

Explicit Prediction of Hail in a Long-Lasting Multicellular Convective System in Eastern China Using Multimoment Microphysics Schemes

LIPING LUO AND MING XUE

Key Laboratory of Mesoscale Severe Weather, Ministry of Education, and School of Atmospheric Sciences, Nanjing University, Nanjing, China, and Center for Analysis and Prediction of Storms, and School of Meteorology, University of Oklahoma, Norman, Oklahoma

KEFENG ZHU AND BOWEN ZHOU

Key Laboratory of Mesoscale Severe Weather, Ministry of Education, and School of Atmospheric Sciences, Nanjing University, Nanjing, China

(Manuscript received 8 October 2017, in final form 26 June 2018)

ABSTRACT

During the afternoon of 28 April 2015, a multicellular convective system swept southward through much of Jiangsu Province, China, over about 7 h, producing egg-sized hailstones on the ground. The hailstorm event is simulated using the Advanced Regional Prediction System (ARPS) at 1-km grid spacing. Different configurations of the Milbrandt–Yau microphysics scheme are used, predicting one, two, and three moments of the hydrometeor particle size distributions (PSDs). Simulated reflectivity and maximum estimated size of hail (MESH) derived from the simulations are verified against reflectivity observed by operational S-band Doppler radars and radar-derived MESH, respectively. Comparisons suggest that the general evolution of the hailstorm is better predicted by the three-moment scheme, and neighborhood-based MESH evaluation further confirms the advantage of the three-moment scheme in hail size prediction. Surface accumulated hail mass, number, and hail distribution characteristics within simulated storms are examined across sensitivity experiments. Results suggest that multimoment schemes produce more realistic hail distribution characteristics, with the three-moment scheme performing the best. Size sorting is found to play a significant role in determining hail distribution within the storms. Detailed microphysical budget analyses are conducted for each experiment, and results indicate that the differences in hail growth processes among the experiments can be mainly ascribed to the different treatments of the shape parameter within different microphysics schemes. Both the differences in size sorting and hail growth processes contribute to the simulated hail distribution differences within storms and at the surface.

1. Introduction

Hailstorms are among the costliest natural disasters in China and many other countries; hailstorms can cause severe injuries and extensive property damage. According to the *Yearbook of Meteorological Disasters in China* (e.g., [China Meteorological Administration 2013](#), 31–43; [2014](#), 37–50; [2015](#), 28–43), hail damage amounts to billions of U.S. dollars annually in China. Improving the prediction of hail, including the size and number of hailstones and the spatial and temporal coverage of hail fall, can help mitigate the impacts of hailstorms through improved warnings. However, the prediction of hailstorms using operational numerical weather prediction (NWP)

models remains a challenge. The explicit prediction of hail at the surface, including the spatial and temporal coverage of hail fall and the hail size distributions, is even more challenging because of the complex microphysical as well as dynamic and thermodynamic processes involved in hail production ([Snook et al. 2016](#); [Labriola et al. 2017](#)).

Our general ability to forecast hail in operational and research settings is still limited ([Moore and Pino 1990](#); [Brimelow et al. 2002](#); [Guo and Huang 2002](#); [Milbrandt and Yau 2006a,b](#), hereafter MY06a and MY06b, respectively; [Brimelow and Reuter 2009](#); [Luo et al. 2017](#), hereafter L17; [Labriola et al. 2017](#)). Existing hail forecast methods include the following four types: (i) hail diagnostics based on observed soundings, (ii) methods using a simple cloud model combined with a hail growth model (e.g., HAILCAST; [Brimelow et al. 2002](#)),

Corresponding author: Ming Xue, mxue@ou.edu

DOI: 10.1175/JAS-D-17-0302.1

© 2018 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) (www.ametsoc.org/PUBSReuseLicenses).

(iii) statistical and machine-learning (ML) hail forecast methods (e.g., random forests, gradient boosting trees, and linear regression), and (iv) predictions using convective-scale NWP models with sophisticated microphysics schemes.

For sounding-based hail diagnostic methods, the most important limitation is the lack of timely soundings. HAILCAST addresses this issue by feeding prognostic model soundings into a time-dependent hail growth model (Brimelow et al. 2002; Adams-Selin and Ziegler 2016). Gagne et al. (2014) statistically validated the hail size and probability forecast skills of ML techniques and HAILCAST, based on 12 hail days from May to June 2014 over the United States, and found that the ML techniques produced smaller size errors compared to HAILCAST and that both approaches, especially HAILCAST, tended to overpredict the maximum hail size. ML methods also showed temporal and spatial offsets with observed hailstorms (Gagne et al. 2014).

Since sophisticated, multimoment microphysics parameterization (MP) schemes in storm-scale NWP models are capable of predicting hydrometeor size distributions within realistic environments, efforts to simulate and predict real hailstorms have been attempted using NWP models in recent years (e.g., MY06a; MY06b; Noppel et al. 2010; Snook et al. 2016; L17). For example, MY06a conducted simulations of a supercell hailstorm that occurred in Canada using the three-moment Milbrandt and Yau (MY) scheme in a mesoscale NWP model. Comparisons with radar observations indicated that the typical supercell structures such as the hook echo, mesocyclone, and suspended overhang region were well reproduced, although the simulated maximum hail size on the ground was underpredicted. Snook et al. (2016) is a more recent example, which evaluated short-term ensemble forecasting of hail for a supercell storm over central Oklahoma using a two-moment MP scheme. They noted that hail prediction might be improved by using more advanced MP schemes, especially via better explicit prediction of the properties of rimed ice.

In the bulk MP schemes typically used in NWP models, the particle size distribution (PSD) of each hydrometeor category is assumed to have a Marshall-Palmer or gamma distribution, and normally, one, two, or three moments of the distribution are predicted for each hydrometeor category. One-moment schemes typically predict only the mixing ratios Q of various hydrometeors (e.g., Lin et al. 1983; Kessler 1995). Two-moment schemes predict the mixing ratios Q and total number concentrations N_t of all or some of the hydrometeors (e.g., Ferrier 1994; Walko et al. 1995; Meyers et al. 1997; Thompson et al. 2004; Milbrandt and Yau 2005a, hereafter

MY05a; Morrison et al. 2005; Morrison and Gettelman 2008; Thompson et al. 2008). To make the gamma distribution (involving three free parameters) fully prognostic, Milbrandt and Yau (2005b, hereafter MY05b) proposed a three-moment scheme by adding a predictive equation for radar reflectivity factor Z of the hydrometeors (related to the sixth moment) to their two-moment scheme.

However, many studies have noted that different bulk MP schemes often produce large differences in various aspects of the simulated storms, including the storm structure, surface accumulated precipitation, and cold pool (e.g., Gilmore et al. 2004; MY06b; Seifert et al. 2006; Morrison et al. 2009; Dawson et al. 2010; Jung et al. 2010; Van Weverberg et al. 2012; Loftus and Cotton 2014; L17). Most of the prior studies have used idealized frameworks, in which the storm environment is horizontally homogeneous, with MY06b, Snook et al. (2016), and L17 being the exceptions. MY06b performed sensitivity experiments of a supercell hailstorm in Canada using different MP schemes. They noted dramatic improvements in the storm structure and the predicted precipitation when switching from a one-moment to a two-moment scheme. More recently, L17 investigated hail forecast skill using various MP schemes for a pulse-type hailstorm in eastern China. They compared simulated total precipitation and maximum estimated size of hail (MESH; Witt et al. 1998) swaths against observations and found that the three-moment MY scheme produced the best forecast.

Hail damage is not simply a function of maximum hail size; cumulative hail mass and number concentration are also important factors (Changnon 1999; Gilmore et al. 2004). Therefore, hail size distribution, cumulative mass, and number concentration predicted by various MP schemes are also worth evaluating in addition to MESH. Moreover, to more robustly evaluate and document the performances and behaviors of different MP schemes, and at the same time to achieve a better understanding of hail production and growth processes, more studies using real data for diverse types of hailstorms that may occur in different storm environments are still needed. Among the existing real-case studies that attempt explicit hail prediction to different degrees, MY06b and Snook et al. (2016) dealt with supercell storms, while L17 dealt with a pulse-type storm. To our knowledge, there has not been a real-case study that focuses on the multicellular type of hailstorm that may organize into mesoscale convective systems (MCSs).

In this study, explicit hail prediction of a long-lasting multicellular hailstorm event that occurred on 28 April 2015 in eastern China is investigated. On that day, the multicellular hailstorms swept southward through most

of Jiangsu Province, China, over about 7 h, producing egg-sized hailstones with diameters of around 20–50 mm on the ground. As in L17, this study employs the Advanced Regional Prediction System (ARPS) model (Xue et al. 2000, 2001, 2003). Simulation experiments are run at a 1-km horizontal grid spacing (instead of 3 km in L17) using one, two, and three-moment MY MP schemes. The goals of this study are twofold. First, the explicit hail forecast skill with different MP schemes, in terms of surface hailstone size distribution, cumulative hail mass, and number concentration, are evaluated. Second, combined with diagnostic analyses of microphysical terms, the reasons for the hail forecast differences between various MP schemes are explored. Surface accumulated precipitation, MESH, and cumulative hail mass and number concentration are examined, and objective MESH evaluation is performed using the fractions skill score (FSS) “neighborhood” technique (Ebert 2009).

The rest of this paper is organized as follows. In section 2, an overview of the 28 April 2015 multicellular hailstorm event is given. The ARPS model setup and hail forecast evaluation metrics are described in sections 3 and 4, respectively. Section 5 compares the explicit hail predictions using different MP schemes and investigates the causes of such differences. Finally, a summary and conclusions are presented in section 6.

2. Case overview

Hailstorms and other forms of severe weather events in the northeastern Asian Pacific coastal regions including eastern and northeastern China are often associated with upper-troposphere cutoff lows (COLs; Tao 1980; Nieto et al. 2005, 2008; Zhang et al. 2008). Most COLs occur over the northeastern part of China and often move southeastward to the coastal region of eastern China (Hu et al. 2010). The COLs can persist for several days and produce high convective instability. The hailstorm studied here occurred during this type of synoptic situation, where an upper-level COL swept from north to south across Jiangsu Province in eastern China, producing severe hail. According to the severe weather reports of the Chinese Meteorological Administration (Fig. 1), parts of Shandong Province, north of Jiangsu, were first hit by intense small-size hailstones (<10 mm) in the morning (around 0900 LST). In the afternoon, multicellular hailstorms formed along the west border of Jiangsu Province and produced a large number of egg-sized hailstones on the ground with the maximum observed size being ~10 cm in Yizheng, China. The hail fall from these storms extended through the western part of Jiangsu Province from north to

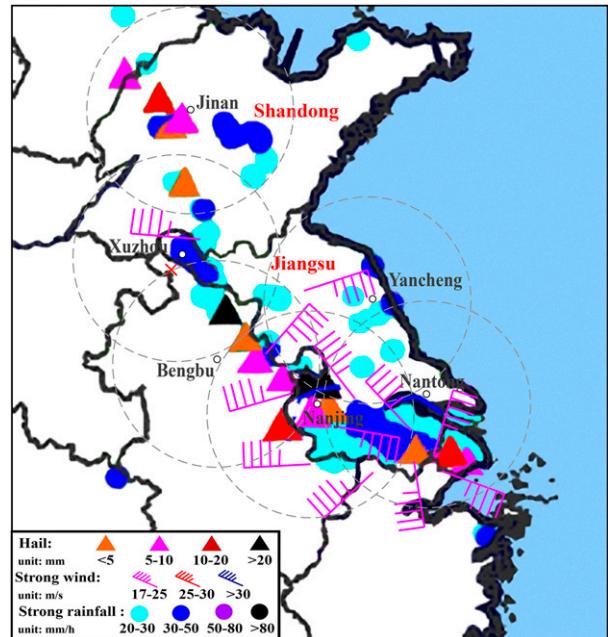


FIG. 1. Map of 24-h reports of severe weather in eastern China, starting from 1500 UTC 27 Apr 2015. Provinces of Shandong and Jiangsu are labeled in red. The open black circles indicate the locations of six operational S-band radars at Jinan, Xuzhou, Bengbu, Yancheng, Nanjing, and Nantong. The large gray dashed circles denote the 230-km range ring for each radar. The red × denotes the location of the extracted sounding shown in Fig. 4. The base map is obtained from the severe weather report maps of the Chinese Meteorological Society.

south, and hail fall was continuous for as long as 7 h from around 0700 through 1400 UTC. In addition, intense lighting and damaging surface winds (~23 m s⁻¹) were also reported.

The multicellular hailstorms were observed by multiple operational S-band Doppler radars, including those at Jinan, Xuzhou, Bengbu, Yancheng, Nanjing, and Nantong, China (see Fig. 1). Figure 2 presents the composite (column maximum) reflectivity from these radars. The life-span of this long-lasting multicellular hailstorm system can be characterized by two episodes. In the first episode, a series of multicellular hailstorms were initiated, and the storms intensified along the northwest border of Jiangsu Province. The storms organized into a northwest–southeast line and moved southeastward (Figs. 2a,b). In the second episode, from 1100 UTC onward, as the hailstorms moved southeastward, the line gradually evolved into a bow-shaped echo, with the middle portion of the line bulging eastward (Figs. 2c–f). By 1400 UTC, the apex of the bow had almost reached the southeast corner of Jiangsu, while the southern portion of the bow was oriented from east to west near the southern border of Jiangsu, with its

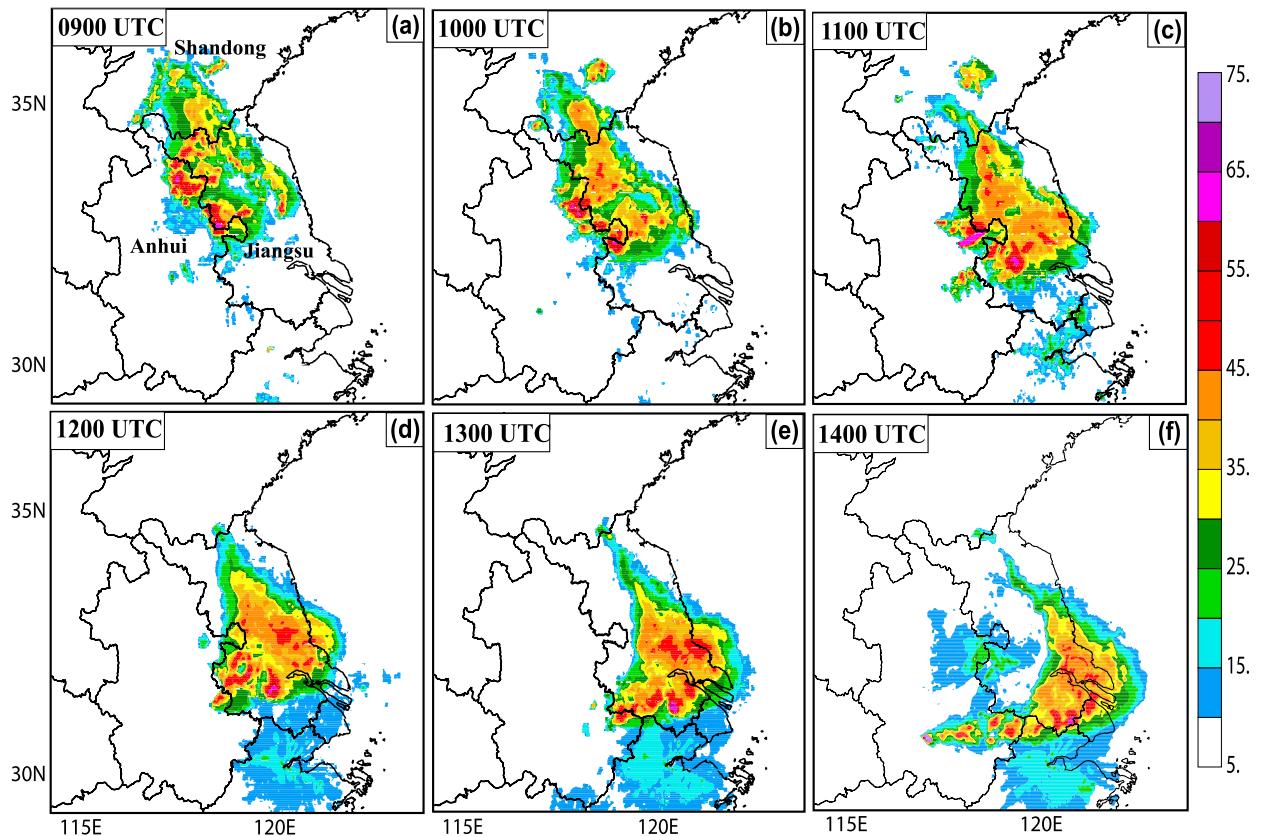


FIG. 2. Composite (column maximum) reflectivity fields of operational radars from (a) 0900 to (f) 1400 UTC at a 1-h interval. Provinces of Shandong, Jiangsu, and Anhui are labeled in (a).

western tail extending well into Anhui Province (Fig. 2f). The system started to weaken 1400 UTC. At all times shown, observed reflectivity exceeds 60 dBZ within some of the storm cells (Fig. 2).

