

# Assimilating Surface Observations in a Four-Dimensional Variational Doppler Radar Data Assimilation System to Improve the Analysis and Forecast of a Squall Line Case

Xingchao CHEN<sup>1,2</sup>, Kun ZHAO<sup>\*1</sup>, Juanzhen SUN<sup>2</sup>, Bowen ZHOU<sup>1</sup>, and Wen-Chau LEE<sup>2</sup>

<sup>1</sup>Key Laboratory for Mesoscale Severe Weather/MOE and School of Atmospheric Science, Nanjing University, Nanjing 210046, China

<sup>2</sup>National Center for Atmospheric Research, Boulder, Colorado 80301, USA

(Received 30 December 2015; revised 24 March 2016; accepted 8 April 2016)

## ABSTRACT

This paper examines how assimilating surface observations can improve the analysis and forecast ability of a four-dimensional Variational Doppler Radar Analysis System (VDRAS). Observed surface temperature and winds are assimilated together with radar radial velocity and reflectivity into a convection-permitting model using the VDRAS four-dimensional variational (4DVAR) data assimilation system. A squall-line case observed during a field campaign is selected to investigate the performance of the technique. A single observation experiment shows that assimilating surface observations can influence the analyzed fields in both the horizontal and vertical directions. The surface-based cold pool, divergence and gust front of the squall line are all strengthened through the assimilation of the single surface observation. Three experiments—assimilating radar data only, assimilating radar data with surface data blended in a mesoscale background, and assimilating both radar and surface observations with a 4DVAR cost function—are conducted to examine the impact of the surface data assimilation. Independent surface and wind profiler observations are used for verification. The result shows that the analysis and forecast are improved when surface observations are assimilated in addition to radar observations. It is also shown that the additional surface data can help improve the analysis and forecast at low levels. Surface and low-level features of the squall line—including the surface warm inflow, cold pool, gust front, and low-level wind—are much closer to the observations after assimilating the surface data in VDRAS.

**Key words:** VDRAS, 4-D data assimilation, radar data, surface data, squall line

**Citation:** Chen, X. C., K. Zhao, J. Z. Sun, B. W. Zhou, and W.-C. Lee, 2016: Assimilating surface observations in a four-dimensional variational Doppler radar data assimilation system to improve the analysis and forecast of a squall line case. *Adv. Atmos. Sci.*, **33**(10), 1106–1119, doi: 10.1007/s00376-016-5290-0.

## 1. Introduction

With the increases in computing power, it is possible to resolve and forecast the features of severe convective systems through a convection-permitting numerical model. One of the major challenges of storm-resolving numerical weather prediction (NWP) models is the accuracy of initial conditions. However, the forecast of meso- or convective-scale weather systems is a highly nonlinear initial value problem; the inaccuracy of the analyzed initial state becomes the fundamental limitation for the prediction of severe convective systems (Schenkman et al., 2011a). The high temporal (<10 min) and spatial (<1.0 km) resolution of Doppler radar makes it the best remote sensing device to probe the variation of hydrometeors and the airflow structures inside the convective system (Lilly, 1990; Sun et al., 1991). Numerous studies have

shown that the assimilation of radar reflectivity and radial velocity data can improve the analysis and subsequent forecast of convective systems (e.g., Sun and Crook, 1998; Weygandt et al., 2002a, 2002b; Dawson and Xue, 2006).

Considering the influence of Earth's curvature, the height between the lowest radar beam (if the lowest radar elevation is 0.5°) and the surface can exceed 1.0 km at 80 km away from the radar site, and could be even more extended under sub-refraction conditions. This blind region to the radar largely lies in the planetary boundary layer (PBL), in which PBL processes such as horizontal convective rolls, dry-lines, and cold-pools can trigger convection under a favorable large-scale environment (Sobash and Stensrud, 2015). In particular, surface-based cold pools produced by the evaporation of precipitation can influence the initiation, development and propagation of convective systems (e.g., Tompkins, 2001; Lima and Wilson, 2008). The intensity of cold pool and environmental vertical wind shear are critical factors in determining a convective system's structure and evolution (e.g., Ro-

\*Corresponding author: Kun ZHAO  
Email: zhaokun@nju.edu.cn

tunno et al., 1988; Weisman et al., 1988; Bryan et al., 2006; Parker, 2010). These PBL processes cannot be resolved well if only radar data are assimilated in NWP. Data from surface observations have been used in the past to detect these processes and will be more valuable if assimilated into NWP models along with radar data.

In the past several years, with the increase and improvement of surface observation networks [e.g., automatic weather stations (AWSs)], surface data at high spatial (<10 km) and temporal (~5 min) resolution have become commonly available in real time. Rapidly evolving surface mesoscale features that are important to the analysis and forecasting of convective systems can be monitored routinely. Therefore, it is expected that assimilating these surface observations into NWP models at frequent intervals can fill the low-level blind region of radar detection and yield a better analysis and forecast, since PBL processes play an essential role in convective systems.

Several studies have been conducted to assimilate surface observations together with radar observations with a convection-permitting model. For example, Zhang et al. (2004) found that when low-level radar observations are missing, the EnKF (ensemble Kalman filter) can provide a better estimate of the storm with the incorporation of surface wind and temperature observations. Dong et al. (2011) indicated that when a radar is far away from a convective system, a positive impact results from assimilating surface observations, as long as the network spacing is less than 20 km. Through assimilating high-resolution Doppler radar radial velocities and in situ surface observations into an EnKF system, Marquis et al. (2014) successfully simulated mesocyclone-scale processes in the Goshen County, Wyoming, and the tornadic supercell of 5 June 2009. Yusouf et al. (2015) developed a multiscale ensemble-based assimilation and prediction system that can assimilate the radar, surface, and other observations together. Using this system, they analyzed the storm-scale features of the 27 April 2011 Alabama severe weather outbreak. Through assimilating radar and surface observations with an EnKF in combination with a double-moment microphysics scheme, Putnam et al. (2014) analyzed the microphysical states and precipitation structures within a mesoscale convective system that passed over western Oklahoma during 8–9 May 2007. Schenkman et al. (2011b) examined the impact of radar and Oklahoma Mesonet data assimilation on the prediction of mesovortices in a tornadic mesoscale convective system (MCS) through a three-dimensional variational data assimilation (3DVAR) system. Results again confirmed the positive impact of surface data in the absence of radar data. Snook et al. (2015) analyzed the same tornadic MCS through an EnKF system and found that assimilating the convective observations (including surface, wind profiler, and upper-air observations) together with radar observations resulted in better meso- and convective-scale features. Similar conclusions, regarding the prediction of heavy rainfall over southern China, were reached in Hou et al. (2013).

Despite positive results, surface observations are still un-

derused in operational data assimilation systems (Pu et al., 2013). To the best of the authors' knowledge, there have been no studies that have assessed the impact of assimilating surface observations along with radar data using a four-dimensional variational (4DVAR) data assimilation system. 4DVAR has been proven to be more accurate than 3DVAR. Moreover, when compared with the EnKF, it allows for the assimilation of asynchronous observations with serially correlated errors by including time correlations (Kalnay et al., 2007). However, the 4DVAR method requires an adjoint model, whose numerical structures must closely follow those in the prognostic model (Tai et al., 2011). Therefore, high computational cost is accrued, and greater complexity expected, in 4DVAR systems.

In this study, we describe a surface data assimilation method in the 4DVAR radar data assimilation system VDRAS (Variational Doppler Radar Analysis System) (Sun and Crook, 1997) developed by NCAR. A squall-line case observed during the Observation, Prediction and Analysis of Severe Convection of China (OPACC) field campaign in 2014 is used to test the impact of the surface data assimilation. The objective is to examine whether additional surface observations can improve the analysis and short-term forecast ability of VDRAS.

The rest of this paper is organized as follows: In section 2, VDRAS and the new development for surface data assimilation are described. An overview of the squall-line case observed in OPACC is introduced in section 3. The experimental design is described in section 4. Results and verification are presented in section 5. A summary and discussion are given in section 6.

## 2. Surface data assimilation in VDRAS

### 2.1. VDRAS description

VDRAS is a convective-scale data assimilation system based on the 4DVAR technique with a cloud model as its constraint. It has been successfully used as the analysis and short-term forecast system of the Beijing Summer Olympics (Sun et al., 2010) and is still in operation at NCAR, the Beijing Meteorological Bureau and Taiwan Central Weather Bureau. The main objective of VDRAS is to produce rapidly updated (in the order of minutes) analysis for severe weather analysis and nowcasting, which is made possible by using a short 4DVAR assimilation window that covers three volumes of radar observations. The main data sources of VDRAS are from Doppler radar and surface networks, while other large-scale data are indirectly incorporated by using a high-resolution mesoscale model's analysis/forecast as the background and boundary conditions. Through the cloud model and the 4DVAR scheme, VDRAS can retrieve the unobserved (by radar) temperature, wind and other microphysical variables by assimilating reflectivity and radial velocity observations from a single or multiple radar networks. Readers are referred to Sun and Crook (1997, 2001) for the system's design and cloud model, Chang et al. (2015) for the recent ad-

dition of an ice physics scheme, and Sun and Crook (2001), Crook and Sun (2004), and Sun et al. (2010) for an evaluation of the system's operational performance.

