ADVANCES IN ATMOSPHERIC SCIENCES, VOL. 30, NO. 2, 2013, 485–502

The Predictability of a Squall Line in South China on 23 April 2007

WU Duochang (吴多常) MENG Zhiyong* (孟智勇), and YAN Dachun (严大春)

Laboratory for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing 100871

(Received 14 April 2012; revised 10 July 2012)

ABSTRACT

This study investigated the predictability of a squall line associated with a quasi-stationary front on 23 April 2007 in South China through deterministic and probabilistic forecasts. Our results show that the squall-line simulation was very sensitive to model error from horizontal resolution and uncertainties in physical parameterization schemes. At least a 10-km grid size was necessary to decently capture this squall line. The simulated squall line with a grid size of 4.5 km was most sensitive to long-wave radiation parameterization schemes relative to other physical schemes such as microphysics and planetary boundary layer. For a grid size from 20 to 5 km, a cumulus parameterization scheme degraded the squall-line simulation (relative to turning it off), with a more severe degradation to grid size <10 km than >10 km.

The sensitivity of the squall-line simulation to initial error was investigated through ensemble forecast. The performance of the ensemble simulation of the squall line was very sensitive to the initial error. Approximately 15% of the ensemble members decently captured the evolution of the squall line, 25% failed, and 60% dislocated the squall line. Using different combinations of physical parameterization schemes for different members can improve the probabilistic forecast. The lead time of this case was only a few hours. Error growth was clearly associated with moist convection development. A linear improvement in the performance of the squall line simulation was observed when the initial error was decreased gradually, with the largest contribution from initial moisture field.

Key words: squall line, predictability, South China, ensemble, moisture

Citation: Wu, D. C., Z. Y. Meng, and D. C. Yan, 2013: The predictability of a squall line in South China on 23 April 2007. *Adv. Atmos. Sci.*, **30**(2), 485–502, doi: 10.1007/s00376-012-2076-x.

1. Introduction

Squall line is a line-shaped mesoscale convective system (MCS). MCSs feature abrupt occurrence and fast movement and are usually associated with severe weather such as strong winds, flash floods, large hail, and even tornadoes. These are major disastrous weather events in warm seasons in China. The current forecast capability of MCSs is very limited, probably due to the lack of observations and lack of understanding of their physical mechanisms. The extent to which an MCS can be forecast accurately, that is, the predictability, is an important question with both scientific and practical significance.

The concept of predictability was first proposed by Thompson (1957) to describe the sensitivity of numerical weather prediction (NWP) to initial and model error. The existence of an upper limit of predictability for a weather system was introduced by Lorenz (1963a, b). This upper limit was proposed to be due to atmospheric instability and turbulence (Kraichnan, 1971; Leith, 1971; Robinson, 1971; Leith and Kraichnan, 1972). Later, more and more attention was given to the error growth mechanism and construction of possible ways to prolong forecast lead time. Mu et al. (2002a, b) classified three predictability problems: (1) for a given maximum allowed prediction error, look for the maximum predictable time; (2) for a given prediction time, look for the minimum prediction error; and (3) for a given maximum allowed prediction error and the prediction time, look for the maximum allowable initial error and the parameter error.

Studies on the predictability of mesoscale systems have mainly focused on relatively larger scales (meso-α, 200–2000 km) and heavy rain events in recent years. Kuo et al. (1995) showed that the simulation of a rapid

*Corresponding author: MENG Zhiyong, zymeng@pku.edu.cn

mesoscale cyclonegenesis that occurred in southeastern United States during 28–29 March 1984 was very sensitive to the horizontal resolution and physical parameterization of the numerical model. Some studies showed that small-scale, small-amplitude, initial error may grow rapidly, reach saturation, and develop upscale to affect the predictability of larger-scale features (Tan et al., 2004; Zhang et al., 2006). Rapid error growth has been shown to be closely related to moist processes (Ehrendorfer et al., 1999; Zhang et al., 2002, 2003) for precipitation events over central Texas in the United States (Zhang et al., 2006) or in South China associated with the meiyu front (Liu and Tan, 2009; Zhu et al., 2009; Liu et al., 2011). Mu and Duan (2003) put forward a conditional nonlinear optimum perturbation method that can capture the initial error components that grow the fastest. This method has been successfully used to target observations of heavy rain event in China (Mu et al., 2007).

Relative to the extensive studies on the predictability of meso-α scale weather systems, studies on the predictability of mesoscale systems with smaller scales, for example, meso-β (20–200 km) or meso-γ (2–20 km), have been very rare. Melhauser and Zhang (2012) examined the predictability of a squall line in North America in 2003 and demonstrated nonlinear error-growth features. A squall line is a meso-β convective system defined as "a line of active thunderstorms, either continuous or with breaks, including contiguous precipitation areas resulting from the existence of the thunderstorms" (Glickman, 2000). Squall lines commonly cause severe disasters in the warm season across a large area of China (Meng and Zhang, 2012; Meng et al., 2013). However, there is a paucity of research on the predictability of squall lines. To what extent can a squall line be predicted deterministically and probabilistically? What could be the lead time of squall-line forecast? How sensitive is squall-line simulation to model and initial error? How and why does the initial error grow? In what way can the predictability of squall-line systems be improved? All of these questions remain unanswered, especially for squall lines in China. The aim of this study was to examine the predictability of a squall line associated with a quasi-stationary front in South China on 23 April 2007 (Meng et al., 2012; Wu and Meng, 2013) with regard to all of these questions.

An overview of the squall line case is presented in section 2. Section 3 examines the sensitivity of the squall line simulation to model error from a deterministic forecast point of view. The sensitivity of the squall line simulation to initial error is explored in section 4 from an ensemble forecast point of view. The possible impact of forecast lead time and a method to improve its predictability are discussed, respectively, in sections 5 and 6. Finally, a summary is given in section 7.

2. Case overview

The squall line of interest formed at 2030 UTC 23 April 2007 near the boundary between Guangxi and Guangdong provinces (Fig. 1a). It moved southeastward to Guangdong at a speed of $\sim$17 m s^{-1} and developed into a very strong and long squall line with a length of $\sim$400 km, spanning the entire Guangdong Province. Bow-shaped segments (Meng et al., 2012), which are usually associated with strong surface winds, formed in the squall line during its development stage. The system lasted for nine hours and caused extensive heavy rain, strong winds, and hail in Guangdong Province. Eight-hour accumulated precipitation of 56.6 mm and wind gusts of 23.8 m s^{-1} (Chen et al., 2008) were observed in the Guangzhou area (Fig. 1).

The squall line was initiated in an environment with a midlevel trough, a surface quasi-stationary front, and a nocturnal low-level jet. A low-pressure center at 850 hPa in southwestern China extended to the east and northeast and merged with a mid-latitude shortwave trough during 1200–1800 UTC (Figs. 2a and c). A stationary front extended from western China all the way to Japan with a large meridional gradient of potential temperature (θ) and an even larger gradient of equivalent potential temperature (θ_e), which resembled a meiyu front (Figs. 2b and d). During 1200–1800 UTC, the southerly component of the 850-hPa wind intensified in Guangxi Province and advected high-θ_e air northward. In the meantime, a channel of northerly flow on the cold side of the front started to strengthen and pushed the quasi-stationary front southward. With the increased convergence between the northerly cold and dry air and southerly warm and moist air, intense convection initiated along the intensifying front and rapidly developed into a squall line around 2030 UTC 23 April 2007. The predictability of this squall line was examined in the following sections using the weather research and forecasting model WRFV3 (Skamarock et al., 2008) with the initial and boundary conditions provided by 6-h 1°×1° final analysis (FNL) of the National Centers for Environmental Prediction (NCEP).

3. Sensitivity of the squall line simulation to model error

3.1 *The control experiment*

The control experiment was initialized at 1200 UTC 23 April 2007 and was integrated for 24 h. It

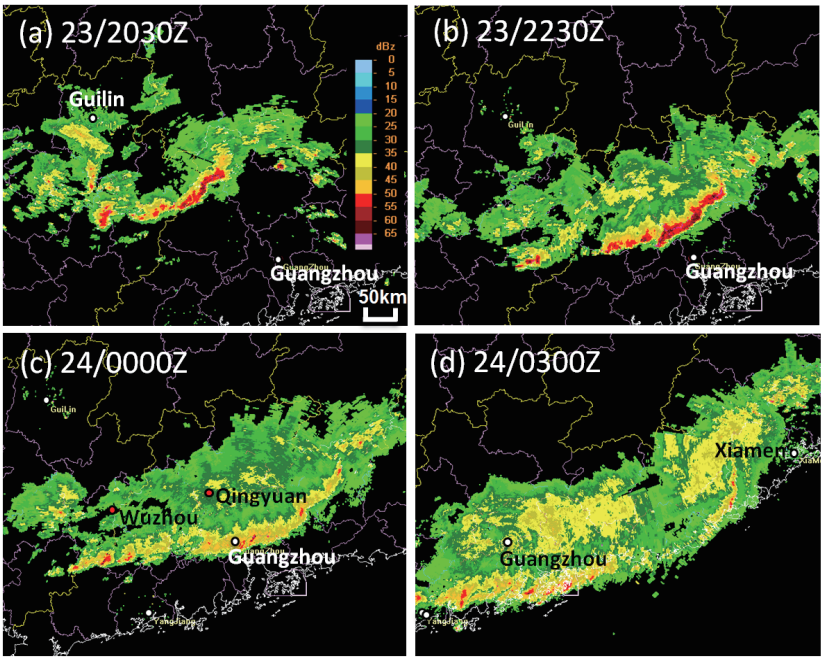

Fig. 1. The observational composite radar reflectivity (the maximum radar reflectivity in vertical column) of the squall line at (a) 2030 UTC 23 April, (b) 2230 UTC 23 April, (c) 0000 UTC 24 April, and (d) 0300 UTC 24 April.