The synoptic patterns associated with this event are shown in Fig. 3. At 0600 UTC (1400 LST) 28 April, the northeastern coastal regions of China were beneath a deep, positively tilted, semipermanent upper-level East Asian trough (EAT). The East Asia upper-tropospheric jet stream (EAJS) was located at the southern periphery of the trough (Fig. 3a). The jet core over land was located at $\sim 30^{\circ}\text{N}$, 120°E , with a maximum wind exceeding 50 m s^{-1} (Fig. 3a). Jiangsu Province (solid red line in Fig. 3a) was located ahead of the EAT and underneath the front-left (more so at earlier times) exit region of the EAJS, where favorable positive vorticity advection from the EAT and the upper-level divergence near the front-left EAJS exit region acted together to destabilize the atmosphere. Moreover, as seen in Figs. 3b–d, from middle to low altitudes, two COLs (denoted C) were embedded within the EAT, with one over the eastern coast of China and the other over the East China Sea.

Strong cold advection is found southeast of the western COL at 500 hPa (Fig. 3b) directly over Jiangsu Province. At 850 hPa (Fig. 3b) and at the surface (Fig. 3d), a prominent convergence line is present between the two cyclonic circulations; this convergence line is also the convergence boundary between the warm (temperature $>28^{\circ}\text{C}$) unstable (CAPE reaching 1500 J kg^{-1}) air mass from the southwest that is partly associated with the southern part of the western cyclonic circulation and the much colder air ($<16^{\circ}\text{C}$) with no CAPE from the northeast that is part of the eastern cyclone. The two low-level cyclones are responsible for setting up a strong convergence zone and moderately high CAPE in the air south of the convergence line. The convergence forcing, coupled with the destabilizing upper-level circulations, creates an environment favorable for intense deep convection.

Because of the lack of observed soundings near the time of hailstorm initiation, a sounding is extracted at 0600 UTC from the NCEP GFS $1^{\circ} \times 1^{\circ}$ final analysis (FNL) at 34°N , 117°E , which is about 30 km southwest of Xuzhou (see Fig. 1). The hailstorm first initiates near Xuzhou approximately 1 h later at around 0700 UTC.

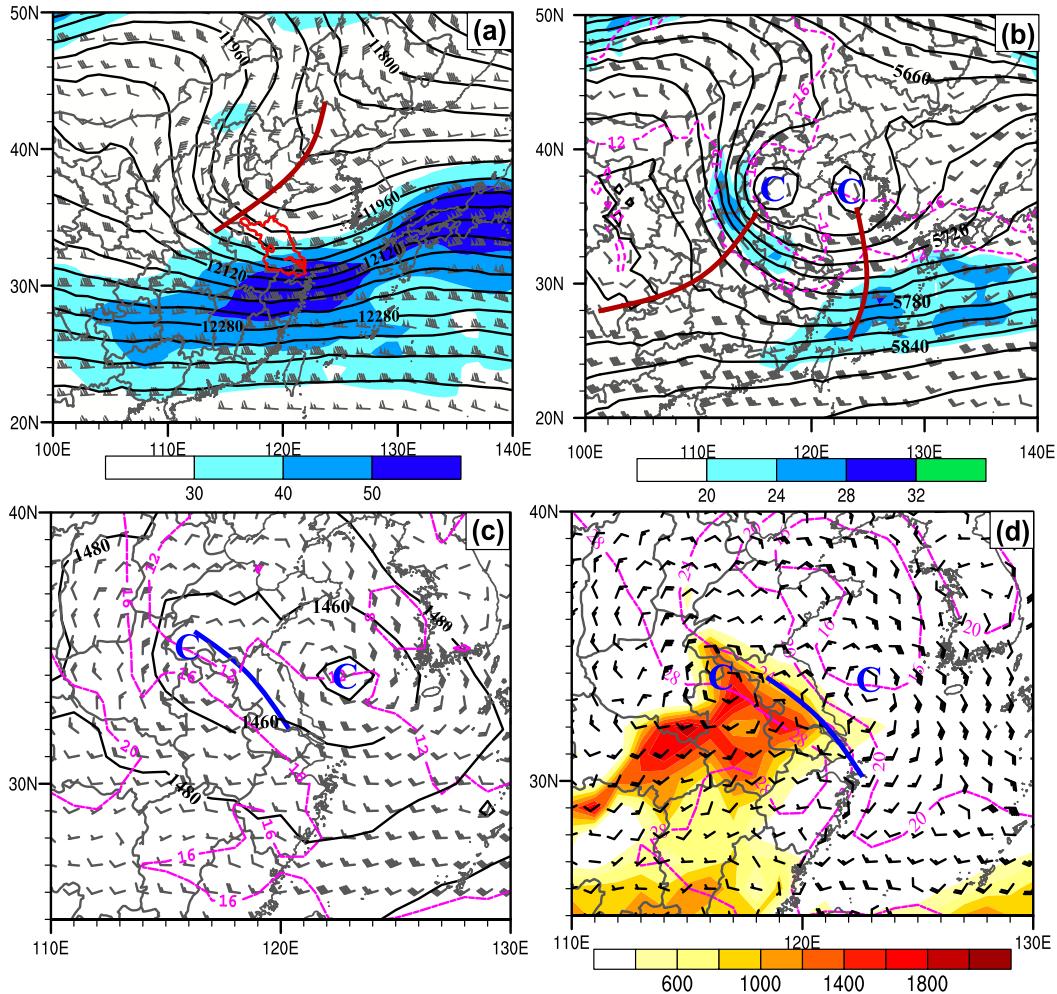


FIG. 3. Synoptic features of (a) 200, (b) 500, (c) 850, and (d) 1000 hPa at 0600 UTC 28 Apr 2015 showing wind barbs (one full barb denotes 2.5 m s^{-1}), temperature (magenta dashed contours; interval of 4°C), and geopotential height (solid black contours; gpm). The shading denotes the horizontal wind speed (m s^{-1}) in (a) and (b) and CAPE (J kg^{-1}) in (d). The thick solid brown lines in (a) and (b) and the thick solid blue lines in (c) and (d) indicate trough lines and shear lines, respectively. Jiangsu Province is outlined by a red solid line in (a), and a blue uppercase C denotes a COL at 500 hPa and cyclonic circulations at 850 and 1000 hPa. The maps are drawn from NCEP GFS FNL data.

The sounding (Fig. 4) has a CAPE of 1433 J kg^{-1} and a convective inhibition of -5 J kg^{-1} . The situation is characterized by strong vertical wind shear, with southeasterly winds below 850 hPa and southwesterly to northwesterly winds within the layers above. The bulk Richardson number is 31.1, and the 0–6-km vertical wind shear is $\sim 24.5 \text{ m s}^{-1}$; these values are generally considered conducive for long-lasting severe convection (Weisman and Klemp 1984). A capping inversion is present between 850 and 800 hPa; this inversion is sufficiently weak that it can be overcome by strong low-level convergence while convection is generally suppressed elsewhere, a situation favoring concentrated, intense, deep convection. Above the

inversion, the air mass is dry and cold, which may be a result of the previously mentioned cold advection. Many previous studies (e.g., Costa et al. 2001; L17) have noted that a dry midlevel layer over a warm moist layer near the surface is favorable for larger hailstones reaching the surface because of reduced melting of hail under such conditions. We note that there are uncertainties involved with extracting a sounding from the FNL dataset; however, this sounding represents the best available source of information about local environmental conditions. Overall, the sounding indicates a conducive environment for deep convection and a high likelihood for the production of large hailstones.

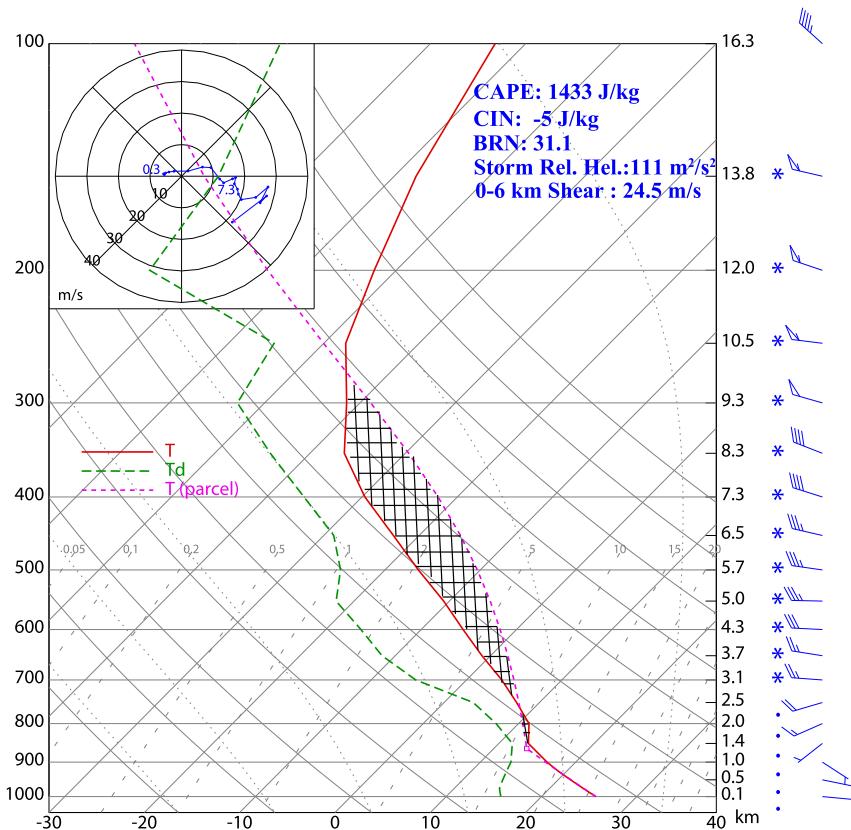


FIG. 4. Skew T plot of a sounding extracted from NCEP FNL data at 34°N, 117°E at 0600 UTC 28 Apr 2015 near Xuzhou.

3. Experiment setup

The Jiangsu hailstorm is simulated using the ARPS model (Xue et al. 2000, 2001, 2003). ARPS is a three-dimensional, nonhydrostatic compressible model using generalized terrain-following coordinates and was designed for regional to storm-scale atmospheric modeling and prediction. All simulations are initialized at 0000 UTC 28 April 2015 and are run for 16 h. The initial condition and boundary conditions at 6-h intervals are obtained from the NCEP final analysis data at $1^\circ \times 1^\circ$ resolution.

Two one-way nested grids at horizontal grid spacings of 3 and 1 km are used (Fig. 5). The 3-km domain covers an area of $1200 \times 1200 \text{ km}^2$ and is centered at 32.5°N, 118.5°E. The 1-km domain is $460 \times 460 \text{ km}^2$ in size and covers almost all of Jiangsu Province. Both domains have 53 vertical levels, which are stretched using a hyperbolic tangent function as described in Xue et al. (1995), with vertical grid spacing varying from 50 m at the surface to nearly 1000 m at the model top; the average vertical grid spacing is 500 m. The upper and lower boundaries are set as rigid walls, and a two-layer soil model is applied to facilitate the calculations of surface fluxes based on the predicted surface temperature and

soil moisture content. Subgrid-scale turbulent mixing is parameterized using a 1.5-order turbulence kinetic energy (TKE) scheme, and radiative processes are parameterized via the NASA Goddard Space Flight Center long- and shortwave radiation schemes. Fourth-order advection is used in both horizontal and vertical directions, and fourth-order computational mixing is applied to suppress numerical noise. More details on the ARPS physics schemes and their settings can be found in Xue et al. (2001, 2003), together with the references for the parameterization schemes.

As noted by previous studies (e.g., Loftus et al. 2014; Snook et al. 2016; L17), hail forecast errors are closely tied to uncertainties within MP schemes. Herein, simulations are conducted using the MY one-, two-, and three-moment schemes (MY05a; MY05b). Morrison and Milbrandt (2011) showed that even two very similar two-moment MP schemes could produce distinct differences in simulated storms because of differences in details of the schemes. For this reason, we choose to limit ourselves to the comparison of MY schemes having the same treatment of microphysical processes but predicting different number of moments or diagnosing

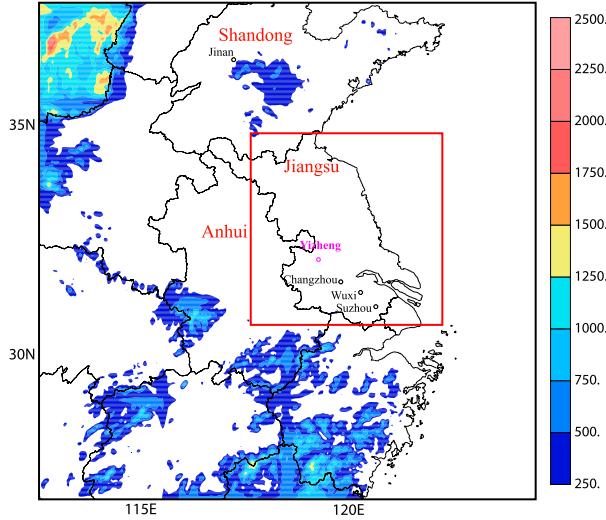


FIG. 5. Model domains of 3- (outer black box) and 1-km (inner red box) grid spacings for the hailstorm simulations. Terrain elevation (m) is plotted in color. Cities of Jinan, Changzhou, Wuxi, and Suzhou where hailstones are reported are labeled in black. The city of Yizheng, where maximum hail size is registered at around 10 cm, is labeled in magenta. Provinces of Shandong, Jiangsu, and Anhui are labeled in red.

one of the DSD parameters. This way, focus is placed on the effects of the number of predicted moments on the hail forecast. In the MY schemes, six distinct hydrometeor categories (i.e., cloud water, cloud ice, rain, snow, graupel, and hail) are included. The PSD of each hydrometeor is represented by a gamma distribution function,

$$N_x(D) = N_{0x} D^{\alpha_x} \exp(-\lambda_x D), \quad (1)$$

where $N_x(D)$ is the total number concentration per unit volume of diameter D for hydrometer category x , α_x is the shape parameter, giving a measure of the spectral width, and N_{0x} and λ_x are the intercept and slope parameters, respectively.

