

Research Article

Impact of 3DVAR Data Assimilation on the Prediction of Heavy Rainfall over Southern China

Tuanjie Hou,^{1,2} Fanyou Kong,² Xunlai Chen,^{2,3} and Hengchi Lei¹

¹ Laboratory of Cloud-Precipitation Physics and Severe Storms, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China

² Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, OK 73072, USA

³ Shenzhen Meteorological Bureau, Shenzhen 518040, China

Correspondence should be addressed to Tuanjie Hou; houtuanjie@gmail.com

Received 26 March 2013; Revised 16 June 2013; Accepted 12 July 2013

Academic Editor: Jidong Gao

Copyright © 2013 Tuanjie Hou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study examines the impact of three-dimensional variational data assimilation (3DVAR) on the prediction of two heavy rainfall events over Southern China by using a real-time storm-scale forecasting system. Initialized from the European Centre for Medium-Range Weather Forecasts (ECMWF) high-resolution data, the forecasting system is characterized by combining the Advanced Research Weather Research and Forecasting (WRF-ARW) model and the Advanced Regional Prediction System (ARPS) 3DVAR package. Observations from Doppler radars, surface Automatic Weather Station (AWS) network, and radiosondes are used in the experiments to evaluate the impact of data assimilation on short-term quantitative precipitation forecast (QPF) skill. Results suggest that extrasurface AWS data assimilation has slight but general positive impact on rainfall location forecasts. Surface AWS data also improve model results of near-surface variables. Radiosonde data assimilation improves the QPF skill by improving rainfall position accuracy and reducing rainfall overprediction. Compared with radar data, the overall impact of additional surface and radiosonde data is smaller and is reflected primarily in reducing rainfall overestimation. The assimilation of all radar, surface, and radiosonde data has a more positive impact on the forecast skill than the assimilation of either type of data only for the two rainfall events.

1. Introduction

Convective storms accompanied with heavy precipitation, hail, and damaging wind occur frequently in summer season in Southern China. To reduce damage from such severe weather, more accurate short-term forecast from convective-scale numerical weather prediction (NWP) models incorporated with robust data assimilation systems have been paid more attention [1–3]. In recent years, several studies have demonstrated that the Advanced Regional Prediction System (ARPS) three-dimensional variational (3DVAR) system is capable of analyzing different data types, by using multiple analysis passes [4–7].

Based on the Advanced Research Weather Research and Forecasting (WRF-ARW) model and the ARPS 3DVAR/Cloud Analysis module, a real-time hourly updated storm-scale forecasting system has been developed collaboratively by the Center for Analysis and Prediction of Storms (CAPS)

in the University of Oklahoma, Shenzhen Meteorological Bureau (SZMB) of China and the Shenzhen Institute of Advanced Technology (SIAT), Chinese Academy of Sciences. The forecasting system, called Hourly Assimilation and Prediction System, or HAPS, has been in daily real-time forecast runs since March 2010. The system was initialized from Global Forecast System (GFS) data and characterized by assimilating reflectivity and radial wind from local Weather Surveillance Radar-1998 Doppler (WSR-98D) radars every hour in real-time forecast in the first phase. During the second phase of upgrading HAPS, not only fine-resolution European Centre for Medium-Range Weather Forecasts (ECMWF) analysis but also forecast data have been used for initial condition (IC) and lateral boundary condition (LBC), and assimilation of surface Automatic Weather Station (AWS) and radiosonde observations were also examined [8].

One way of improving precipitation forecast is thought to use as much available data as possible. Better mesoscale analyses can be provided from frequent intermittent assimilation of surface data [9] or combination of surface and radiosonde data within and above the model planetary boundary layer [10]. Kain et al. [11] demonstrated that assimilation of radar data and other conventional observations had a positive impact on convection-scale model forecasts. Ha et al. [12] found that the assimilation of both Doppler radar and surface data had a more positive impact on quantitative precipitation forecast (QPF) skill than the assimilation of radar data or surface data only. Considering the consistent positive impact of radar data assimilation shown in recent studies [13–16], emphasis is placed on evaluating the effect of surface AWS and radiosonde data assimilation. AWS data have the advantage of better temporal and spatial resolutions, while radiosonde data can provide multilayer meteorological information, both of which should be used effectively in NWP models. As the purpose of deploying HAPS forecast system is to improve convective-scale precipitation forecasts, experiments in the study are focused on the data assimilation results from the convective-scale domain only.

A description of the model and the observation network is given in Section 2. In Section 3, two heavy rainfall events and corresponding experimental design are introduced. In Section 4, impacts of data assimilation on precipitation distribution, near-surface variables, and the QPF skill are evaluated. Finally, summary and conclusions are given in Section 5.

2. Model Description and Observation Network

2.1. Model Description. The HAPS forecast system, using the WRF-ARW (V3.3.1) model as the forecast model and the ARPS 3DVAR package to provide data assimilation, consists of a mesoscale domain with 12 km horizontal grid spacing and a one-way nested convective-scale domain at 4 km grid spacing (Figure 1). The model is initiated from the ECMWF high-resolution data which have a horizontal resolution of 0.25 deg and 20 vertical pressure levels with model top at 10 hPa. The mesoscale domain that covers Southern China region produces 48 h forecasts every 12 h (initiated at 0000 and 1200 UTC) to serve as background and LBC for the convective-scale forecasts. The convective-scale domain produces 12 h forecasts every hour (e.g., initiated at 0000, 0100, 0200 UTC, etc.).

Both model domains have 51 vertical levels. The RRTM longwave radiation [17], Goddard shortwave radiation [18], MYJ planetary boundary layer [19, 20], and Noah land surface model [21] are used in both domains in the real-time system. The Eta microphysics and Kain-Fritsch cumulus parameterization [22] are used in the 12 km domain, while the more complicated New Thompson scheme [23] is used in the 4 km domain.

The 3DVAR method minimizes a cost function that includes the background, observation, and mass conservation constraint terms. The cost function [4] can be

written as

$$J(\mathbf{x}) = \frac{1}{2} (\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}^b) + \frac{1}{2} [\mathbf{H}(\mathbf{x}) - \mathbf{y}^o]^T \times \mathbf{R}^{-1} [\mathbf{H}(\mathbf{x}) - \mathbf{y}^o] + J_c(\mathbf{x}), \quad (1)$$

where the first term on the right-hand side measures the departure of the analysis vector \mathbf{x} from the background \mathbf{x}^b , weighed by the inverse of the background error covariance matrix \mathbf{B} . The analysis variables include three wind components (u , v , and w), potential temperature, pressure, and water vapor mixing ratio. The second term measures the departure of \mathbf{x} , projected into observation space by H , from the observation vector \mathbf{y}^o . The analyzed radial velocity V_r can be approximated as

$$V_r = \frac{(X - X_o) u + (Y - Y_o) v + (Z - Z_o) w}{r}, \quad (2)$$

where u , v , and w are wind components in Cartesian coordinates (X, Y, Z) , (X_o, Y_o, Z_o) is the radar location, and r is the distance between radar site and data point less than 100 km. The third term $J_c(\mathbf{x})$ represents dynamic or equation constraints. The detailed description of ARPS 3DVAR can be found in [4, 24].

Observation errors can be generated from different sources, such as instrument noise, preprocessing, and representative error. Considering the random noise in radial velocity and the calibration error in reflectivity, Xu et al. [25] discussed the effect of such observational errors. They found that the four-dimensional variational data assimilation technique was not very sensitive to the observational errors, although larger errors in some retrieved fields did occur after increasing errors in radial velocity and reflectivity. Gong et al. [26] discussed partition of observation and background error covariances with the Bessel fitting function and the data from three dense radiosonde observation regions including East Asia, North America, and Europe. For surface AWS data assimilation, Xu et al. [27] conducted experiments to test the impact of the actual elevation difference between observation sites and numerical model surface. Their studies all demonstrated the importance of estimating observation error statistics but can only be classified as sensitivity experiments. To represent the observations accurately from instruments in China, more systematic experiments are needed. Therefore, due to lack of reliable statistics on error information for China instruments, observation errors are assumed to be uncorrelated and specified according to estimated errors for the various observational platforms [7].

A multiscale analysis procedure is included in the ARPS 3DVAR package. It means that multiple analysis passes with each one including different data types and using different filter scales can be used for data assimilation. The filter scale, defined as radius in the package, is guided by the density of observation sites. Considering the distances of those observation sites shown in Figure 1, the horizontal influence radiiuses for radar, AWS, and radiosonde data are selected as 20, 50, and 400 km, respectively. The vertical influence radius is four in grid points.

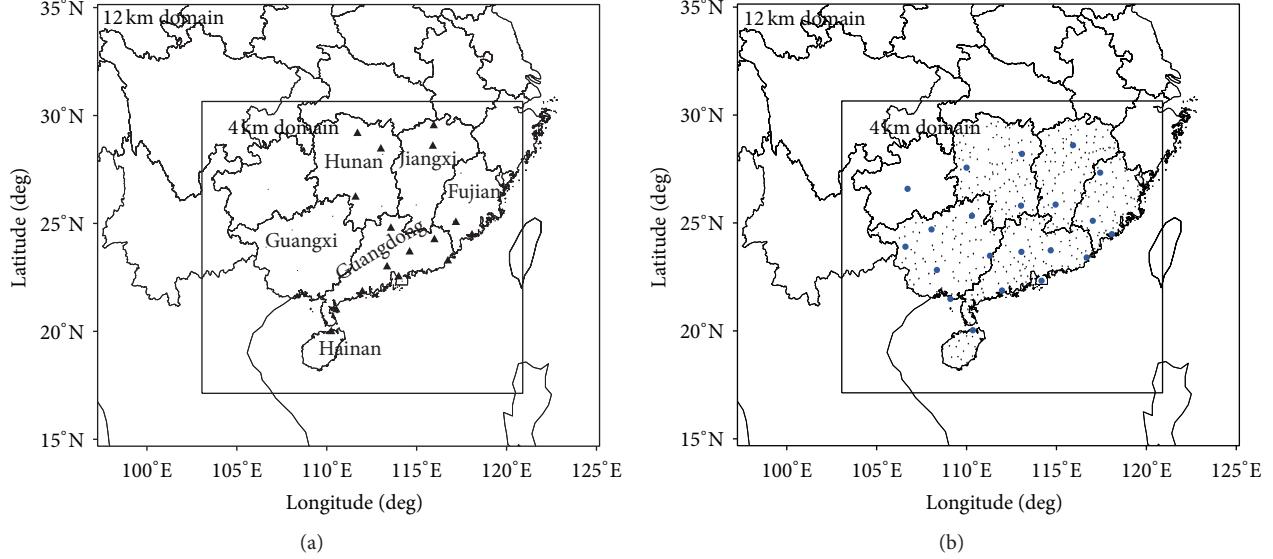


FIGURE 1: Model domains and locations of (a) radar (black triangles) sites, and (b) AWS (small black dots), and radiosonde (big blue dots) sites.