In this study, the operational WRF forecast during OPACC at 4 km spatial resolution is used to provide the background and boundary conditions. The operational WRF forecast is performed over a single domain that covers most parts of China with  $1409 \times 1081$  horizontal grid points and 27 vertical levels. The WRF model employs the ACM2 boundary layer scheme (Pleim, 2007), the Pleim–Xiu surface model (Pleim, 2007), the CAM longwave and shortwave radiation schemes, and the WRF Morrison double-moment microphysics scheme (Morrison et al., 2009). In addition, the VAD (velocity azimuth display) wind profile (Klazura and Imy, 1993) for each radar has been computed and blended with the WRF background to produce a gridded mesoscale analysis using a Barnes interpolation technique (Barnes, 1964). This mesoscale analysis is used as a first guess in the first analysis cycle. Each analysis cycle contains a 15-min assimilation window and a 5-min short forecast to provide background for the next cycle. In each cycle, an optimal initial state between the background field and radar observations is calculated by minimizing a cost function ( $J$ ),

$$J = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_b)^T + \frac{1}{2} \sum_{\sigma, t} [\eta_v(v_r - v_{0,r})^2 + \eta_q(q_r - q_{0,r})^2] + J_p + J_{mb}. \quad (1)$$

In Eq. (1), the summation is conducted over space ( $\sigma$ ) and time ( $t$ ),  $\mathbf{x}_0$  represents the model state variables at the beginning of the current cycle,  $\mathbf{x}_b$  stands for the background information, and  $\mathbf{B}$  is the covariance matrix of background error. The weighting coefficients  $\eta_v$  and  $\eta_q$  stand for the inverse of the observational error variance for radial wind and rainwater, respectively. Uncorrelated observation error is assumed. The model-generated Doppler radial velocity and rainwater mixing ratio are represented by  $v_r$  and  $q_r$ , while their observed counterparts are expressed by  $v_{0,r}$  and  $q_{0,r}$ .  $J_p$  is a penalty term for additional constraints to preserve the temporal and spatial smoothness of the analysis.  $J_{mb}$  stands for the background penalty term that measures the difference between the 4DVAR analysis and the mesoscale analysis to ensure that the 4DVAR radar analysis is not too far from the mesoscale analysis in the region outside the radar coverage. The observed rainwater mixing ratio is estimated from radar reflectivity by the following formula, which assumes a Marshall–Palmer raindrop size distribution (Sun and Crook, 1997):

$$q_r = \frac{1}{\rho} 10^{[(Z-43.1)/17.5]}, \quad (2)$$

where  $Z$  is the reflectivity in dBZ and  $\rho$  is the density of air. A complete description of VDRAS can be found in Sun and Crook (2001), and Crook and Sun (2004).

Many previous studies have shown that VDRAS can retrieve the meteorological variables of convective storms through the depth of the troposphere (e.g., Sun and Zhang,

2008). For real-time applications, however, VDRAS is usually implemented to retrieve only low-level wind and temperature, due to limited computation resources Sun et al. (2010). In this paper, we retrieve the meteorological variables of the squall line up to 15 km, with a focus on the improvements at the low levels.

## 2.2. Improvements to the surface data assimilation in VDRAS

In its original version, surface observations are used to improve the retrieval capability of VDRAS. However, they are used to correct the WRF analysis/forecast prior to the 4DVAR radar data assimilation using a 3D analysis method in the mesoscale background analysis step (Crook and Sun, 2004). Surface observations are interpolated horizontally onto the VDRAS grid using the Barnes interpolation scheme, and then vertically combined with the VAD-corrected WRF background data (described in section 2.1) using a least-squares fitting method. The reason for such a design is that, in the past, most of the available surface observations were from the conventional synoptic network. The spatial and temporal resolutions of the conventional surface observations are much sparser and less frequent than radar observations, so that they only stand for the large-scale environmental conditions. This two-step approach can provide a more detailed analysis of the convective-scale features from the radar observations by using a small assimilation window and a small length scale for the background error covariance represented in a recursive filter (Hayden and Purser, 1995; Sun and Zhang, 2008).

With the development of AWS surface observation networks, mesoscale and in some cases convective-scale processes, can be captured. The temporal resolution of surface observations can reach 5 min, which is similar to the duration of radar volume scan. Concurrent assimilation of surface and radar observations in 4DVAR is expected to yield a better mesoscale analysis and short-term forecast than the original two-step approach in which the surface observations are firstly interpreted to the background analysis, and then the radar data assimilated through the 4DVAR scheme. In the current work, surface temperature and wind are assimilated with radar reflectivity and radial velocity simultaneously, with the following cost function:

$$J = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_b)^T + \frac{1}{2} \sum_{\sigma, t} [\eta_v(v_r - v_{0,r})^2 + \eta_q(q_r - q_{0,r})^2 + \eta_T(T_s - T_{0,s})^2 + \eta_U(U_s - U_{0,s})^2 + \eta_V(V_s - V_{0,s})^2] + J_p + J_{mb}. \quad (3)$$

Three additional terms are included in Eq. (3), representing surface temperature ( $T_s$ ) and the  $x$  and  $y$  components of the surface winds ( $U_s, V_s$ ). The observed surface 10 m winds ( $U_{10m}, V_{10m}$ ) from AWSs are extrapolated to the lowest model level ( $U_{0,s}$  and  $V_{0,s}$ ) based on an empirical power law relation in the surface layer (Peterson and Hennessey, 1978),

$$\frac{U_{10m}}{U_{0,s}} = \left(\frac{10}{Z_s}\right)^\alpha, \quad (4)$$

where  $Z_s$  is the height of the lowest model grid level, set to 100 m in this study, which is around the top of the surface layer. The exponent  $\alpha$  is an empirically derived coefficient that depends on the local stability. For neutral conditions,  $\alpha$  is approximately 0.143 (Hsu et al., 1994). Estimated temperature at the lowest model level ( $T_{0,s}$ ) is extrapolated from the AWS 2 m temperature using a lapse rate of  $6.5^\circ\text{C km}^{-1}$ . The weighting coefficients  $\eta_T, \eta_U$  and  $\eta_V$  stand for the inverse of the respective observation error variances. In this study, all surface weighting coefficients are taken as constants (set to 1.0) for simplicity. The optimal weighting coefficients will require further study. Corresponding to the new cost function, the adjoint model of VDRAS is modified and the accuracy of the adjoint model verified following the method in Sun and Crook (1997).

### 3. Overview of the squall-line case

The squall-line case studied in this paper is selected from the OPACC field campaign. OPACC is a 5-year project focused on studying the dynamical and microphysical characteristics of convective weather systems over eastern China, as well as the observation techniques and forecasting methods for such systems. The first field campaign of OPACC was conducted from 1 June to 3 August 2014 over the Yangtze River–Huaihe River (YHR) basin. Several quasi-linear convective systems and two squall lines were observed on 30 July 2014 during the intensive observing period (IOP) 10 of OPACC. A tropical cyclone was situated over the South China Sea at that time. Its circulation transported warm and moist air from the ocean inland to the YHR basin, setting up a moist and unstable environment that was favorable for convective initiation. Figure 1 shows the merged composite radar reflectivity from 0830 to 1600 UTC 30 July 2014, using six operational radars (FYRD, BBRD, NJRD, HFRD, TLRD and AQRD in Fig. 2). The merged composite radar reflectivity is computed at each grid point by taking the maximum reflectivity among all six radars. A convective system was initiated near 0330 UTC over the northern part of the observation domain in Fig. 1 (not shown). It organized into a west–east oriented squall line (the first squall line) with extreme rainfall and high surface wind 5 hours later (Fig. 1a). In the next three hours, this squall line moved to the central part of the domain and broke into several north–south oriented quasi-linear MCSs around 1130 UTC (Fig. 1b). By 1430 UTC, a new squall line had been triggered on the east side of these MCSs (Fig. 1c). The second squall line matured at around 1600 UTC and formed a bowing structure with a maximum reflectivity of greater than 55 dBZ (Fig. 1d). Extreme surface wind (up to  $19.2 \text{ m s}^{-1}$ ) and precipitation (up to  $56.6 \text{ mm h}^{-1}$ ) were also observed at this time.