The three-moment scheme is used as the control experiment, and three other experiments are performed using the one- and two-moment MY schemes. Table 1 summarizes the key parameters of all experiments. Two of the simulations, called FixA and DiagA, use variants of the MY two-moment schemes with different treatment of the shape parameter. FixA sets the shape parameter to a default constant value of 0 for all hydrometeor categories, while in DiagA, the shape parameters of hail and other hydrometeors are diagnosed from the mean mass diameter of the corresponding categories based on Eqs. (12) and (13) of MY05a. The one-moment MY scheme is also tested with intercept and shape parameters set to their default constant values. The above configurations are similar to

TABLE 1. List of 3- and 1-km simulations.

Expt	Description
CNTL	Three-moment; full version of scheme
DiagA	Two-moment; diagnosis of relation for $\alpha_x = f(D_{mx})$, $x \in (\text{rain } r, \text{ ice } i, \text{ snow } s, \text{ graupel } g, \text{ hail } h)$
FixA	Two-moment; fixed $\alpha_x = 0$
Single	One-moment; $\alpha_x = 0$, $N_{ic} = 1 \times 10^8 \text{ m}^{-3}$, $N_{0r} = 8 \times 10^6 \text{ m}^{-4}$, $N_{0s} = 3 \times 10^6 \text{ m}^{-4}$, $N_{0g} = 4 \times 10^5 \text{ m}^{-4}$, $N_{0h} = 4 \times 10^4 \text{ m}^{-4}$

those of L17 except for the inclusion of the 1-km nested grid. Because the 1-km grid is expected to simulate the hailstorm better, we focus on results of the 1-km grid in this paper. We also examined the forecasts from the 3-km grid (not shown); the dominant cells simulated on the two grids are found to be generally similar, although some differences exist in storm intensity at small scales.

4. Evaluation metrics for hail prediction

Three metrics for explicit hail prediction are used to evaluate hail forecast skill for the various MP schemes within the sensitivity experiments. They are MESH (Witt et al. 1998), maximum hail size D_{max} (MY06a), and surface accumulated hail number concentration (SAHNC). In addition, an objective neighborhood-based evaluation technique, the fractions skill score, is used to verify the simulated MESH against the radar-derived counterpart. L17 examined the accumulated surface precipitation and MESH fields based on the simulations of a pulse hailstorm but not the D_{max} and SAHNC. MESH and D_{max} were also examined in Snook et al. (2016) for a supercell-storm case.

a. MESH

As described in L17, the MESH algorithm uses a weighted integration of radar reflectivity exceeding 40 dBZ above the melting level to obtain an estimate of the maximum size of hail occurring at the surface. Following L17, reflectivity datasets from multiple radars are interpolated to the model grid to derive MESH. Since the MESH algorithm was only configured for hail sizes larger than 19 mm (Witt et al. 1998), and Cintineo et al. (2012) and L17 only evaluated MESH down to the size of 21 and 19 mm, respectively, MESH values below 20 mm are excluded in this study. More details about the MESH algorithm can be found in L17.

We note that since the MESH algorithm relies entirely upon the weighted integration of radar reflectivity exceeding 40 dBZ above the 0°C level to estimate hail size at the surface, there may exist some biases within

the derived MESH swath (e.g., [Cintineo et al. 2012](#); [Ortega et al. 2009](#)). Because no other high-quality, high-resolution observation of hail size is available, herein, we choose to use the high-resolution radar-derived MESH for verification of the hail simulations.

b. Maximum hail size

The maximum hail size D_{\max} ([MY06a](#)) is defined as the largest hail size for which the total number concentration of hail particles greater than a diameter is equal to the prespecified total number concentration N_{THRE} . For example, if D_{\max} is 40 mm, the total number concentration of hailstones larger than 40 mm is N_{THRE} . The D_{\max} parameter serves to identify the instantaneous presence of large hail within the storm. Following [MY06a](#), a threshold value of $N_{\text{THRE}} = 10^{-4} \text{ m}^{-3}$ is adopted here.

c. SAHNC

Given that the accumulated hail number is also important for hail prediction, SAHNC is proposed as a new parameter to estimate the surface accumulated number concentration of hail larger than a particular size. The SAHNC parameter is not only useful for identifying the surface accumulated hail size distribution but also helpful in understanding storm evolution. SAHNC is defined as an integration of the flux of large hail $R_h(D)$ at 60-s intervals during hail fall from T_0 to T_1 ,

$$N_{D_h}(D) = \int_D^{\infty} N_h(D^*) dD^*, \quad (2)$$

$$V_h(D) = \gamma a_h D^{b_h} e^{-f_h}, \quad (3)$$

$$R_h(D) = N_h(D) V_h(D), \quad (4)$$

$$\text{SHNAC}(D) = \int_{T_0}^{T_1} R_h(D) dt, \quad (5)$$

where $N_{D_h}(D)$ is the total number concentration of hail larger than diameter D and the size distribution of hail is described by a gamma distribution function as Eq. (1). Terminal fall velocity at the surface for a hailstone with diameter D is given by Eq. (3), where $\gamma = (\rho_0/\rho)^{1/2}$ is the density correction factor with ρ_0 being the reference density of 1.225 kg m^{-3} ([Foote and Toit 1969](#)) and ρ being the air density; a_h , b_h , and f_h are set to be 206.89, 0.6384, and 0.0, respectively, following [Ferrier \(1994\)](#).

d. Neighborhood-based hail forecast evaluation

As reviewed in [Casati et al. \(2008\)](#), evaluation of forecasts from high-resolution models has been a subject of active research in recent years, and various evaluation metrics have been developed. Objective evaluation of hail forecasts is still very challenging, partly because of the lack of high-quality, high-resolution hail observations

([Snook et al. 2016](#)). In this study, the FSS neighborhood technique ([Roberts and Lean 2008](#)) is applied to MESH fields derived from multiple radar observations and the simulations to examine hail size forecast skill using different MP schemes. Distinct from traditional point-by-point evaluation techniques, FSS compares fractional coverage of forecasts against that of observations within a neighborhood centered at each grid point. By varying the neighborhood size and the MESH threshold, scale-dependent forecast skill can be assessed. Following [Roberts and Lean \(2008\)](#), FSS is defined as

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

where FBS is the fractional Brier score, given as

$$\text{FBS} = \frac{1}{N} \sum_{i=1}^N [\langle P_{F(i)} \rangle_s - \langle P_{O(i)} \rangle_s]^2. \quad (7)$$

In Eqs. (6) and (7), N is the total number of grid boxes in the predefined neighborhood (within a given radius); $P_{F(i)}$ and $P_{O(i)}$ are the fractional areas at the i th neighborhood of forecast and observation, respectively. They are analogous to the probability that a given neighborhood contains values larger than the prespecified threshold. Therefore, FSS compares fractional coverage over a neighborhood of given size rather than values at each grid box, and FSS values range from 0 to 1. A score of 1 signifies a forecast perfectly matching the observation within a specific neighborhood for a given intensity threshold, while 0 signifies a complete mismatch. A forecast is considered to be skillful when the FSS value exceeds $\text{FSS}_{\text{useful}}$, which is defined as ([Roberts and Lean 2008](#))

$$\text{FSS}_{\text{useful}} = 0.5 + \frac{f_{\text{obs}}}{2}, \quad (8)$$

where f_{obs} is an average of observed fraction within the entire domain. By calculating FSS at a variety of spatial scales (neighborhood sizes) and MESH thresholds, one can determine how the forecast skill varies with spatial scale and at which scale a forecast has useful skill for a given MESH threshold.

5. Results

In this section, results of the simulations are presented. First, to validate the simulations of the multicell hailstorm system, simulated composite (column maximum) radar reflectivity is compared with corresponding radar observations. Explicit hail forecast skills using

various MP schemes are then evaluated, in terms of MESH, surface accumulated solid water mass, and SAHNC. Neighborhood-based FSSs for simulated MESH are calculated against radar-derived MESH. The differences in hail distribution characteristics within storms simulated using different MP schemes are also investigated. To understand the reasons behind the differences among various MP schemes for hail prediction, microphysical budget analyses are performed.

a. Simulated storm evolution

As discussed in section 2, the life span of this multicellular hailstorm can be characterized by two episodes between 0700 and 1400 UTC 28 April 2015. Figures 6 and 7 show simulated composite reflectivity fields from the experiments and radar observations from one time during each episode, at 1100 and 1400 UTC, respectively. Comparisons with the radar observations (Figs. 6e, 7e) indicate that the time and location of the hailstorm's initiation along the northwest border of Jiangsu Province, as well the later organization into a large bow-shaped echo, are well reproduced in the simulations (Figs. 6a–d, 7a–d). The direction of movement of the simulated storms is also in general agreement with observations.

Although the storms' evolution and motion are generally reproduced quite well in the simulations, there exist significant storm intensity differences among the experiments using various MP schemes. Experiment Single underpredicts the reflectivity magnitude, having few instances of reflectivity exceeding 60 dBZ, and the stratiform precipitation region (<35 dBZ) in Single is larger compared to simulations using multimoment schemes (Figs. 6a, 7a). This result differs from those of some previous studies (e.g., Morrison et al. 2009; Bryan and Morrison 2012; Baba and Takahashi 2014); their studies noted that the stratiform precipitation region in idealized two-dimensional squall-line simulations was smaller using a one-moment than a two-moment scheme, and they attributed it to decreased rain evaporation rates in the two-moment schemes in the trailing stratiform region. A possible reason for this behavior in our case may be that for the multimoment scheme, the size-sorting mechanism may allow larger-sized hailstones to fall rapidly toward the ground, giving less time for hail mass advection downwind of the updraft and hence the smaller stratiform precipitation region.

The reflectivity magnitudes from the multimoment schemes are more or less overpredicted when compared with radar observations (Figs. 6b–e, 7b–e). Thus, storm intensities produced by the multimoment schemes are not too different based on radar reflectivity, while the hail prediction skill of each scheme exhibits significant discrepancies (see Figs. 8–12). Moreover, in all cases, the

westward extension of the reflectivity toward Anhui Province is underpredicted, and the simulated system exhibits slower southward movement (possibly because of uncertainties in the initial condition), resulting in the entire system being displaced almost 80 km northward compared to the observations by 1400 UTC.

b. Hailstone forecast and evaluation

As the main goal of this study is to evaluate the hail forecast skills of various MP schemes, predicted hail size distribution features, including hailstone size, mass, and number concentration, will be examined in this section.

1) RADAR-BASED HAIL FORECAST EVALUATION USING MESH

Swaths of MESH derived from the forecasts and from radar observations between 0600 and 1600 UTC at 5-min intervals are presented in Fig. 8. There is one primary, nearly continuous MESH swath in the radar observations (Fig. 8e), with MESH values exceeding 40 mm in many locations. Along the radar-indicated MESH swath, egg-sized hailstones (with diameters of 20–50 mm) were reported in several cities, including Yizheng, Changzhou, Wuxi, and Suzhou, China. Among them, maximum MESH values exceed 80 mm over Yizheng, which coincides with the location of largest hail reported during this storm—a report from Yizheng of hail over 100 mm in diameter. Compared with radar-derived MESH, the MESH swath derived from experiment Single exhibits smaller maximum hail sizes, with several scattered cores of MESH values of no more than 35 mm predicted (Fig. 8a). In FixA and DiagA, which use two-moment schemes, MESH swaths exhibit significant differences from observations. Maximum MESH in FixA is highly overestimated; the MESH swath from this experiment exhibits large areas of MESH exceeding 70 mm (Fig. 8b). In DiagA, MESH is significantly underestimated, not exceeding 30 mm at any point (Fig. 8c). The MESH swath of the control (CNTL) appears to match well with the observed swath, with the MESH values in the range of 40–50 mm within narrow cores (Fig. 8d). A narrow core of high MESH values (over 80 mm) is present within the primary MESH swath in CNTL (Fig. 8d). This is generally consistent with the maximum registered hail size in Yizheng, although the swath is displaced tens of kilometers northward from the observations because of the overall northeastward position error of the storms.

FSSs at different neighborhood radii and MESH thresholds and the corresponding FSS_{useful} values are presented in Fig. 9. The scale at which FSS exceeds FSS_{useful} can be considered the “skillful scale” of forecast (Roberts and Lean 2008).

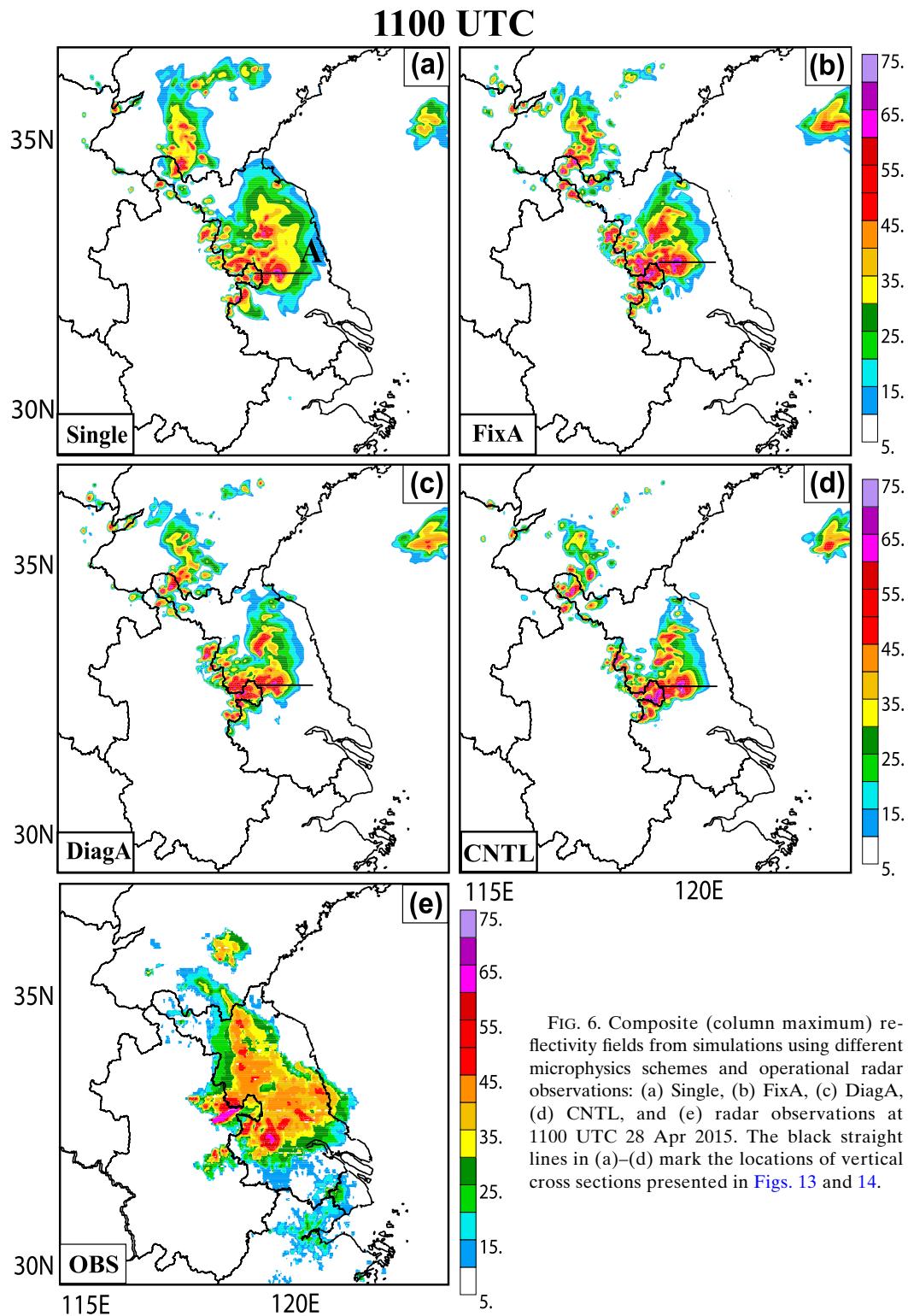


FIG. 6. Composite (column maximum) reflectivity fields from simulations using different microphysics schemes and operational radar observations: (a) Single, (b) FixA, (c) DiagA, (d) CNTL, and (e) radar observations at 1100 UTC 28 Apr 2015. The black straight lines in (a)–(d) mark the locations of vertical cross sections presented in Figs. 13 and 14.