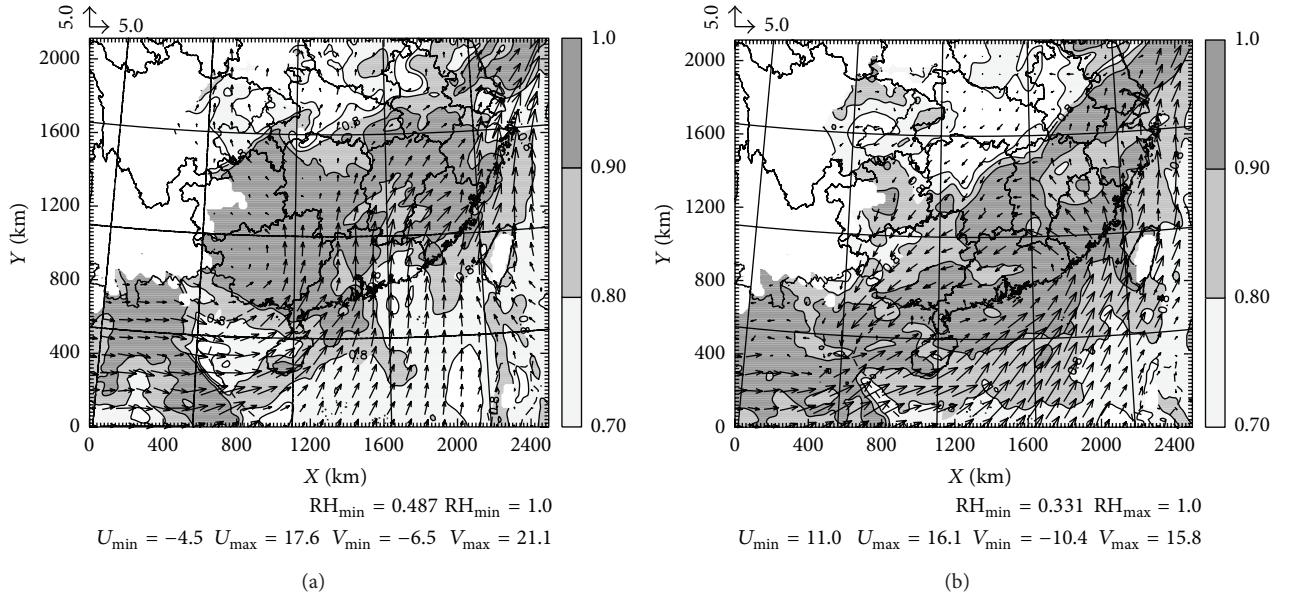


FIGURE 2: 850 hPa wind (arrows, $m s^{-1}$) and relative humidity (shading) at (a) 0000 UTC June 29, 2011 and (b) 0000 UTC July 15, 2011.

2.2. Observation Network. The available observations for this study are from WSR-98D radars, surface AWS, radiosonde, and rain gauge measurements. There are 16 radars in the 4 km domain (triangles in Figure 1(a)), including eight in Guangdong Province and eight in other four provinces. The WSR-98D radars are S-band radars with the similar characteristics to those of the Weather Surveillance Radar-1998 Doppler (WSR-88D) radars in the US operational Doppler radar network [28]. Volumetric radar data from these radars are available every 6 min for 9 elevation angles, with resolutions of radar reflectivity and velocity on the radials of being 1.0 km and 250 m [29]. Raw radar data must be processed through

automated quality control and averaged into data “columns” by using the program *88D2arps*.

The AWS network in China has been used primarily in observing weather conditions but not yet been applied to NWP models [27]. The surface AWS data include near-surface temperature, dewpoint, and wind components with the time interval of every 5 min. Data from 437 surface AWS sites (Figure 1(b)) with a horizontal distance of 3–10 km are used in the study. In addition, data from 21 radiosonde sites (blue dots in Figure 1(b)) in the 4 km domain is also used in data assimilation. These sites are part of the radiosonde network stations distributed throughout China,

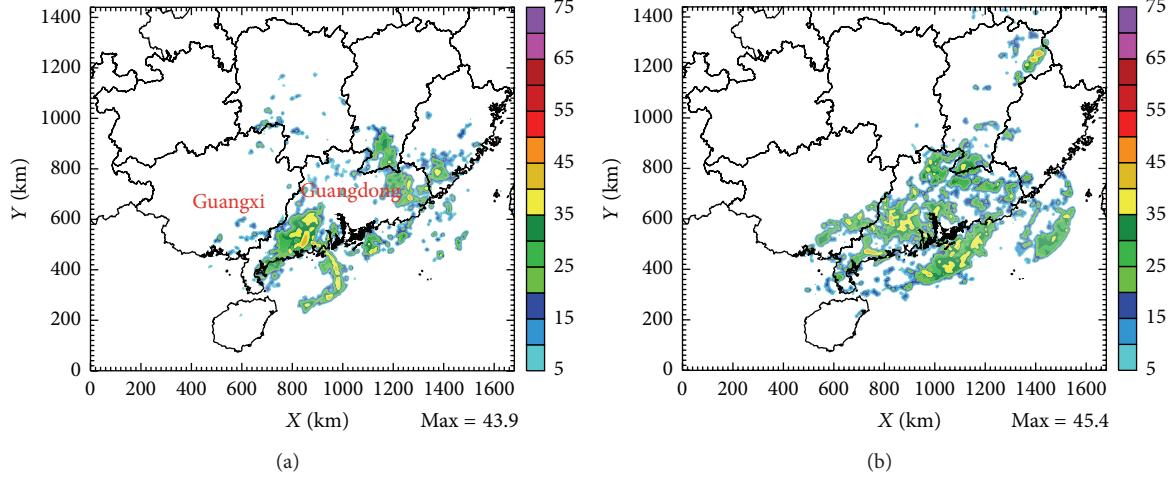


FIGURE 3: Composite reflectivity from observations at (a) 1200 UTC June 29, 2011 and (b) 1200 UTC July 15, 2011.

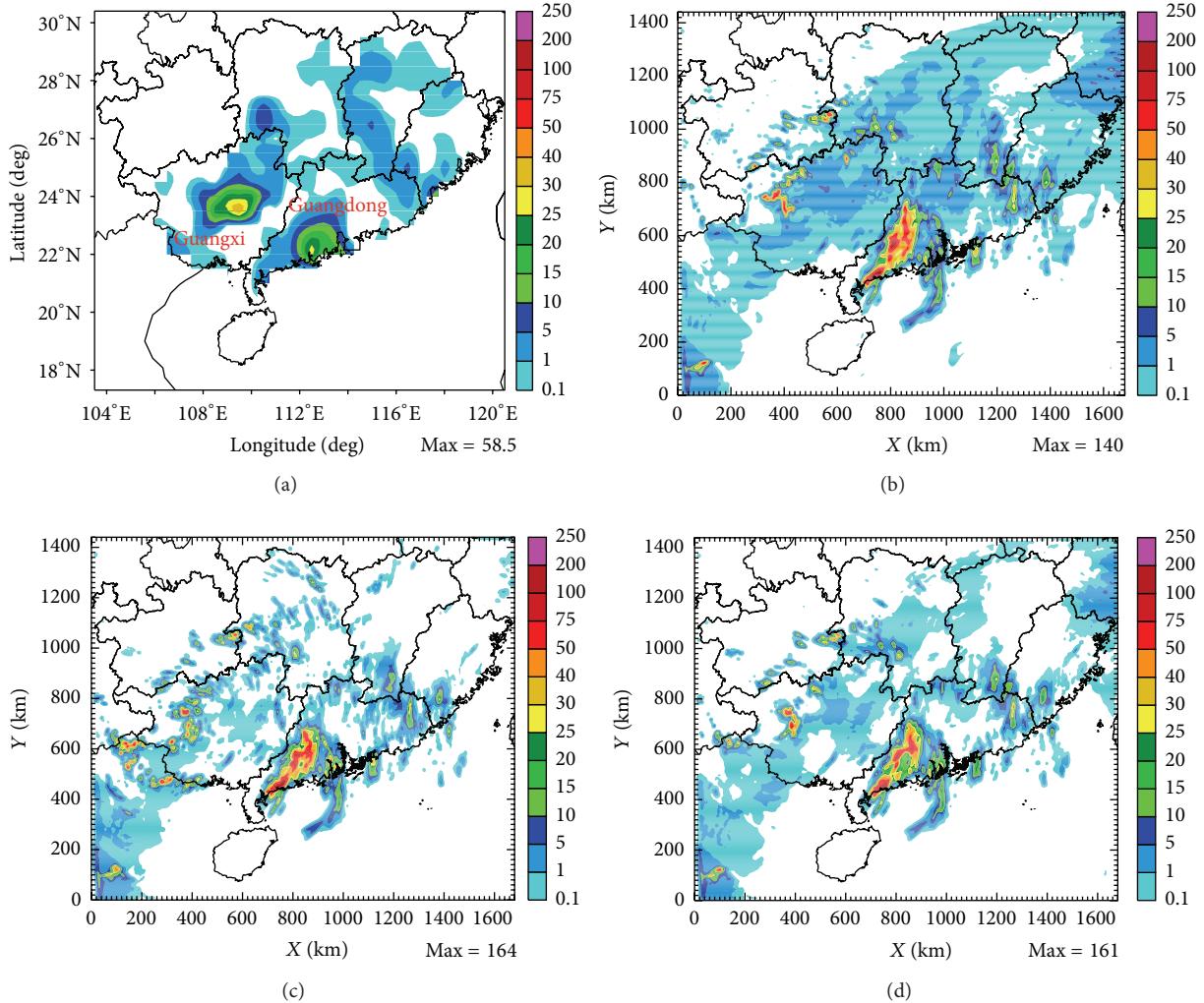


FIGURE 4: 3 h accumulated precipitation from (a) observations, (b) CNTL, (c) RADSND, and (d) ALLDATA between 1200 and 1500 UTC June 29, 2011 for case 20110629.

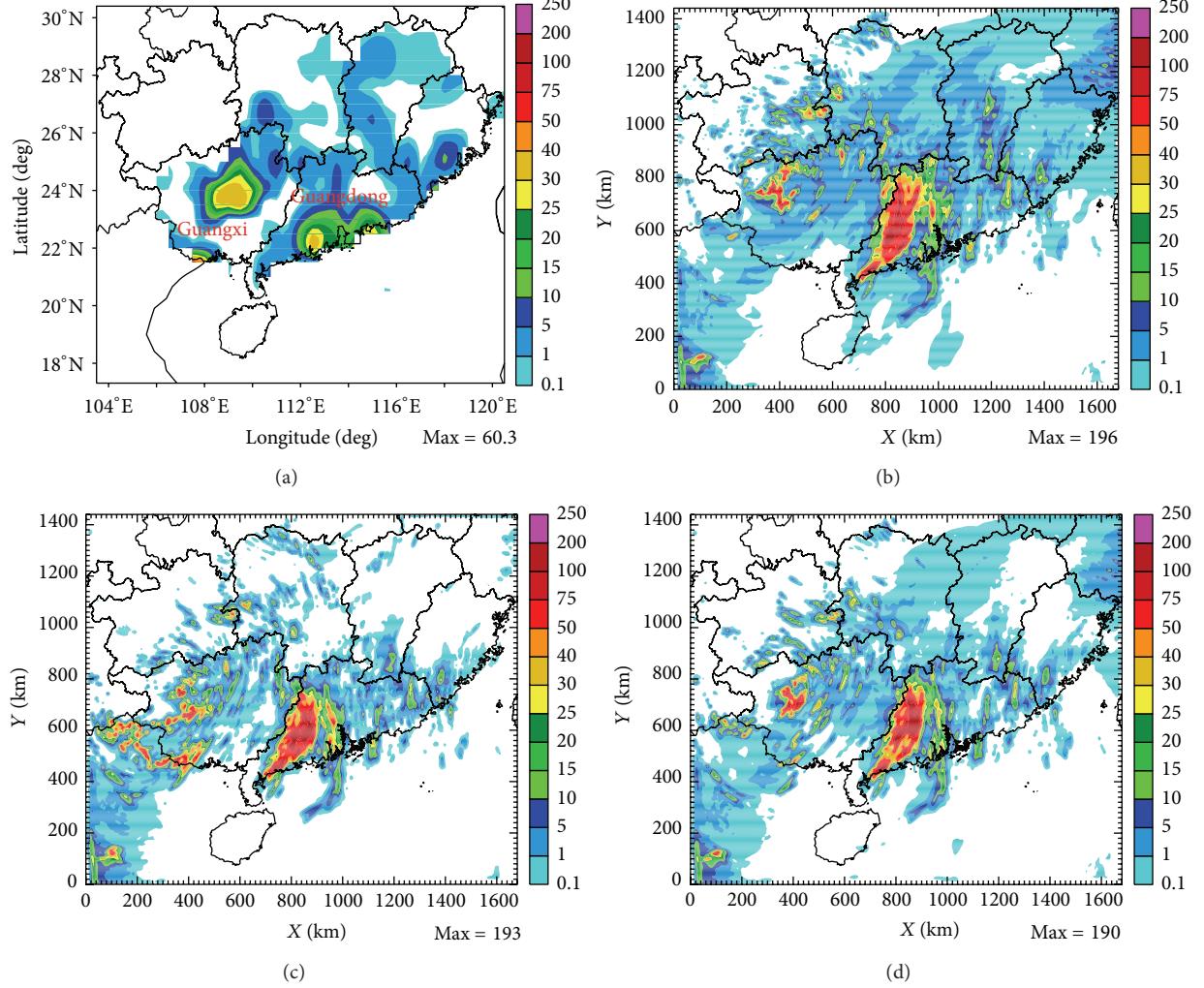


FIGURE 5: 6 h accumulated precipitation from (a) observations, (b) CNTL, (c) RADSND, and (d) ALLDATA between 1200 and 1800 UTC June 29, 2011 for case 20110629.