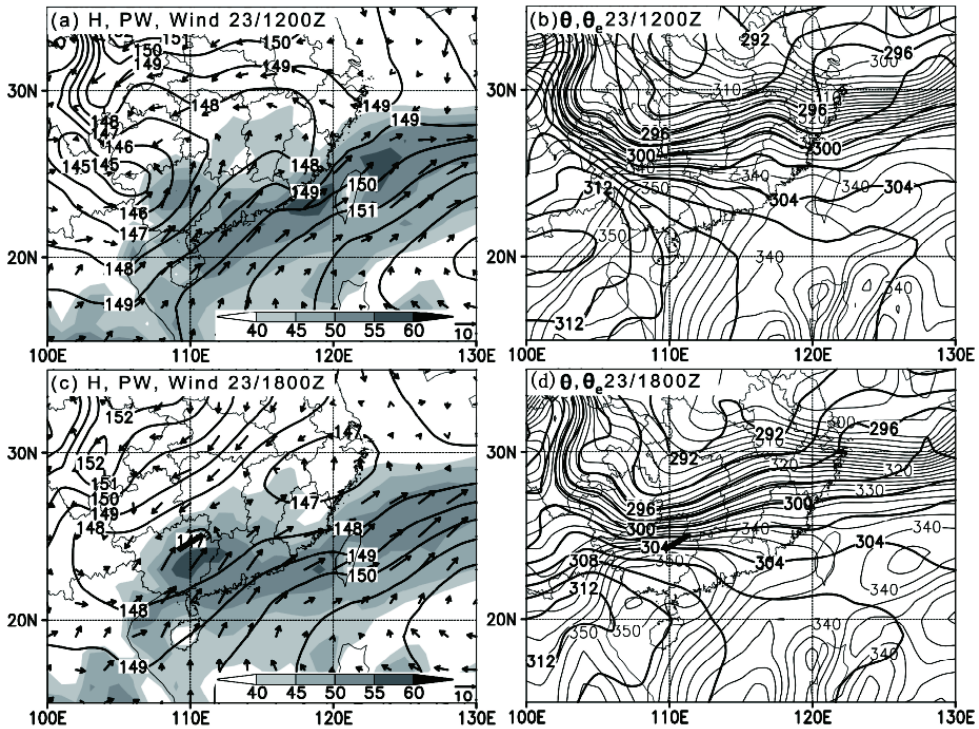

Fig. 2. Environmental features of the squall line based on NCEP FNL analysis. Left panels highlights the evolution of geopotential height (black contour every 20 m), ground-relative wind (m s^{-1}) at 850 hPa, and column-integrated precipitable water (shaded, kg m^{-2}) at (a) 1200 UTC 23 April, (c) 1800 UTC 23 April. (b) and (d) are the same as (a) and (c) but for potential temperature (heavy contour every 2 K) and equivalent potential temperature (thin contour, every 2 K) at 850 hPa. The heavy short lines in (c) and (d) denote the position where the squall line formed 2.5 h later.

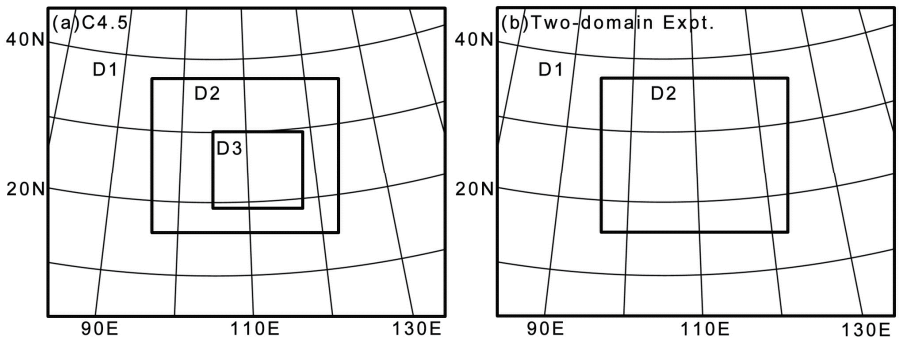

Fig. 3. Model domains for (a) the control experiment C4.5 and (b) the experiments with two domains.

utilized three two-way nested domains with horizontal grid spaces of 40.5, 13.5, and 4.5 km, respectively (Fig. 3a). Domain 1 (D1) used 150 (lon) ×120 (lat) horizontal grid points, which covered almost all of China. The two nested domains used 205 (lon)×172 (lat) horizontal grid points (D2) and 316 (lon) ×265 (lat) horizontal grid points (D3), respectively. Each domain had 35 vertical layers, with a model top of 10 hPa. The single-moment six-class microphysics scheme with graupel WSM6 (Hong et al., 2004), the Yonsei State University (YSU) planetary boundary layer (PBL) scheme (Hong et al., 2006), and the rapid radiative transfer model (RRTM) longwave radiation scheme were used for all the three domains. The Grell-Devenyi cumulus scheme (Grell and Devenyi, 2002) was used for D1 and D2. No cumulus parameterization was used for D3. The control experiment is hereafter referred to as C4.5 (Table 1).

The control experiment well captured the formation, evolution, and structure features of the squall line. The hourly evolution of the leading edge of the simulated convective line was very close to the observation data (Figs. 4a, b). The simulated composite radar reflectivity in D3 at 0000 UTC 24 April 2007 (Fig. 4c) was close to the observation data (Fig. 1c) in terms of intensity, structure, and location except for a slightly larger extent. The bow-shaped feature was also captured successfully. The result of the control experiment demonstrates the capability of the model in simulating this squall line event. Thus we were able to use C4.5 as a benchmark to examine the predictability of the squall line in terms of the simulation performance at 0000 UTC 24 April, when the squall line matured.

3.2 *Impact of physical parameterization schemes*

Uncertainty in physical parameterization schemes is a major source of model error (Orrell, 2003). Physical parameterization aims to account for aggregate effect of sub-grid physical processes that are lacking in current knowledge and/or observation data. Uncertain-

Table 1. Experimental designs for deterministic forecasts. The bold Italic scheme is the difference of the experiment from C4.5.

Expt.	Grid Size (km)	Cumulus schemes For D1 and D2	Microphysics schemes	PBL schemes	Longwave radiation schemes
C4.5	4.5	Grell-Devenyi for D1 & D2	WSM6	YSU	rrtm
KFcum	***4.5***	***Kain-Fritsch*** for D1 & D2	WSM6	YSU	rrtm
THmps	4.5	Grell-Devenyi for D1 & D2	***Thompson***	YSU	rrtm
MRFpbl	4.5	Grell-Devenyi for D1 & D2	WSM6	***MRF***	rrtm
EHSlws	4.5	Grell-Devenyi for D1 & D2	WSM6	YSU	***EHS forcing***
R20_ None	***20***	Grell-Devenyi for D1, ***None*** for D2	WSM6	YSU	rrtm
R20_Grell	***20***	Grell-Devenyi for D1 & D2	WSM6	YSU	rrtm
R13.5_None	***13.5***	Grell-Devenyi for D1, ***None*** for D2	WSM6	YSU	rrtm
R13.5_Grell	***13.5***	Grell-Devenyi for D1 & D2	WSM6	YSU	rrtm
R10_ None	***10***	Grell-Devenyi for D1, ***None*** for D2	WSM6	YSU	rrtm
R10_Grell	***10***	Grell-Devenyi for D1 & D2	WSM6	YSU	rrtm
R5_ None	***5***	Grell-Devenyi for D1, ***None*** for D2	WSM6	YSU	rrtm
R5_ Grell	***5***	Grell-Devenyi for D1 & D2	WSM6	YSU	rrtm

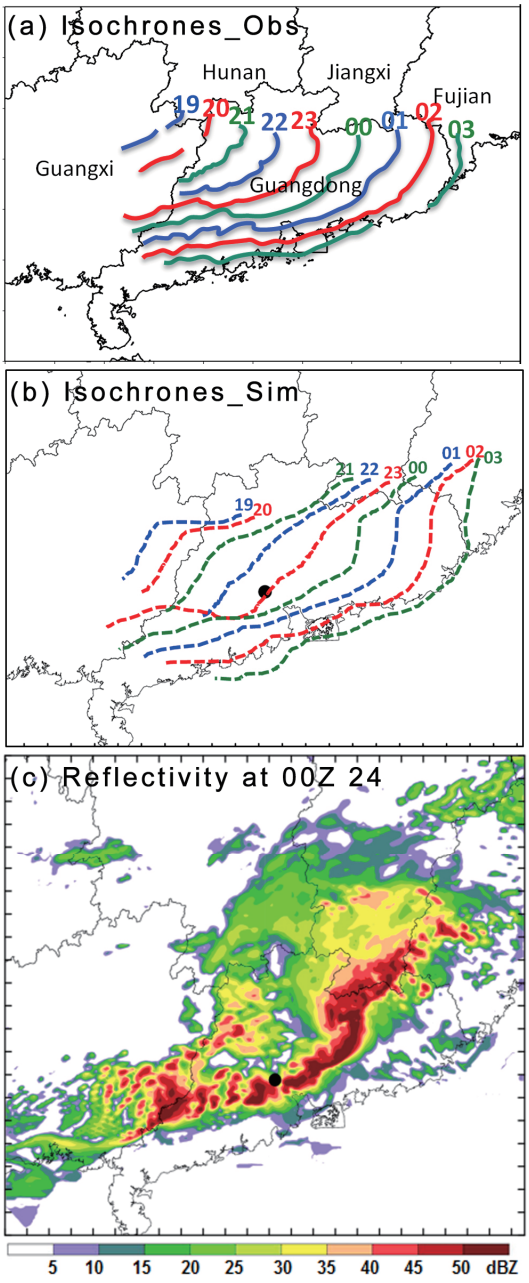

Fig. 4. The isochrones of (a) the observed and (b) the simulated squall line. The numbers corresponding to each line denote the time in hour (UTC) of the day. (c) gives the simulated composite radar reflectivity at 0000 UTC 24 in C4.5.

ties of physical parameterization schemes come from deployed assumptions, empirical formulas, and coefficients. Four experiments, including KFcum, THmps, MRFpbl, and EHSlws, were constructed to explore the sensitivity of the squall line simulation to physical parameterization schemes (Table 1). These experiments were identical to the control experiment C4.5 except that a different cumulus scheme, microphysics scheme, PBL scheme, or long-wave radiation scheme were used, respectively.

Significantly different features were observed in the simulated composite radar reflectivity of D3 at 0000 UTC 24 April among these four experiments. The simulated squall line was not very sensitive to the changes of cumulus parameterization scheme in the coarser domains D1 and D2 (Fig. 5a). When the microphysics scheme was changed from WSM6 to Thompson, the convections became weaker; thus the convective line was narrower, with a few broken points in the middle of the squall line (Fig. 5b). Greater sensitivity was observed when a different PBL scheme was used (Fig. 5c). Three patches of convections lined up with large gaps; thus, the simulation failed to generate a continuous squall line. The greatest sensitivity in the simulation was obtained when a different long-wave radiation scheme was applied (Fig. 5d). The southern part displayed only scattered convective cells, though the northern part was well simulated.