1400 UTC

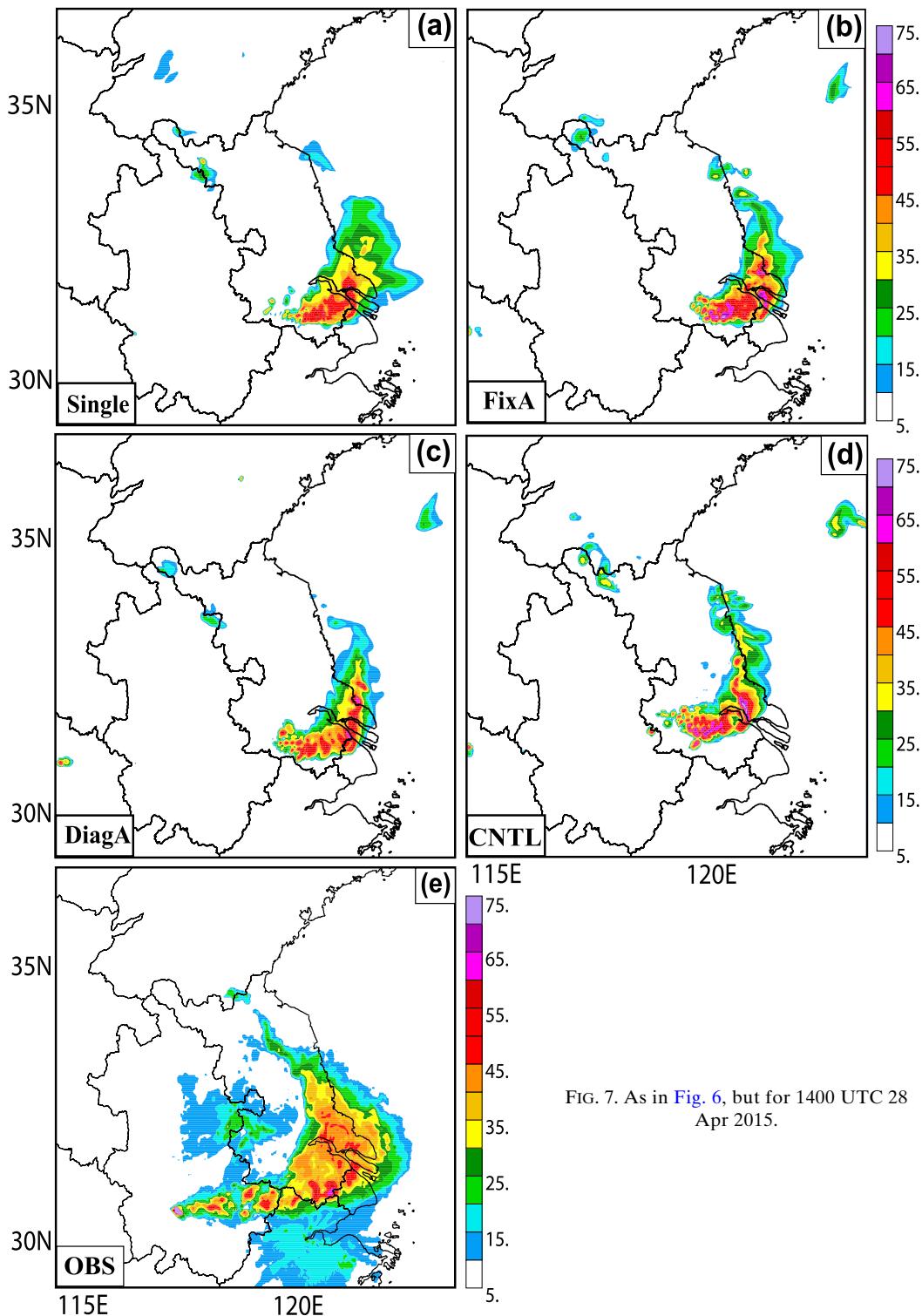


FIG. 7. As in Fig. 6, but for 1400 UTC 28 Apr 2015.

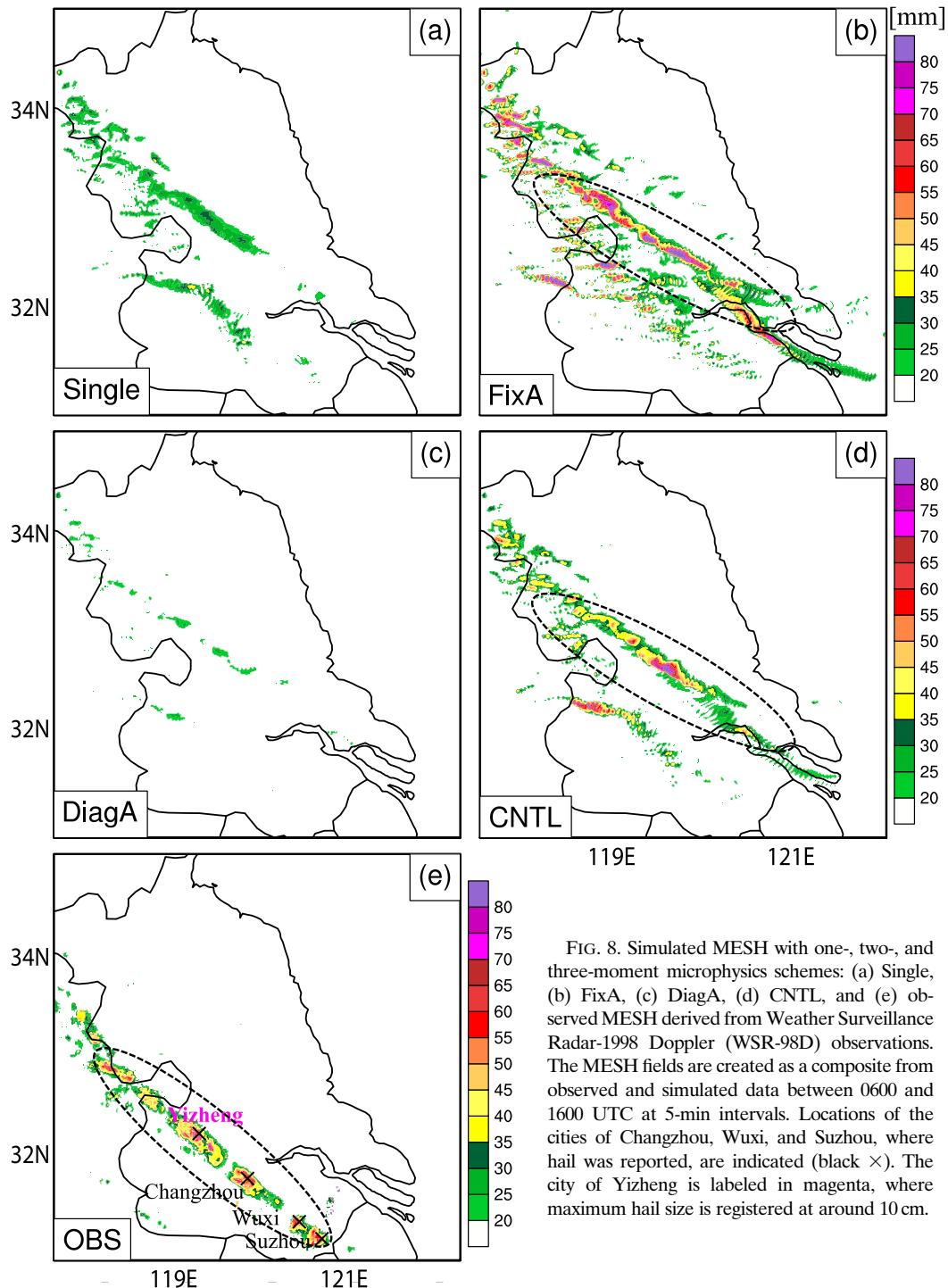


FIG. 8. Simulated MESH with one-, two-, and three-moment microphysics schemes: (a) Single, (b) FixA, (c) DiagA, (d) CNTL, and (e) observed MESH derived from Weather Surveillance Radar-1998 Doppler (WSR-98D) observations. The MESH fields are created as a composite from observed and simulated data between 0600 and 1600 UTC at 5-min intervals. Locations of the cities of Changzhou, Wuxi, and Suzhou, where hail was reported, are indicated (black \times). The city of Yizheng is labeled in magenta, where maximum hail size is registered at around 10 cm.

FSSs of FixA and CNTL with MESH thresholds of 30 and 40 mm and Single with MESH thresholds of 30 mm increase with increasing neighborhood radius, and higher scores are achieved at the lower MESH threshold (Fig. 9). For MESH thresholds of 30 and 40 mm (Figs. 9a,b), experiment CNTL achieves useful

skill for neighborhood radii of 47 and 50 km and larger, respectively. The large neighborhood radius for useful skill is likely due in large part to the northeastward displacement error of the model storms. The comparisons of FSSs from the experiments further confirm that the three-moment scheme outperforms others in terms

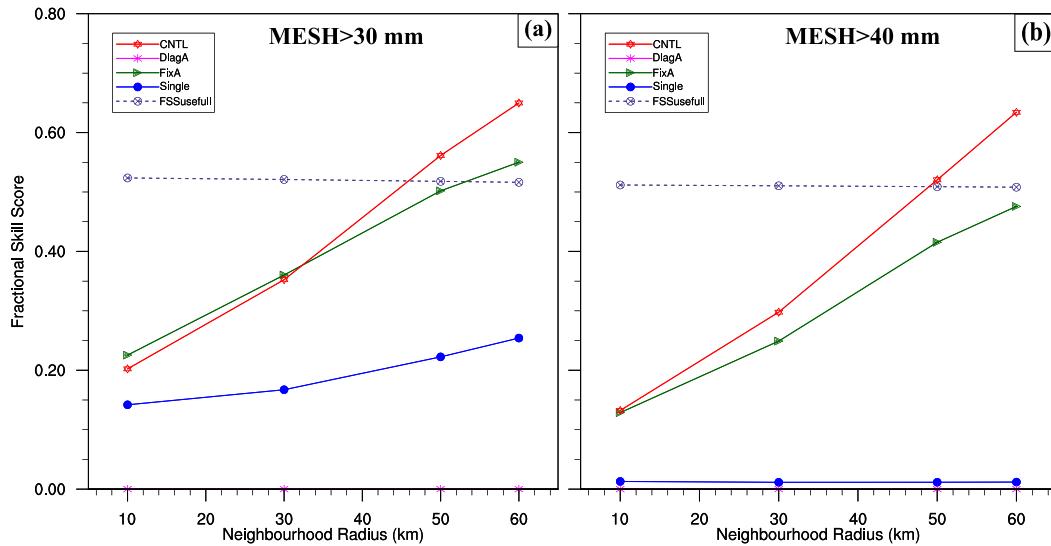


FIG. 9. Fraction skill scores of MESH from 0600 to 1600 UTC for simulations Single, FixA, DiagA, and CNTL with MESH exceeding (a) 30 and (b) 40 mm, over neighborhood radii of 10, 30, 50, and 60 km. The FSS_{useful} value is also shown.

of hail forecast skill, especially for hail exceeding 40 mm, as indicated by the notably higher FSSs at all spatial scales larger than 10 km (Fig. 9b). Experiment Single has no skill at any spatial scale for hailstones larger than 40 mm, as indicated by the zero FSSs at all neighborhood radii (Fig. 9b). Similarly, DiagA shows no skill for predicting hailstones larger than 30 mm (Fig. 9a).

2) SURFACE ACCUMULATED HAIL MASS AND NUMBER CONCENTRATION

Since there are no in situ hail-count observations, the simulated hail mass and number concentration fields are compared among the experiments in light of the registered hail reports. Surface accumulated hail mass fields from the experiments, calculated between 0600 and 1600 UTC throughout almost the entire life span of the hailstorm, are shown in Fig. 10. The corresponding SAHNC, derived from the model output within the same period at 60-s intervals for hail diameter thresholds of 30 and 40 mm, are shown in Figs. 11 and 12.

The storms in Single produce a northwest–southeast-oriented swath of hail mass, with peak values of around 80 mm, mainly concentrated within a narrow band approximately 10 km in width (Fig. 10a). In DiagA, the accumulated hail mass amounts are the smallest among all experiments, with peak values below 40 mm (Fig. 10c). The width of the primary hail mass swath in FixA is the widest (around 20 km) among the experiments. Peak values in FixA and CNTL are similar (~80 mm), and the hail mass predicted by CNTL appears to be more concentrated along a relatively straight path (Fig. 10d).

Distributions of SAHNC from the experiments generally coincide with their hail mass distributions. For example, SAHNC of Single with hail size larger than 30 mm (SAHNC30) within the primary narrow hail mass band is around $10^5\text{--}10^6\text{ m}^{-2}$ (Fig. 11a), and the number concentration of hail larger than 40 mm (SAHNC40) is around 10^3 m^{-2} (Fig. 12a). In contrast, SAHNC30 of DiagA is around $10^2\text{--}10^3\text{ m}^{-2}$, approximately one to two orders of magnitude smaller than other experiments (Fig. 11c). Only a few small patches of SAHNC40 are presented in Fig. 12c. Although the magnitudes of SAHNC30 and SAHNC40 from FixA are similar to those from CNTL, FixA produces more large hailstones at the surface than other experiments and predicts a swath almost twice the width of that in CNTL.