TABLE 1: Experiment description.

Experiments	Description
CNTL	Radar data assimilation at start time in 4 km domain
RADAWS	CNTL + AWS data assimilation at start time in 4 km domain
RADSND	CNTL + radiosonde data assimilation at start time in 4 km domain
ALLDATA	CNTL + AWS and radiosonde data assimilation at start time in 4 km domain

with improvement in instrumentation model and correction method during the last 50 years [30, 31]. The conventional radiosonde data are available at 0000 and 1200 UTC, including temperature, dewpoint, and wind profiles from 850 to 100 hPa. The total number of rain gauges in the 4 km domain is 1353 (figure not shown). They are used for quantitative precipitation verification in later sections.

3. Overview of Rainfall Events and Experimental Design

3.1. Overview of Rainfall Events. Two heavy rainfall events associated with southwest monsoons are selected in this study, as the southwestern summer monsoons are typical synoptic settings producing rainfall over Southern China in June and July. The first rainfall event (case 20110629) was from June 28, 2011 to June 30, 2011, affecting several provinces in Southern China with heavy rain and severe flooding. The second rainfall event (case 20110715), occurred from July 15, 2011 to July 18, 2011, characterized by nonuniformity and extremely high rainfall rates in localized areas.

Figure 2 shows the 850 hPa wind field and relative humidity at 0000 UTC June 29 and 0000 UTC July 15, 2011 from the data of ECMWF analysis for the two events. At 0000 UTC June 29, 2011 (Figure 2(a)), warm and moist air is transported from coastal areas near Southern China to inland regions through the strong southwest winds, forming a southwest-northeast water vapor band. The high relative humidity of

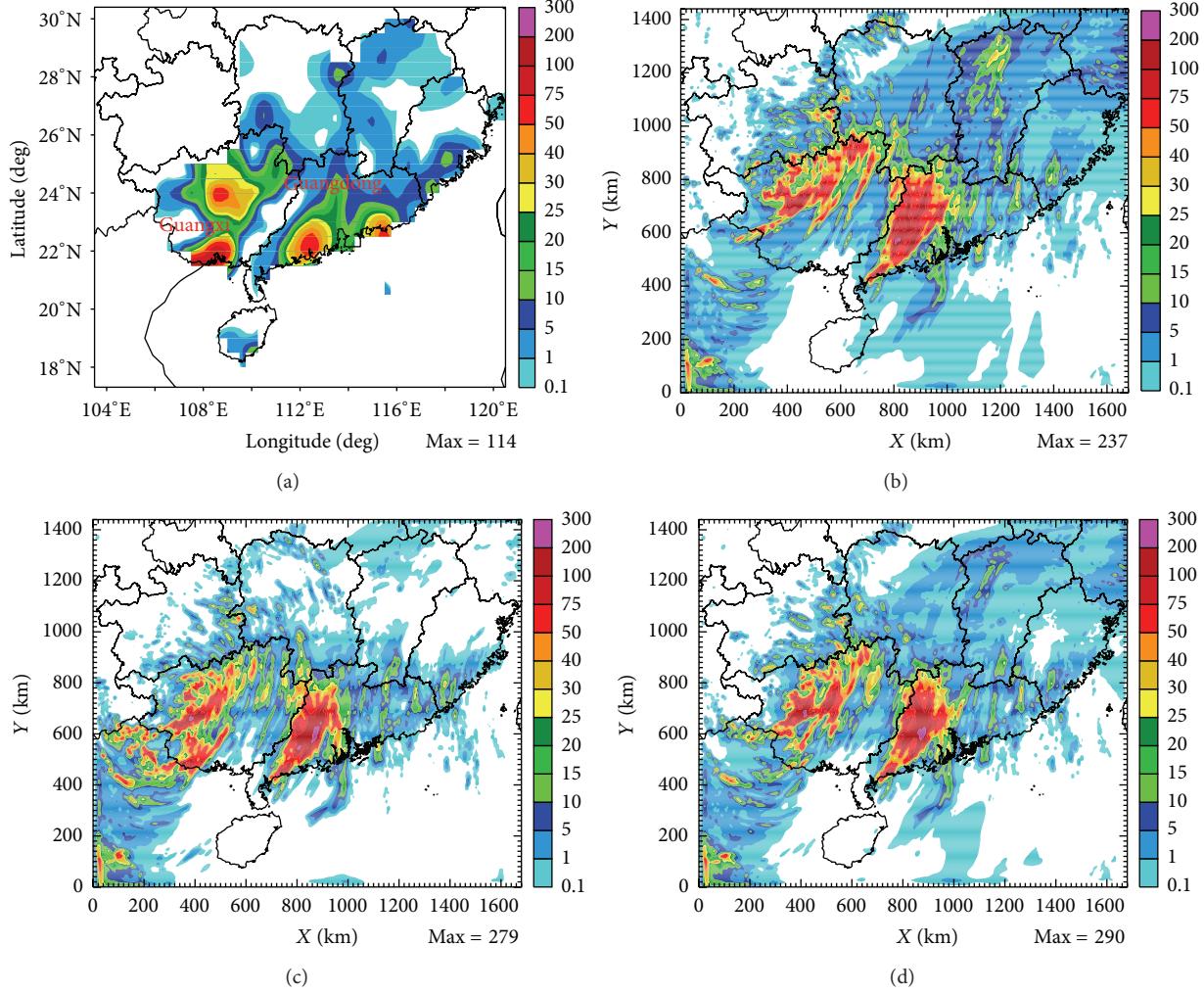


FIGURE 6: 12 h accumulated precipitation from (a) observations, (b) CNTL, (c) RADSND, and (d) ALLDATA between 1200 UTC 29 June and 0000 UTC June 30, 2011 for case 20110629.

over 80% in several provinces, combined with the gradual deeper trough at higher levels, contributes to the formation of heavy rainfall. Similar to the first case, southwest winds and a water vapor band at 0000 UTC July 15, 2011 appear at coastal areas of Southern China (Figure 2(b)). However, the moist air band in this event is somewhat narrower, due to the wind shear from southwest to southeast over eastern Guangdong and Fujian. Therefore, the low pressure and vapor transport mainly affect the coastal provinces over Southern China.

To have a clear picture of echo structures, Figure 3 shows the observed composite reflectivity at 1200 UTC June 29 and 1200 UTC July 15, 2011, respectively. Composite reflectivity is defined as the maximum reflectivity in the vertical column. The observed reflectivity is derived from the ARPS 3DVAR analysis by using available radar data (with radar station locations shown in Figure 1(a)). At 1200 UTC June 29, 2011 (Figure 3(a)), the main echo regions are at southwest Guangdong and nearby coastal areas, with the maximum value of 43.9 dBZ. In addition, scattered weaker cells can also be seen over surrounding areas of Guangdong.

In comparison, larger reflectivity regions with the similar magnitude cover almost all the Guangdong Province at 1200 UTC July 15, 2011 (Figure 3(b)). Consequently, both events caused severe waterlogging in multiple cities of Guangdong Province.

3.2. Experimental Design. One 12 km and 4 km forecast pair for each rainfall event is conducted. For case 20110629, 12 km forecast is initiated at 0000 UTC 29 June 2011 and runs for 48 h to 0000 UTC 1 July 2011. The 12 h forecast in the 4 km domain is from 1200 UTC 29 June to 0000 UTC 30 June 2011. For the other rainfall event, the 12 km forecast is from 0000 UTC 15 July to 0000 UTC 17 July 2011, and the 4 km run is from 1200 UTC 15 July to 0000 UTC 16 July 2011.

Data assimilation experiments are carried out in the convective-scale domain without cycling. As previous studies have demonstrated that radar data assimilation is useful in improving precipitation pattern and short-term QPF skill, the impact of radar data assimilation will not be examined again in the study. The focus of the present study is to

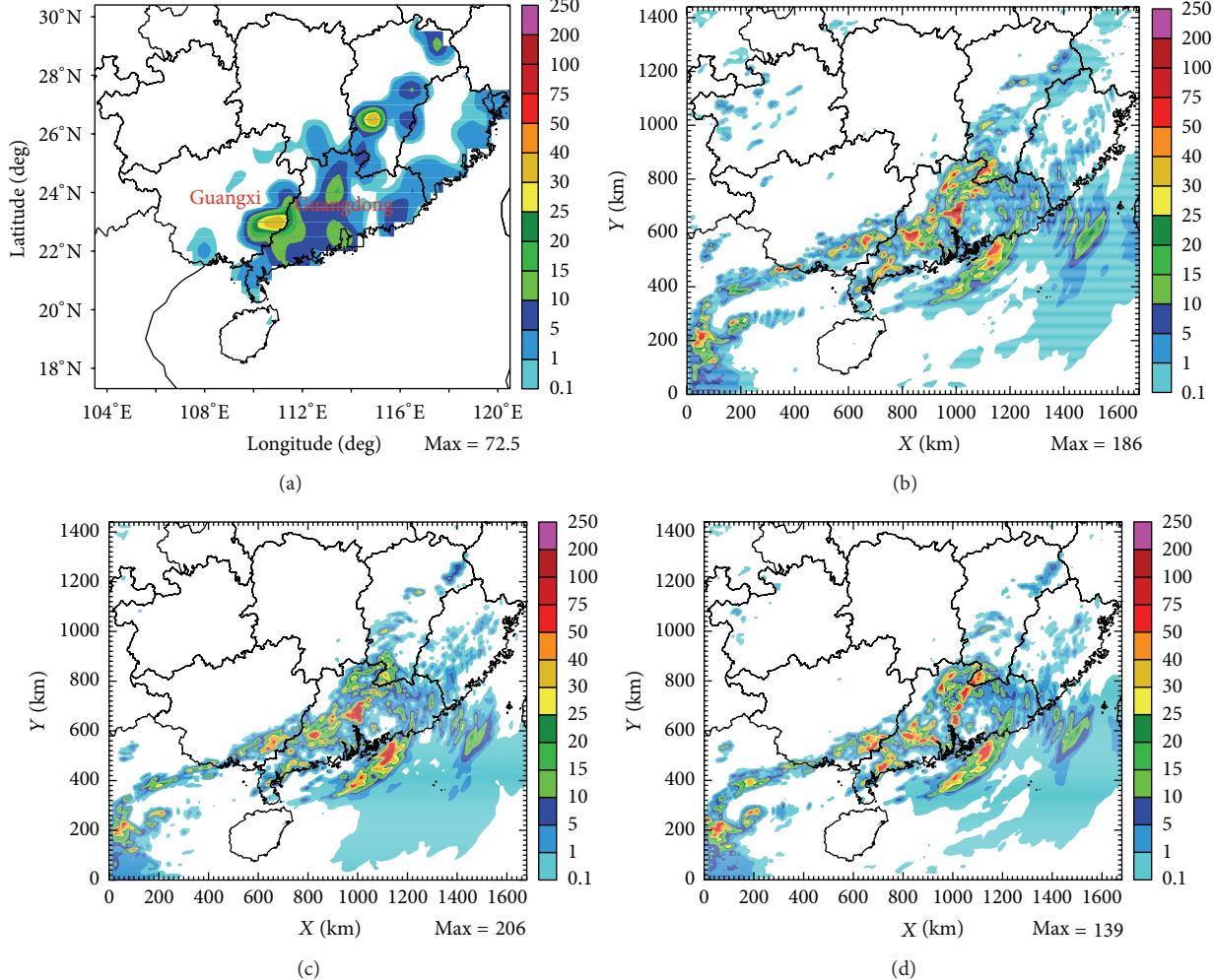


FIGURE 7: 3 h accumulated precipitation from (a) observations, (b) CNTL, (c) RADSND, and (d) ALLDATA between 1200 and 1500 UTC July 15, 2011 for case 20110715.

