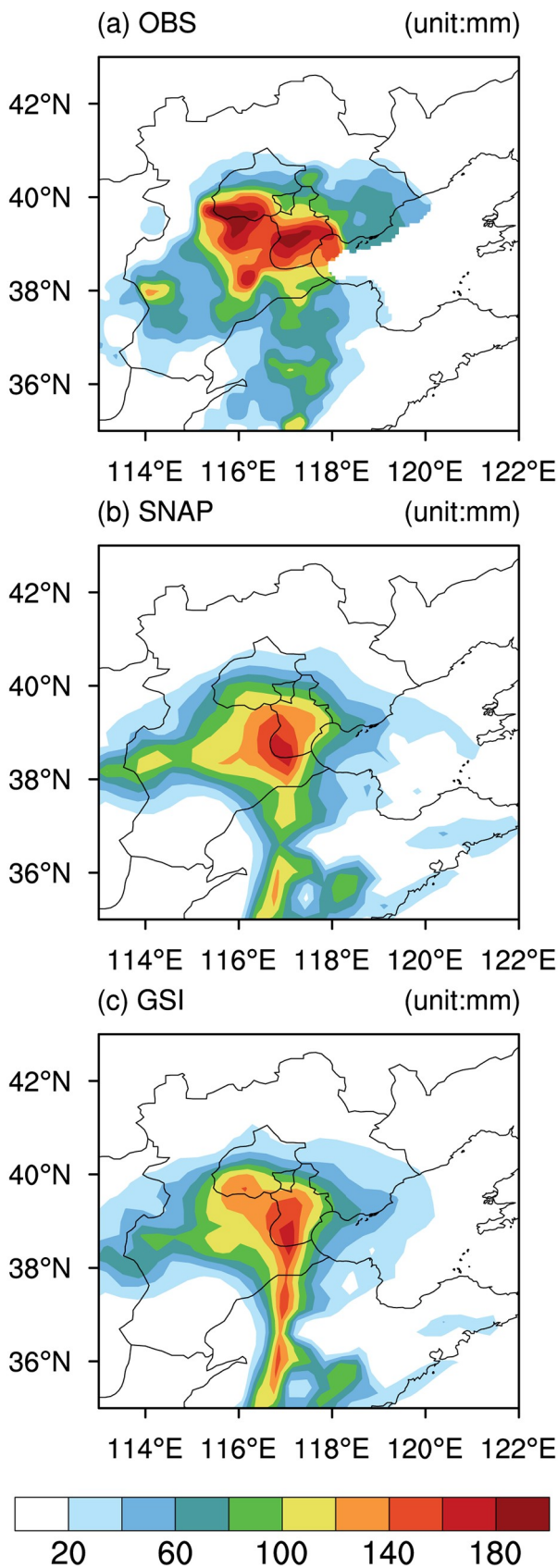
Fig. 13. The 12-h accumulated precipitation forecast from 1800 UTC 19 to 0600 UTC 20 July 2016 (unit: mm): observations (a) OBS; forecasts (b) SNAP and (c) GSI.

UTC 20 July 2016 in North China. Figure 13a shows the precipitation observations (OBS) obtained from hourly precipitation observations of 2380 national observation stations; the amount of 12-h accumulated precipitation exceeded 140 mm. Figures 13b and c show the 12-h precipitation forecasts of SNAP and GSI, respectively. It can be seen from Fig. 13 that the precipitation forecast intensity (Figs. 13b and c) was weaker than the observed precipitation (Fig. 13a). Heavy rainfall mainly occurred in the Taihang Mountains, the south-central part of Beijing, and Tianjin. SNAP was better than GSI for predicting the location of precipitation. Furthermore, the RMSEs and spatial CCs between the observations and precipitation forecasts of SNAP/GSI were 30.20/ 30.39 and 0.82/0.81 respectively, which quantitatively showed that the performance of SNAP was superior to GSI. Figure 14 shows the 12-h accumulated precipitation classification ETS values for thresholds of 5, 15, 30, 70 and 100 mm. It can be seen that, except for the threshold of 70 mm and 100 mm, SNAP outperformed GSI for the other thresholds shown.