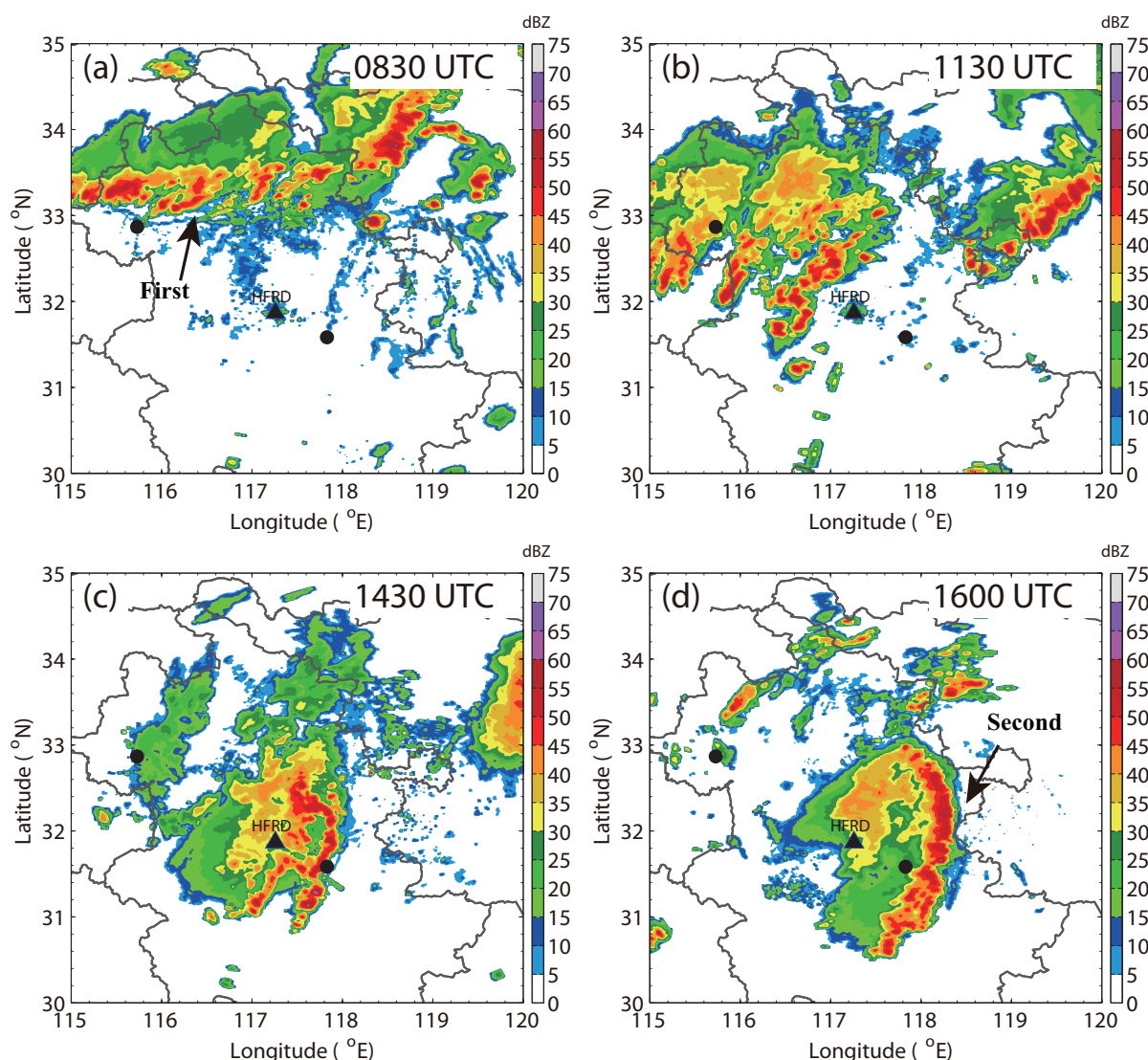
### 4. Experimental design

Two single observation experiments (SP-AWS and SP-NOAWS) were designed to investigate the impact of assim-

ilating single-point surface observations on the VDRAS analysis in both the horizontal and vertical directions. The domain of the single observation experiments is shown by a red dashed box in Fig. 2. Reflectivity and radial velocity from the Hefei radar (HFRD in Fig. 2) are assimilated together with the surface wind and temperature observations from the AWS (red dot in Fig. 2) in SP-AWS, whereas HFRD data alone are assimilated in SP-NOAWS. The domain of the two experiments has  $80 \times 80$  horizontal grid points and 50 vertical levels. The horizontal resolution of these two experiments is 3 km and the vertical grid size is 200 m, with the lowest model level at 100 m above the surface. The data assimilation procedure starts at 1300 UTC 30 July 2014, and ends at 1620 UTC. This time period includes the initiation and development stages of the second squall line documented in section 3 (Figs. 1c and d). Altogether, eleven analysis cycles are included. Each cycle has a 15 min 4DVAR time window, followed by a 5-min forecast, to provide background for the next cycle. The short forecasts between two assimilation windows are added such that no observations will be used twice (Sun and Zhang, 2008). The typical volumetric scan for radars is 5–6 min and the time resolution of AWS observations is 5 min; two to three radar volumetric scans and three surface observations are included in each assimilation window. Increments (SP-AWS minus SP-NOAWS) of different model variables (temperature, pressure and horizontal wind) are analyzed in section 5.1.

Three assimilation experiments (RAD-ONLY, SURF-MESO and SURF-4DVAR) are conducted to examine whether including the AWS observations can improve VDRAS with multiple radar assimilations. The model domain for these three experiments is larger than the single observation experiment, in order to include all seven radars. It has  $140 \times 140$  horizontal grid points and 50 vertical levels (see Fig. 2). The grid spacing is the same as the single observation experiments. In RAD-ONLY, surface data are not used either in the background analysis step or in the 4DVAR cost function. Radar reflectivity and radial velocity data from six S-band operational radars (FYRD, BBRD, NJRD, HFRD, TLRD and AQRD) and one C-band research radar developed by Nanjing University (NJU-CPOL in Fig. 2) are assimilated through the 4DVAR cost function shown in Eq. (1), using the VAD-corrected WRF 4-km data as the mesoscale background. In the experiment SURF-MESO, the original surface data assimilation method is applied, i.e., the AWS data are used in the background analysis prior to the 4DVAR radar data assimilation, as described in section 2. In SURF-4DVAR, the background analysis is the same as RAD-ONLY but additional AWS data are assimilated in the 4DVAR cycles using the cost function in Eq. (3).

The three assimilation experiments start at 0700 UTC 30 July 2014, and end at 1940 UTC. The period extends from the initiation of the first quasi-linear MCS to the dissipation of the second bowing squall line (Fig. 1). Time lengths of assimilation windows are 15 min, followed by a 5-min forecast—the same as in the single observation experiments. A 1-h forecast is performed at the end of each analysis cycle to further



**Fig. 1.** Mosaics of composite radar reflectivity at (a) 0830, (b) 1130, (c) 1430 and (d) 1600 UTC. Provincial borders are superimposed as black contours. The location of Hefei radar is represented by the black triangle, and the two surface stations used for independent verification are marked by black dots.

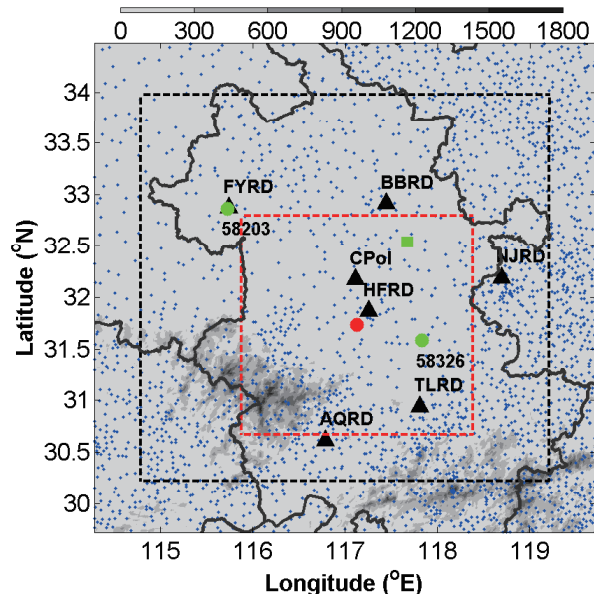
investigate the influence of surface data assimilation on the forecast. Data from two AWSs (green dots in Fig. 2) are not assimilated in SURF-MESO or SURF-4DVAR, and are used for verification. A wind profiler (green square in Fig. 2) is also used to examine the effects of surface data assimilation on low-level wind fields.

## 5. Results and verification

### 5.1. Single AWS observation experiment

Figure 3 shows the analysis increments of temperature, pressure and horizontal wind fields on the lowest model level (100 m) between SP-AWS and SP-NOAWS at 1500 UTC 31 July 2014. The second squall line evolved into a bowing structure on the east side of the dissipating quasi-linear MCS at this time. The assimilation of a single surface observation produces near-surface cooling around the AWS station (Fig.

3a) and intensifies the surface-based cold pool strength of the old MCS. The second squall line is located at the leading edge of this intensified cold pool. Spatial inhomogeneity can be found with a stronger cooling over the region covered by the new squall line (around 117.8°E), which represents the intensification of the cold pool produced by the second squall line after surface data are assimilated. Though the pressure field has not been assimilated through the 4DVAR scheme directly, it is adjusted through the cloud model and its adjoint model in the 4DVAR system during the assimilation cycles, wherever the pressure responds to the analyzed temperature. The analysis increment of the pressure field has a similar spatial distribution pattern as the temperature field. Increases in pressure are found over the old and new cold pool area produced by the original MCSs and the second squall line, respectively. Analysis increments in  $U$  and  $V$  wind components are shown in Figs. 3b and c. The spatial inhomogeneity of the wind fields is much clearer than for the temperature and pressure fields.