evaluate the impact of surface AWS and radiosonde data (experiments shown in Table 1). Hence, the basic experiment (or control experiment, CNTL) is a WRF-ARW forecast with radar data assimilation conducted at the initiation time in the 4 km domain. Additional experiments assimilate radar data plus surface AWS and radiosonde data, which are named as RADAWS and RADSND, respectively. To examine the impact of all the combined data, experiments with radar, surface AWS, and radiosonde data assimilated together (ALLDATA) at the initiation time of the 4 km domain are also performed for the two cases.

4. Results

4.1. Impact of Data Assimilation on Precipitation Distribution.

To show the impact of data assimilation on precipitation distribution, 3, 6, and 12 h accumulated precipitations from both observations and three experiments (CNTL, RADSND, and ALLDATA) for the two rainfall cases are provided in this section. Model results are from forecasts in the convective-scale domain. The precipitation distribution from RADAWS

is not shown as little difference is detected compared with that from CNTL. However, the positive impact of AWS data assimilation still can be seen from near-surface variables, which will be presented in Section 4.2.

Figure 4 presents the first 3 h accumulated precipitation between 1200 and 1500 UTC June 29, 2011 from both observations and experiments (CNTL, RADSND, and ALLDATA). According to observations (Figure 4(a)), there are primarily two precipitation regions, one in the central Guangxi Province and the other in the southern Guangdong Province. The observed maximum 3 h rainfall is 58.5 mm. Compared with observations, the main precipitation region in Guangdong generated from experiment CNTL (Figure 4(b)) is a little north and overestimated with the maximum value of up to 140 mm. And the rainfall region in Guangdong becomes much smaller in size and is displaced northwestward, as no radar data are available in Guangxi. Further radiosonde (Figure 4(c)) as well as radiosonde and AWS (Figure 4(d)) data assimilation reduces the size of rainfall overestimation region in Guangdong but still not captures the rainfall region in Guangxi.

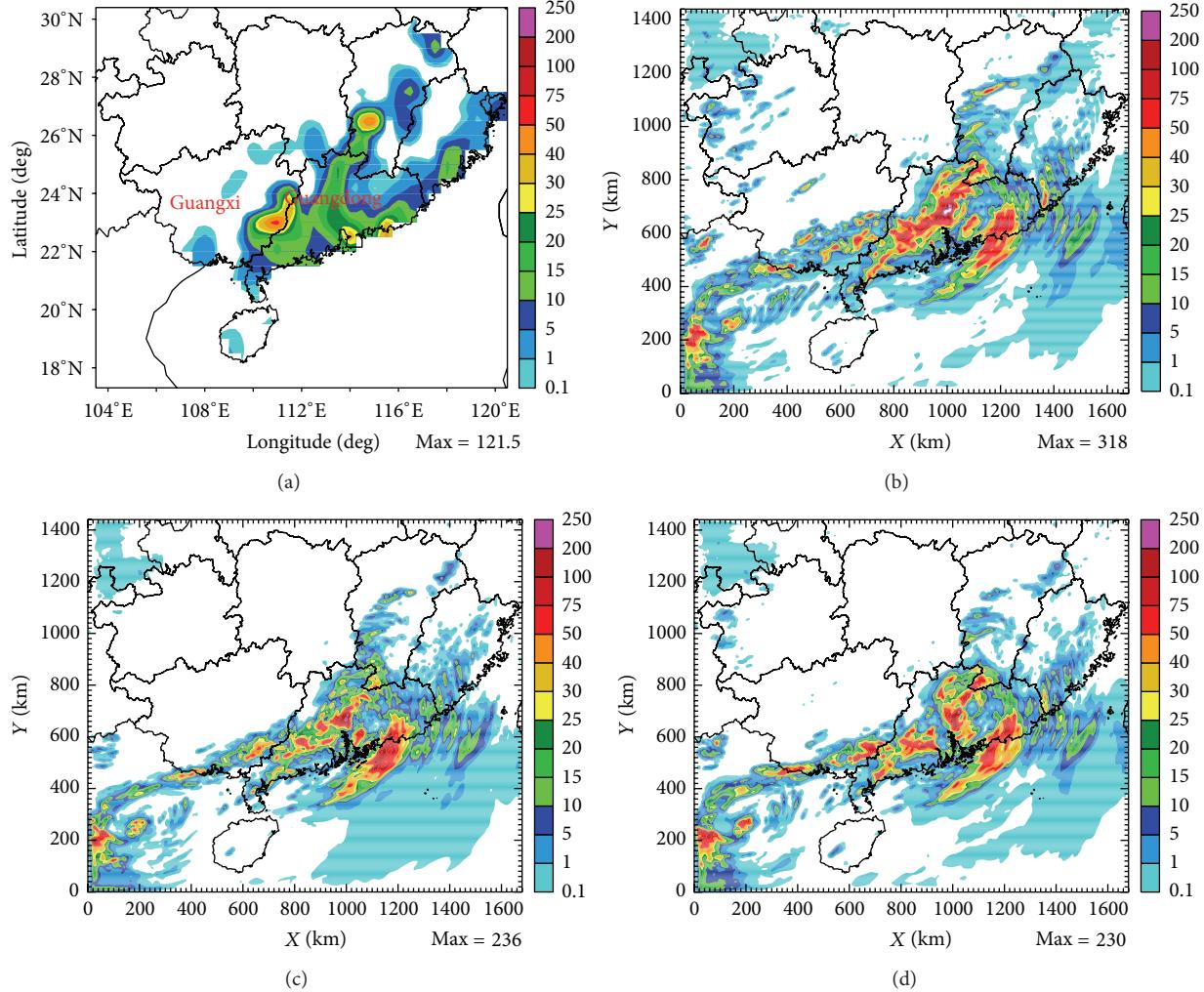


FIGURE 8: 6 h accumulated precipitation from (a) observations, (b) CNTL, (c) RADSND, and (d) ALLDATA between 1200 and 1800 UTC July 15, 2011 for case 20110715.

At 1800 UTC (the time of 6 h forecast), the observed precipitation pattern (Figure 5(a)) is similar to that of 3 h before, with two main precipitation regions. The model forecasts from CNTL, RADSND, and ALL DATA (Figures 5(b)–5(d)) also show one main precipitation region in Guangdong and are different from each other in the small scattered precipitation regions in Guangxi.

Rainfall continues increase during the latter 6 h forecast. The 12 h rainfall records from rain gauge observations (Figure 6(a)) suggest that the two precipitation regions become larger both in size and the central value, characterized by one region with the maximum value of 114 mm in Guangxi and the other rainband with the southwest-northeast direction in Guangdong. Results from radar data assimilation (Figure 6(b)) show two precipitation regions, although the direction of the precipitation region in Guangdong is not close enough to the coastline. The experiment RADSND (Figure 6(c)) improves the precipitation pattern in Guangxi by increasing the precipitation coverage. Compared with observations and experiment CNTL, experiment

ALLDATA generally improves the forecast by moving the precipitation region in Guangxi southward although still have the problem of rainfall overestimation.

Figures 7–9 present the first 3, 6, and 12 h accumulated precipitations from both observations and three experiments for case 20110715. For the first 3 h between 1200 and 1500 UTC July 15, 2011, Figure 7(a) shows a southwest-northeast rainband over southern coastal provinces, with the maximum value of 72.5 mm. The precipitation patterns from CNTL, RADSND, and ALLDATA (Figures 7(b)–7(d)) are quite close to each other; all are successful in generating the basic rainband. Overall, the experiment ALLDATA is better in terms of the maximum precipitation amount, even though it is higher than that from observations. At 1800 UTC July 15, 2011, the observed 6 h accumulated precipitation areas (Figure 8(a)) are still over Guangdong and its surrounding regions, but the rainfall amount increases with the maximum value of 121.5 mm. At that time, differences among experiments can be seen, as more detailed structures of the precipitation patterns are generated from RADSND

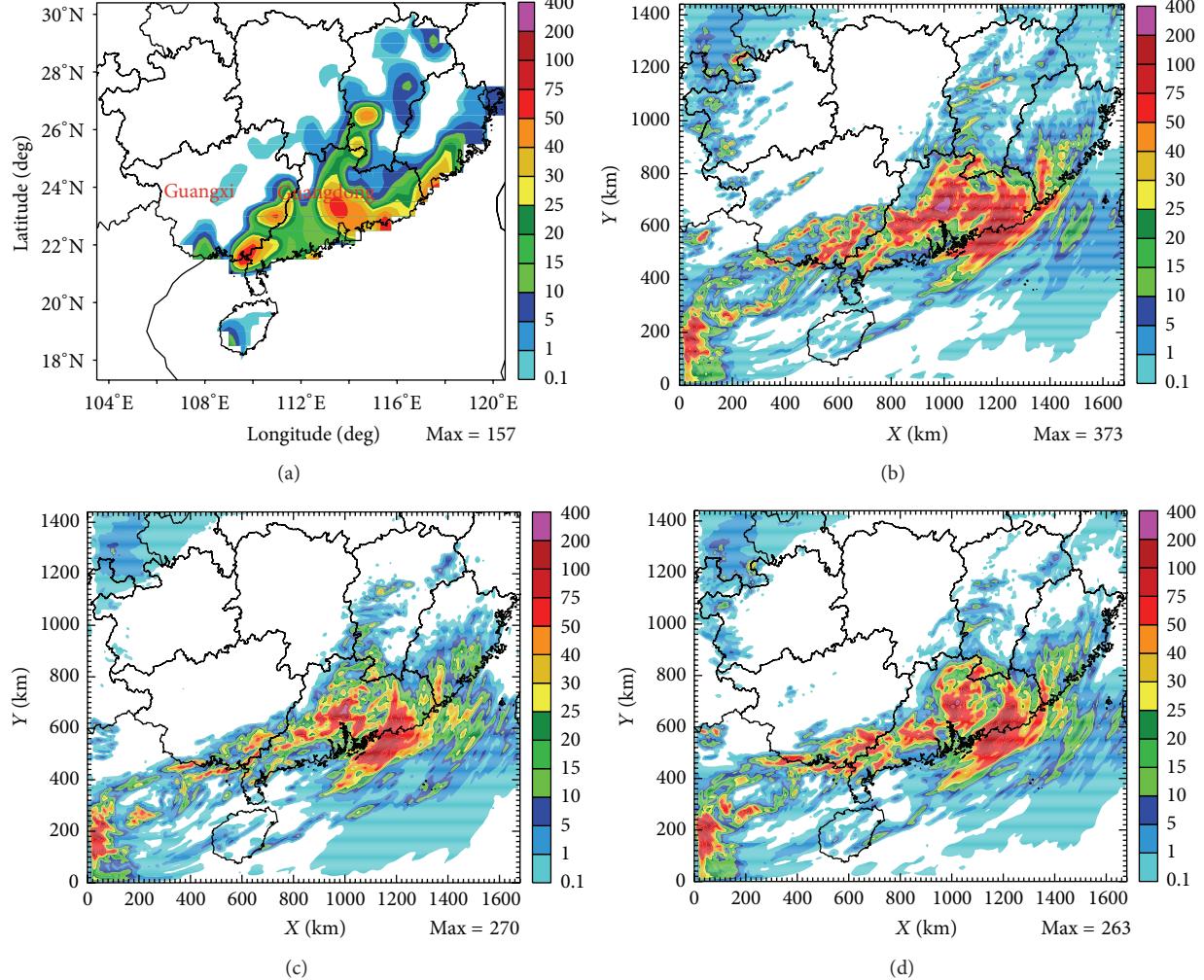


FIGURE 9: 12 h accumulated precipitation from (a) observations, (b) CNTL, (c) RADSND, and (d) ALLDATA between 1200 UTC July 15 and 0000 UTC July 16, 2011 for case 20110715.