Ensemble members are very important for the ensemble-based data assimilation methods, which use the linear combination of ensemble perturbations to express the analysis increment. Therefore, the generation and updating strategy of ensemble perturbations are of great importance. In this part, the time period from 0300 UTC 19 to 1500 UTC 20 July 2016, characterized by heavy rainfall events, was selected by the ensemble spread test, which had six assimilation windows. Figure 15 shows time series of ensemble spread during the test period for the u/v horizontal wind components, perturbation potential temperature T, and water vapor mixing ratio q state variables. It can be seen that the ensemble spread did not decrease with the increase in forecast time, and showed a cyclic characteristic in each assimilation window. For the u and q variables, the ensemble spread in an assimilation window was reduced. However, for the v and T variables, the ensemble spread was first smaller, and then larger in each assimilation window, which may be due to the nonlinearity of the numerical model. The assimilation res-
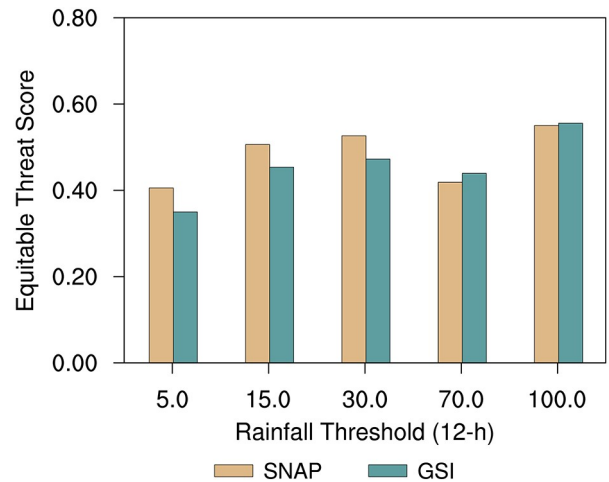

Fig. 14. The ETS of 12-h cumulative precipitation classifications from 1800 UTC 19 to 0600 UTC 20 July 2016.
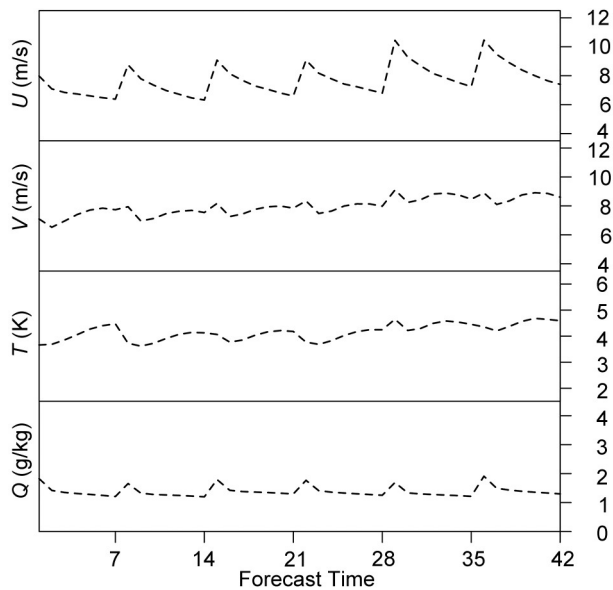
Fig. 15. Time series of the ensemble spread during six cycle assimilation windows from 0300 UTC 19 to 1500 UTC 20 July 2016, for U: u wind, V: v wind, T: perturbation potential temperature, and Q: water vapor mixing ratio state variables of cases of heavy rainfall.

ults showed that the generation and update schemes of ensemble perturbations adopted in this study were effective.

5. Summary and concluding remarks

This paper describes a newly developed, SNAP, based on the multigrid NLS-4DVar assimilation framework and the GSI-based quality control and observation operator modules, which was evaluated with the WRF-ARW numerical forecast model. The particular advantages of SNAP are as follows:

• It can effectively absorb multiple-source (conventional, radar, and satellite) observations.