The differences between these simulations and C4.5 were also reflected quantitatively by difference total energy (DTE). The DTE is defined as shown in the work of Zhang et al. (2003):

$$\mathrm{DTE} = \frac{1}{2}\sum\left(U'^2_{ijk} + V'^2_{ijk} + \kappa T'^2_{ijk}\right) ,$$

where U', V', T' denote the wind component and temperature differences between two simulations, $\kappa = c_p/T_r$ (c_p is the heat capacity at constant pressure and T_r is the reference temperature), and i, j, and k denote the index of grid point in x, y, and σ dimensions. The evolution of DTE calculated over D3 between C4.5 and the four sensitivity experiments is plotted in Fig. 6. The differences started to increase significantly when the convections were initiated near 1600 UTC, and they saturated near 0000 UTC 24 April when the squall matured. The difference between experiments with different long-wave radiation schemes (C4.5 and EHSlws) grew most rapidly and was much larger than others, probably because the squall line formed during the night. The smallest sensitivities were found among different cumulus schemes used in coarser domains. This result indicates that the simulation of this squall line had much greater sensitivity to the uncertainties in longwave radiation scheme than those in cumulus, microphysics, and PBL schemes.

3.3 *Impact of model resolution*

Model resolution is an important factor that affects the simulation accuracy of any scale of weather

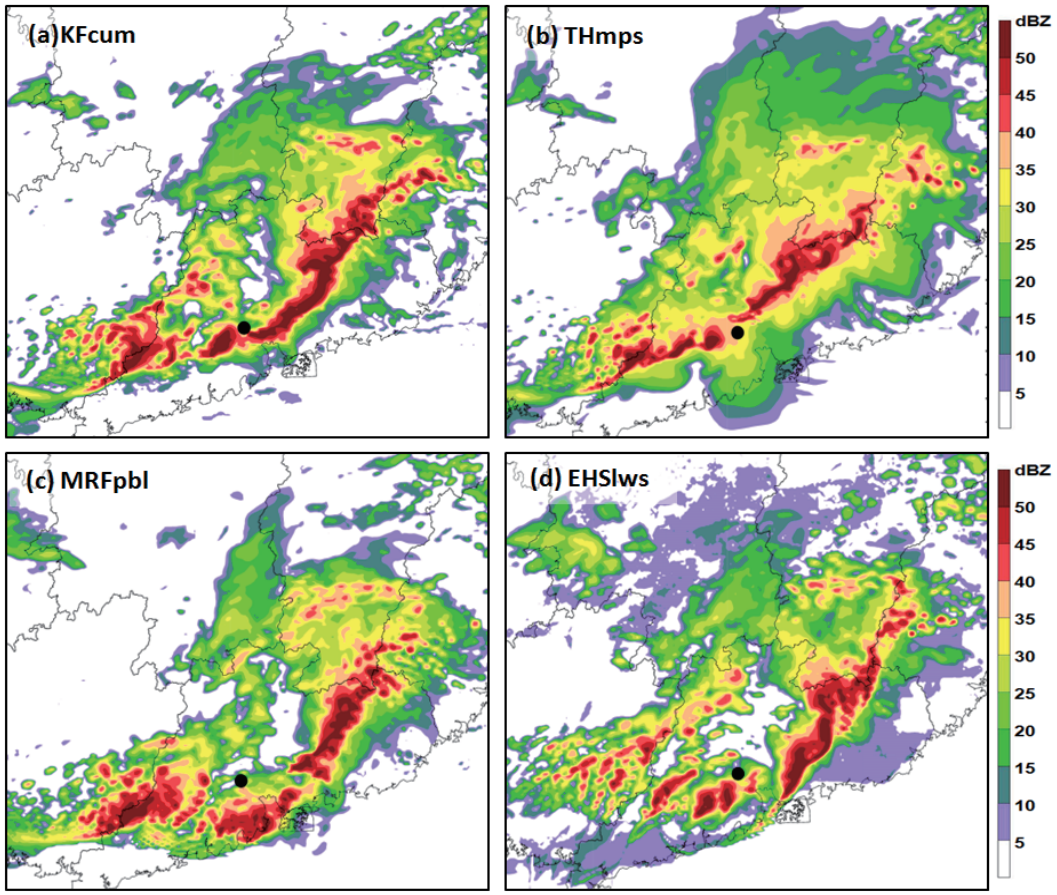

Fig. 5. The simulated composite radar reflectivity at 0000 UTC 24 April of experiments (a) KFcum, (b) THmps, (c) MRFpbl, and (d) EHSlws.

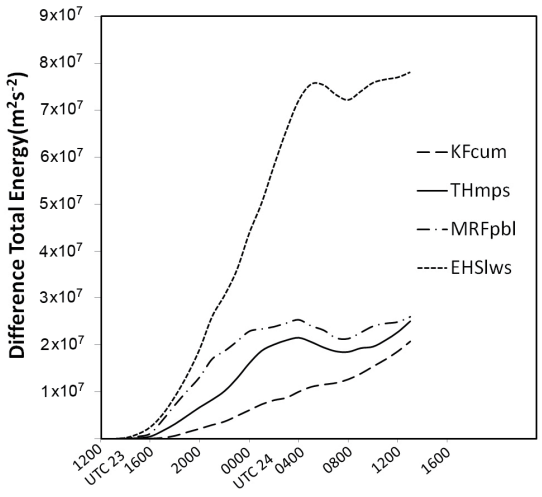

Fig. 6. The evolution of DTE ($m^2\ s^{-2}$) between C4.5 and KFcum (dashed), THmps (solid), MRFpbl (dot-dashed) and EHSlws (dotted).

system. It is especially important for the simulation of MCSs, which involves complicated multi-scale interactions. To examine the sensitivity of this squall line event to model resolution, because there was not much room to change grid size for D3 in the control experiment C4.5, several two-domain experiments were performed with D1 and D2, covering the same area as those in C4.5 (Fig. 3b). D2 used a grid size varying gradually from 20 km to 13.5 km, 10 km, and 5 km. D1 and D2 used the same physical parameterization schemes, except for the cumulus scheme for D2. Considering the uncertainty in the impact of the cumulus parameterization scheme (used or not) for a grid size between 5 and 20 km, two sets of experiments were performed: The first set used the same cumulus parameterization scheme as that in C4.5 for D2, named R20_Grell, R13.5_Grell, R10_Grell, and R5_Grell (Table 1). The second set used no cumulus parameterization scheme for D2, named as R20_None, R13.5_None, R10_None, and R5_None (Table 1).

Result shows that the model simulation was very sensitive to horizontal resolution. Generally speaking, when cumulus parameterization was turned off, a higher resolution produced a better simulated squall

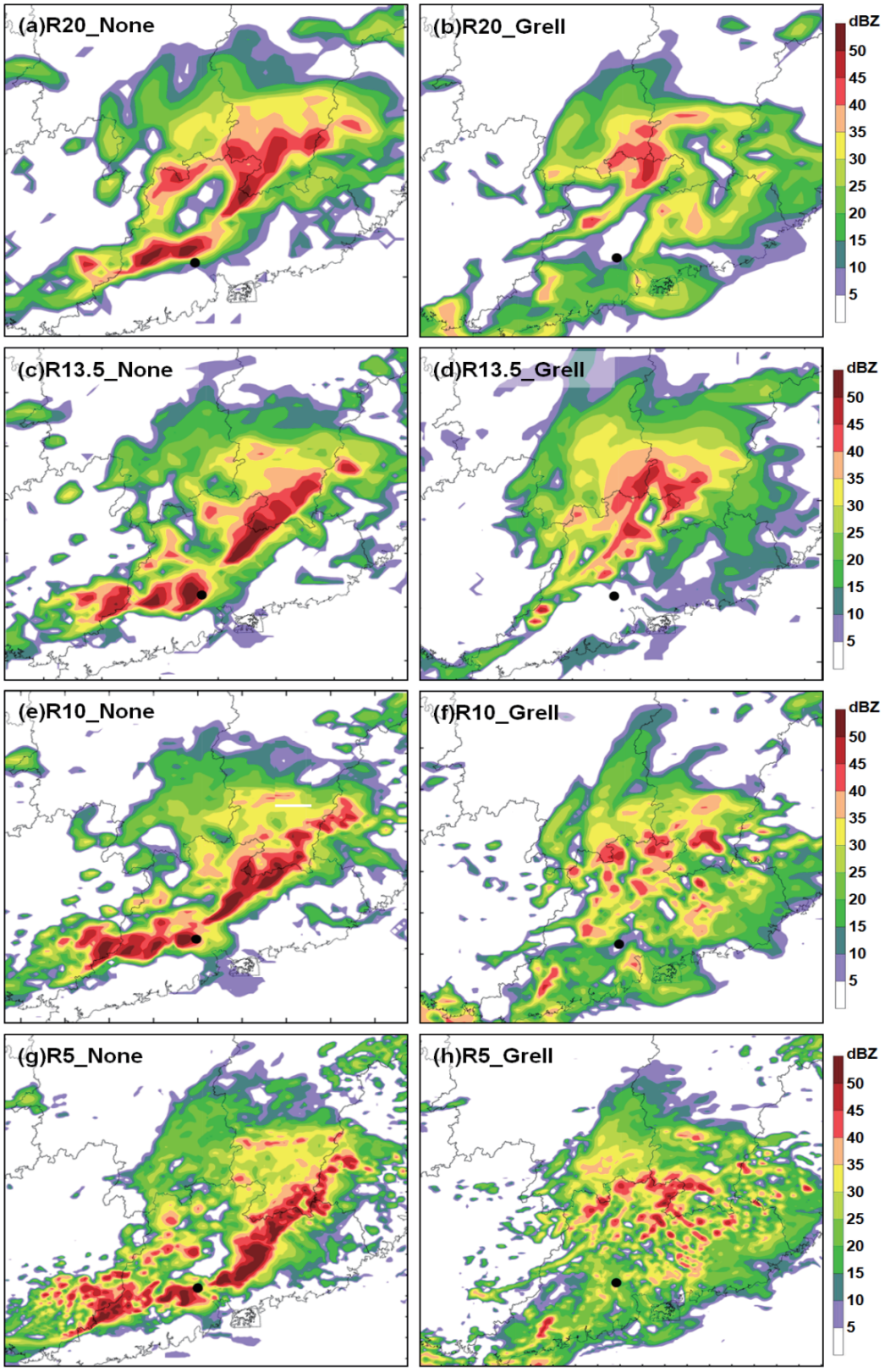

Fig. 7. The simulated composite radar reflectivity at 0000 UTC 24 April for experiments (a) R20_None, (b) R20_Grell, (c) R13.5_None, (d) R13.5_Grell, (e) R10_None, (f) R10_Grell, (g) R5_None, and (h) R5_Grell.

line. In R20_None (Fig. 7a), the simulation only showed a cluster of convections with two line patterns, one in south Jiangxi Province and one in south Guangdong Province. When the grid size decreased to 13.5 km, a line of separated convections were captured at the correct location, but a weak line remained behind it (Fig. 7c). When the grid size decreased to 10 km, a continuous line was produced with features close to those of the observations (Fig. 7e). With a further decrease of grid size to 5 km, the simulation was almost the same as that of C4.5, though the latter had a transitioning domain 2 with a grid size of 13.5 km (Fig. 7g). It also captured the bow-shaped structure.