and ALLDATA (Figures 8(c) and 8(d)). The discrepancies in precipitation distribution become larger by the time of 12 h forecasts. The 12 h accumulated precipitation distribution (Figure 9(a)) is characterized by separate localized precipitation centers, mainly with one at the southwest edge of Guangdong and the other one in central Guangdong. In comparison with experiment CNTL (Figure 9(b)), RASND (Figure 9(c)) shows positive impact by reducing the size of rainfall overestimation regions and the value of the maximum rainfall amount. Experiment ALLDATA (Figure 9(d)) further improves forecasts by splitting the one rainfall region over central Guangdong into smaller separate precipitation centers, which are closer to observations.

On the whole, forecasts with radar data assimilation are successful in producing the general precipitation patterns but not satisfying in forecasting accurate locations of precipitation centers and rainfall amounts. Addition of radiosonde data assimilation can either reduce the rainfall areal coverage overestimation or the precipitation maxima. Combination of

all the three data types is helpful in producing more detailed precipitation spatial distribution.

4.2. AWS Data Assimilation Impact on Near-Surface Variables. Surface observations, although they provide data of only one level, have the advantage of better temporal and spatial resolutions with both thermal and wind information. Root mean square errors (RMSEs) for 2 m temperature and relative humidity as well as 10 m wind components are used to evaluate the impact of AWS data assimilation in this section. Interpolation from model grid data to observation locations is performed over the AWS sites.

The RMSE is defined as

$$\text{RMSE} = \left(\frac{1}{N} \sum_{i=1}^N (F_i - O_i)^2 \right)^{1/2}, \quad (3)$$

where F_i represents variables, including temperature, relative humidity, and wind components from model results, and O_i is

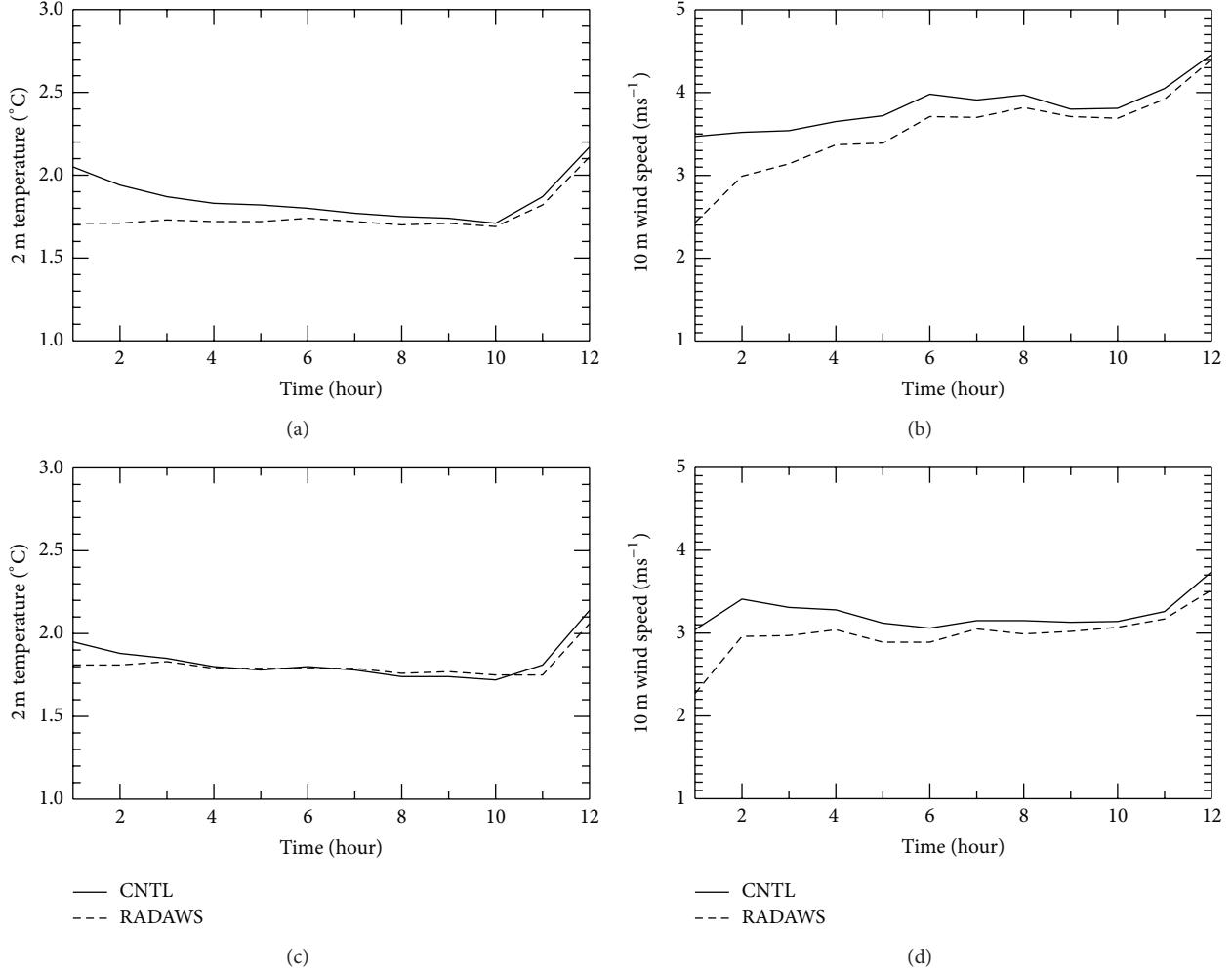


FIGURE 10: Time series of RMSEs for (a) 2 m temperature and (b) 10 m wind speed during the 12 h forecast for case 20110629 valid from 1300 UTC June 29 to 0000 UTC June 30 in the 4 km domain and RMSEs for (c) 2 m temperature and (d) 10 m wind speed during the 12 h forecast for case 20110715 valid from 1300 UTC July 15 to 0000 UTC July 16 in the 4 km domain.

for the same variables from AWS data. N is the total number of observations and is equal to or less than 437, depending on the availability of data.

Figure 10 presents the time series of RMSEs for 2 m temperature and 10 m wind speed from experiments CNTL and RADAWS for both cases. For case 20110629, compared with results in experiment CNTL, the RMSE in experiment RADAWS for temperature (Figure 10(a)) is reduced from 2.1 to 1.7°C at 1 h forecast, and for 10 m wind speed (Figure 10(b)) is reduced from 3.5 to 2.4 ms⁻¹ at 1 h forecast. For case 20110715 (Figures 10(c) and 10(d)), both RMSEs for 2 m temperature and 10 m wind speed are reduced, although the positive impact on temperature only lasts for 3 h.

Overall, AWS data assimilation has positive impact on near-surface variables. Ha et al. [12] pointed out that surface data contributed to the occurrence of intensified low-level winds. Stensrud et al. [32] found that the assimilation of only surface observations was beneficial to the creation of more realistic mesoscale features associated with mesoscale convective systems. High-resolution surface data do have the

potential of improving storm-scale convective features [33]. Hence, although our focus is on improving QPF skill, it is still encouraging to add AWS data assimilation in the real-time forecasting system.

4.3. Data Assimilation Impact on QPF Skill. To evaluate the QPF skills of above data assimilation experiments, the bias score [34] and Equitable Threat Score (ETS, [35]) for 1, 3, and 6 h accumulated precipitations with thresholds of 1, 5 and 10 mm from the convective-scale domain are calculated. Similar to the calculation of the RMSE, interpolation from model grid data to station locations is first performed before calculating the biases and ETSs. For each case, the number of available sites is approximately 1000.

The bias is defined as

$$\text{bias} = \frac{F}{O}, \quad (4)$$

where F is the number of events that are forecast and O is the number of the events that occurred. For a specified

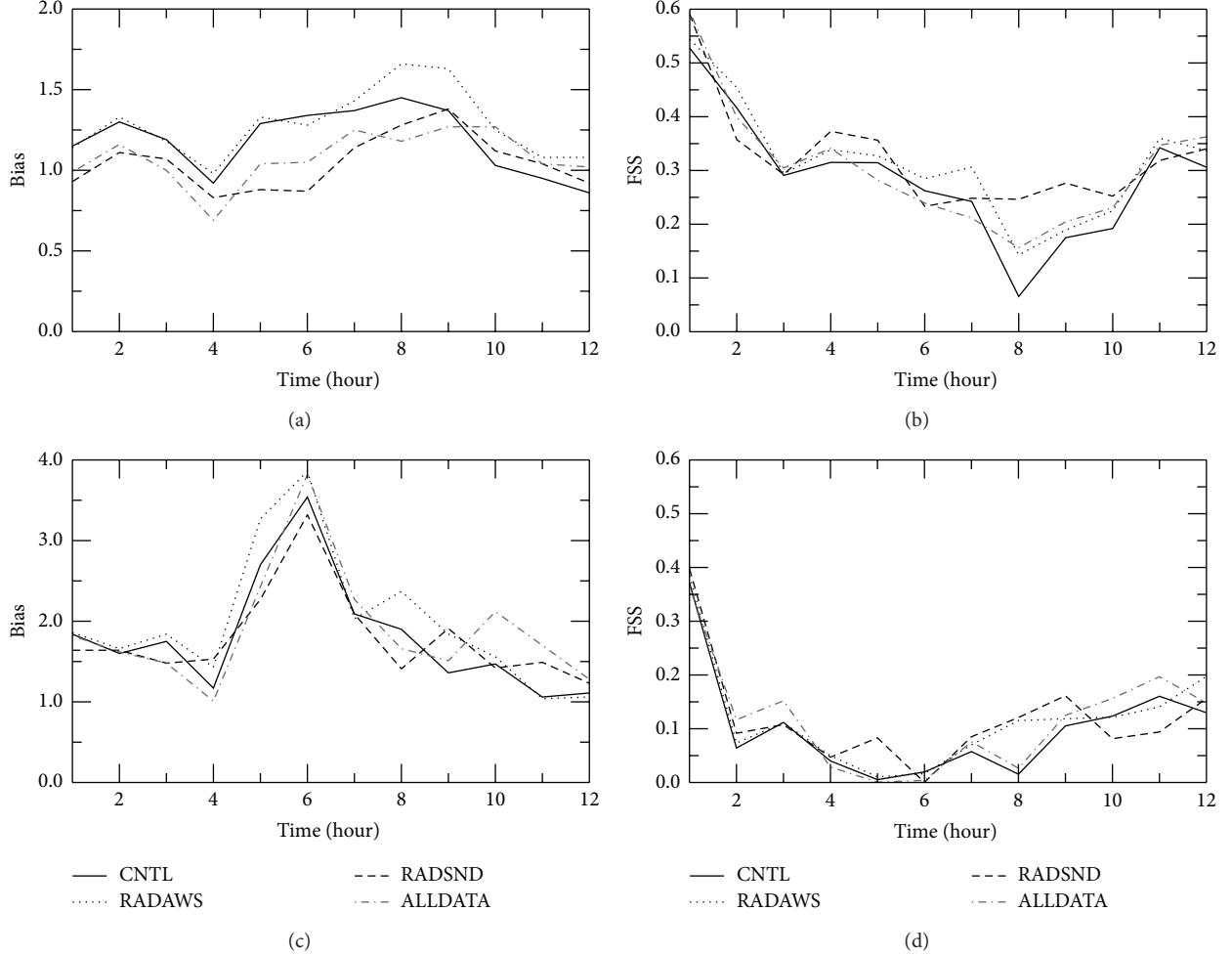


FIGURE 11: Bias score and FSS for hourly accumulated precipitation from experiments CNTL, RADAWS, RADSND, and ALLDATA for case 20110629. (a) Bias and (b) FSS with the threshold of 1 mm and (c) bias and (d) FSS with the threshold of 5 mm.

threshold, a perfect forecast would have a bias of 1, while values of bias less than and greater than one represent rainfall underprediction and overprediction, respectively.