• It makes full use of GSI-based data-processing (including quality control and thinning) and observation operator modules to generate observation innovation.

• The multigrid NLS-4DVar assimilation framework can sequentially revise multiscale errors and accelerate iterative convergence, thus improving the assimilation accuracy and computational efficiency.

• The application of the fast localization scheme simplifies the complicated localization process and makes it possible for the NLS-4DVar method to be applied operationally.

In the case evaluation experiments, compared to observations, CTRL produced a strong false heavy precipitation center, and the weak precipitation area of observations was strengthened. SNAP eliminated the false heavy precipitation center by assimilating the conventional observations, which effectively weakened the false heavy precipitation, and the position of the heavy precipitation also improved. The analysis increment was in good agreement with the pre-

cipitation forecast, which indicates that SNAP can effectively absorb observation information, improve the initial field, and further improve the precipitation forecast. In the one-week cycle assimilation experiments, the averaged RMSEs of SNAP were slightly lower than those of GSI for the u/v wind components, T, and q, as a whole. Furthermore, precipitation verification experiments showed that SNAP outperformed GSI.

At present, in SNAP, the assimilation of radar, satellite, and other unconventional observations is still in progress. At the same time, for the localization scheme of multigrid NLS-4DVar, $\rho_{m,(i)}$ and $\rho_{o,(i)}$ at the ith grid scale are extracted from the finest grid scale without the multiscale localization strategy. In future work, the coupling between the multigrid NLS-4DVar assimilation framework and the multiscale localization strategy will be studied. In addition, how to choose an accurate localization radius adaptively and robustly plays a vital role in building a mature assimilation system. SNAP urgently needs the development of such an adaptive localization scheme, and this work is ongoing. Assimilation of multiscale observations and a big data–driven NLS-4DVar (Tian and Zhang, 2019) consisting of two ensembles, a prepared historical "big data" ensemble and a small "online" ensemble, is also underway and will be introduced in future papers. Bias correction with the NLS-4DVar method is also being investigated for satellite data assimilation.

Electronic supplementary material: Supplementary material is available in the online version of this article at https://doi.org/10.1007/s00376-020-9252-1.

REFERENCES

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903, https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.

Benjamin, S. G., and Coauthors, 2004: An hourly assimilation-forecast cycle: The RUC. *Mon. Wea. Rev.*, **132**, 495–518, https://doi.org/10.1175/1520-0493(2004)132<0495:AHACTR>2.0.CO;2.

Benjamin, S. G., and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The rapid refresh. *Mon. Wea. Rev.*, **144**, 1669–1694, https://doi.org/10.1175/MWR-D-15-0242.1.

Briggs, W., V. E. Henson, and S. F. McCormic, 2000: *A Multigrid Tutorial*. 2nd ed., SIAM, 95-109.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and

B. He, 2010a: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: Description and single-observation experiments. *Mon. Wea. Rev.*, **138**, 1550–1566, https://doi.org/10.1175/2009MWR3157.1.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010b: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: One-month experiments with real observations. *Mon. Wea. Rev.*, **138**, 1567–1586, https://doi.org/10.1175/2009MWR3158.1.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.*, **129**, 569–585, https://doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.

Chen, S. H., and W. Y. Sun, 2002: A one-dimensional time dependent cloud model. *J. Meteor. Soc. Japan*, **80**, 99–118, https://doi.org/10.2151/jmsj.80.99.

Clayton, A. M., A. C. Lorenc, and D. M. Barker, 2013: Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Quart. J. Roy. Meteor. Soc.*, **139**, 1445–1461, https://doi.org/10.1002/qj.2054.

Courtier, P., J. N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**, 1367–1387, https://doi.org/10.1002/qj.49712051912.