Forecasts of accumulated hail mass and number can also be cross-referenced with hail reports and photographs from the event to infer their level of accuracy. More hail photographs of this event can be found online (<http://news.eastday.com/c/20150428/u1a8688796.html>). In some areas, photographs and reports indicate that the depth of surface accumulated hail exceeded 10 cm, which is more or less consistent with the surface accumulated solid precipitation of CNTL (Fig. 10d); it produces a concentrated hail mass band of over 8 cm in depth. If we assume hail to be spherical and transfer the reported hail depth and size to hail number accumulated at the surface, this corresponds to a value of $10^3\text{--}10^4\text{ m}^{-2}$ for hail larger than 4 cm. This is also better captured by CNTL than other experiments, although there is still overestimation in some areas. Based on these

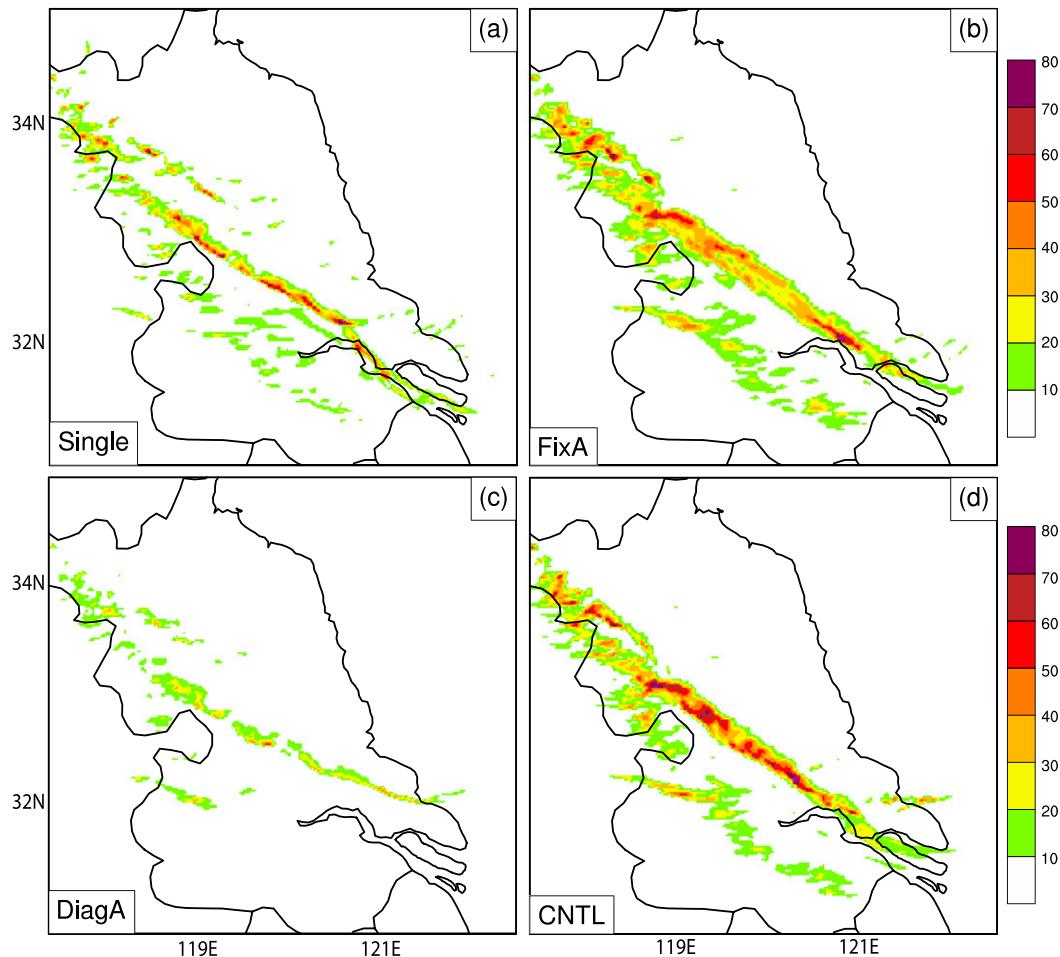


FIG. 10. Surface accumulated solid precipitation in water-equivalent depth (mm) between 0600 and 1600 UTC from simulations (a) Single, (b) FixA, (c) DiagA, and (d) CNTL.

inferences, the surface accumulated hail mass and number predicted by CNTL appear to be more accurate compared with other experiments.

3) HAILSTONE DISTRIBUTION CHARACTERISTICS WITHIN STORMS

Given the significant differences in the predicted surface hail size distributions among various MP schemes, we next examine hail distribution properties within simulated storms. Microphysical fields, including the hail mass content Q_h , total hail number concentration N_{th} , maximum hail size D_{max} , and reflectivity Z are examined at 1100 UTC, when the cells are vigorous and well developed (see Figs. 13 and 14). Vertical cross sections are taken from west to east, passing through the primary hail mass core of the simulated cells; the locations are indicated by the thick black lines in Figs. 6a–d. Hail size spectra at some typical points in sensitivity experiments are also examined (Fig. 15).

The hail distribution characteristics within the storms from various experiments exhibit large differences (Fig. 13). For example, FixA and DiagA produce copious hail mass aloft, with a peak Q_h of 13 g m^{-3} (Fig. 13b: $x \sim 200\text{--}216 \text{ km}$, $z \sim 7\text{--}9 \text{ km}$; Fig. 13d: $x \sim 212\text{--}217 \text{ km}$, $z \sim 8\text{--}10 \text{ km}$), while the peak Q_h in CNTL and Single is only about 7 g m^{-3} (Figs. 13a,e: $x \sim 198\text{--}202 \text{ km}$, $z \sim 4\text{--}6 \text{ km}$). The N_{th} fields from the two-moment simulations are generally similar to CNTL: all have larger N_{th} values ($>10^3 \text{ m}^{-3}$) in the storm anvil region above the freezing level. The magnitudes of the peak N_{th} in FixA are one to two orders higher than the peak N_{th} values in CNTL, especially in the rear part of the cell (Fig. 13b: $x \sim 165\text{--}204 \text{ km}$, $z \sim 2\text{--}7 \text{ km}$). More discussions on the overprediction of the moments in FixA are provided in section 5c through budget analyses of hail source and sink terms. The N_{th} field diagnosed by Single is considerably different from the other runs; the peak in N_{th} is much smaller ($<10^{2.5} \text{ m}^{-3}$). Since N_{th} in

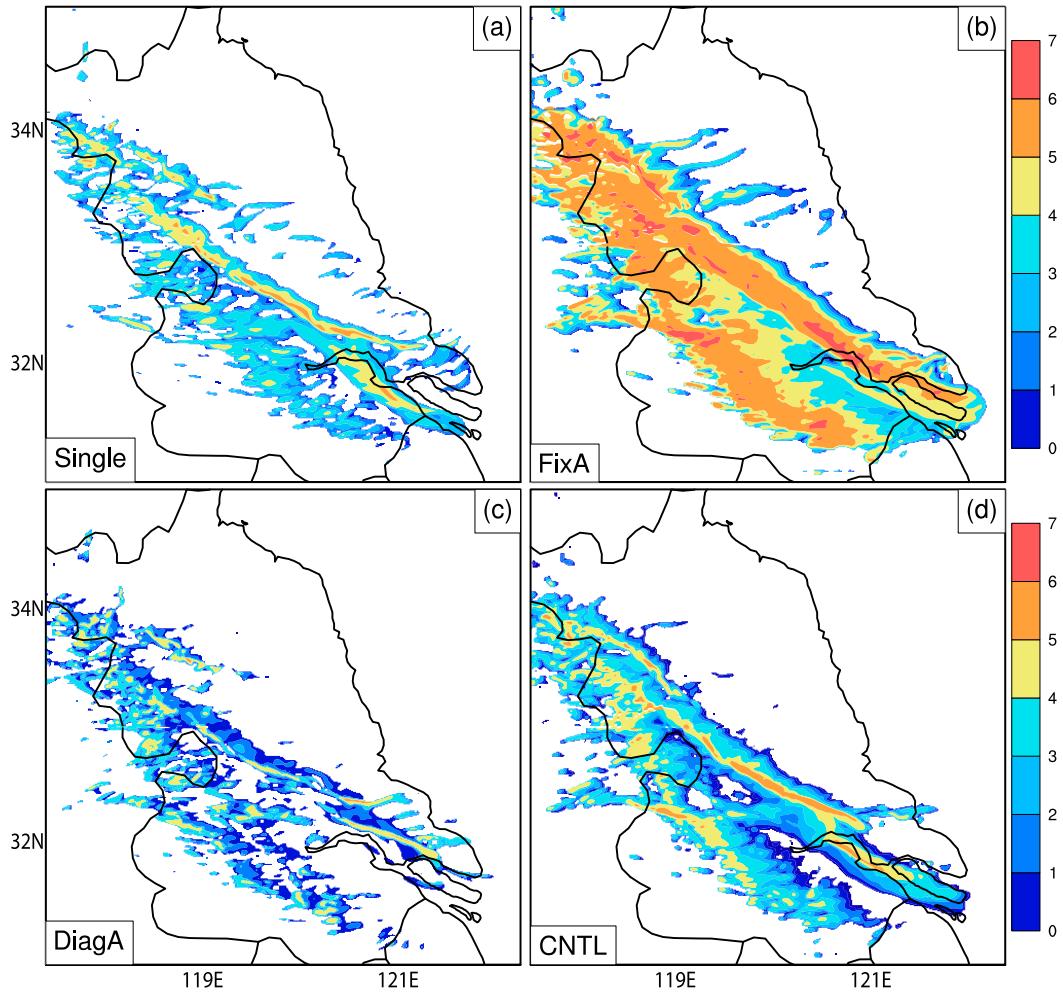


FIG. 11. SAHNC at 60-s intervals between 0600 and 1600 UTC 28 Apr 2015, derived from simulations (a) Single, (b) FixA, (c) DiagA, and (d) CNTL with hail diameter larger than 30 mm. SAHNC is in a base-10 logarithmic scale [i.e., $\log_{10}(\text{SAHNC}); \text{m}^{-2}$].

Single is a monotonic function of Q_h , the peaks of N_{th} and Q_h are collocated.

The D_{max} and reflectivity of Single (Fig. 14a) are also monotonically related to Q_h , and their peak cores are also collocated. Since only the mixing ratio is predicted in Single, all the diagnosed moments sediment at the mass-weighted fall speed, precluding any size sorting. In experiments using multimoment schemes, D_{max} and reflectivity generally increase toward the surface, and the high D_{max} and reflectivity columns are located almost directly below the corresponding Q_h cores (Figs. 13, 14b,d,e), consistent with a size-sorting process. To investigate the effect of size sorting on hail distribution, additional sets of experiments with the size-sorting effect suppressed in FixA_nosize_sorting and CNTL_nosize_sorting using the two- and three-moment schemes were conducted. In these experiments, size

sorting for hydrometeor species is disallowed by forcing all predicted moments to sediment at the mass-weighted fall speed. We examined the microphysical fields from the experiments with size sorting disabled and found that they exhibited substantial similarities to experiment Single. The Q_h and N_{th} fields in these experiments display a broader region of relatively weak gradients over most of the forward flank above the melting layer, with smaller D_{max} values (Figs. 13, 14c,f). This strongly suggests that the size-sorting effect plays an important role in controlling the hail distribution characteristics within the storm. In multimoment MY schemes, different moment-weighted terminal velocities enable size sorting of particles, leading to more realistic hail distribution properties in the vertical.

Hail size spectra within the storms for each experiment at 1100 UTC, in the main updraft region of each

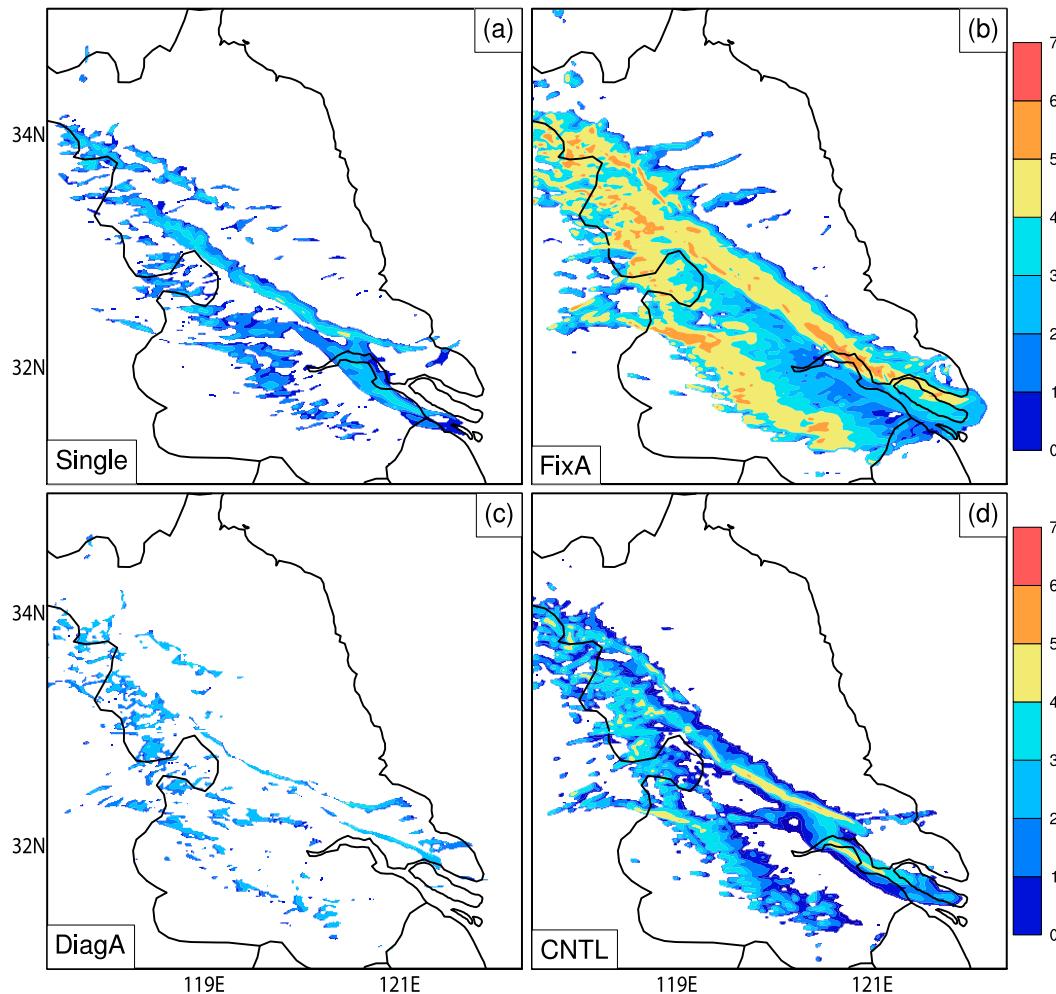


FIG. 12. As in Fig. 11, but for hail diameter exceeding 40 mm.

case (as indicated in Fig. 14), are plotted in Fig. 15. In FixA and Single, $\alpha_h = 0$; the DSD curves (black and green in Fig. 15) are exponential. FixA suffers from excessive size sorting, since Q_h sediments faster than N_{th} when α_h is fixed at 0 (as can be seen in Fig. 1 of MY05a), consistent with the unrealistically large D_{max} at the low levels (Fig. 14b). In contrast, the diagnosed α_h in DiagA is about 2.8 at this point (see Fig. 16), and size sorting is more limited. The hail size spectrum appears to be artificially narrowed in DiagA compared to CNTL, causing a shift of spectrum distribution toward smaller sizes. Figure 15b shows that, in CNTL, the hail size spectrum becomes broader as height decreases (Fig. 15b), corresponding to decreasing slope parameter. Meanwhile, D_{max} increases quickly as the ground is approached (Fig. 14e).