The ETS is defined as

$$\text{ETS} = \frac{C - H}{F + O - C - H}, \quad (5)$$

where $H = (F \cdot O)/N$, C is the number of the events that are correctly forecast, and F and O are defined as above. For a perfect forecast, the ETS is 1. A larger ETS indicates a more skillful forecast.

The neighborhood-based Fractions Skill Score (FSS) is also used for the verification. It is defined as [36]

$$\text{FSS} = 1 - \frac{(1/N) \sum_{i=1}^N [P_{F(i)} - P_{O(i)}]^2}{(1/N) \left(\sum_{i=1}^N P_{F(i)}^2 - \sum_{i=1}^N P_{O(i)}^2 \right)}, \quad (6)$$

where N is the number of grid points within the neighborhood of grid point i , $P_{F(i)}$ and $P_{O(i)}$ are the neighborhood probabilities at the i th grid box in the model forecast and observed fraction fields, respectively. A perfect forecast

would have a FSS of 1, while a score of 0 means no skill. The radius of influence in the study is specified as 25 km.

Figure 11 shows the bias score and FSS for hourly accumulated precipitation with thresholds of 1 and 5 mm from all the experiments for case 20110629. For the verifications with the threshold of 1 mm, the biases (Figure 11(a)) in RADSND and ALLDATA are generally smaller than those in CNTL and RADAWS, and the experiment RADAWS has the highest bias errors during the latter 5 h. The FSS (Figure 11(b)) are rather mixed, as additional AWS or radiosonde data assimilation produce better forecasts than experiment CNTL during certain hours, but no experiment is consistently better than the others throughout the 12 h forecast. The results from the verifications with the threshold of 5 mm (Figures 11(c) and 11(d)) are also mixed, including the higher degree of overprediction from RADAWS, smaller bias errors in the first few hours but worse results later from ALLDATA and fluctuations of FSS from all experiments.

Comparison of bias score and ETS for 6 h accumulated precipitation from all the experiments is shown in Figure 12. RADAWS produces larger biases than CNTL both in the

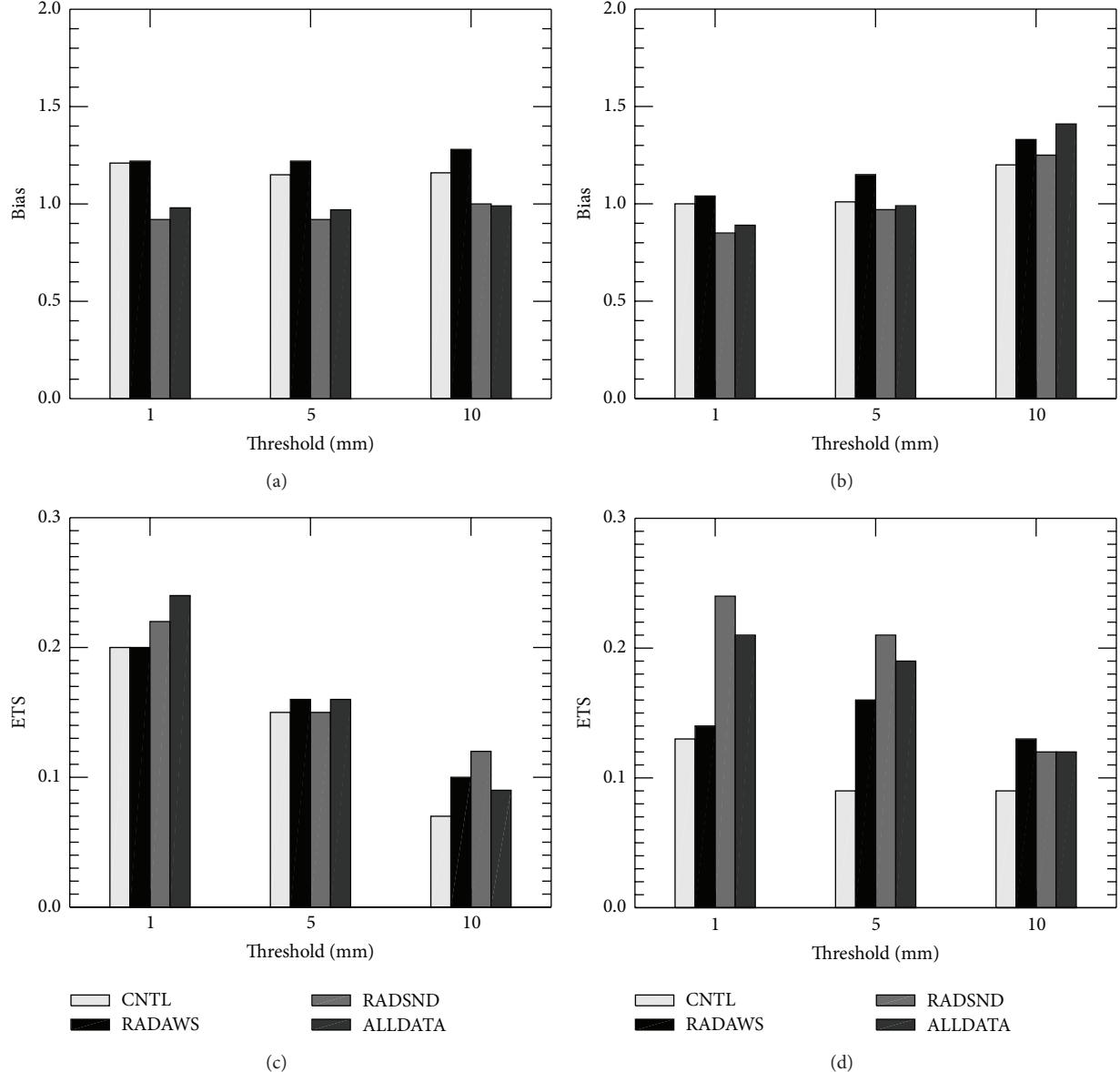


FIGURE 12: Bias score and ETS for 6 h accumulated precipitation with the thresholds of 1, 5, and 10 mm from the four experiments for case 20110629. Bias for (a) the first 6 h and (b) the latter 6 h precipitations and ETS for (c) the first 6 h and (d) the latter 6 h precipitations.

first and latter 6 h forecasts, indicating that addition of AWS data increases the degree of rainfall overprediction for this case. The biases of RADSND and ALLDATA are very similar, showing lower errors with all thresholds in the first 6 h and with the 1 and 5 mm thresholds in the latter 6 h. ETS in RADAWS are higher than those in CNTL, indicating that additional AWS data assimilation improves forecasts in terms of position accuracy. In comparison, RADSND and ALLDATA generally exhibit higher ETS than CNTL. In addition, RADSND and ALL DATA also show more skill than RADAWS with the 1 mm threshold, but not consistent with the other two thresholds.

Figure 13 shows the bias score and FSS for hourly accumulated precipitation with thresholds of 1, 5, and 10 mm from all the experiments for case 20110715. After an overview of

the biases in the left column, we can see that biases are clustered into two groups, including CNTL and RADAWS with larger biases, as well as RADSND and ALLDATA with lower bias errors. That means addition of radiosonde data reduces the degree of rainfall overprediction. FSS shown in Figure 13 also suggest that CNTL and RADAWS have similar skills, while RADSND and ALLDATA exhibit generally but not very consistent higher scores.

Figure 14 shows the bias score and ETS for the first and latter 6 h accumulated precipitations with the three thresholds from all the experiments for case 20110715. It can be seen that RADSND and ALLDATA exhibit lower biases than CNTL and RADAWS (Figures 14(a) and 14(b)), which is consistent with precipitation distribution as shown before. For ETS in the first 6 h (Figure 14(c)), RADAWS produces

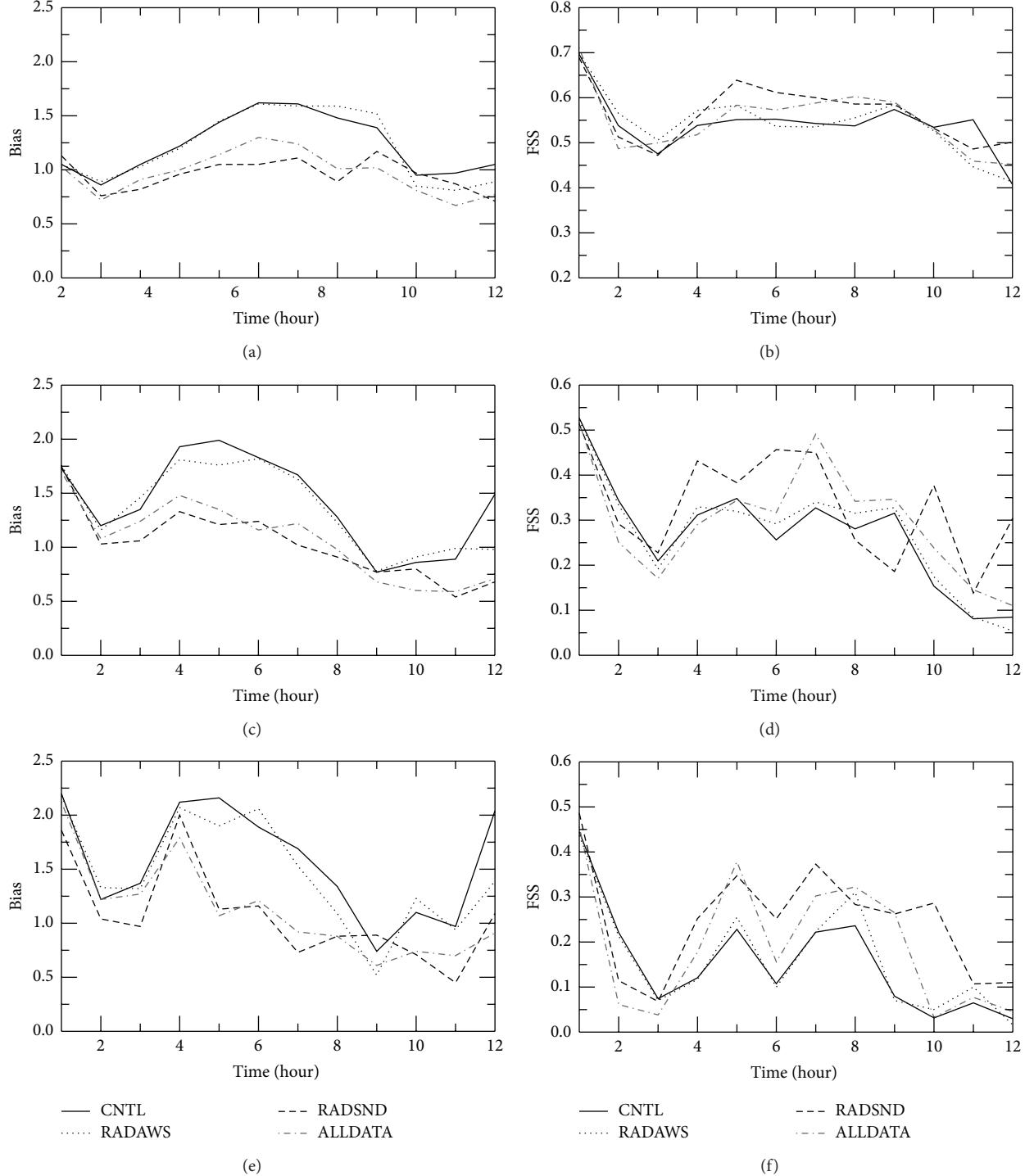


FIGURE 13: Bias score and FSS for hourly accumulated precipitation for case 20110715 from experiments CNTL, RADAWS, RADSND, and ALLDATA. (a) Bias and (b) FSS with the threshold of 1 mm, (c) bias and (d) FSS with the threshold of 5 mm, and (e) bias and (f) FSS with the threshold of 10 mm.