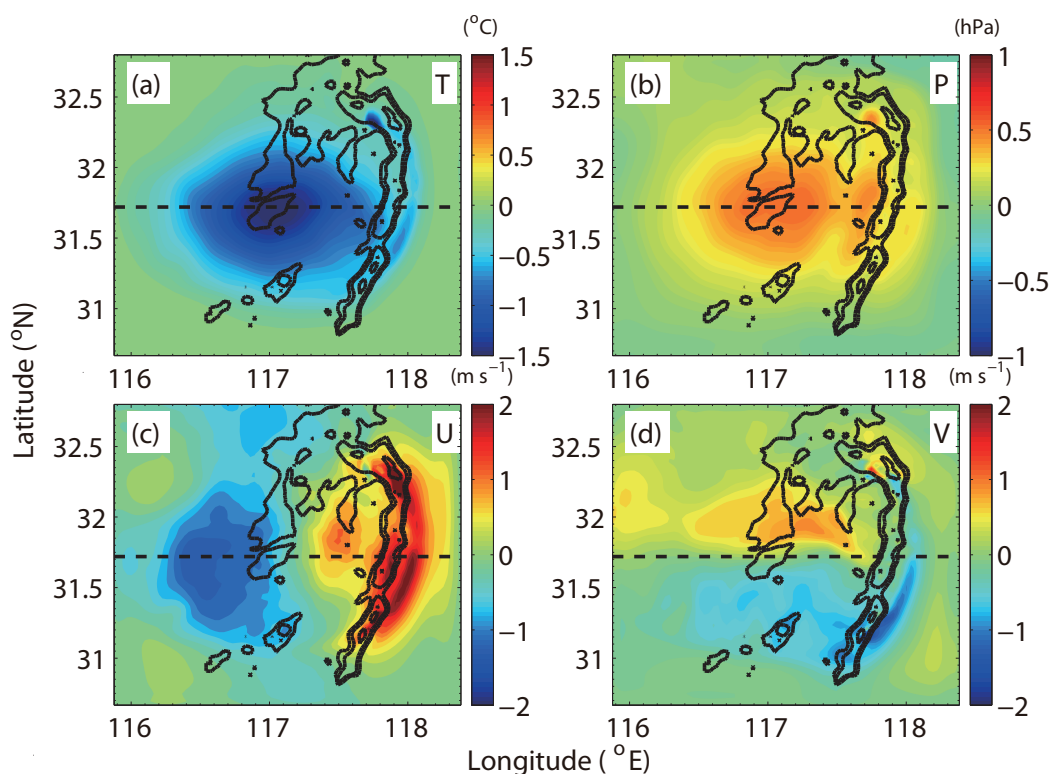
**Fig. 2.** Analysis and forecast domains for the real-data assimilation experiments (black dashed box) and the single observation experiment (red dashed box). The locations of the six S-band operational radars and one C-band research radar (CPol) are marked by the black triangles. The blue dots represent AWSs. Two AWSs (ID: 58203 and 58326) and a wind profiler used for verification are marked by the green dots and square. The AWS used in the single observation experiment is marked by the red dot. Orography (units: m) is shown in gray-scale, and provincial borders by black lines.

Incremental divergence can be found under the quasi-linear MCS, corresponding to the region with decreasing temperature and increasing pressure. Assimilation of the single-point surface observations also increases the strength of the outflow to the east of the new squall line.

Vertical cross sections along the black dashed lines in Fig. 3 are presented in Fig. 4. Two convective systems are seen from the radar reflectivity contours, the old quasi-linear MCS in the center and the newly developed squall line on the right. As in Fig. 3, strengthening of the cold pool under the MCS is indicated by decreased temperature in Fig. 4a, and increased pressure in Fig. 4b, below 1.5 km. Meanwhile, the incremental strengthening of the squall line's cold pool is not clear over this cross section. This squall line was formed shortly before 1500 UTC; its cold pool is still very shallow at this time and can only be found at the lowest levels (Fig. 3b). Incremental low-level and high-level divergence of the MCS in the center can also be found in Figs. 4c and d. The gust front produced by the squall line is also intensified at around 118°E (Fig. 4c).

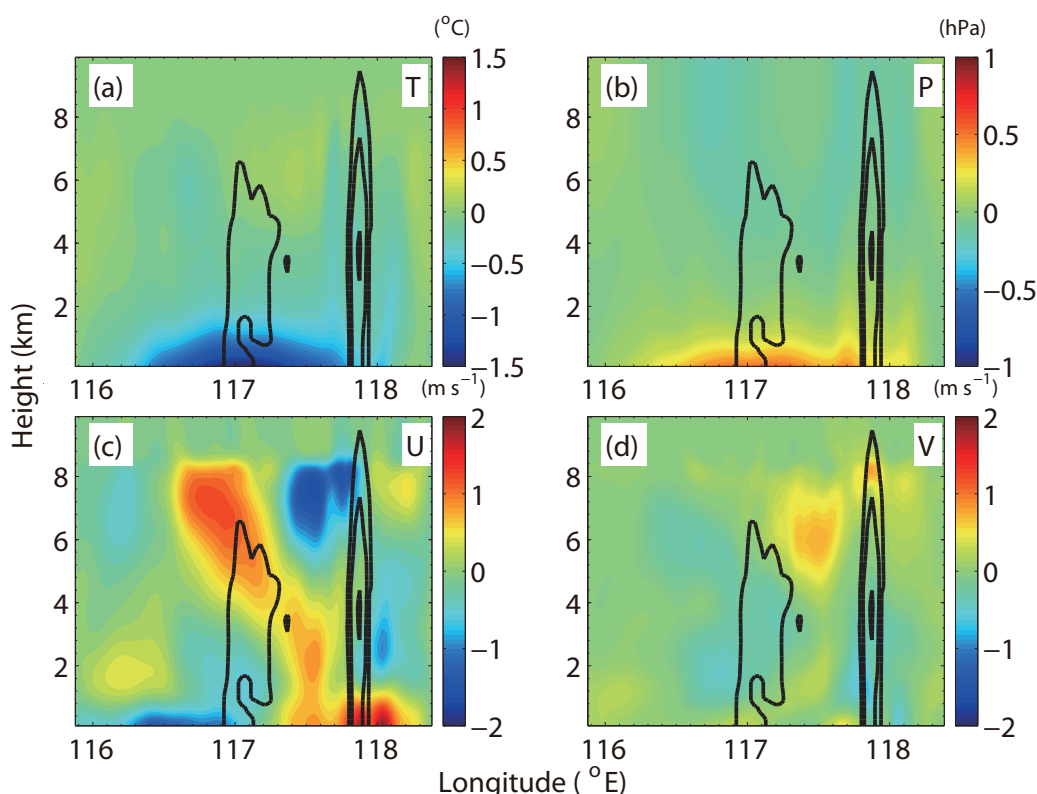
## 5.2. VDRAS analyses of the real data experiments

The AWS observed surface wind and temperature fields at 1500 UTC are shown in Fig. 5a, with the observed composite radar reflectivity superimposed as black contours. As seen in Fig. 1, this is the time when the convective systems have moved from the northern part to the central part of the domain. Heavy rainfall occurred over the path of the system, so the surface temperature over these areas is much colder



**Fig. 3.** Horizontal distributions of the analysis increments (SP-AWS minus SP-NOAWS) of (a) temperature, (b) pressure, (c)  $U$  wind and (d)  $V$  wind, on the lowest model level (100 m) at 1500 UTC 31 July 2014. Composite radar reflectivity is shown in black contours, starting from 30 dBZ and with an interval of 10 dBZ.





**Fig. 4.** Vertical cross sections of the analysis increments of (a) temperature, (b) pressure, (c)  $U$  wind and (d)  $V$  wind, along the black dashed line in Fig. 3. Observed radar reflectivity is shown in black contours, starting from 30 dBZ and with an interval of 10 dBZ.

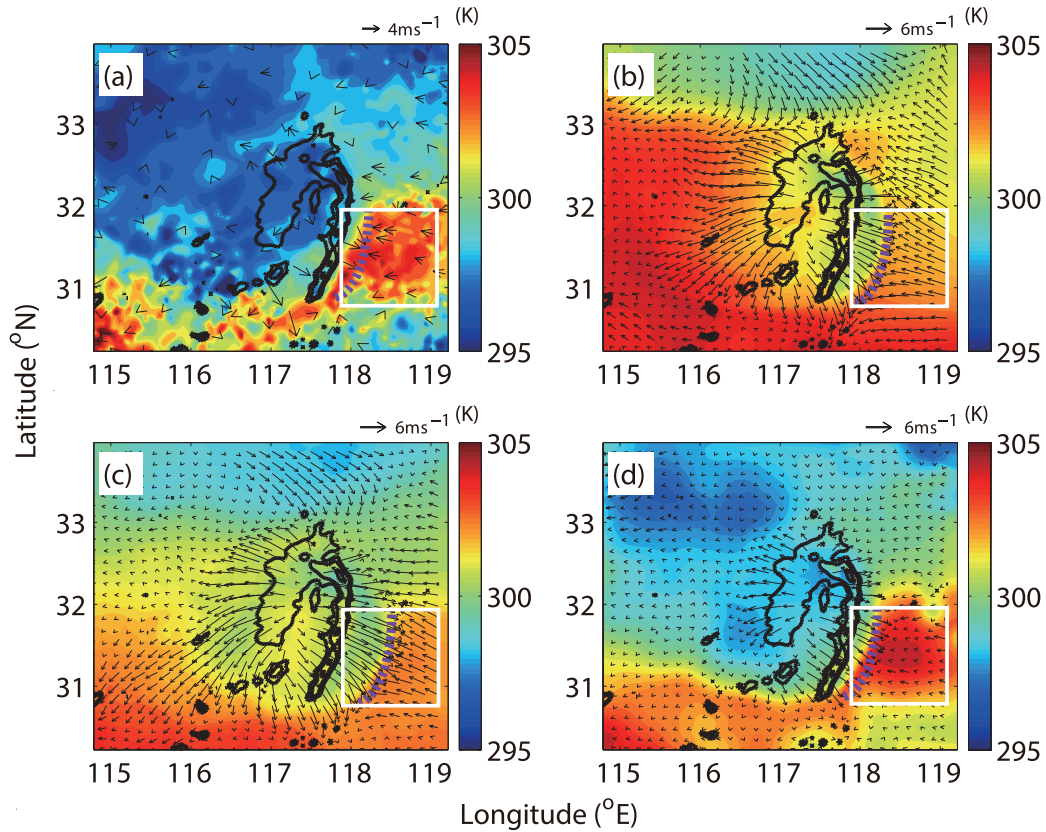
than in the southern part. Based on the surface wind observations, a clear gust front, marked by the black dashed line in Fig. 5a, and warm surface inflow, occur over the southeast side of the squall line.