Furthermore, given that the shape parameter has significant effects on sedimentation and microphysical growth rates (Milbrandt and McTaggart-Cowan 2010; Mansell 2010; Dawson et al. 2014) and size sorting can

also affect size spectra of hydrometeors (MY05a), diagnostic analyses are performed to assess the differences in α_h among the experiments. The horizontally and temporally averaged α_h values within the storm, with the one standard deviation interval shaded, are plotted for each experiment between 0600 and 1600 UTC in Fig. 16. The mean α_h of CNTL decreases significantly from ~ 3.2 near the surface to ~ 0.2 near the melting layer (approximately 4 km above the surface). The decrease is almost linear with height up to approximately 2.5 km, and the decrease continues above 4 km. This shape parameter profile agrees with previous studies (MY06a; MY06b), which noted that large α_h mainly occurred below 600 hPa, with near-zero values above 600 hPa. The smaller α_h above the freezing level may partly result from the creation of hail via freezing of raindrops, which adds numerous small particles to the hail distribution (MY06b). Below the freezing level, smaller hail particles tend to melt quickly, increasing α_h . In any case, the α_h

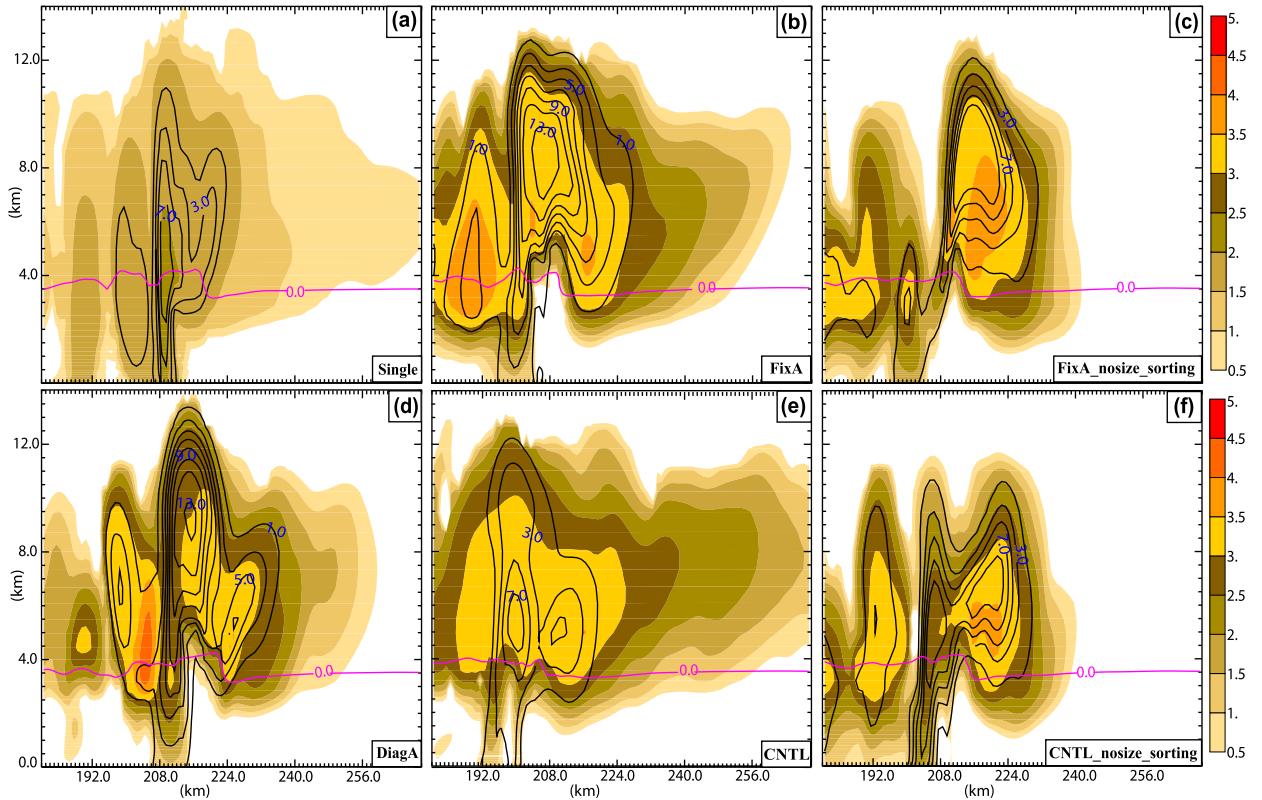


FIG. 13. Vertical cross sections of the mass content (black contours; g m^{-3}) and total number concentration of hail using a base-10 logarithmic scale [i.e., $\log_{10}(N_{\text{th}})$; shaded; m^{-3}], through the hail mass cores of the primary hail-producing storm A at 1100 UTC (Figs. 6a–d), for experiments of (a) Single, (b) FixA, (c) FixA_nosize_sorting, (d) DiagA, (e) CNTL, and (f) CNTL_nosize_sorting. The locations of the vertical slices are marked by solid black lines in Figs. 6a–d. The contour of 0°C (magenta) is also overlaid.

profile of CNTL indicates its variation with height from 0 to 5, suggesting that using a fixed α_h value is inappropriate.

Although the diagnosed α_h in DiagA exhibits a vertical profile with the correct trend, it differs quite significantly from that of CNTL. The diagnosed α_h does not decrease quickly with height and maintains high values (around 2.9) at upper levels (Fig. 16). These result directly from the α_h diagnostic formula used [see Eq. (12) in MY05a] that keeps the diagnosed value within a range of 2.8 to 4.5 for the hail with mean mass diameter below 8 mm. This diagnostic relation was obtained using a one-dimensional model where only the sedimentation process was considered (MY05b); it appears to be inaccurate compared to values produced by the three-moment scheme.

c. Hail production budget analyses

To gain additional physical insights regarding the differences between MP schemes, budget analyses of microphysical processes responsible for hail production within the simulated storms are performed. The hail production terms are integrated over the entire 1-km

simulation domain from surface to model top, according to the Q_h and N_{th} prediction equations for the MY scheme [see Eqs. (A7) and (A13) in MY05b]. Figure 17 shows the time series of Q_h source and sink terms for experiments Single, FixA, DiagA, and CNTL, while Fig. 18 shows the time series of N_{th} source and sink terms for FixA, DiagA, and CNTL. In Single, only the mixing ratio is predicted, and the total number concentration is diagnosed and monotonically related to the mixing ratio (MY05b). The main microphysics processes contributing to Q_h production are hail collection of rain (colqrh) and cloud water (colqch), and hail melting to rain (meltqh; Fig. 17). The main processes contributing to N_{th} production are ice and rain collision to produce hail (nclirh), graupel conversion to hail (ncong), and hail melting to rain (nmelth; Fig. 18). Other processes, for example, hail collection of snow or ice, are minimal and are therefore not shown.

The terms colqrh, colqch, and meltqh in FixA are significantly larger than those in other experiments between 0600 and 1500 UTC, with peak rates of approximately 3000, 8000, and $-10\,200$ kiloton min^{-1} at

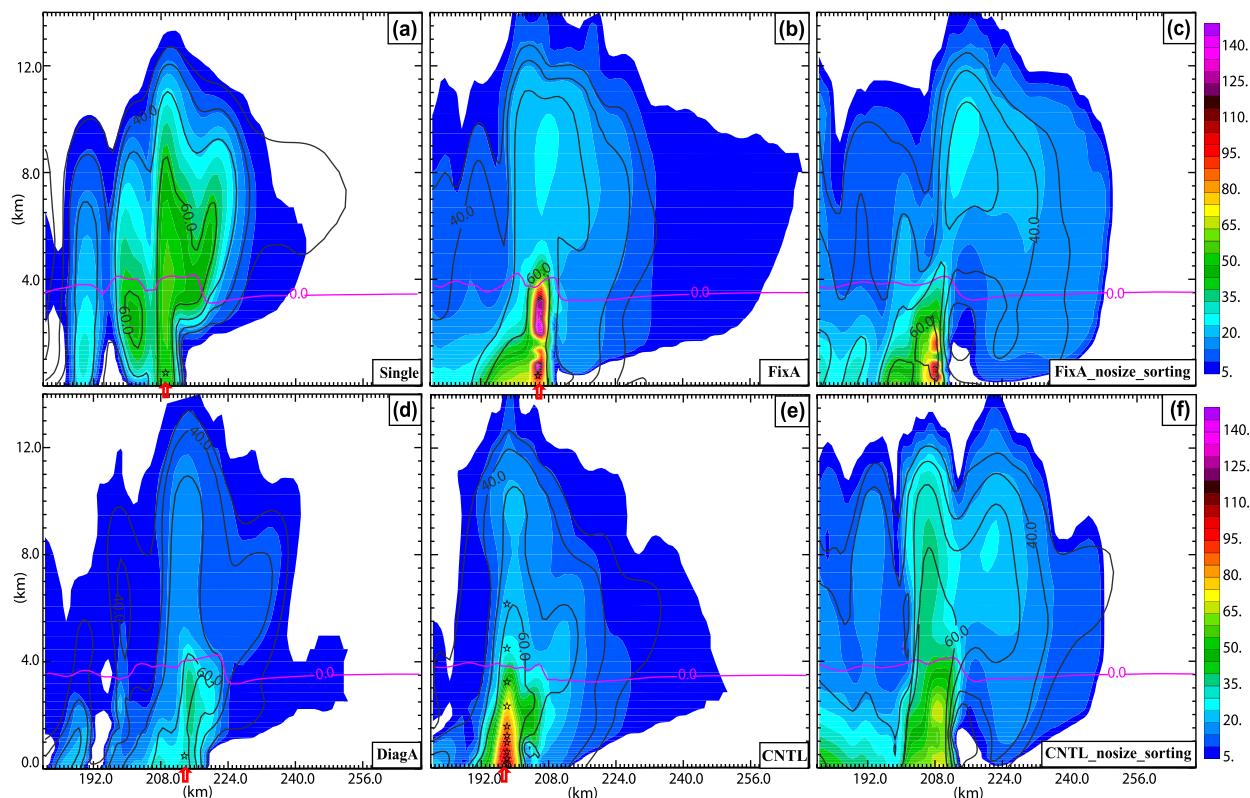


FIG. 14. As in Fig. 13, but for D_{\max} (shaded; mm). The reflectivity contours of 30, 40, 50, and 60 dBZ (black) and temperature contours of 0°C (magenta) are also overlaid. The black hollow stars refer to the approximate locations of the hail size distributions for the experiments in Fig. 15.

0930 UTC, respectively. The term n_{clrh} in FixA is also larger than that in other experiments (Figs. 17 and 18). These are consistent with the higher peak values of Q_h and N_{th} within the hailstorm in FixA as mentioned earlier (Fig. 13b). In other experiments, hail mass and number production rates from the corresponding processes are smaller, especially the hail mass production rates in DiagA and Single (Fig. 17). Generally speaking, the main differences in hail production processes (viz., the dominant microphysical processes responsible for hail mass and number growth— colqch , colqrh , meltqh and n_{clrh} , n_{congh} , n_{melth} , respectively) among the experiments match the differences in their corresponding predictions of hail within storms and hail accumulation at the surface, as discussed earlier.

Time-averaged vertical profiles of the dominant microphysical processes for hail mass and number growth are plotted in Figs. 19 and 20. The profiles are averaged horizontally over points within the hailstorm between 0600 and 1500 UTC using data at 1-min intervals. It is noted that hail collection of rain (colqrh) is still large below the melting level in the MY scheme, although it is more than cancelled out by melting (meltqh). This

appears to be a peculiarity of the MY scheme in that it explicitly allows for collection of rain by hail that is actively melting below the melting level. To analyze the results more physically, terms colqrh and meltqh are combined to yield a new term, net_melting (Fig. 19).

The profiles exhibit generally similar vertical patterns for the experiments, with negative values of net_melting and n_{melth} occurring below ~ 3 km (around the height of melting level), positive colqch , n_{clrh} , and n_{congh} between ~ 3 and ~ 8 km (Figs. 19, 20). Above the melting level, N_{th} increases rapidly because of n_{clrh} and n_{congh} at the magnitude of 10^{14} min^{-1} , causing the increase in λ_h [see Eq. (A5) in MY05a] and decrease in D_{\max} as shown in Fig. 14. Significant differences still exist among the experiments. In Single, both colqch and net_melting terms are much smaller than their counterparts in other experiments (Fig. 19). For experiments using multimoment MP schemes, compared to the peak values of CNTL, FixA has larger peak values of colqch , net_melting , n_{clrh} , and n_{melth} (Figs. 19 and 20). These larger peak values appear to reflect rapid growth of hail within the cloud and more efficient melting of hail below the melting layer in FixA, which seems to be adding to the effect of

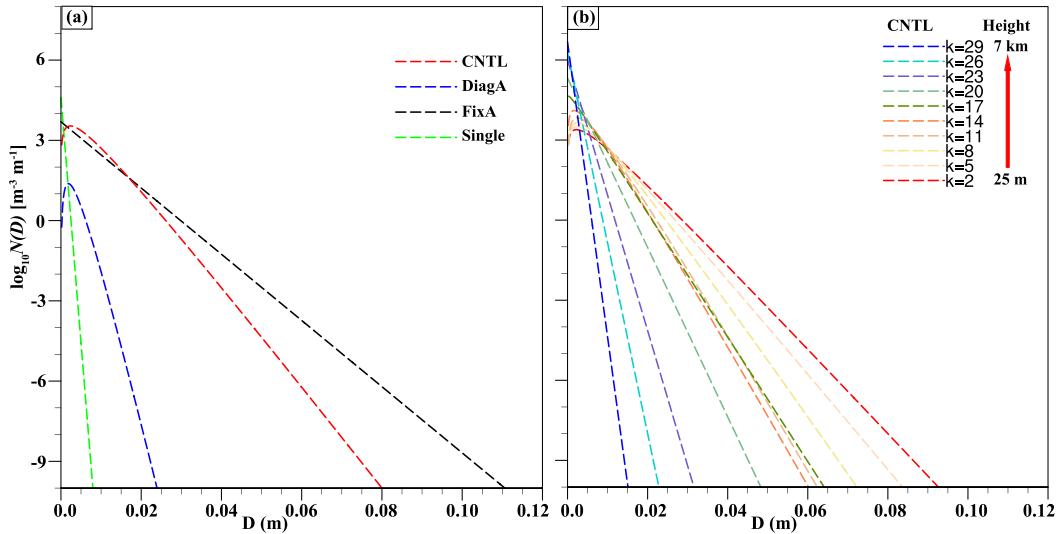


FIG. 15. Hail size distributions at 1100 UTC (a) from Single, FixA, DiagA, and CNTL at $z = 251$ m corresponding to the black star locations indicated in Figs. 14a, 14b, 14d, and 14e and (b) from various levels of CNTL within layers between 25 m and 7 km corresponding to the black star locations in Fig. 14e.

size sorting by increasing the mean diameter of hail that has yet to fully melt. In DiagA, as the hail size spectrum is artificially narrowed (see Fig. 15a), hail with small mean diameter tends to melt more efficiently (as indicated by the higher net_melting rates in Fig. 19) and have lower terminal velocities; these effects combine to result in less hail accumulating at the surface (see Figs. 11c, 12c).