Dennis, J. E., and R. B. Schnabel, 1996: Numerical Methods for Unconstrained Optimization and Nonlinear Equations (Classics in Applied Mathematics). *SIAM*, 378 pp.

Dudhia, J., 1989: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107, https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**(C5), 10 143–10 162, https://doi.org/10.1029/94JC00572.

Evensen, G., 2003: The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, **53**, 343–367, https://doi.org/10.1007/s10236-003-0036-9.

Evensen, G., 2007: *Data Assimilation-The Ensemble Kalman Filter*. Springer, 157–176.

Gauthier, P., M. Tanguay, S. Laroche, S. Pellerin, and J. Morneau, 2007: Extension of 3DVAR to 4DVAR: Implementation of 4DVAR at the Meteorological Service of Canada. *Mon. Wea. Rev.*, **135**, 2339–2364, https://doi.org/10.1175/MWR3394.1.

Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter-3D variational analysis scheme. *Mon. Wea. Rev.*, **128**, 2905–2919, https://doi.org/10.1175/1520-0493(2000)128<2905:AHEKFV>2.0.CO;2.

Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129**, 2776–2790, https://doi.org/10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2.

Hong, S. Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, https://doi.org/10.1175/MWR3199.1.

Houtekamer, P. L., and H. L. Mitchell, 2005: Ensemble Kalman filtering. *Quart. J. Roy. Meteor. Soc.*, **131**, 3269–3289, https://doi.org/10.1256/qj.05.135.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811, https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.

Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.*, **133**, 604–620, https://doi.org/10.1175/MWR-2864.1.

Hu, M., and Coauthors, 2018: Grid-point Statistical Interpolation (GSI) User's Guide Version 3.7. Developmental Testbed Center. Available from http://www.dtcenter.org/com-GSI/users/docs/index.php.

Hunt, B. R., Kostelich, E. J., Ott, E., Szunyogh, I., 2007: Efficient data assimilation for spatio temporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**(1–2), 112–126, https://doi.org/10.1016/j.physd.2006.11.008.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W. S. Wu, and S. Lord, 2009: Introduction of the GSI into the NCEP global data assimilation system. *Wea. Forecasting*, **24**, 1691–1705, https://doi.org/10.1175/2009WAF2222201.1.

Kuhl, D. D., T. E. Rosmond, C. H. Bishop, J. McLay, and N. L. Baker, 2013: Comparison of hybrid ensemble/4DVar and 4DVar within the NAVDAS-AR data assimilation framework. *Mon. Wea. Rev.*, **141**, 2740–2758, https://doi.org/10.1175/MWR-D-12-00182.1.

Lewis, J. M., and J. C. Derber, 1985: The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus A: Dynamic Meteorology and Oceanography*, **37**, 309–322, https://doi.org/10.3402/tellusa.v37i4.11675.

Li, W., Y. F. Xie, S. M. Deng, and Q. Wang, 2010: Application of the multigrid method to the two-dimensional doppler radar radial velocity data assimilation. *J. Atmos. Oceanic Technol.*, **27**(2), 319–332, https://doi.org/10.1175/2009JTECHA1271.1.

Li, Z. J., Y. Chao, J. C. McWilliams, and K. Ide, 2008: A three-dimensional variational data assimilation scheme for the Regional Ocean Modeling System: Implementation and basic experiments. *J. Geophy. Res.*, **113**, C05002, https://doi.org/10.1029/2006JC004042.

Li, Z. J., Y. Chao, J. D. Farrara, and J. C. McWilliams, 2013: Impacts of distinct observations during the 2009 Prince William Sound field experiment: A data assimilation study. *Cont. Shelf Res.*, **63**, S209–S222, https://doi.org/10.1016/j.csr.2012.06.018.

Liao, J., and Coauthors, 2018: Pre-process and data selection for assimilation of conventional observations in the CMA global atmospheric reanalysis. *Advances in Meteorological Science and Technology*, **8**(1), 133–142, https://doi.org/10.3969/j.issn.2095-1973.2018.01.018.