The results were quite different when the cumulus parameterization was turned on. All of the simulations with a grid size from 20 to 5 km were much worse than those without the cumulus parameterization (Figs. 7b, d, f, and h). Experiments with a grid size >10 km captured a short line, but it was far north of the observations (Figs. 7b and d), while those <10 km only produced scattered cells in a more extensive stratiform area (Figs. 7f and h). Cumulus parameterization had a larger negative impact at a grid size <10 km than >10 km. Similar results were obtained when a different cumulus scheme was used (not shown).

These results indicate that the impact of grid size on the simulated squall line is associated with other aspects of model configuration. Whether a cumulus parameterization scheme is used or not for a grid size between 5 and 20 km may have a great effect on the simulation results. Therefore, a resolution of at least 10 km is necessary to successfully capture the squall line without using a cumulus scheme. Zhang et al. (2002) proposed that a better simulation resulting from using a higher resolution is due to a better representation of the moist physics. If latent-heat release is turned off, no apparent improvement is obtained with the increase of the horizontal resolution. Our results suggest that the better representation of the moist physics using a higher resolution is more likely associated with the explicit rather than implicit precipitating processes.

4. Sensitivity of squall line simulation to initial error

In addition to the model error from horizontal resolution and physical parameterization schemes as discussed in previous sections, initial error is another major source of forecast error. Considering the great sensitivity of squall-line simulation to grid size and physical parameterization and other potential errors, probabilistic (ensemble) forecasts may represent a better way to forecast MCSs. To more thoroughly investigate the sensitivity of squall-line simulation to the initial error, we used ensemble forecasting by randomly perturbing the initial field of the control experiment with the perturbations of zero mean and a standard deviation that represents the analysis error of the initial fields. The predictability of the squall line was examined by determining the percentage of members with good simulations of the squall line in the whole ensemble and their sensitivity to different NWP aspects.

4.1 *The control ensemble experiment*

The control ensemble experiment (S12Z) used an ensemble of 40 members initiated by randomly perturbing the FNL analysis at 1200 UTC 23 April 2007, which was the initial field of the control deterministic experiment C4.5. The perturbations were generated by randomly sampling the background error covariance of WRF-3DVAR (Skamarock et al., 2008). The standard deviations of the initial ensemble were roughly 1 K for temperature, 2 m s^{-1} for winds, and 0.5 g kg^{-1} for the water vapor mixing ratio. The boundary fields were perturbed in the same way. The model domain and physical parameterization setup for each member was the same as that in C4.5.

The result shows that the simulated squall line was very sensitive to the initial error. Large ensemble spread was observed in the simulated composite radar reflectivity at 0000 UTC 24 (Fig. 8). To determine the predictability of the squall line, the forecast ensemble at 0000 UTC 24 were divided into three types. Type 1 included those members that captured the squall line close to the observations (e.g., member 10 or M10). The criteria used to determine whether a squall line was captured using the methods of Parker and Johnson (2000) and Meng et al. (2012). Namely, if a 100-km-long 40-dBZ composite radar reflectivity band lasted no less than three h in a simulation, the simulation was considered to have captured a squall line. If the center of the 40 dBZ band of the simulated squall line was located within 200 km of the observed squall line at 0000 UTC 24, it was considered as "close to observation". Type 2 included those members that captured a squall line but with a location error >200 km (e.g., M09). Type 3 included those members in which no squall line formed (e.g., M25). Type 2 and Type 3 used the same criteria as those used for Type 1 to determine whether a squall line was captured. As shown in Table 2, 15% of the members captured a squall line close to the observations, while 60% of the members captured the squall lines, but with a large location error. The members that failed to capture the squall line accounted for 25% of the members.

The mean formation time, duration, and maximum

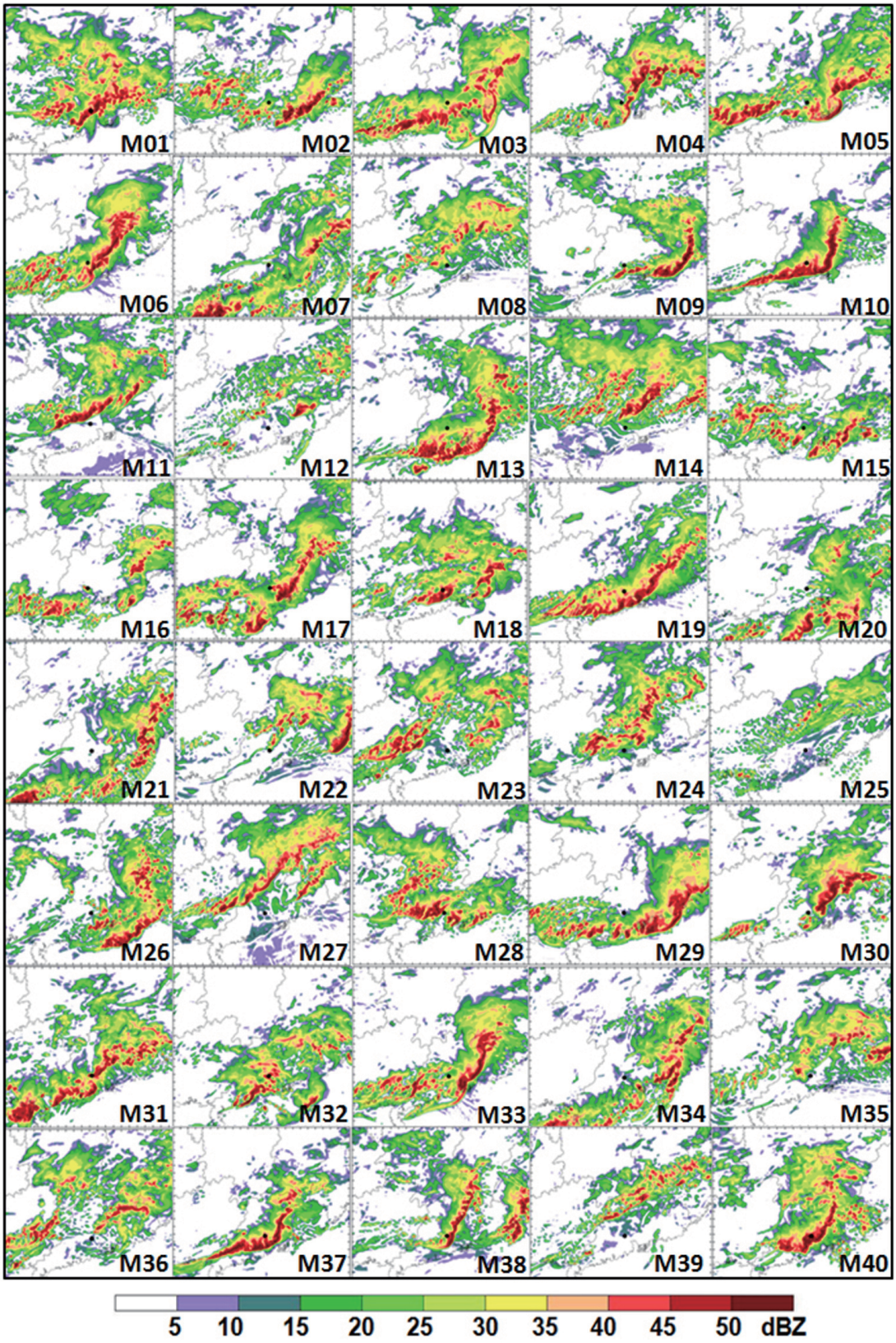

Fig. 8. The simulated composite radar reflectivity at 0000 UTC 24 April of the 40 ensemble members in experiment S12Z.

Table 2. The percentage of the three types of ensemble members in different ensemble experiments.

	Single scheme		Multi-scheme
	S12Z	S00Z	M12Z
Squall lines close to Obs	15%	0	15%
Squall lines with bad location	60%	27.5%	70%
No squall line	25%	72.5%	15%

length of the simulated squall lines in Type 1 and 2, individually and totally (denoted "ave"), were compared to observations (Fig. 9). The formation time was defined as the moment when the 40-dBZ band of the squall line first reached 100 km long. The duration was the period that the 40-dBZ band of the squall line remained longer than 100 km. The maximum length denoted the maximum length of the 40-dBZ band of the squall line during its duration (Meng et al., 2012). Our results show that the simulated squall lines formed generally earlier than the observations, especially in Type 2 which was 2-h ahead of the observations. The averaged duration of the simulated squall lines was 0.5 h shorter than the observations. Type 2 had an even shorter duration than Type 1. The large location error in Type 2 was likely because the squall lines formed and weakened earlier relative to the observations. The maximum lengths of the squall lines of both Type 1 and 2 were larger than the observations. The larger average of Type 1 was due to an outlier in which the simulated squall line had a maximum length of 650 km with a tremendous bow shape.

The growth of initial error can be quantitatively shown in terms of the evolution of the ensemble spread. Figure 10 shows the horizontal distribution of a vertically averaged ensemble spread of temperature at different times from 1400 UTC 23 April to 0400 UTC 24 April. In the early stage, the ensemble spread of ∼1 K was confined to northern Guangxi (Fig. 10a). The ensemble spread increased in both magnitude and extent with the formation and development of the squall line; it also moved with the squall line (Figs. 10b–e). A maximum value of 2.5 K was reached at 0000 UTC 24 April (Fig. 10e). The ensemble spread decreased in both magnitude and scale after 0300 UTC 24 April, when the squall line began to weaken (Fig. 10f). This result indicates that moist convection is conducive to the growth of the ensemble spread.