6. Summary and conclusions

This study evaluates the ability of different MP schemes within storm-scale NWP model runs at a 1-km grid spacing to explicitly predict hail in a long-lasting multicellular hailstorm that occurred in Jiangsu Province, China, on 28 April 2015. The hailstorm started within a strong low-level convergence between two low-level cyclones that are underneath an upper-level cutoff low within a deep coastal trough. The environment, associated with a weak warm, moist PBL-capping inversion, featured strong 0–6-km wind shear, moderate CAPE, and a very low CIN. The midlevels were relatively dry. Such an environment is generally conducive to deep convection that tends to produce large hailstones.

The simulations employed one-, two-, and three-moment MY MP schemes (MY05a; MY05b). Two variants of two-moment schemes were used: one in which the shape parameters of hydrometeors were fixed at zero and the other in which the shape parameters were diagnosed as a monotonically increasing function of the mean mass-weighted diameter of hydrometeor particles. Evaluations were performed against

available observations, including radar reflectivity, radar-derived maximum estimated size of hail (MESH), and available severe weather reports. Furthermore, neighborhood-based fractions skill scores (FSSs) were calculated for the simulated MESH fields for objective evaluation.

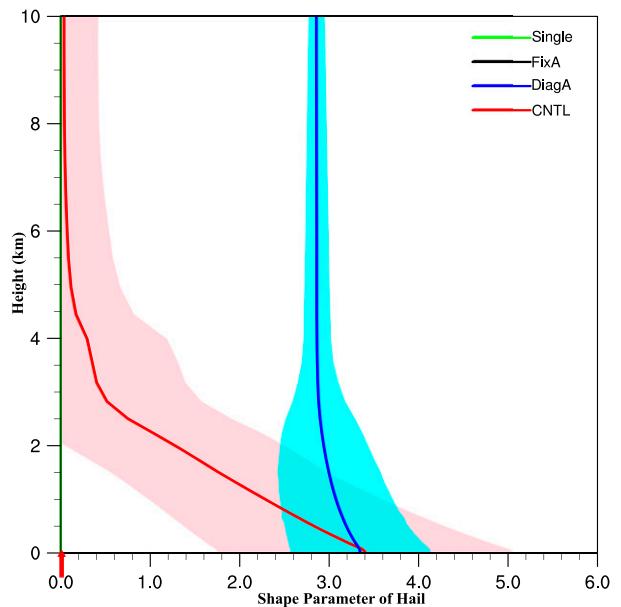


FIG. 16. Profiles of the shape parameters of experiments Single, FixA, DiagA, and CNTL. The profiles are horizontal averages within the simulated hailstorms over the 1-km domain of the simulations between 0600 and 1600 UTC at a 1-min interval. The shaded areas denote plus and minus one standard deviation from the average profiles.

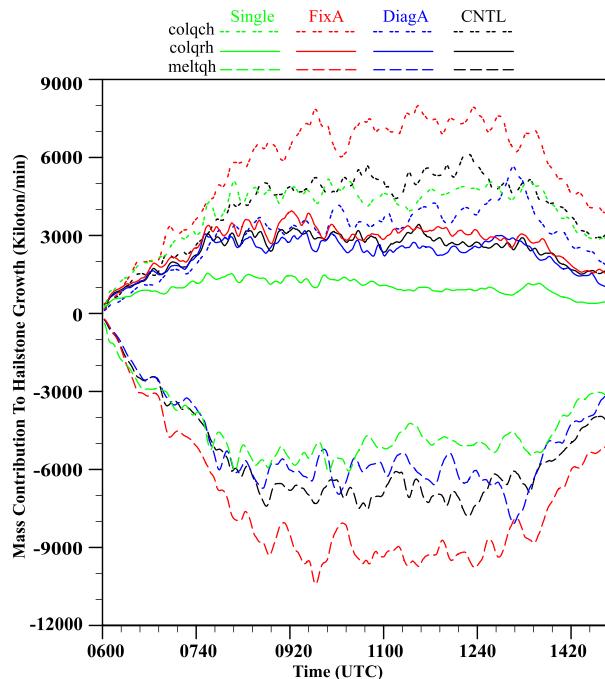


FIG. 17. Time series of microphysical tendency terms in the hail mass equation for experiments Single, FixA, DiagA, and CNTL between 0600 and 1500 UTC. The line color denotes the experiment, while the line pattern denotes microphysical processes responsible for hail growth. The microphysical processes include hail collection of cloud water (colqch), hail collection of rain (colqrh), and hail melting to rain (meltqh). Hail production from other processes, including hail collection of snow or ice, are minimal and are not plotted.

Evaluations against observed radar reflectivity indicate that the time and location of the hailstorm initiation and the later organization of storm cells into a large bow echo are reasonably reproduced by all experiments. Compared with radar observations, experiment FixA, which uses a fixed shape parameter of zero, substantially overpredicts the magnitudes of reflectivity, and as a result produces unrealistically high MESH values (with maxima exceeding 70 mm) compared with radar-derived MESH (which has maxima of only 40–50 mm). In contrast, in Single and DiagA that use a single-moment and diagnostic shape parameter scheme, respectively, reflectivity and MESH fields are underpredicted in both intensity and extent. CNTL using three-moment MY scheme produces MESH swaths that agree more closely with radar-derived MESH swaths than other experiments, and neighborhood-based MESH evaluations further show that the three-moment scheme has notably higher fractional skill scores at all spatial scales compared to the other schemes, especially for large hailstones.

Surface accumulated hail mass, number, and hail distribution characteristics within storms are intercompared

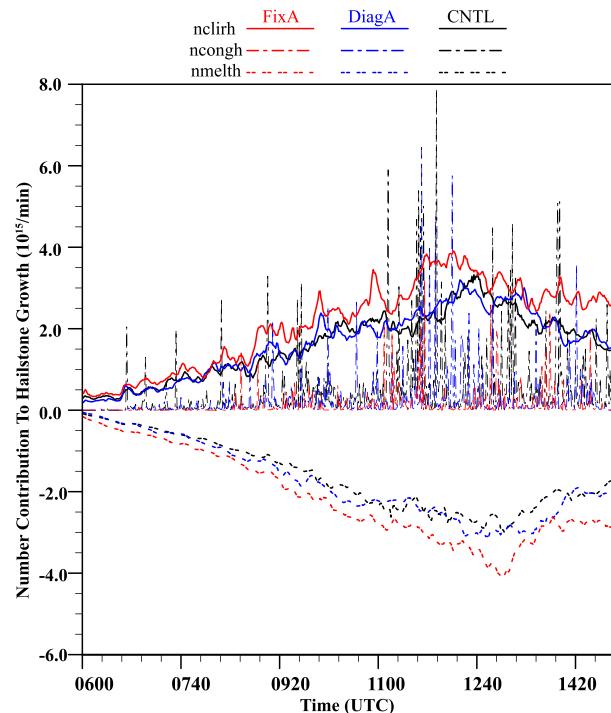


FIG. 18. As in Fig. 17, but for the main source and sink terms in the hail number concentration equation for experiments FixA, DiagA, and CNTL. The line color denotes the experiment, while the line pattern denotes microphysical processes responsible for hail production. The microphysical processes include ice and rain collision to produce hail (nclirh), graupel conversion to hail (ncongh), and hail melting to rain (nmelth). Other processes are minimal and are not plotted.

among the experiments. Results suggest that FixA produces significant amounts of large hail accumulated over a much wider swath than CNTL. For Single and DiagA, the peak SAHNC values are about two orders of magnitude smaller than those of CNTL, especially in DiagA, where almost no hail larger than 40 mm is produced. Examinations of hail distributions within storms indicate that since all the moments of a given hydrometeor type are monotonically related to the mixing ratio (which is the only moment predicted) in Single, no size sorting can occur. For multimoment schemes, different moment-weighted terminal velocities allow for size sorting of particles, making it possible to reproduce more realistic PSDs within the storm. However, substantial differences in the hail size distributions are still present within storms simulated using different multimoment schemes. For example, FixA, which uses a two-moment scheme with a fixed α_x value of zero, suffers from excessive size sorting, which leads to an unrealistic shift in hail DSD toward larger hailstones during sedimentation. On the other hand, the diagnostic α_x used in DiagA is at least 2.8, resulting in a hail size spectrum that appears to be artificially narrowed compared to

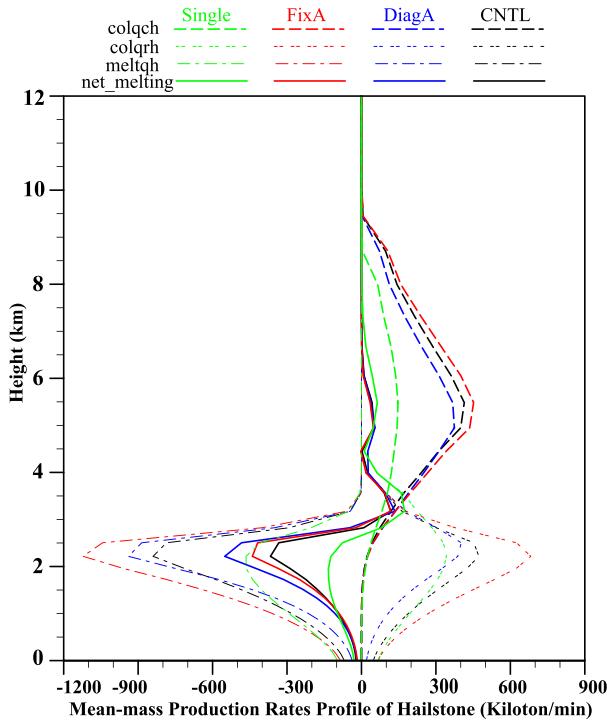


FIG. 19. Vertical profiles of the mean hail production rates for hail collection of cloud (colqch), hail collection of rain (colqrh), hail melting to rain (meltqh), and net melting of hail (net_melting) for experiments Single, FixA, DiagA, and CNTL. The profiles are averaged horizontally over points within the hailstorm at 1-min intervals between 0600 and 1500 UTC 28 Apr 2015.

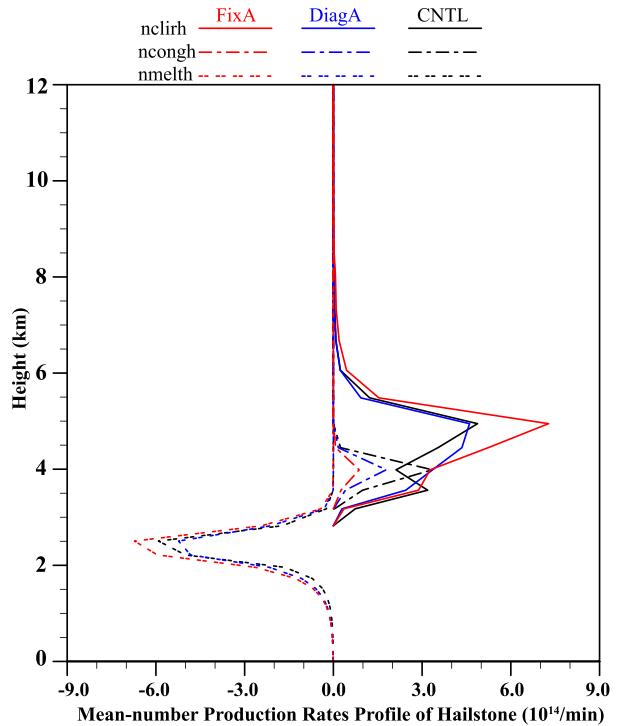


FIG. 20. As in Fig. 19, but for vertical profiles of the mean hail number production rates for ice and rain collision to produce hail (nclirh), graupel conversion to hail (ncongh), and hail melting to rain (nmelth) for experiments FixA, DiagA, and CNTL.

CNTL; it causes a spectrum shift toward smaller hailstones and yields smaller D_{max} and SAHNC. These results indicate that, although excessive size sorting is more limited in DiagA, the specific α_x diagnostic relation derived from the sedimentation-only one-dimensional model appears inaccurate. Therefore, more accurate diagnostic relations for α_x may need to be derived, using perhaps output from full three-moment simulations. In fact, our preliminary results using this approach are encouraging, and more complete results will be reported in a separate paper.

Furthermore, budget analyses of Q_h and N_{th} source and sink terms suggest that, above the melting level, colqch and nclirh are dominantly responsible for hail mass and number growth, respectively. Below the melting level, meltqh and nmelth dominate. However, it is noted that in the MY scheme, large colqrh below the melting level appears to be unphysical, since it allows hail to collect rain efficiently even when hail is actively melting. The differences in hail growth processes among different experiments are closely linked to the treatment of the shape parameter in different MP schemes, which further leads to the differences in the predicted surface

accumulated hail mass, SAHNC, and hail distribution within the simulated storms.

In the end, we note that there are many other possible configurations for one- or two-moment MP schemes in terms of the choice of fixed or variable intercept and shape parameters, which can be further evaluated in the future. We also note that because of the lack of reliable observations of surface accumulated hailstones, evaluations of explicit hail prediction in this paper carry a certain degree of uncertainty. In situ observations of microphysical processes, as well as hail size distributions, are needed for more reliable evaluations.

Acknowledgments. This work was primarily supported by the National 973 Fundamental Research Program of China (2013CB430103). Liping Luo was supported by Nanjing University for her extended visit at CAPS, University of Oklahoma. The work was also supported by the National Science Foundation of China (Grant 41405100), Foundation of China Meteorological Administration special (Grant GYHY201506006). We gratefully acknowledge the High Performance Center (HPCC) of Nanjing University for doing the numerical calculations in this paper on its IBM Blade cluster system. The second author acknowledges the support of

the National Thousand Person Plan hosted at Nanjing University, and the support of NSF Grant AGS-1261776 for the Severe Hail, Analysis, Representation, and Prediction (SHARP) project. Nathan Snook and Jonathan Labriola are thanked for proofreading and improving the manuscript. Suggestions and comments from three anonymous reviewers also improved our paper. The NCEP FNL data can be downloaded freely online (<http://rda.ucar.edu/datasets/ds083.2/>). The radar dataset is provided by the Climate Data Center at National Meteorological Information Center of the China Meteorological Administration, and the processed data used in this paper are available online (<https://pan.baidu.com/s/1kV0WrTP>; password: 293v).