The estimated surface wind and temperature fields at 1500 UTC from the three assimilation experiments are shown in Figs. 5b–d. Surface wind and temperature are extrapolated from the model temperature and wind on the lowest grid level at 100 m. When only radar data are used in the 4DVAR assimilation, large differences can be found between the AWS observation (Fig. 5a) and the RAD-ONLY analysis (Fig. 5b). The surface temperature of the northern and central domain in the RAD-ONLY analysis is much warmer than the AWS observations, especially for the region without the radar coverage. On the other hand, the surface warm inflow on the southeast side of the squall line (shown by the white box in Fig. 5) is colder than observed. Surface-based cold pools of the old quasi-linear MCS and the bowing squall line are weaker. However, surface outflow from the cold pools is much stronger than in the AWS observations, and the position of the gust front is also further ahead of the squall line. In the RAD-ONLY experiment, the surface wind analysis is a combination of the upper-level radar radial velocity observations and the background field through the spatial correlation implied in the cloud model, which also results in the temperature increment in the region with radar data through the cross-variable correlation in the cloud model. However, in the re-

gion without radar observations, both the wind and temperature analyses are mainly from the WRF background. Without the surface observations, the errors in the background cannot be corrected. That is the main reason for the large temperature error in the environment surrounding the convective system. It is also noted that in the region with radar observations, the temperature still shows a few differences from the AWS observations, which can be partly attributed to the inadequate sample of radar data at low levels and the uncertainty of the cross-variable correlation in the cloud model.

In SURF-MESO, the AWS observations are used during the background analysis step. This results in an improvement of the surface temperature analysis compared to RAD-ONLY (Fig. 5c). Nevertheless, significant differences can still be found in the surface temperature, the southeast inflow, the cold-pool outflow and the position of the gust front. When AWS observations are assimilated together with radar data in the SURF-4DVAR experiment, the modeled surface temperature and wind fields are much closer to the observations (Fig. 5d). The northern and central domain is much colder, while southeast surface inflow is much warmer than that in the other two experiments. The strength of the cold pool, southeast inflow, surface outflow and the location of the gust front are also close to the AWS observations.

In order to quantitatively compare the analysis between the three experiments and the AWS observations, the gridded surface temperature and wind are interpolated to the AWS



**Fig. 5.** Surface temperature (color-shaded) and wind vectors (black arrows) at 1500 UTC 30 July 2014 from (a) AWS observations, (b) RAD-ONLY, (c) SURF-MESO and (d) SURF-4DVAR. The position of the gust front is shown as a purple dashed line, and the observed composite radar reflectivity—starting from 30 dBZ and with an interval of 10 dBZ—is superimposed in all subfigures as black solid lines. The surface inflow region on the east side of the squall line is shown by a white box.

locations. The mean temperature (MTD) and wind (MVD) differences are calculated:

$$\text{MTD} = \frac{1}{N} \sum_{k=1}^N \sqrt{(T_{\text{aws}} - T)^2}, \quad (5)$$

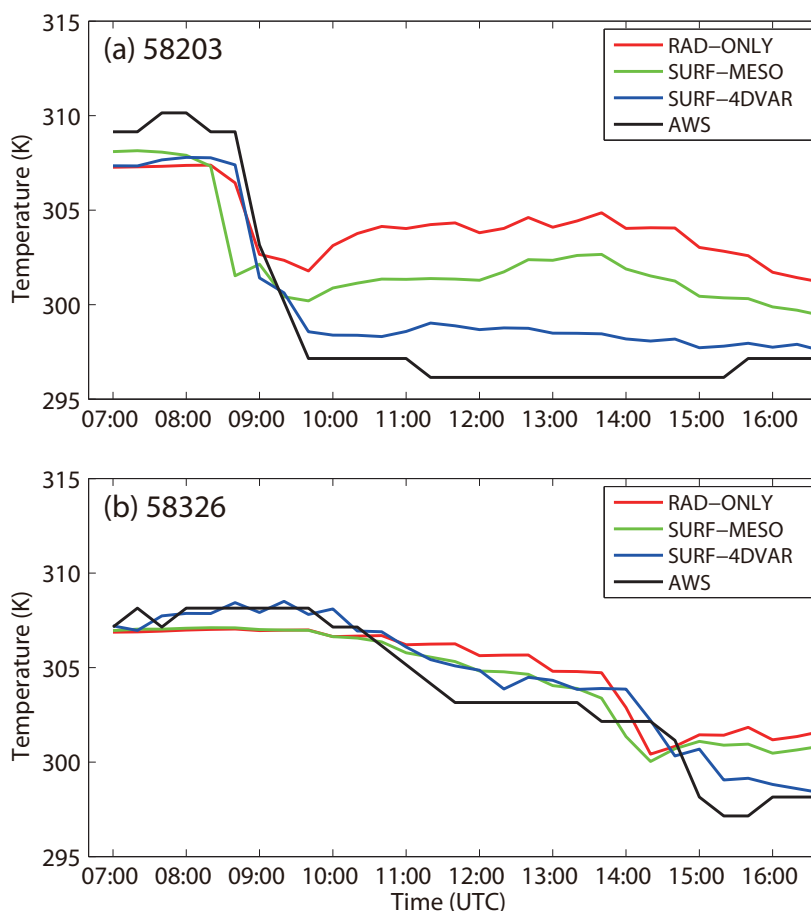
$$\text{MVD} = \frac{1}{N} \sum_{k=1}^N \sqrt{(U_{\text{aws}} - U)^2 + (V_{\text{aws}} - V)^2}, \quad (6)$$

where  $N$  represents the number of AWSs.  $T_{\text{aws}}$ ,  $U_{\text{aws}}$  and  $V_{\text{aws}}$  stand for the observed surface temperature and winds from each AWS. The extrapolated model surface temperature and winds at the AWS locations are represented by  $T$ ,  $U$  and  $V$ . The MTDs at 1500 UTC in RAD-ONLY, SURF-MESO and SURF-4DVAR are 3.49 K, 2.38 K and 1.23 K, and the MVDs are  $3.91 \text{ m s}^{-1}$ ,  $2.52 \text{ m s}^{-1}$  and  $2.13 \text{ m s}^{-1}$ , respectively. Comparison of SURF-MESO with RAD-ONLY suggests that assimilating surface observations can help improve the near-surface fields of VDRAS. The further improvement of SURF-4DVAR over SURF-MESO demonstrates the advantages of the newly developed 4DVAR surface data assimilation methodology in comparison with the original methodology.

Two independent AWSs (Nos. 58203 and 58326—marked by green dots in Fig. 2) are chosen to further examine the improvement of VDRAS analysis using the new methodology in SURF-4DVAR. Station 58203 is located in the northwest. It recorded the surface temperature change during the passage of the first west–east oriented squall line at around 0840 UTC. Similarly, station 58326 is chosen for its spatial proximity to the second squall line. In Fig. 6a, a dramatic decrease in surface temperature ( $\sim 13 \text{ K}$ ) between 0830 and 0930 UTC is found at station 58203 after the passage of the squall line. In RAD-ONLY, the timing of the surface temperature drop is close to the observations, while the amplitude ( $\sim 4.5 \text{ K}$ ) is much smaller. For SURF-MESO, the amplitude of the temperature drop is 7.3 K—closer to the observations than RAD-ONLY. The timing of the initial temperature drop, however, is 20 minutes earlier than both the observations and the RAD-ONLY experiment. In SURF-4DVAR, the surface temperature decreases by about 10.3 K—even closer to the observed magnitude. The timing of the event also agrees well with the observations.