slightly better forecast than CNTL. When compared with RADAWS, RADSND and ALLDATA improve forecast skills slightly only for the thresholds of 5 and 10 mm. In the latter 6 h (Figure 14(d)), ALLDATA exhibit the highest ETS among all the experiments for both light and heavy rainfall.

In summary, additional surface AWS data assimilation improves the location of rainfall forecasting slightly; however, it may increase the degree of rainfall overestimation sometimes. In comparison, radiosonde data assimilation improves the QPF skill mainly by reducing rainfall overprediction.

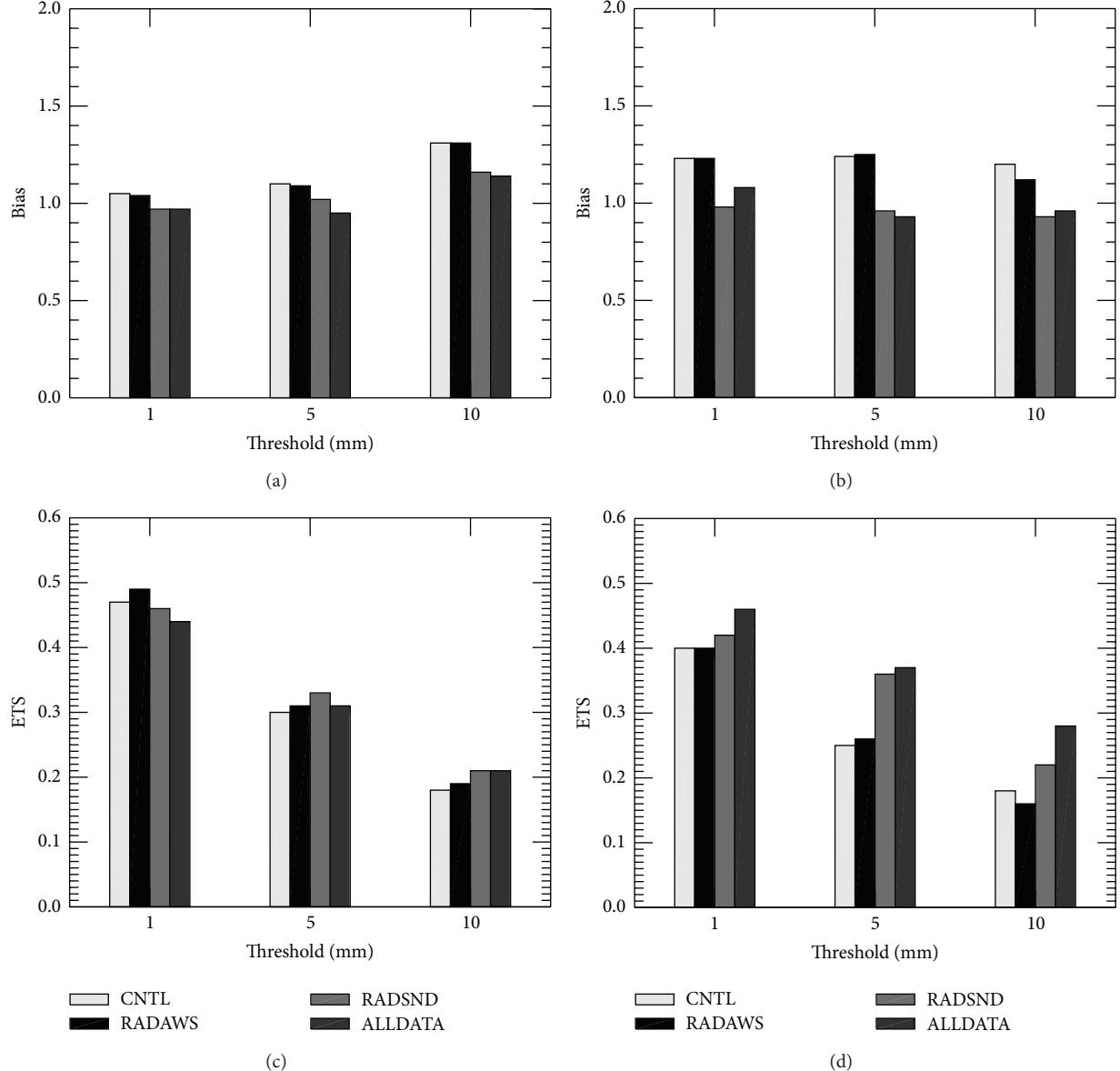


FIGURE 14: Bias score and ETS for 6 h accumulated precipitation with the thresholds of 1, 5, and 10 mm for case 20110715 from the four experiments. Bias for (a) the first 6 h and (b) the latter 6 h precipitations and ETS for (c) the first 6 h and (d) the latter 6 h precipitations.

Therefore, combination of the surface AWS and radiosonde data assimilation has the overall positive impact of improving rainfall position accuracy and reducing the degree of over-prediction. Dawson II and Xue [33] found that additional surface data in the initial conditions produced positive but not significant impact on position error forecast of the mesoscale convective system. The study of Sheng et al. [28] demonstrated that surface and upper-air data assimilation on the 6 km domain forecast was beneficial but rather small. On the whole, our results are in agreement with these previous studies. To compare the impact of surface AWS and radiosonde data with that of radar data more clearly, Figure 15 shows the bias score and ETS for 12 h accumulated precipitation from experiments CNTL and ALLDATA, as

well as forecasts without any data assimilation (short for NOASSIM) for both cases.

The biases scores of NOASSIM for both cases (Figures 15(a) and 15(c)) are less than 0.7 for the 1 mm threshold and even less than 0.5 for the 5 and 10 mm thresholds. In contrast, the bias scores of CNTL are too large, all over the perfect value of 1. Further addition of surface AWS and radiosonde data from experiment ALLDATA suggests the best forecasts in terms of bias errors, as they are closer to 1. The ETS from all experiments (Figures 15(b) and 15(d)) suggest that radar data assimilation improves the forecasts evidently and surface AWS and radiosonde data have the general positive but much smaller impact. The lower ETS in ALLDATA than that in CNTL with the 1 mm threshold (Figure 15(d)) is speculated

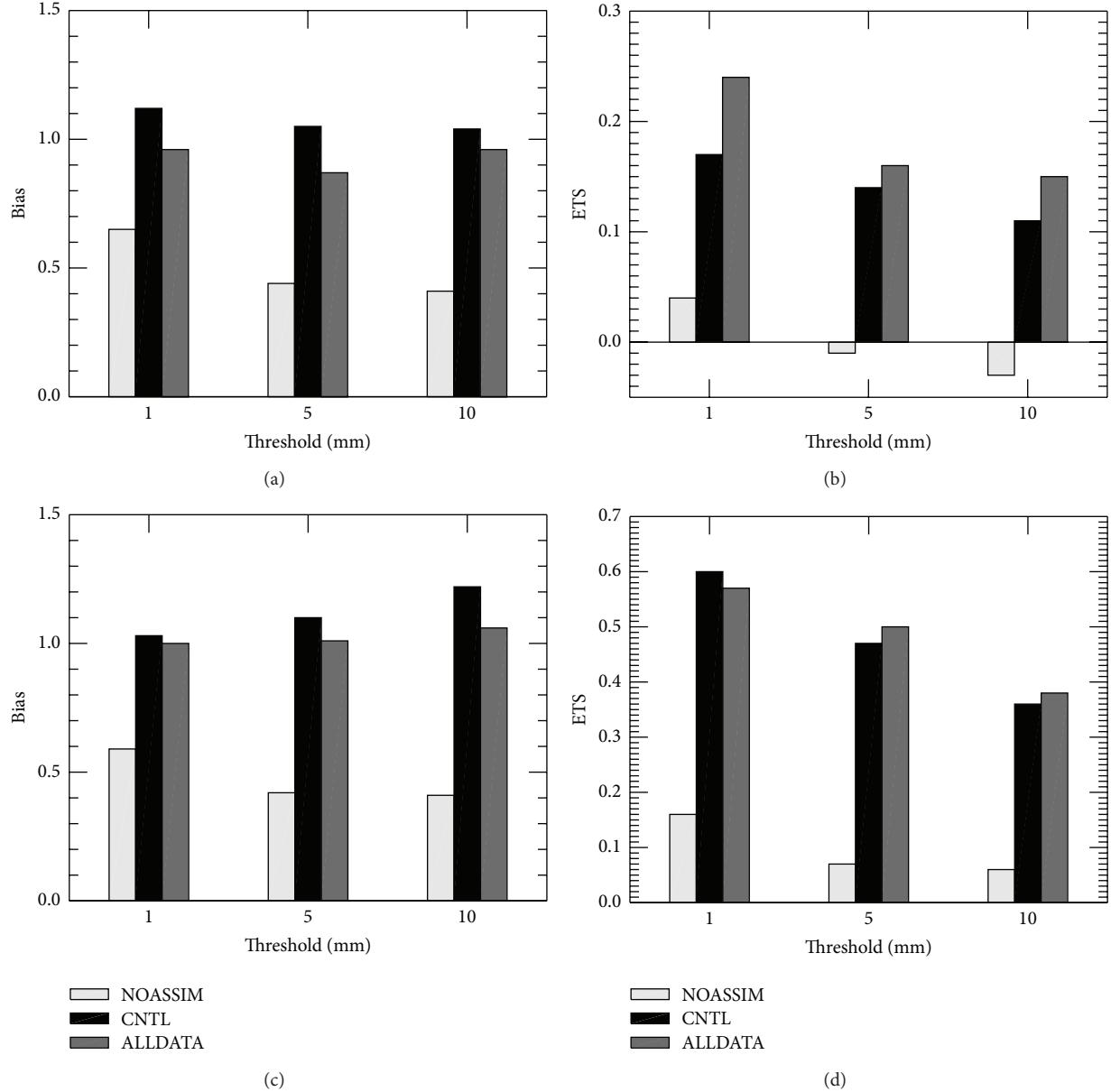


FIGURE 15: Bias score and ETS for 12 h accumulated precipitation with the thresholds of 1, 5, and 10 mm from NOASSIM, CNTL, and ALLDATA. (a) Bias and (b) ETS for case 20110629 and (c) bias and (d) ETS for case 20110715.

from the slight negative impact of radiosonde data, as shown before in Figure 14(c). Overall, additional surface AWS and radiosonde data assimilation reduce rainfall overprediction from radar data assimilation. Even though they have smaller impact than radar data in terms of ETS, they still should be considered in convective-scale forecasts.