The association between error growth and moist convection development can also be seen from the evolution of the vertical distribution of horizontally averaged ensemble spread of different variables at 3-h intervals from 1200 UTC 23 to 1200 UTC 24 April (Fig. 11). Similar to the evolution of horizontal distribution of the ensemble spread of T (Fig. 10), the ensemble spread became larger with time until 0000 UTC 24 (Fig. 11, black line) and fell back persistently thereafter for latitudinal wind component u, temperature T, and water vapor mixing ratio q_v. This feature can be clearly seen from the evolution of the domain averaged ensemble spread as shown by the dots and circles marked on the X-axis in the same color convection. During the developing stage from 1200 UTC 23 to 0000 UTC 24, a peak of ensemble spread appeared at ∼2 km above the ground (Figs. 11a–c, solid lines), which was likely driven by moist processes. Larger peaks were also seen near the tropopause for u and T. Apparent jumps between 1500 and 1800 UTC 23 April were seen around the peak levels. For example, the ensemble spread of u near 11 km jumped from 3.2 to 4.8 m s^{-1}, while that of T near 12 km jumped from 1 to 2 K, and that of q_v near 2 km jumped from 1 to 1.3 g kg^{-1}. These jumps appeared during the formation process of the squall line; thus, they are likely a result of the differences in the development of the simulated squall line in different members. When the squall line started to weaken after 0000 UTC 24, the ensemble spread decreased gradually except for the temperature ensemble spread lower than ∼2 km. The ensemble spread of T under 2 km still increased after the squall line started to weaken until ∼0900 UTC April 24, then it started to decrease. This feature was likely associated with the development of the cold pool because the cold pool reached its peak sometime after the storm began to weaken, with the downdraft gradually dominating the vertical motion.

4.2 *Impact of initial error*

The high sensitivity of the squall line simulation to the different initial perturbations was clearly shown in the control ensemble forecast. We further investigated the key factor governing the squall-line forecast capability of the numerical model to determine how the difference between the initial fields of good simulation and poor simulation led to significantly different performances. M10 from Type 1 was chosen as a good member that captured a squall line generally close to the observations, while M25 from Type 3 was chosen as a poor member that failed to produce the squall line over the whole forecast period.

The impact of the difference in the initial fields was examined in a way similar to that of Melhauser and

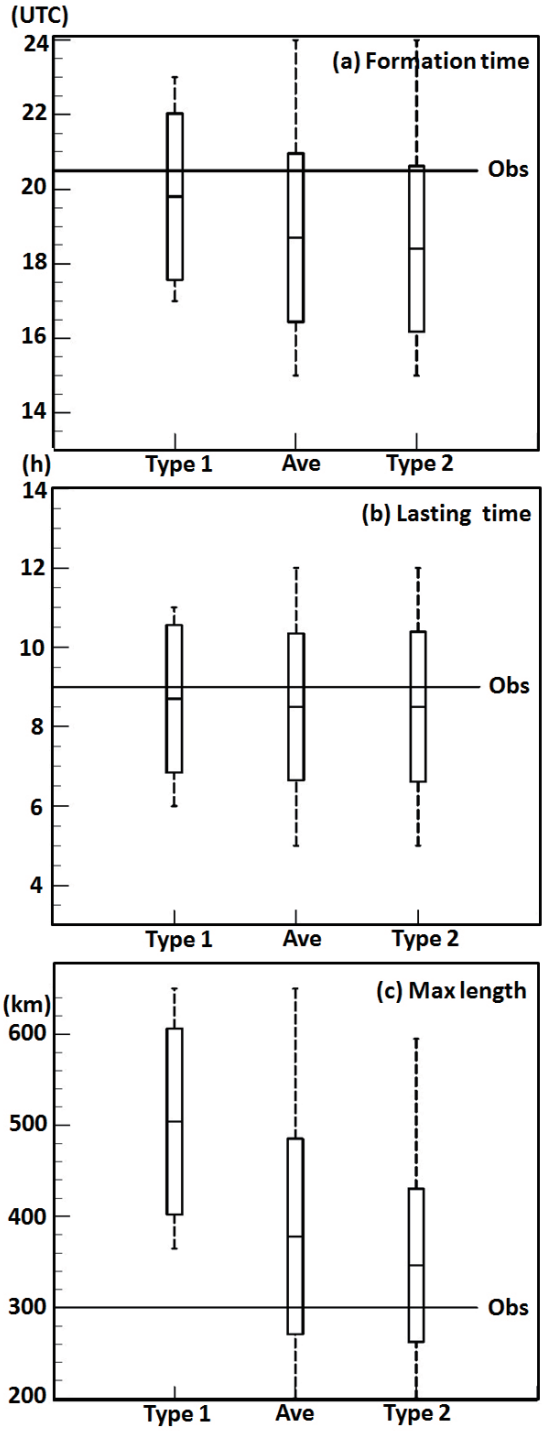

Fig. 9. The box-and-whisker plot of (a) the formation time (UTC 23 April), (b) the duration (hour), and (c) the maximum length (km) of the simulated squall lines of Type 1, Type 2 and the summation of Type 1 and Type 2 (referred to as Ave) in comparison to the observation (heavy line). The short line in the rectangle denotes the mean. The rectangle denotes one standard deviation range. The largest and smallest values are marked above and below the rectangle.

Zhang (2012). The difference between the initial fields of M10 and M25 was first equally split into 10 parts. By adding these 10 parts gradually to the initial field of the poor member (M25) and rerunning the model, the sensitivity of the squall-line simulation to the better and better initial quality was investigated in more detail. Experiment All_1 added one 10% part of all model variables (including wind components u and v, temperature T, water vapor mixing ratio q_v, vertical velocity w, column mass of dry air mu, and perturbation geopotential PH) to the initial field of M25. Experiment All_2 added a second 10% part from the remaining nine 10% parts to the initial field of All_1. In other words, two 10% parts were added to the initial fields of M25. The same procedure was used to produce experiments All_3 to All_9, with the initial field becoming closer and closer to the initial field of the good member M10.

With the gradual improvement of the initial field, the performance of the squall line simulation improved gradually. The simulated composite radar reflectivity at 0000 UTC 24 April (Fig. 12) of experiments All_1 to All_9 changed from no squall line at all gradually to a squall line much closer to that in M10. Each additional increment in the initial condition made an obvious improvement in the squall line simulation. The two convection clusters in northern and southern Guangdong Province simulated by M25 first turned into two line segments individually (All_3). Then the two lines extended toward each other and merged into one line when more 10% parts were added to the initial field (All_5). With further improvement of the initial field, the simulated line became closer and closer to that of M10, with more detailed structures captured, such as the development of the bowed feature (All_7 to All_9). This gradual or linear improvement in the simulation from the linearly improved initial field was somewhat different from that of Melhauser and Zhang (2012). They observed a sudden change in the simulation of a squall-line case in North America when they gradually added those equally split initial differences. This result indicates that the sensitivity of squall line simulation to initial error could be linear or nonlinear and may change from case to case.

A further step was then taken to investigate the contribution of different variables to this linear improvement in the squall-line simulation. UV_1–9, T_1–9, Q_v_1–9 were performed by individually adding the differences of wind, T, and q_v part by part to the initial field of M25. The result shows that adding only the differences of wind (Figs. 13a–c) or T (Figs. 13d–f) led to much less improvement to the simulation relative to adding in only q_v (Fig. 14). To confirm the important contribution of moisture field, NoQ_v_1–9 was per-

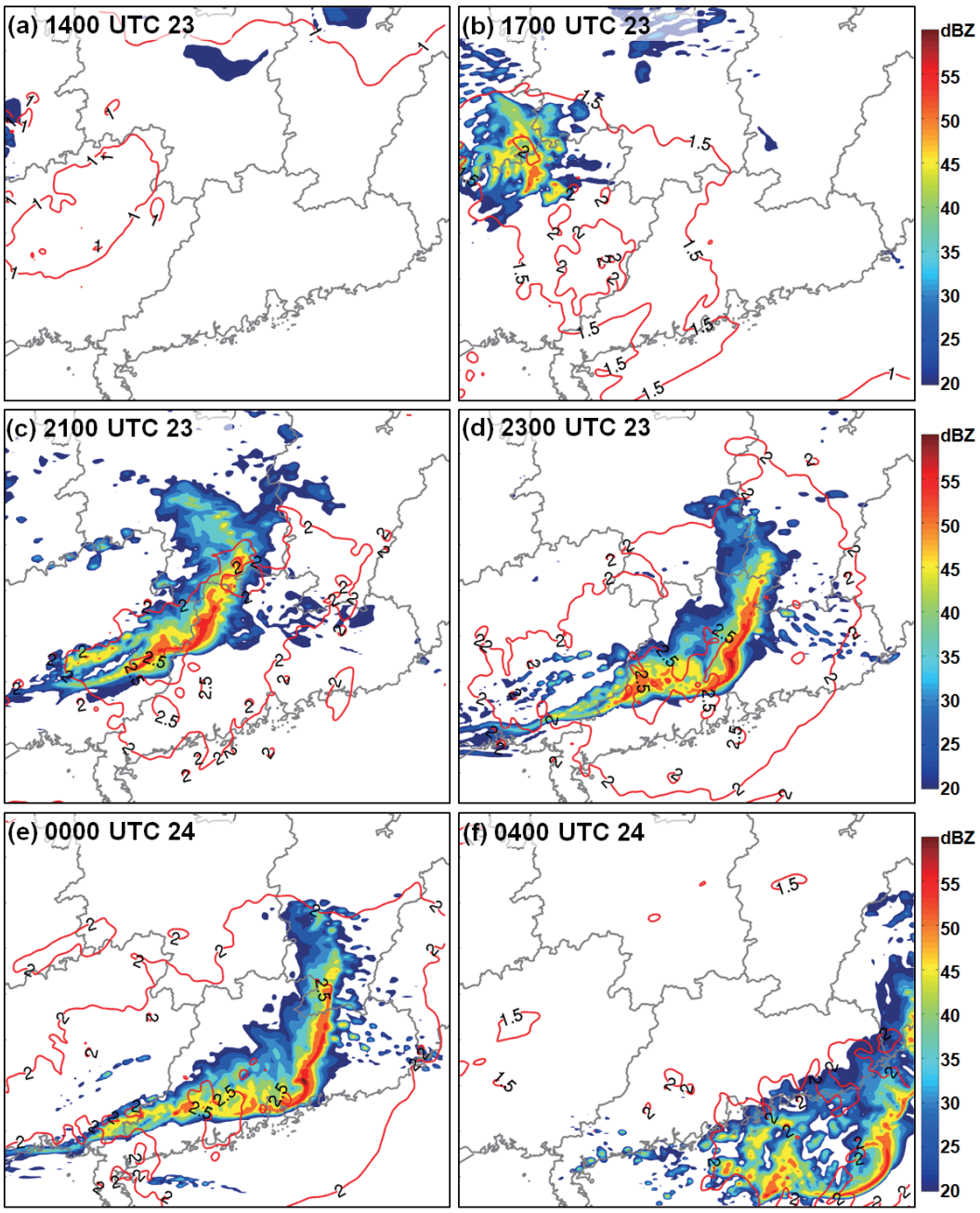

Fig. 10. The horizontal distribution of the ensemble spread of temperature (red contour, K) at (a) 1400 UTC 23 April, (b) 1700 UTC 23 April, (c) 2100 UTC 23 April, (d) 2300 UTC 23 April, (e) 0000 UTC 24 April, and (f) 0400 UTC 24 April. The shading denotes the simulated composite radar reflectivity of M10 (a good member) at the corresponding times.

formed by adding the differences of all variables except q_v, (Figs. 13g–i). NoQ$_v$_1–9 showed only slightly better results than UV_1–9 and T_1–9, but it still failed to capture a continuous squall line, even when the initial condition of all variables except for moisture was almost the same as those of M10 in NoQ$_v$_9. This result indicates that moisture may be a more sensitive factor that affects the predictability of squall lines than other variables.