In Fig. 6b, before the initiation of the second squall line, simulated surface temperatures at station 58326 are close to one another, and in general agree with the observations. Bet-





**Fig. 6.** Time series of observed and analyzed surface temperature at AWS No. (a) 58203 and (b) 58326.

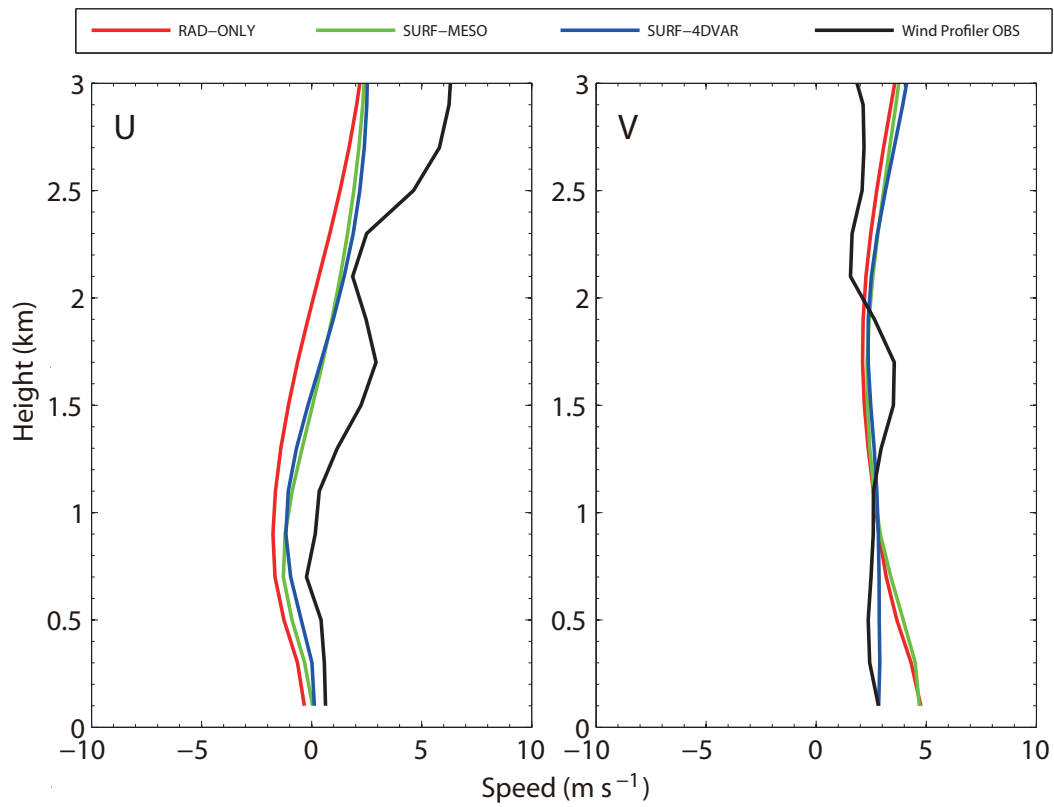
ter results are found in SURF-4DVAR over the other two. After the passage of the squall line, surface temperature decreases in all three experiments. The magnitude and timing of the surface temperature drop in SURF-4DVAR is the closest to the observations among the three. The MTDs of this surface station are 1.82 K, 1.42 K and 0.91 K for RAD-ONLY, SURF-MESO and SURF-4DVAR, respectively. The comparisons with the two representative stations again confirm the improvement of VDRAS analysis by assimilating surface observations together with radar data.

Besides the improvements of the near-surface fields, assimilating surface observations may also improve the analysis results at higher levels through the 4DVAR scheme. To investigate whether the surface data assimilation can improve the VDRAS analysis in the PBL and above, observations from a nearby wind profiler (green square in Fig. 2) are compared for further validation. Observed and analyzed vertical profiles of  $U$  and  $V$  below the 3 km altitude at 1500 UTC are presented in Fig. 7. The vertical profiles of  $U$  in SURF-MESO and SURF-4DVAR are closer to the observations than those in RAD-ONLY. A slight improvement is achieved in SURF-4DVAR in the lowest 1 km for  $U$ , while a significant improvement is found for  $V$ , where the analysis results are closest to

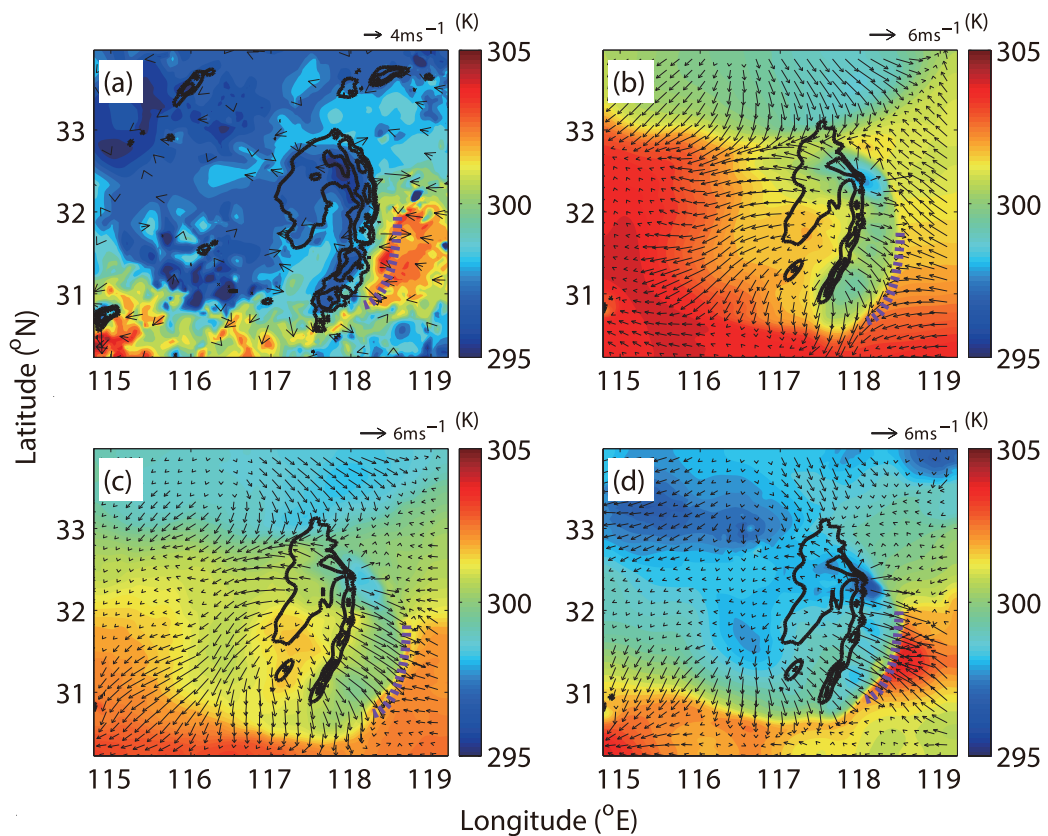
the observations. There is little difference between the three experiments above 1 km. The MVDs below 3 km are also calculated for each experiment, and they are  $2.69 \text{ m s}^{-1}$ ,  $2.15 \text{ m s}^{-1}$  and  $1.82 \text{ m s}^{-1}$  in RAD-ONLY, SURF-MESO and SURF-4DVAR, respectively. The comparison results show that assimilating surface data through the new 4DVAR scheme can improve the analyzed low-level wind profiles in VDRAS, with the most significant improvement found within the PBL.

### 5.3. Evaluation of 1-h nowcasts

To validate the improvement of the very short-term forecasting skill of VDRAS through assimilating the surface data, a 1-h nowcast is launched from each analysis cycle in all three assimilation experiments (RAD-ONLY, SURF-MESO and SURF-4DVAR). Observed surface temperature, surface wind and composite radar reflectivity at 1600 UTC are compared with the 1-h nowcast fields starting from the analysis at 1500 UTC in Fig. 8. Overall, the forecasted reflectivity, surface wind and temperature from SURF-4DVAR are the best among three experiments. The one-hour forecasted temperature field in RAD-ONLY at 1600 UTC (Fig. 8b) is much warmer than observed for most parts of the domain, just as the analyzed surface field is at 1500 UTC (Fig. 5b). The cold



**Fig. 7.** Vertical profiles of  $U$  and  $V$  from wind profiler observations (black line) and VDRAS analysis at 1500 UTC 30 July 2014.



**Fig. 8.** As Fig. 5, but for the 1-h forecasted field at 1600 UTC.

pool is too weak and the surface outflow on the west side of this squall line is too strong. In SURF-MESO, the surface temperature is closer to the AWS observations than the RAD-ONLY forecast (Fig. 8c), but the cold pool is still too weak and the outflow is again too strong. Compared with the other two experiments, the 1-h forecasted surface fields in SURF-4DVAR are much closer to the observations (Fig. 8d). The northern and central parts of the OPACC domain are colder than the southern domain, the cold pool is stronger, the surface outflow is weaker, and the surface inflow on the east side of the squall line is warmer. The position of the gust front ahead of the squall line is also close to the AWS observations. Quantitatively, the MTDs of the RAD-ONLY, SURF-MESO and SURF-4DVAR experiments are 4.44 K, 2.48 K and 1.26 K, and the MVDs are  $4.09 \text{ m s}^{-1}$ ,  $2.42 \text{ m s}^{-1}$  and  $2.29 \text{ m s}^{-1}$ , respectively. All these results indicate improved forecast ability with the additional assimilation of the surface observations in VDRAS. It is noteworthy that the cold pool for the squall line (Fig. 8a) remains close to the leading edge of the convection in the observations at 1600 UTC, while it appears to surge quickly outward in the 1-h forecasts of the three experiments. This disagreement may be induced by the lack of PBL processes in VDRAS, which can result in a stronger cold pool in the 1-h forecast.