REFERENCES

- Adams-Selin, R. D., and C. L. Ziegler, 2016: Forecasting hail using a one-dimensional hail growth model within WRF. *Mon. Wea. Rev.*, **144**, 4919–4939, <https://doi.org/10.1175/MWR-D-16-0027.1>.
- Baba, Y., and K. Takahashi, 2014: Dependency of stratiform precipitation on a two-moment cloud microphysical scheme in mid-latitude squall line. *Atmos. Res.*, **138**, 394–413, <https://doi.org/10.1016/j.atmosres.2013.12.009>.
- Brimelow, J. C., and G. W. Reuter, 2009: Explicit forecasts of hail occurrence and expected hail size using the GEM-HAILCAST system with a rainfall filter. *Wea. Forecasting*, **24**, 935–945, <https://doi.org/10.1175/2009WAF2222138.1>.
- , —, and E. R. Poolman, 2002: Modeling maximum hail size in Alberta thunderstorms. *Wea. Forecasting*, **17**, 1048–1062, [https://doi.org/10.1175/1520-0434\(2002\)017<1048:MMHSIA>2.0.CO;2](https://doi.org/10.1175/1520-0434(2002)017<1048:MMHSIA>2.0.CO;2).
- Bryan, G. H., and H. Morrison, 2012: Sensitivity of a simulated squall line to horizontal resolution and parameterization of microphysics. *Mon. Wea. Rev.*, **140**, 202–225, <https://doi.org/10.1175/MWR-D-11-00046.1>.
- Casati, D. B., and Coauthors, 2008: Forecast verification: Current status and future directions. *Meteor. Appl.*, **15**, 3–18, <https://doi.org/10.1002/met.52>.
- Changnon, S. A., 1999: Data and approaches for determining hail risk in the contiguous United States. *J. Appl. Meteor.*, **38**, 1730–1739, [https://doi.org/10.1175/1520-0450\(1999\)038<1730:DAAFDH>2.0.CO;2](https://doi.org/10.1175/1520-0450(1999)038<1730:DAAFDH>2.0.CO;2).
- China Meteorological Administration, 2013: *Yearbook of Meteorological Disasters in China* (in Chinese). China Meteorological Press, 222 pp.
- , 2014: *Yearbook of Meteorological Disasters in China* (in Chinese). China Meteorological Press, 238 pp.
- , 2015: *Yearbook of Meteorological Disasters in China* (in Chinese). China Meteorological Press, 240 pp.
- Cintineo, J. L., T. M. Smith, V. Lakshmanan, H. E. Brooks, and K. L. Ortega, 2012: An objective high-resolution hail climatology of the contiguous United States. *Wea. Forecasting*, **27**, 1235–1248, <https://doi.org/10.1175/WAF-D-11-00151.1>.
- Costa, S., P. Mezzasalma, V. Levizzani, P. P. Alberoni, and S. Nanni, 2001: Deep convection over northern Italy: Synoptic and thermodynamic analysis. *Atmos. Res.*, **56**, 73–88, [https://doi.org/10.1016/S0169-8095\(00\)00091-0](https://doi.org/10.1016/S0169-8095(00)00091-0).
- Dawson, D. T., II, M. Xue, J. A. Milbrandt, and M. Yau, 2010: Comparison of evaporation and cold pool development between single-moment and multimoment bulk microphysics schemes in idealized simulations of tornadic thunderstorms. *Mon. Wea. Rev.*, **138**, 1152–1171, <https://doi.org/10.1175/2009MWR2956.1>.
- , E. R. Mansell, Y. Jung, L. J. Wicker, M. R. Kumjian, and M. Xue, 2014: Low-level Z_{DR} signatures in supercell forward flanks: The role of size sorting and melting of hail. *J. Atmos. Sci.*, **71**, 276–299, <https://doi.org/10.1175/JAS-D-13-0118.1>.
- Ebert, E. E., 2009: Neighborhood verification: A strategy for rewarding close forecasts. *Wea. Forecasting*, **24**, 1498–1510, <https://doi.org/10.1175/2009WAF2222251.1>.
- Ferrier, B. S., 1994: A double-moment multiple-phase four-class bulk ice scheme. Part I: Description. *J. Atmos. Sci.*, **51**, 249–280, [https://doi.org/10.1175/1520-0469\(1994\)051<0249:ADMMPF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1994)051<0249:ADMMPF>2.0.CO;2).
- Foote, G. B., and P. S. D. Toit, 1969: Terminal velocity of raindrops aloft. *J. Appl. Meteor.*, **8**, 249–253, [https://doi.org/10.1175/1520-0450\(1969\)008<0249:TVORA>2.0.CO;2](https://doi.org/10.1175/1520-0450(1969)008<0249:TVORA>2.0.CO;2).
- Gagne, D. J., II, A. McGovern, and M. Xue, 2014: Machine learning enhancement of storm-scale ensemble precipitation forecasts. *Wea. Forecasting*, **29**, 1024–1043, <https://doi.org/10.1175/WAF-D-13-00108.1>.
- Gilmore, M. S., J. M. Straka, and E. N. Rasmussen, 2004: Precipitation uncertainty due to variations in precipitation particle parameters within a simple microphysics scheme. *Mon. Wea. Rev.*, **132**, 2610–2627, <https://doi.org/10.1175/MWR2810.1>.
- Guo, X., and M. Huang, 2002: Hail formation and growth in a 3D cloud model with hail-bin microphysics. *Atmos. Res.*, **63**, 59–99, [https://doi.org/10.1016/S0169-8095\(02\)00019-4](https://doi.org/10.1016/S0169-8095(02)00019-4).
- Hu, K., R. Lu, and D. Wang, 2010: Seasonal climatology of cut-off lows and associated precipitation patterns over northeast China. *Meteor. Atmos. Phys.*, **106**, 37–48, <https://doi.org/10.1007/s00703-009-0049-0>.
- Jung, Y., M. Xue, and G. Zhang, 2010: Simulations of polarimetric radar signatures of a supercell storm using a two-moment bulk microphysics scheme. *J. Appl. Meteor. Climatol.*, **49**, 146–163, <https://doi.org/10.1175/2009JAMC2178.1>.
- Kessler, E., 1995: On the continuity and distribution of water substance in atmospheric circulations. *Atmos. Res.*, **38**, 109–145, [https://doi.org/10.1016/0169-8095\(94\)00090-Z](https://doi.org/10.1016/0169-8095(94)00090-Z).
- Labriola, J., N. Snook, Y. Jung, B. Putnam, and M. Xue, 2017: Ensemble hail prediction for the storms of 10 May 2010 in south-central Oklahoma using single- and double-moment microphysical schemes. *Mon. Wea. Rev.*, **145**, 4911–4936, <https://doi.org/10.1175/MWR-D-17-0039.1>.
- Lin, Y. L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Appl. Meteor.*, **22**, 1065–1092, [https://doi.org/10.1175/1520-0450\(1983\)022<1065:BPOTSF>2.0.CO;2](https://doi.org/10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2).
- Loftus, A. M., and W. R. Cotton, 2014: A triple-moment hail bulk microphysics scheme. Part II: Verification and comparison with two-moment bulk microphysics. *Atmos. Res.*, **150**, 97–128, <https://doi.org/10.1016/j.atmosres.2014.07.016>.
- , —, and G. G. Carrió, 2014: A triple-moment hail bulk microphysics scheme. Part I: Description and initial evaluation. *Atmos. Res.*, **149**, 35–57, <https://doi.org/10.1016/j.atmosres.2014.05.013>.
- Luo, L., M. Xue, K. Zhu, and B. Zhou, 2017: Explicit prediction of hail using multimoment microphysics schemes for a hailstorm of 19 March 2014 in eastern China. *J. Geophys. Res. Atmos.*, **122**, 7560–7581, <https://doi.org/10.1002/2017JD026747>.

- Mansell, E. R., 2010: On sedimentation and advection in multi-moment bulk microphysics. *J. Atmos. Sci.*, **67**, 3084–3094, <https://doi.org/10.1175/2010JAS3341.1>.
- Meyers, M. P., R. L. Walko, J. Y. Harrington, and W. R. Cotton, 1997: New RAMS cloud microphysics parameterization. Part II: The two-moment scheme. *Atmos. Res.*, **45**, 3–39, [https://doi.org/10.1016/S0169-8095\(97\)00018-5](https://doi.org/10.1016/S0169-8095(97)00018-5).
- Milbrandt, J. A., and M. K. Yau, 2005a: A multimoment bulk microphysics parameterization. Part I: Analysis of the role of the spectral shape parameter. *J. Atmos. Sci.*, **62**, 3051–3064, <https://doi.org/10.1175/JAS3534.1>.
- , and —, 2005b: A multimoment bulk microphysics parameterization. Part II: A proposed three-moment closure and scheme description. *J. Atmos. Sci.*, **62**, 3065–3081, <https://doi.org/10.1175/JAS3535.1>.
- , and —, 2006a: A multimoment bulk microphysics parameterization. Part III: Control simulation of a hailstorm. *J. Atmos. Sci.*, **63**, 3114–3136, <https://doi.org/10.1175/JAS3816.1>.
- , and —, 2006b: A multimoment bulk microphysics parameterization. Part IV: Sensitivity experiments. *J. Atmos. Sci.*, **63**, 3137–3159, <https://doi.org/10.1175/JAS3817.1>.
- , and R. McTaggart-Cowan, 2010: Sedimentation-induced errors in bulk microphysics schemes. *J. Atmos. Sci.*, **67**, 3931–3948, <https://doi.org/10.1175/2010JAS3541.1>.
- Moore, J. T., and J. P. Pino, 1990: An interactive method for estimating maximum hailstone size from forecast soundings. *Wea. Forecasting*, **5**, 508–525, [https://doi.org/10.1175/1520-0434\(1990\)005<0508:AIMFEM>2.0.CO;2](https://doi.org/10.1175/1520-0434(1990)005<0508:AIMFEM>2.0.CO;2).
- Morrison, H., and A. Gettelman, 2008: A new two-moment bulk stratiform cloud microphysics scheme in the Community Atmosphere Model, version 3 (CAM3). Part I: Description and numerical tests. *J. Climate*, **21**, 3642–3659, <https://doi.org/10.1175/2008JCLI2105.1>.
- , and J. Milbrandt, 2011: Comparison of two-moment bulk microphysics schemes in idealized supercell thunderstorm simulations. *Mon. Wea. Rev.*, **139**, 1103–1130, <https://doi.org/10.1175/2010MWR3433.1>.
- , J. A. Curry, and V. I. Khvorostyanov, 2005: A new double-moment microphysics parameterization for application in cloud and climate models. Part I: Description. *J. Atmos. Sci.*, **62**, 1665–1677, <https://doi.org/10.1175/JAS3446.1>.
- , 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, <https://doi.org/10.1175/2008MWR2556.1>.
- Nieto, R., and Coauthors, 2005: Climatological features of cutoff low systems in the Northern Hemisphere. *J. Climate*, **18**, 3085–3103, <https://doi.org/10.1175/JCLI3386.1>.
- , M. Sprenger, H. Wernli, R. M. Trigo, and L. Gimeno, 2008: Identification and climatology of cut-off lows near the tropopause. *Ann. N. Y. Acad. Sci.*, **1146**, 256–290, <https://doi.org/10.1196/annals.1446.016>.
- Noppel, H., U. Blahak, A. Seifert, and K. D. Beheng, 2010: Simulations of a hailstorm and the impact of CCN using an advanced two-moment cloud microphysical scheme. *Atmos. Res.*, **96**, 286–301, <https://doi.org/10.1016/j.atmosres.2009.09.008>.
- Ortega, K. L., T. M. Smith, K. L. Manross, K. A. Scharfenberg, A. Witt, A. C. Kolodziej, and J. J. Gourley, 2009: The Severe Hazards Analysis and Verification Experiment. *Bull. Amer. Meteor. Soc.*, **90**, 1519–1530, <https://doi.org/10.1175/2009BAMS2815.1>.
- Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, <https://doi.org/10.1175/2007MWR2123.1>.
- Seifert, A., A. Khain, A. Pokrovsky, and K. D. Beheng, 2006: A comparison of spectral bin and two-moment bulk mixed-phase cloud microphysics. *Atmos. Res.*, **80**, 46–66, <https://doi.org/10.1016/j.atmosres.2005.06.009>.
- Snook, N., Y. Jung, J. Brotzge, B. Putnam, and M. Xue, 2016: Prediction and ensemble forecast verification of hail in the supercell storms of 20 May 2013. *Wea. Forecasting*, **31**, 811–825, <https://doi.org/10.1175/WAF-D-15-0152.1>.
- Tao, S. Y., 1980: *Heavy Rain in China* (in Chinese). Science Press, 224 pp.
- Thompson, G., R. M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.*, **132**, 519–542, [https://doi.org/10.1175/1520-0493\(2004\)132<0519:EFOWPU>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<0519:EFOWPU>2.0.CO;2).
- , P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115, <https://doi.org/10.1175/2008MWR2387.1>.
- Van Weverberg, K., A. M. Vogelmann, H. Morrison, and J. A. Milbrandt, 2012: Sensitivity of idealized squall-line simulations to the level of complexity used in two-moment bulk microphysics schemes. *Mon. Wea. Rev.*, **140**, 1883–1907, <https://doi.org/10.1175/MWR-D-11-00120.1>.
- Walko, R. L., W. R. Cotton, M. Meyers, and J. Harrington, 1995: New RAMS cloud microphysics parameterization part I: The single-moment scheme. *Atmos. Res.*, **38**, 29–62, [https://doi.org/10.1016/0169-8095\(94\)00087-T](https://doi.org/10.1016/0169-8095(94)00087-T).
- Weisman, M. L., and J. B. Klemp, 1984: The structure and classification of numerically simulated convective storms in directionally varying wind shears. *Mon. Wea. Rev.*, **112**, 2479–2498, [https://doi.org/10.1175/1520-0493\(1984\)112<2479:TSACON>2.0.CO;2](https://doi.org/10.1175/1520-0493(1984)112<2479:TSACON>2.0.CO;2).
- Witt, A., M. D. Eilts, G. J. Stumpf, J. Johnson, E. D. W. Mitchell, and K. W. Thomas, 1998: An enhanced hail detection algorithm for the WSR-88D. *Wea. Forecasting*, **13**, 286–303, [https://doi.org/10.1175/1520-0434\(1998\)013<0286:AEHDAF>2.0.CO;2](https://doi.org/10.1175/1520-0434(1998)013<0286:AEHDAF>2.0.CO;2).
- Xue, M., K. Droegemeier, V. Wong, A. Shapiro, and K. Brewster, 1995: Advanced Regional Prediction System (ARPS) version 4.0 user's guide. University of Oklahoma Center for Analysis and Prediction of Storms Rep., 380 pp., <http://www.caps.ou.edu/ARPS>.
- , —, and —, 2000: The Advanced Regional Prediction System (ARPS)—A multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part I: Model dynamics and verification. *Meteor. Atmos. Phys.*, **75**, 161–193, <https://doi.org/10.1007/s007030070003>.
- , and Coauthors, 2001: The Advanced Regional Prediction System (ARPS)—A multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteor. Atmos. Phys.*, **76**, 143–165, <https://doi.org/10.1007/s007030170027>.
- , D. Wang, J. Gao, K. Brewster, and K. K. Droegemeier, 2003: The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation. *Meteor. Atmos. Phys.*, **82**, 139–170, <https://doi.org/10.1007/s00703-001-0595-6>.
- Zhang, C., Q. Zhang, and Y. Wang, 2008: Climatology of hail in China: 1961–2005. *J. Appl. Meteor. Climatol.*, **47**, 795–804, <https://doi.org/10.1175/2007JAMC1603.1>.