5. Sensitivity of squall-line simulation to lead time

How early a storm can be forecast for an allowable forecast error is an important question both in predictability research and disaster preparedness practice.

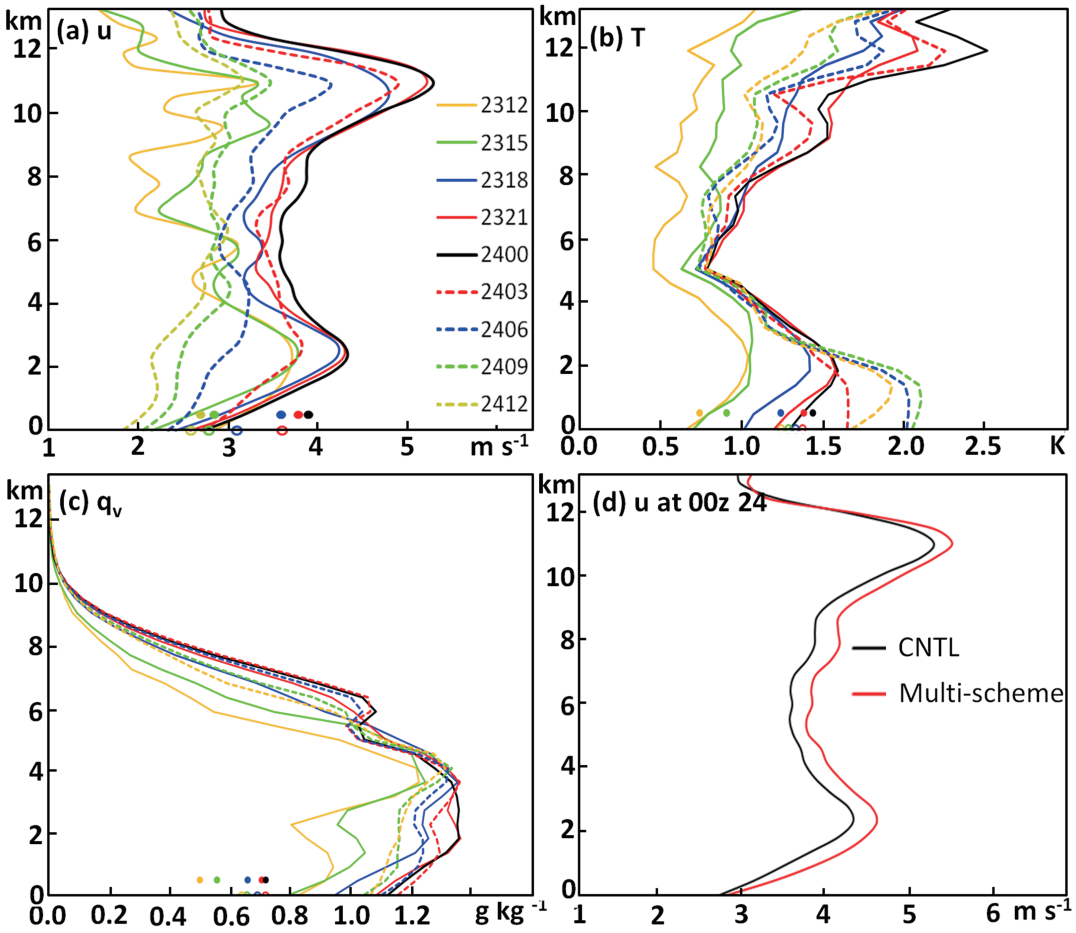

Fig. 11. The vertical distribution of horizontally averaged ensemble spread of (a) u, (b) T, (c) q_v every 3 h in S12Z from 1200 UTC 23 to 1200 UTC 24, and (d) u at 0000 UTC 24 in S12Z and M12Z. The dots and circles with the same color convention marked on the X-axis in (a), (b), (c) denote the domain averaged ensemble spread for the corresponding variables at the corresponding times. (Dots are corresponding to solid lines, while circles are corresponding to dashed lines).

The sensitivity of the ensemble forecast performance to different lead times was explored in this study by starting the ensemble forecast 12 h earlier than the control ensemble forecast at 0000 UTC 23 April (referred to as S00Z).

The ensemble forecasts were classified into the same three types as in S12Z based on the simulated squall-line features at 0000 UTC 24 April. The result shows that starting the model 12-h earlier significantly worsened the forecast (Table 2). None of the 40 members captured a squall line close to the observations. Only 27.5% of the members generated a squall line with large location error. No squall line formed in the remaining 72.5% of the ensemble. We also tried starting the model 6 h earlier than the control experiment at 0600 UTC 23 April, but an even worse result emerged. Its poor quality was likely due to the fact that many fewer radiosonde observations were ingested in the analysis at 0600 UTC. This result suggests that error accumulation during a longer lead time may result in a larger uncertainty in the environment, which may substantially decrease the probability of a good forecast. This experiment indicates that a proper lead time to predict a squall line is $\leqslant$10 h.

6. A possible way to improve the predictability of the squall line

As mentioned in section 3.2, physical parameterization is a major error source of NWP. The WRF model has several schemes from which to choose for each type of the physical parameterization. Each scheme has its own advantages and limits. Studies show that none of the schemes works consistently better than

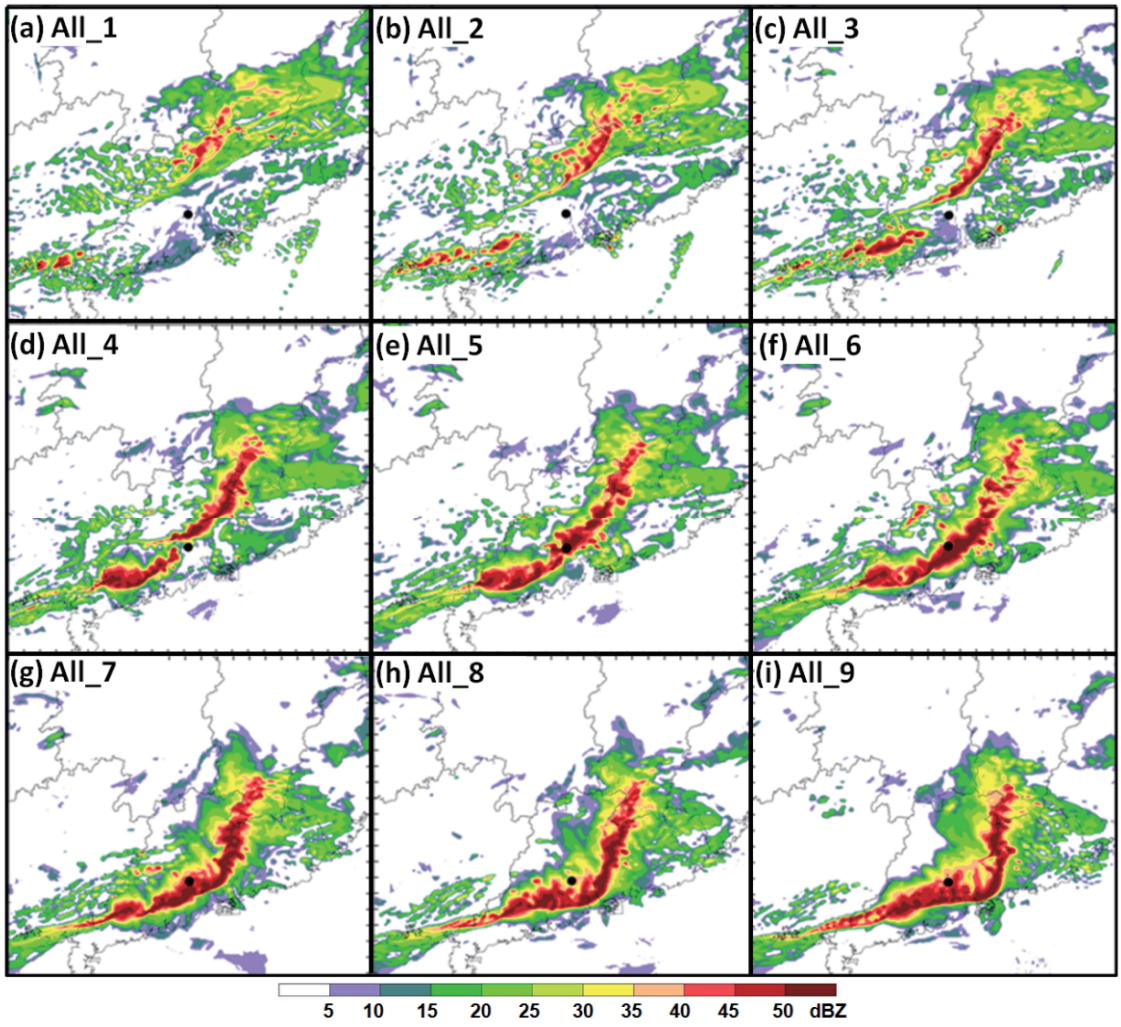

Fig. 12. The simulated composite radar reflectivity at 0000 UTC 24 April in experiments All_1–9.

other schemes (Wang and Seaman, 1997). Consequently, using different parameterization schemes for different members (hereafter referred to as a multi-scheme ensemble) could be a good way to account for model uncertainties. Stensrud et al. (2000) first used a multi-scheme ensemble for a mesoscale ensemble forecast. Fujita et al. (2007) found that an ensemble using a combination of different physical parameterization schemes produced a larger ensemble spread of thermodynamic variables than an ensemble generated by perturbing only the initial field with all members using the same parameterization schemes. Meng and Zhang (2007, 2008a, b, 2011) utilized the multi-scheme ensemble in an ensemble Kalman filter (EnKF) data assimilation; they successfully accounted for model error and apparently improved the performance of the EnKF. We sought to determine the extent to which using a multi-scheme ensemble may affect the performance of squall-line ensemble forecast. To this end, experiment M12Z was performed using the same initial ensemble as that in S12Z, but this experiment was integrated with different combinations of physical parameterization schemes for different members (Table 3). Several schemes available in WRF for cumulus, microphysics, PBL, and long-wave radiation parameterizations were chosen to be assigned as equally as possible to the 40 members. The schemes in boldface italics were used in the control ensemble forecast S12Z.