5. Summary and Conclusions

This study is part of effort to improve the forecast skill of the HAPS system for the Shenzhen Meteorological Bureau, with the main focus on evaluating the impact of 3DVAR data assimilation on short-term QPF by considering different observation data sources from Doppler weather radars,

surface AWS, and conventional radiosondes. Data assimilation experiments for the June 29 and July 15, 2011 rainfall events over Southern China are carried out. Since the positive impact of radar data assimilation has been demonstrated in previous studies, this study is emphasized on evaluating additional surface AWS and radiosonde data assimilation impacts on short-term QPF skill.

Extrasurface AWS data assimilation has slight but general positive impact on rainfall location forecasts. The positive impact of AWS data can also be shown from the forecast of near-surface variables. At the same time, it cannot be denied that AWS data assimilation may increase the degree of rainfall overprediction sometimes. AWS data are available at the start time of every hour and can be provided timely,

being recommended in the daily real-time forecasting system.

Radiosonde data assimilation improves the QPF skill in terms of improving rainfall position accuracy and reducing rainfall overprediction. Although radiosonde data have relatively coarser resolution and are available every 12 h, they are also beneficial to convective-scale precipitation forecasts. Combination of the surface AWS and radiosonde data assimilation has the overall positive impact of improving rainfall position accuracy and reducing the degree of overprediction.

Radar data assimilation has the obvious positive impact on precipitation forecasts in terms of ETS but also produces higher bias errors for overprediction. Compared with radar data, the overall impact of additional surface and radiosonde data is smaller and is reflected primarily in reducing rainfall overestimation. Therefore, the assimilation of all radar, surface and radiosonde data has a more positive impact on the QPF skill than the assimilation of either type of data only.

Only two cases are investigated in this study. To improve the precipitation forecast more effectively, more research efforts are needed, including more case studies and further quality control of observation data. It should be noted that cycling experiments are also performed with a 6 h AWS data cycle window at 1 h intervals in the 12 km domain, with mixed or even negative impacts. AWS data quality control and improvement of the cycling process might need to be further looked into in the future work. In addition, higher model resolution of 1 km or even sub-100-meter should also be examined in future studies.

Acknowledgments

The research was supported by a grant to CAPS from Shenzhen Meteorological Bureau (SZMB). All numerical experiments and analyses were performed on the SZMB supercomputer facility. The authors benefited from discussions with Dr. Jidong Gao for using the ARPS 3DVAR package. The work was also supported by the National Basic Research Program of China (973 Program, Grant no. 2013CB430105), the Knowledge Innovation Program of the Chinese Academy of Sciences (Grant no. KZCX2-EW-203), and the National Natural Science Foundation of China (Grant no. 41105095).

References

- [1] M. Xue, D. Wang, J. Gao, K. Brewster, and K. K. Droegemeier, "The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation," *Meteorology and Atmospheric Physics*, vol. 82, no. 1-4, pp. 139–170, 2003.
- [2] F. Kong, K. K. Droegemeier, and N. L. Hickmon, "Multiresolution ensemble forecasts of an observed tornadic thunderstorm system. Part II: storm-scale experiments," *Monthly Weather Review*, vol. 135, no. 3, pp. 759–782, 2007.
- [3] A. J. Clark, S. J. Weiss, J. S. Kain et al., "An overview of the 2010 Hazardous Weather Testbed experimental forecast program spring experiment," *Bulletin of the American Meteorological Society*, vol. 93, no. 1, pp. 55–74, 2012.
- [4] J. Gao, M. Xue, K. Brewster, and K. K. Droegemeier, "A three-dimensional variational data analysis method with recursive filter for Doppler radars," *Journal of Atmospheric and Oceanic Technology*, vol. 21, no. 3, pp. 457–469, 2004.
- [5] M. Hu, M. Xue, and K. Brewster, "3DVAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part I: cloud analysis and its impact," *Monthly Weather Review*, vol. 134, no. 2, pp. 675–698, 2006.
- [6] K. Zhao and M. Xue, "Assimilation of coastal Doppler radar data with the ARPS 3DVAR and cloud analysis for the prediction of Hurricane Ike (2008)," *Geophysical Research Letters*, vol. 36, no. 12, Article ID L12803, 2009.
- [7] A. D. Schenkman, M. Xue, A. Shapiro, K. Brewster, and J. Gao, "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," *Monthly Weather Review*, vol. 139, no. 1, pp. 224–246, 2011.
- [8] T. Hou, F. Kong, and X. Chen, "Evaluation of 3DVAR data assimilation and planetary boundary layer scheme sensitivities for two heavy rainfall cases in Southern China," in *Proceedings of the 13th Weather Research and Forecasting Model Workshop*, p. 76, Boulder, Colo, USA, 2012.
- [9] F. H. Ruggiero, K. D. Sashegyi, R. V. Madala, and S. Raman, "The use of surface observations in four-dimensional data assimilation using a mesoscale model," *Monthly Weather Review*, vol. 124, no. 5, pp. 1018–1033, 1996.
- [10] D. R. Stauffer, N. L. Seaman, and F. S. Binkowski, "Use of four-dimensional data assimilation in a limited-area mesoscale model Part II: effects of data assimilation within the planetary boundary layer," *Monthly Weather Review*, vol. 119, no. 3, pp. 734–754, 1991.
- [11] J. S. Kain, M. Xue, M. C. Coniglio et al., "Assessing advances in the assimilation of radar data and other mesoscale observations within a collaborative forecasting-research environment," *Weather and Forecasting*, vol. 25, no. 5, pp. 1510–1521, 2010.
- [12] J.-H. Ha, H.-W. Kim, and D.-K. Lee, "Observation and numerical simulations with radar and surface data assimilation for heavy rainfall over central Korea," *Advances in Atmospheric Sciences*, vol. 28, no. 3, pp. 573–590, 2011.
- [13] Q. Xiao, Y.-H. Kuo, J. Sun et al., "Assimilation of Doppler radar observations with a regional 3DVAR system: impact of Doppler velocities on forecasts of a heavy rainfall case," *Journal of Applied Meteorology*, vol. 44, no. 6, pp. 768–788, 2005.
- [14] Q. Xiao and J. Sun, "Multiple-radar data assimilation and short-range quantitative precipitation forecasting of a squall line observed during IHOP_2002," *Monthly Weather Review*, vol. 135, no. 10, pp. 3381–3404, 2007.
- [15] J. Gao and D. J. Stensrud, "Assimilation of reflectivity data in a convective-scale, cycled 3DVAR framework with hydrometeor classification," *Journal of the Atmospheric Sciences*, vol. 69, no. 3, pp. 1054–1065, 2012.
- [16] S. J. Rennie, S. L. Dance, A. J. Illingworth, S. P. Ballard, and D. Simonin, "3D-Var assimilation of insect-derived Doppler radar radial winds in convective cases using a high-resolution model," *Monthly Weather Review*, vol. 139, no. 4, pp. 1148–1163, 2011.
- [17] E. J. Mlawer, S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, "Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave," *Journal of Geophysical Research D*, vol. 102, no. 14, pp. 16663–16682, 1997.

- [18] J. Dudhia, "Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model," *Journal of the Atmospheric Sciences*, vol. 46, no. 20, pp. 3077–3107, 1989.
- [19] Z. I. Janjic, "The step-mountain coordinate: physical package," *Monthly Weather Review*, vol. 118, no. 7, pp. 1429–1443, 1990.
- [20] Z. I. Janjic, "Nonsingular implementation of the Mellor-Yamada level 2.5 scheme in the NCEP Meso Model," 2001, <http://www.emc.ncep.noaa.gov/officenotes/newernotes/on437.pdf>.
- [21] F. Chen and J. Dudhia, "Coupling and advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: model implementation and sensitivity," *Monthly Weather Review*, vol. 129, no. 4, pp. 569–585, 2001.
- [22] J. S. Kain and J. Kain, "The Kain—Fritsch convective parameterization: an update," *Journal of Applied Meteorology*, vol. 43, no. 1, pp. 170–181, 2004.
- [23] G. Thompson, P. R. Field, R. M. Rasmussen, and W. D. Hall, "Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: implementation of a new snow parameterization," *Monthly Weather Review*, vol. 136, no. 12, pp. 5095–5115, 2008.
- [24] M. Hu, M. Xue, J. Gao, and K. Brewster, "3DVAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part II: impact of radial velocity analysis via 3DVAR," *Monthly Weather Review*, vol. 134, no. 2, pp. 699–721, 2006.
- [25] X. Xu, L. Liu, and G. Zheng, "Dynamical and microphysical retrieval from simulated Doppler radar observations using the 4DVAR assimilation technique," *Acta Meteorologica Sinica*, vol. 19, no. 2, pp. 160–173, 2005.
- [26] J. Gong, L. Wei, S. Tao et al., "Accurate estimation and application of 3D error covariance structures in global data assimilation. Part I: accurate estimation of error covariance in observation space," *Acta Meteorologica Sinica*, vol. 64, no. 6, pp. 669–683, 2006 (Chinese).
- [27] Z. Xu, J. Gong, J. Wang et al., "A study of assimilation of surface observational data in complex terrain. Part I: influence of the elevation difference between model surface and observation site," *Chinese Journal of Atmospheric Sciences*, vol. 31, no. 2, pp. 222–232, 2007 (Chinese).
- [28] C. Sheng, S. Gao, and M. Xue, "Short-range prediction of a heavy precipitation event by assimilating Chinese CINRAD-SA radar reflectivity data using complex cloud analysis," *Meteorology and Atmospheric Physics*, vol. 94, no. 1-4, pp. 167–183, 2006.
- [29] H. Liu, J. Xue, J. Gu, and H. Xu, "Radar data assimilation of the GRAPES model and experimental results in a typhoon case," *Advances in Atmospheric Sciences*, vol. 29, no. 2, pp. 344–358, 2012.
- [30] Y. Guo and Y. Ding, "Long-term free-atmosphere temperature trends in China derived from homogenized in situ radiosonde temperature series," *Journal of Climate*, vol. 22, no. 4, pp. 1037–1051, 2009.
- [31] Y. Guo and Y. Ding, "Impacts of reference time series on the homogenization of radiosonde temperature," *Advances in Atmospheric Sciences*, vol. 28, no. 5, pp. 1011–1022, 2011.
- [32] D. J. Stensrud, N. Yussouf, D. C. Dowell, and M. C. Coniglio, "Assimilating surface data into a mesoscale model ensemble: cold pool analyses from spring 2007," *Atmospheric Research*, vol. 93, no. 1-3, pp. 207–220, 2009.
- [33] D. T. Dawson II and M. Xue, "Numerical forecasts of the 15-16 June 2002 southern plains mesoscale convective system: impact of mesoscale data and cloud analysis," *Monthly Weather Review*, vol. 134, no. 6, pp. 1607–1629, 2006.
- [34] R. A. Anthes, "Regional models of the atmosphere in middle latitudes," *Monthly Weather Review*, vol. 111, no. 6, pp. 1306–1335, 1983.
- [35] J. T. Schaefer, "The critical success index as an indicator of warning skill," *Weather and Forecasting*, vol. 5, no. 4, pp. 570–575, 1990.
- [36] N. M. Roberts and H. W. Lean, "Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events," *Monthly Weather Review*, vol. 136, no. 1, pp. 78–97, 2008.