Comparing the temperature time series of the observa-

tions at the two AWS stations (Figs. 9a and b) with that of multiple 1-h forecasts from the continuous analyses for the three experiments, SURF-4DVAR performs much better at capturing the surface temperature drop after the passage of the two squall lines. The MTDs of station 58203 (58326) are 5.61 K (2.2 K), 3.96 K (1.81 K) and 1.49 K (1.45 K) in RAD-ONLY, SURF-MESO and SURF-4DVAR respectively. Figure 10 compares the observed and forecasted vertical profiles of  $U$  and  $V$  at 1600 UTC. Both the  $U$  and  $V$  components from SURF-4DVAR agree better with the wind profiler observations than the rest. The primary improvement is found again in the PBL below 1 km—similar to the analysis results in Fig. 7. The MVDs below 3 km are  $5.35 \text{ m s}^{-1}$ ,  $4.46 \text{ m s}^{-1}$  and  $3.46 \text{ m s}^{-1}$  for RAD-ONLY, SURF-MESO and SURF-4DVAR, respectively. These validation results show that assimilating surface data together with radar observations not only improves the representation of the surface winds and temperature, but also the low-level dynamic and thermodynamic fields, especially within the PBL.

Compared with the temperature and wind fields, the 1-h forecast of the precipitation is not as significantly improved by the surface data assimilation. Quantitatively, the RMSE of composite radar reflectivity is 11.11 dBZ in SURF-4DVAR—slightly smaller than the 11.39 dBZ and 11.35 dBZ in RAD-ONLY and SURF-MESO for the 1-h nowcast from 1500

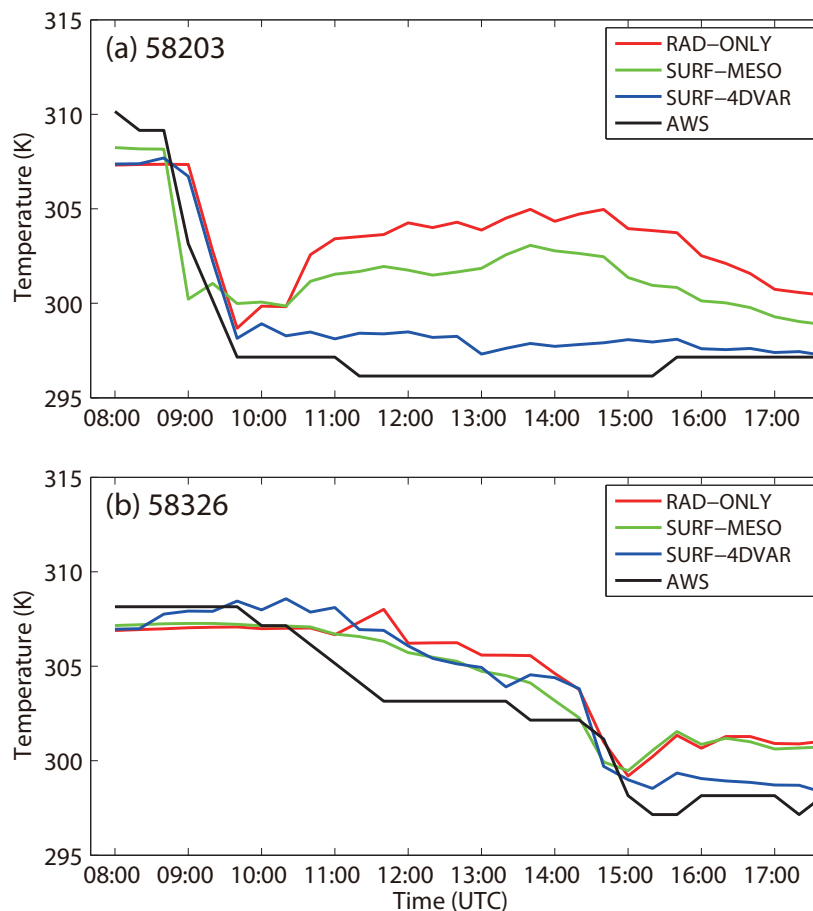


Fig. 9. As in Fig. 6, but for the observed and 1-h forecasted surface temperature.

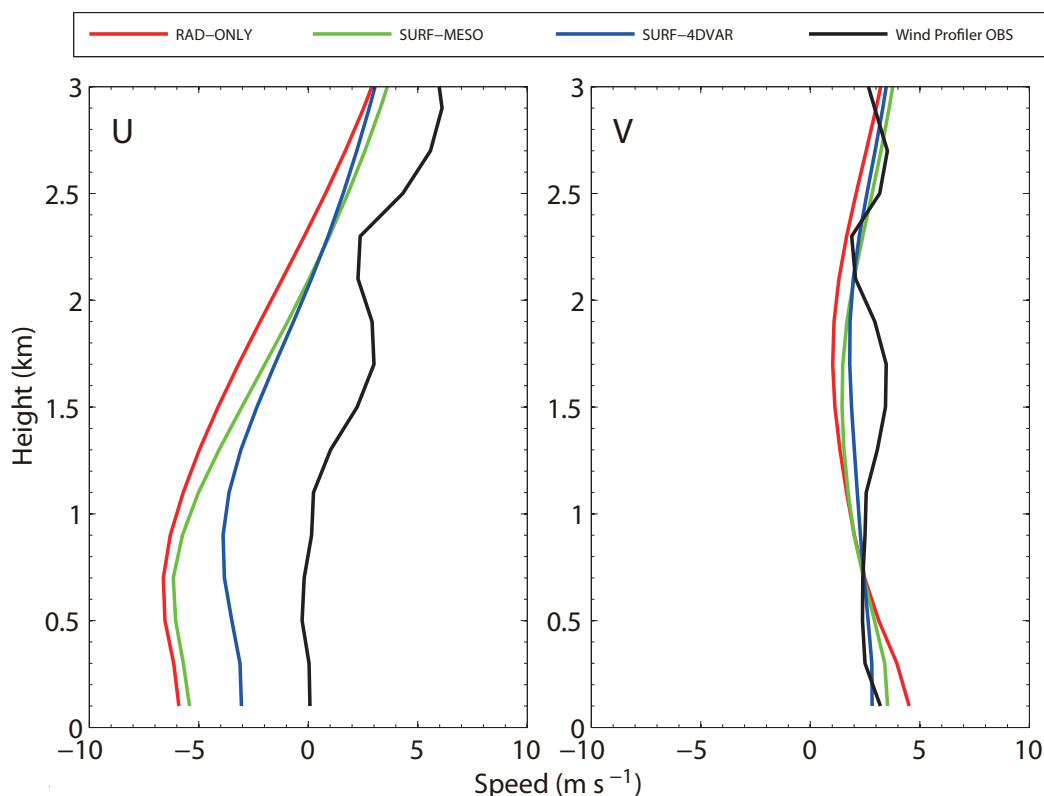


Fig. 10. As in Fig. 7, but for the observed and 1-h forecasted wind profiles at 1600 UTC.

UTC. The differences between these forecasts are not significant. The reasons for the small differences need to be further investigated.

## 6. Summary and discussion

This paper explores the effect of using an improved surface observation assimilation scheme on the analysis and nowcasting ability of VDRAS. Different from the original version, where surface temperature and wind observations are mainly integrated into the model background using a simple interpolation scheme, the current method can assimilate high spatial and temporal resolution surface observations together with the radar reflectivity and radial velocity in the 4DVAR scheme. Surface observations are interpolated to the model bottom level and assimilated every 5 min. The objective of this study was to show whether the new scheme, through minimizing the cost function with the additional surface observation terms in the 4DVAR-based VDRAS, can obtain better low-level dynamic and thermodynamic analysis. The squall-line case of 30 July 2014, observed during IOP 10 of the OPACC field campaign, is used to demonstrate the performance of the new scheme. A first set of experiments assimilates observations from a single AWS station. Results show that the surface-based divergence, cold pool and outflow of the squall line are all strengthened towards the observed values. Some improvements are exhibited in the vertical profiles of wind, mostly within the PBL. A second

set of three experiments (RAD-ONLY: only radar observations are assimilated; SURF-MESO: radar observations are assimilated with surface data blended with the WRF background; and SURF-4DVAR: radar and surface observations are assimilated together) is conducted to examine the impact of the surface data assimilation on the analysis and forecast of VDRAS. Among the three experiments, the analyzed surface temperature and wind in SURF-4DVAR show the best agreement with the AWS observations. The low-level wind profiles, especially those within the PBL, are also improved upon surface data assimilation. Similar improvements within the PBL are found in the 1-h forecast starting from the analysis fields.