Using a multi-scheme ensemble apparently improved the performance of the ensemble forecast. The percentage of the members that totally missed the squall line decreased from 25% to only 15%, while the percentage of the members that captured a squall line increased from 60% to 70%, though no change occurred in the percentage of members of Type 1 (Table 2). The better performance of M12Z was likely due

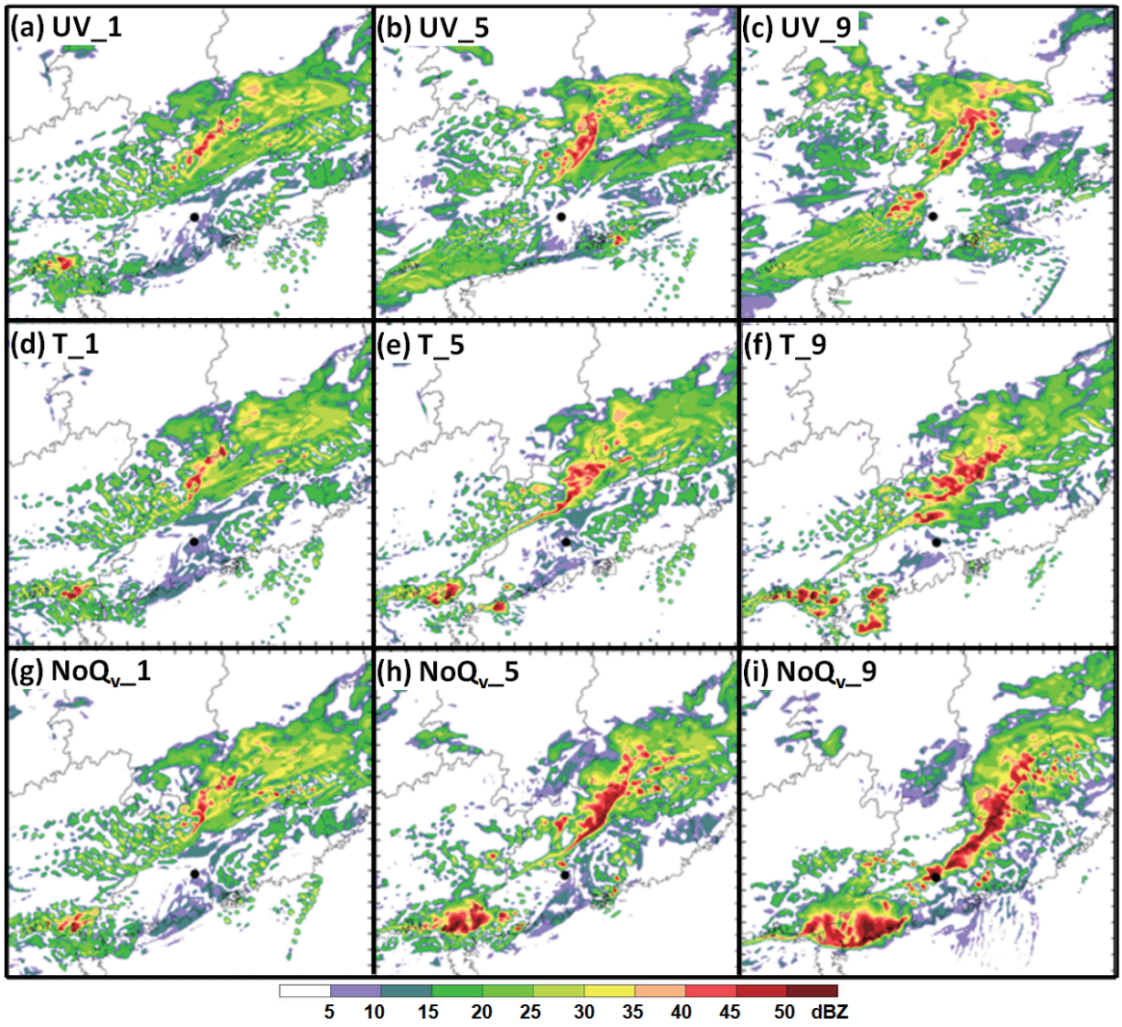

Fig. 13. The same as Fig. 12 but for experiments UV_1, 5, 9, T_1, 5, 9 and NoQ$_v$_1, 5, 9.

to the resulting larger and more reasonable ensemble spread (Fig. 11d), which helped to increase the probability of the ensemble to produce a better estimation of the truth.

7. Summary and discussion

This study investigated the predictability of a squall line in South China on 23 April 2007 through deterministic and probabilistic (ensemble) forecasts using WRFV3.

The impact of model error on the squall line simulation used a deterministic forecast approach. Our results show that the impact of uncertainties in physical parameterization schemes on the squall line simulation was two folds. First, the simulated squall line with a grid size of 4.5 km was most sensitive to different longwave radiation schemes relative to different cumulus, microphysics, and PBL schemes. Second, whether cumulus parameterization scheme was turned on or off made a considerable difference for a grid size varying from 20 km down to 5 km. Using a cumulus parameterization scheme degraded the simulation of the squall line with a grid size from 20 km to 5 km. The degradation to the simulation with a grid size <10 km was more severe than that with a coarser resolution. When cumulus parameterization was turned off, the higher the grid size was, the better the simulation was, with more and more detailed structures captured. Our results show that at least a 10-km grid size was necessary to reasonably capture the squall line in this event.

The sensitivity of the squall-line simulation to the initial error was investigated from a probabilistic point of view through ensemble forecasting. By perturbing the initial field with a standard deviation representing the analysis error, a 40-member, 24-h ensemble forecast was performed, with a lead time of eight h. Approximately 15% of the ensemble members decently

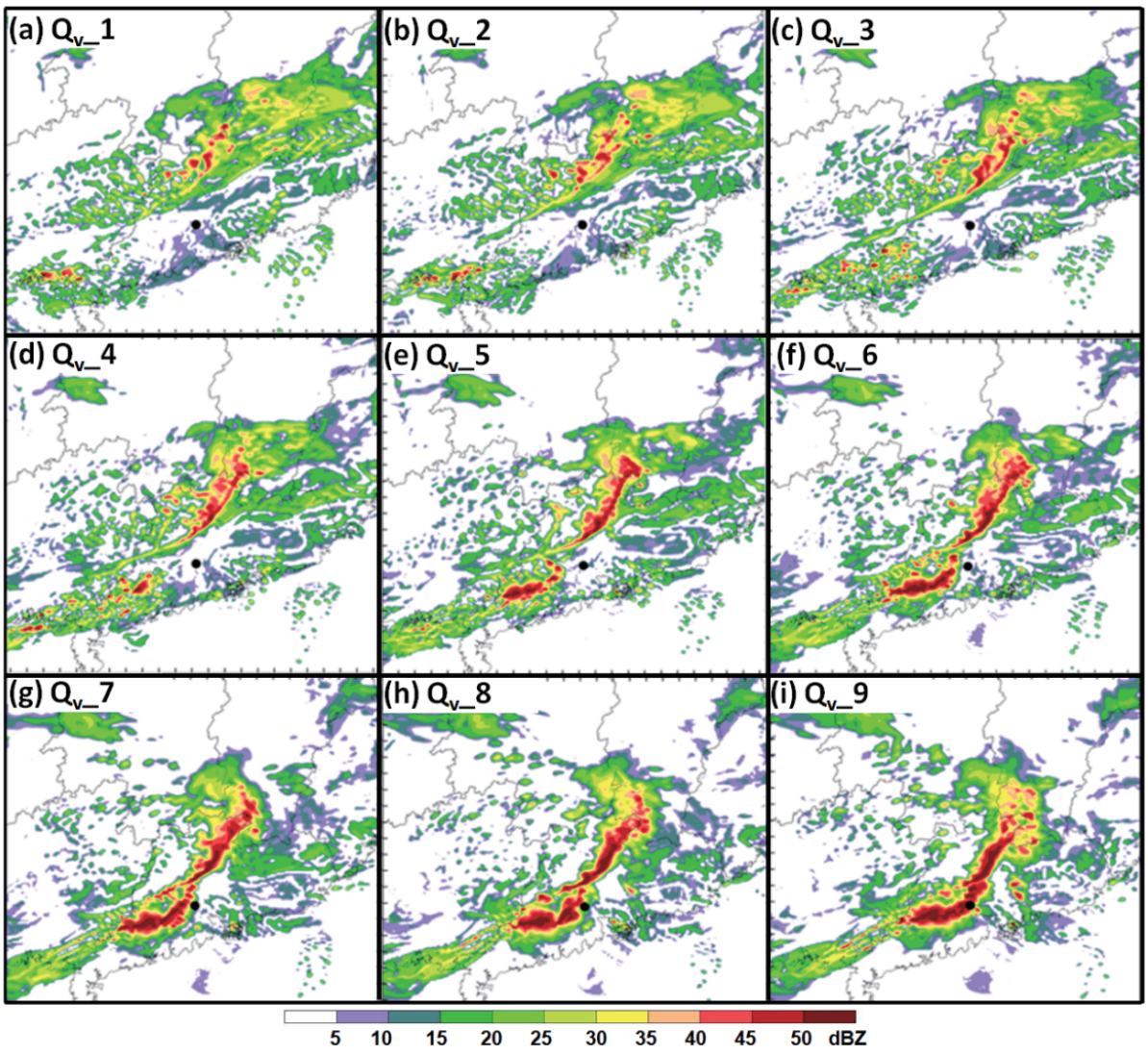

Fig. 14. The same as Fig. 12 but for experiments Q_v_1–9.