Lin, Y. L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteorol.*, **22**, 1065–1092, https://doi.org/10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2.

Lorenc, A. C., 2003a: Modelling of error covariances by 4D-Var data assimilation. *Quart. J. Roy. Meteor. Soc.*, **129**, 3167–3182, https://doi.org/10.1256/qj.02.131.

Lorenc, A. C., 2003b: The potential of the ensemble Kalman filter for NWP—A comparison with 4D-VAR. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183–3203, https://doi.org/10.1256/

qj.02.132.

Lorenc, A. C., 2013: Recommended nomenclature for EnVar data assimilation methods. Research Activities in Atmospheric and Oceanic Modelling, WGNE. [Available from http://www.wcrp-climate.org/WGNE/BlueBook/2013/individual-articles/01_Lorenc_Andrew_EnVar_nomenclature.pdf]

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**(D14), 16 663–16 682, https://doi.org/10.1029/97JD00237.

Pan, Y. J., K. F. Zhu, M. Xue, X. G. Wang, M. Hu, S. G. Benjamin, S. S. Weygandt, and J. S. Whitaker, 2014: A GSI-based coupled EnSRF–En3DVar hybrid data assimilation system for the operational rapid refresh model: Tests at a reduced resolution. *Mon. Wea. Rev.*, **142**, 3756–3780, https://doi.org/10.1175/MWR-D-13-00242.1.

Pan, Y. J., M. Xue, K. F. Zhu, and M. J. Wang, 2018: A prototype regional GSI-based EnKF-variational hybrid data assimilation system for the Rapid Refresh forecasting system: Dual-resolution implementation and testing results. *Adv. Atmos. Sci.*, **35**(5), 518–530, https://doi.org/10.1007/s00376-017-7108-0.

Qiu, C. J., A. M. Shao, Q. Xu, and L. Wei, 2007: Fitting model fields to observations by using singular value decomposition: An ensemble-based 4DVar approach. *J. Geophy. Res.*, **112**, D11105, https://doi.org/10.1029/2006JD007994.

Rabier F., and Coauthors, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.*, **126**, 1143–1170, https://doi.org/10.1002/qj.49712656415.

Rosmond, T., and L. Xu, 2006: Development of NAVDAS-AR: Non-linear formulation and outer loop tests. *Tellus A: Dynamic Meteorology and Oceanography*, **58**, 45–58, https://doi.org/10.1111/j.1600-0870.2006.00148.x.

Rutledge, S. A., and P. V. Hobbs, 1984: The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. XII: A diagnostic modeling study of precipitation development in narrow cold-frontal rainbands. *J. Atmos. Sci.*, **20**, 2949–2972, https://doi.org/10.1175/1520-0469(1984)041<2949:TMAMSA>2.0.CO;2.

Skamarock, W. C., and J. B. Klemp, 2008: A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *J. Comput. Phys.*, **227**, 3465–3485, https://doi.org/10.1016/j.jcp.2007.01.037.

Tian, X. J., and Z. H. Xie, 2012: Implementations of a square-root ensemble analysis and a hybrid, localization into the POD-based ensemble 4DVar. *Tellus A: Dynamic Meteorology and Oceanography*, **64**, 18375, https://doi.org/10.3402/tellusa.v64i0.18375.

Tian, X. J., and X. B. Feng, 2015: A non-linear least squares enhanced POD-4DVar algorithm for data assimilation. *Tellus A: Dynamic Meteorology and Oceanography*, **67**, 25340, https://doi.org/10.3402/tellusa.v67.25340.

Tian, X. J., and H. Q. Zhang, 2019: A big data-driven nonlinear least squares four‐dimensional variational data assimilation method: Theoretical formulation and conceptual evaluation. *Earth and Space Science*, **6**, 1430–1439, https://doi.org/10.1029/2019EA000735.

Tian, X. J., Z. H. Xie, and A. G. Dai, 2008: An ensemble-based explicit four-dimensional variational assimilation method. *J. Geophy. Res.*, **113**, D21124, https://doi.org/10.1029/2008JD010358.