The current study reveals the potential benefits of assimilating surface data in a convection-permitting model in addition to radar data assimilation. Assimilating surface data can fill in the near-surface observational gap of radar data and improve the representation of near-surface and PBL conditions, which are essential for accurate analysis and forecasting of convective systems. Needless to say, more case studies are still needed to confirm the impact of the surface data assimilation, especially on the precipitation forecast. The idea presented in this study will be applied to mesoscale data assimilation systems, such as WRF data assimilation systems, in the future, to evaluate the applicability of the 4DVAR assimilation combining surface and radar observations to other systems.

**Acknowledgements.** This research was primarily supported by the National Fundamental Research 973 Program of China (Grant No. 2013CB430101), the National Natural Science Foundation of China (Grant Nos. 41275031, 41322032 and 41475015), the Social Commonwealth Research Program (Grant Nos. GYHY201506004 and GYHY201006007), and the Program for New Century Excellent Talents in Universities of China. The author Xingchao CHEN thanks NCAR for hosting his visit, and the graduate school administration in Nanjing University for supporting his visit.

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## REFERENCES

- Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteor.*, **3**, 396–409.
- Bryan, G. H., J. C. Knierel, and M. D. Parker, 2006: A multimodel assessment of RKW theory's relevance to squall-line characteristics. *Mon. Wea. Rev.*, **134**, 2772–2792.
- Chang, S.-F., Y.-C. Liou, J.-Z. Sun, and S.-L. Tai, 2015: The implementation of the ice-phase microphysical process into a four-dimensional Variational Doppler Radar Analysis System (VDRAS) and its impact on parameter retrieval and quantitative precipitation nowcasting. *J. Atmos. Sci.*, **73**, 1015–1038.
- Crook, N. A., and J. Z. Sun, 2004: Analysis and forecasting of the low-level wind during the Sydney 2000 forecast demonstration project. *Wea. Forecasting*, **19**, 151–167.
- Dawson, D. T., and M. Xue, 2006: Numerical forecasts of the 15–16 June 2002 Southern Plains mesoscale convective system: impact of mesoscale data and cloud analysis. *Mon. Wea. Rev.*, **134**, 1607–1629.
- Dong, J. L., M. Xue, and K. Droegemeier, 2011: The analysis and impact of simulated high-resolution surface observations in addition to radar data for convective storms with an ensemble Kalman filter. *Meteorol. Atmos. Phys.*, **112**, 41–61.
- Hayden, C. M., and R. J. Purser, 1995: Recursive filter objective analysis of meteorological fields: Applications to NESDIS operational processing. *J. Appl. Meteor.*, **34**, 3–15.
- Hou, T. J., F. Y. Kong, X. L. Chen, and H. C. Lei, 2013: Impact of 3DVAR data assimilation on the prediction of heavy rainfall over Southern China. *Advances in Meteorology*, Article ID 129642, <http://dx.doi.org/10.1155/2013/129642>.
- Hsu, S. A., E. A. Meindl, and D. B. Gilhousen, 1994: Determining the power-law wind-profile exponent under near-neutral stability conditions at sea. *J. Appl. Meteor.*, **33**, 757–765.
- Kalnay, E., H. Li, T. Miyoshi, S.-C. Yang, and J. Ballabrera-Poy, 2007: 4-D-Var or ensemble Kalman filter? *Tellus A*, **59**, 758–773.
- Klazura, G. E., and D. A. Imy, 1993: A description of the initial set of analysis products available from the NEXRAD WSR-88D system. *Bull. Amer. Meteor. Soc.*, **74**, 1293–1311.
- Lilly, D. K., 1990: Numerical prediction of thunderstorms—has its time come? *Quart. J. Roy. Meteor. Soc.*, **116**, 779–798.
- Lima, M. A., and J. W. Wilson, 2008: Convective storm initiation in a moist tropical environment. *Mon. Wea. Rev.*, **136**, 1847–1864.
- Marquis, J., Y. Richardson, P. Markowski, D. Dowell, J. Wurman, K. Kosiba, P. Robinson, and G. Romine, 2014: An investigation of the Goshen county, Wyoming, tornadic supercell of 5 June 2009 using EnKF assimilation of mobile mesonet and radar observations collected during VORTEX2. Part I: Experiment design and verification of the EnKF analyses. *Mon. Wea. Rev.*, **142**, 530–554.
- Morrison, H., G. Thompson, and V. Tatarskii, 2009: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. *Mon. Wea. Rev.*, **137**, 991–1007.
- Parker, M. D., 2010: Relationship between system slope and updraft intensity in squall lines. *Mon. Wea. Rev.*, **138**, 3572–3578.
- Peterson, E. W., and J. P. Hennessey Jr., 1978: On the use of power laws for estimates of wind power potential. *J. Appl. Meteor.*, **17**, 390–394.
- Pleim, J. E., 2007: A combined local and nonlocal closure model for the atmospheric boundary layer. Part II: Application and evaluation in a mesoscale meteorological model. *Journal of Applied Meteorology and Climatology*, **46**, 1396–1409.
- Pu, Z. X., H. L. Zhang, and J. Anderson, 2013: Ensemble Kalman filter assimilation of near-surface observations over complex terrain: Comparison with 3DVAR for short-range forecasts. *Tellus A*, **65**, 19620.
- Putnam, B. J., M. Xue, Y. Jung, N. Snook, and G. F. Zhang, 2014: The analysis and prediction of microphysical states and polarimetric radar variables in a mesoscale convective system using double-moment microphysics, multinet radar data, and the ensemble Kalman filter. *Mon. Wea. Rev.*, **142**, 141–162.
- Rotunno, R., J. B. Klemp, and M. L. Weisman, 1988: A theory for strong, long-lived squall lines. *J. Atmos. Sci.*, **45**, 463–485.
- Schenkman, A. D., M. Xue, A. Shapiro, K. Brewster, and J. D. Gao, 2011a: The analysis and prediction of the 8–9 May 2007 Oklahoma tornadic mesoscale Convective system by assimilating WSR-88D and CASA radar data using 3DVAR. *Mon. Wea. Rev.*, **139**, 224–246.
- Schenkman, A. D., M. Xue, A. Shapiro, K. Brewster, and J. D. Gao, 2011b: Impact of CASA radar and Oklahoma mesonet data assimilation on the analysis and prediction of tornadic mesovortices in an MCS. *Mon. Wea. Rev.*, **139**, 3422–3445.
- Snook, N., M. Xue, and Y. Jung, 2015: Multiscale EnKF assimilation of radar and conventional observations and ensemble forecasting for a tornadic mesoscale convective system. *Mon. Wea. Rev.*, **143**, 1035–1057.
- Sobash, R. A., and D. J. Stensrud, 2015: Assimilating surface mesonet observations with the EnKF to improve ensemble forecasts of convection initiation on 29 May 2012. *Mon. Wea. Rev.*, **143**, 3700–3725.
- Sun, J. Z., and N. A. Crook, 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. Z., and N. A. Crook, 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. Z., and N. A. Crook, 2001: Real-time low-level wind and



- temperature analysis using single WSR-88D data. *Wea. Forecasting*, **16**, 117–132.
- Sun, J. Z., and Y. Zhang, 2008: Analysis and prediction of a squall line observed during IHOP using multiple WSR-88D observations. *Mon. Wea. Rev.*, **136**, 2364–2388.
- Sun, J. Z., D. W. Flicker, and D. K. Lilly, 1991: Recovery of three-dimensional wind and temperature fields from simulated single-Doppler radar data. *J. Atmos. Sci.*, **48**, 876–890.
- Sun, J. Z., M. X. Chen, and Y. C. Wang, 2010: A frequent-updating analysis system based on radar, surface, and mesoscale model data for the Beijing 2008 forecast demonstration project. *Wea. Forecasting*, **25**, 1715–1735.
- Tai, S.-L., Y.-C. Liou, J.-Z. Sun, S.-F. Chang, and M.-C. Kuo, 2011: Precipitation forecasting using Doppler radar data, a cloud model with adjoint, and the weather research and forecasting model: Real case studies during SoWMEX in Taiwan. *Wea. Forecasting*, **26**, 975–992.
- Tompkins, A. M., 2001: Organization of tropical convection in low vertical wind shears: The role of cold pools. *J. Atmos. Sci.*, **58**, 1650–1672.
- Weisman, M. L., J. B. Klemp, and R. Rotunno, 1988: Structure and evolution of numerically simulated squall lines. *J. Atmos. Sci.*, **45**, 1990–2013.
- Weygandt, S. S., A. Shapiro, and K. K. Droegemeier, 2002a: Retrieval of model initial fields from single-Doppler observations of a supercell thunderstorm. Part II: Thermodynamic retrieval and numerical prediction. *Mon. Wea. Rev.*, **130**, 454–476.
- Weygandt, S. S., A. Shapiro, and K. K. Droegemeier, 2002b: Retrieval of model initial fields from single-Doppler observations of a supercell thunderstorm. Part I: Single-Doppler velocity retrieval. *Mon. Wea. Rev.*, **130**, 433–453.
- Yussouf, N., Dowell, D. C., Wicker, L. J., Knopfmeier, K. H., & Wheatley, D. M., 2015: Storm-scale data assimilation and ensemble forecasts for the 27 April 2011 severe weather outbreak in Alabama. *Mon. Wea. Rev.*, **143**(8), 3044–3066.
- Zhang, F. Q., C. Snyder, and J. Z. Sun, 2004: Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, **132**, 1238–1253.