Table 3. The distribution of different physical schemes in the multi-scheme ensemble experiment. The scheme in bold Italic is the one used in C4.5.

No. of members using a cumulus scheme and the scheme used	No. of members using a microphysics scheme and the scheme used	No. of members using a PBL scheme and the scheme used			No. of members using a long-wave scheme and the scheme used	
13, Kain-Fritsch	4, Lin	1,***YSU***	2,ETA	1,MRF	3,***rrtm***	1,cam
	4, Thompson	1,YSU	2,ETA	1,MRF	3,rrtm	1,cam
	5, ***WSM6 with graupel***	2,YSU	2,ETA	1,MRF	3,rrtm	2,cam
13, Betts-Miller	4, Lin	1,YSU	2,ETA	1,MRF	3,rrtm	1,cam
	4, Thompson	1,YSU	2,ETA	1,MRF	3,rrtm	1,cam
	5, WSM6 with graupel	2,YSU	2,ETA	1,MRF	3,rrtm	2,cam
14, ***Grell − Devenyi***	4, Lin	1,YSU	2,ETA	1,MRF	3,rrtm	1,cam
	5, Thompson	2,YSU	2,ETA	1,MRF	3,rrtm	2,cam
	5, WSM6 with graupel	2,YSU	2,ETA	1,MRF	3,rrtm	2,cam

captured the evolution of the squall line. Approximately 25% of the ensemble members failed to capture the squall line, whereas the remaining 60% of the ensemble captured the squall line but with a large location error. The growth of the ensemble spread was clearly associated with moist convection development in terms of both time and space. A linear improvement in the squall-line simulation was observed when the initial error decreased gradually, which was different from the nonlinear improvement that was seen in a squall line case in North America. Our results also show that a better simulation of the squall line was mainly due to the improvement of moisture field relative to other variables.

Using different combinations of physical parameterization schemes for different members can apparently improve the probabilistic forecast, likely through a better representation of model uncertainties in the ensemble. The percentage of the members that totally missed the squall line decreased from 25% to only 15%. Furthermore, a different lead time may substantially affect the probabilistic forecast. If the model was initialized 12- or 6-h earlier, the performance of the ensemble became much worse. None of the ensemble members was reasonably simulated, which indicates that the lead time of this squall-line case in the current model configuration was only a few hours.

Notably, the result of this work was based on only one squall-line case. The predictability of squall lines is likely to exhibit flow dependence similar to that of severe heavy rain events. Because the predictability of squall lines, especially in China, has not been well studied, we believe that the sensitivities of the simulated squall line to the model and initial errors, the error growth features, the general performance of probabilistic forecast examined in this study may contribute to the general knowledge of squall line forecasting. In addition, the investigation of the proper lead time and possible model improvements may help to improve the predictability of squall-line systems.

Acknowledgements. We thank Fuqing ZHANG from Pennsylvania State University for discussions that benefited this study. The radar mosaics were produced using the software provided by Prof. Liping LIU from Chinese Academy of Meteorological Sciences. This work was supported by National Key Basic Research Program of China (Grant No. 2013CB430104), National Natural Science Foundation (Grant Nos. NSFC41075031 and NSFC40921160380), and the R&D Special Fund for Public Welfare Industry (meteorology) (Grant No. GYHY200906025).

REFERENCES

Chen, R., S. Gu, and Y. Huang, 2008: Doppler radar characteristics of a severe squall line event. *Guangdong Meteorology*, **30**(1), 20–23. (in Chinese)

Ehrendorfer, M., R. M. Errico, and K. D. Raeder, 1999: Singular-vector perturbation growth in a primitive equation model with moist physics. *J. Atmos. Sci.*, **56**, 1627–1648.

Fujita, T., D. J. Stensrud, and D. C. Dowell, 2007: Surface data assimilation using an ensemble Kalman filter approach with initial condition and model physics uncertainties. *Mon. Wea. Rev.*, **135**, 1846–1868.

Glickman, T. S., Ed., 2000: *Glossary of Meteorology.* 2nd ed. Amer. Meteor. Soc., 855pp. [Available online at http://amsglossary.allenpress.com/glossary/search?id=squall-line1.]

Grell, G. A., and D. Devenyi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Lett.*, **29**(14), doi: 10.1029/2002GL015311.

Hong, S., J. Dudhia, and S. Chen, 2004: A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Mon. Wea. Rev.*, **132**, 103–120.

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

Kraichnan, R. H., 1971: Inertial-range transfer in two- and three-dimensional turbulence. *J. Fluid Mech.*, **47**, 525–535.

Kuo, Y. H., J. R. Gyakum, and Z. Guo, 1995: A case of rapid continental mesoscale cyclogenesis. Part I: Model sensitivity experiments. *Mon. Wea. Rev.*, **123**, 970–997.

Leith, C. E., 1971: Atmospheric predictability and two-dimensional turbulence. *J. Atmos. Sci.,* **28**, 145–161.

Leith, C. E., and R. H. Kraichnan, 1972: Predictability of turbulent flows. *J. Atmos. Sci.*, **29**, 1041–1058.

Liu, J., and Z. Tan, 2009: Mesoscale predictability of Meiyu heavy rainfall. *Adv. Atmos. Sci.*, **26**, 438–450, doi: 10.1007/s00376-009-0438-9.

Liu J., Z. Tan, and S. Gu, 2011: Flow dependent mesoscale predictability of Meiyu heavy rainfall. *Chinese J. Atmos. Sci.*, **35**, 912–926. (in Chinese)

Lorenz, E. N., 1963a: Deterministic nonperiodic flow. *J. Atmos. Sci.,* **20**, 130–141.

Lorenz, E. N., 1963b: The predictability of hydrodynamic flow. *Transactions of the New York Academy of Sciences*, **25B**, 409–432.

Melhauser, C., and F. Zhang, 2012: Practical and intrinsic predictability of severe and convective weather at the mesoscales. *J. Atmos. Sci.*, **69**, 3350–3371.

Meng, Z., and F. Zhang, 2011: Limited-area ensemble-based data assimilation. *Mon. Wea. Rev.*, **139**, 2025–2045.

Meng, Z., and F. Zhang, 2008a: Test of an ensemble-Kalman filter for mesoscale and regional-scale data

assimilation. Part III: Comparison with 3DVar for a real-data case study. *Mon. Wea. Rev.*, **136**, 522–540.

Meng, Z., and F. Zhang, 2008b: Test of an ensemble-Kalman filter for mesoscale and regional-scale data assimilation. Part IV: Comparison with 3DVar in a month-long experiment. *Mon. Wea. Rev.*, **136**, 3671–3682.

Meng, Z., and F. Zhang, 2007: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part II: Imperfect model experiments. *Mon. Wea. Rev.*, **135**, 1403–1423.

Meng, Z., and Y. Zhang, 2012: On the Squall Lines Preceding Landfalling Tropical Cyclones in China. *Mon. Wea. Rev.*, **140**, 445–470.

Meng, Z., D. Yan, and Y. Zhang, 2013: General features of squall lines in East China . *Mon. Wea. Rev.*, doi: 10.1175/MWR-D-12-00208.1. (in press)

Meng, Z., F. Zhang, P. Markowski, D. Wu, and K. Zhao, 2012: A modeling study on the development of bowing structure and associated rear Inflow within a squall line over South China. *J. Atmos. Sci.*, **69**, 1182–1207.

Mu, M., and W. Duan, 2003: A new approach to studying ENSO predictability: Conditional nonlinear optimal perturbation. *Chinese Science Bulletin*, **48**, 1045–1047.

Mu, M., W. Duan, and J. Wang, 2002a: The predictability problems in numerical weather and climate prediction. *Adv. Atmos. Sci.*, **19**, 191–204.

Mu, M., J. Li, J. Chou, W. Duan, and J. Wang, 2002b: Theoretical research on the predictability of climate system. *Climatic and Environmental Research*, **7**, 227–235. (in Chinese)

Mu, M., H. Wang, and F. Zhou, 2007: A preliminary application of conditional nonlinear optimal perturbation to adaptive observation. *Chinese J. Atmos. Sci.*, **31**, 1102–1112. (in Chinese)

Orrell, D., 2003: Model error and predictability over different timescales in the Lorenz '96 systems. *J. Atmos. Sci.*, **60**, 2219–2228.

Parker, M. D., and R. H. Johnson, 2000: Organizational modes of midlatitude mesoscale convective systems. *Mon. Wea. Rev.*, **128**, 3413–3436.

Robinson, G. D., 1971: The predictability of a dissipative flow. *Quart. J. Roy. Meteor. Soc.*, **93**, 409–418.

Skamarock, W. C., and Coauthors, 2008: A description of the advanced research WRF version 3. NCAR Tech. Note TN-475_STR, 113pp.

Stensrud, D. J., J.-W. Bao, and T. T. Warner, 2000: Using initial condition and model physics perturbations in short-range ensemble simulations of mesoscale convective systems. *Mon. Wea. Rev.*, **128**, 2077–2107.

Tan, Z., F. Zhang, R. Rotunno, and C. Snyder, 2004: Mesoscale predictability of moist baroclinic waves: Experiments with parameterized convection. *J. Atmos. Sci.*, **61**, 1794–1804.

Thompson, P. D., 1957: Uncertainty of initial state as a factor in the predictability of large scale atmospheric flow patterns. *Tellus*, **9**, 275–295.

Wang, W., and N. L. Seaman, 1997: A comparison study of convective parameterization schemes in a mesoscale model. *Mon. Wea. Rev.*, **125**, 252–278.

Wu, D., and Z. Meng, 2013: On the movement and mesoscale surface structure of a squall line on 23 April 2007 in Guangdong. *Journal of Natural Science of Peking University*, accepted.

Zhang, F., C. Snyder, and R. Rotunno, 2002: Mesoscale predictability of the 'surprise' 24–25 January 2000 snowstorm. *Mon. Wea. Rev.*, **130**, 1617–1632.

Zhang, F., C. Snyder, and R. Rotunno, 2003: Effects of moist convection on mesoscale predictability. *J. Atmos. Sci.*, **60**, 1173–1185.

Zhang, F., A. M. Odins, and J. W. Nielsen-Gammon, 2006: Mesoscale predictability of an extreme warm-season precipitation event. *Wea. Forecasting*, **21**, 149–166.

Zhu, B., W. Lin, and Y. Zhang, 2009: Impacts of initial perturbations on predictability of a heavy rain in South China. *Chinese J. Atmos. Sci.*, **33**, 1333–1347. (in Chinese)