Tian, X. J., Z. H. Xie, and Q. Sun, 2011: A POD-based ensemble four-dimensional variational assimilation method. *Tellus A: Dynamic Meteorology and Oceanography*, **63**, 805–816, https://doi.org/10.1111/j.1600-0870.2011.00529.x.

Tian, X. J., H. Q. Zhang, X. B. Feng, and Y. F. Xie, 2018: Nonlinear least squares En4DVar to 4DEnVar methods for data assimilation: Formulation, analysis, and preliminary evaluation. *Mon. Wea. Rev.*, **146**, 77–93, https://doi.org/10.1175/MWR-D-17-0050.1.

Wang, B., J. J. Liu, S. D. Wang, W. Cheng, L. Juan, C. S. Liu, Q. N. Xiao, and Y. H. Kuo, 2010: An economical approach to four-dimensional variational data assimilation. *Adv. Atmos. Sci.*, **27**, 715–727, https://doi.org/10.1007/s00376-009-9122-3.

Wu, W. S., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905–2916, https://doi.org/10.1175/1520-0493(2002)130<2905:TDVAWS>2.0.CO;2.

Xie, Y. F., S. E. Koch, J. A. McGinley, S. Albers, and N. Wang, 2005: A sequential variational analysis approach for mesoscale data assimilation. *Preprints, 21st Conf. on Weather Analysis and Forecasting/17th Conf. on Numerical Weather Prediction*, Washington, DC, Amer. Meteor. Soc.

Xie, Y., S. Koch, J. McGinley, S. Albers, P. E. Bieringer, M. Wolfson, and M. Chan, 2011: A space-time multiscale analysis system: A sequential variational analysis approach. *Mon. Wea. Rev.*, **139**, 1224–1240, https://doi.org/10.1175/2010MWR3338.1.

Xu, Q., 1996: Generalized adjoint for physical processes with parameterized discontinuities. Part I: Basic issues and heuristic examples. *J. Atmos. Sci.*, **53**(8), 1123–1142, https://doi.org/10.1175/1520-0469(1996)053<1123:GAFPPW>2.0.CO;2.

Zhang, F. Q., Y. H. Weng, J. A. Sippel, Z. Y. Meng, and C. H. Bishop, 2009: Cloud-resolving hurricane initialization and prediction through assimilation of Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **137**, 2105–2125, https://doi.org/10.1175/2009MWR2645.1.

Zhang, H. Q., and X. J. Tian, 2018a: An efficient local correlation matrix decomposition approach for the localization implementation of ensemble-based assimilation methods. *J. Geophy. Res.*, **123**, 3556–3573, https://doi.org/10.1002/2017JD027999.

Zhang, H. Q., and X. J. Tian, 2018b: A multigrid nonlinear least squares four-dimensional variational data assimilation scheme with the advanced research weather research and forecasting model. *J. Geophy. Res.*, **123**, 5116–5129, https://doi.org/10.1029/2017JD027529.

Zhang, H. Q., 2019: Improvement and application of nonlinear least squares ensemble four-dimensional variational assimilation method. PhD dissertation, Institute of Atmospheric Physics, Chinese Academy of Sciences, 136–139. (in Chinese)

Zhang, M., and F. Q. Zhang, 2012: E4DVar: Coupling an ensemble Kalman filter with four-dimensional variational data assimilation in a limited-area weather prediction model. *Mon. Wea. Rev.*, **140**, 587–600, https://doi.org/10.1175/MWR-D-11-00023.1.

Zhu, K. F., Y. J. Pan, M. Xue, X. G. Wang, J. S. Whitaker, S. G. Benjamin, S. S. Weygandt, and M. Hu, 2013: A regional GSI-based ensemble Kalman filter data assimilation system for the rapid refresh configuration: Testing at reduced resolution. *Mon. Wea. Rev.*, **141**, 4118–4139, https://doi.org/10.1175/MWR-D-13-00039.1.