Preprints, 23nd Conf. on Weather Analysis and Forecasting/19th Conf. Num. Wea. Pred. Amer. Meteor. Soc., Omaha, Nebraska, June 1-5, 2009.

16A.3

A REAL-TIME STORM-SCALE ENSEMBLE FORECAST SYSTEM: 2009 SPRING EXPERIMENT

Fanyou Kong[1]*, Ming Xue[1,2], Kevin W. Thomas[1], Yunheng Wang[1],
Keith Brewster[1], Jidong Gao[1], Kelvin K. Droegemeier[1,2],
Jack Kain[4], Steven J. Weiss[3], David Bright[3], Michael C. Coniglio[3], and Jun Du[5]
[1]Center for Analysis and Prediction of Storms, and [2]School of Meteorology,
University of Oklahoma, Norman, OK 73072
[3]NOAA/NMS/NCEP Storm Prediction Center
[4]NOAA National Severe Storm Laboratory, Norman, OK 73072
[5]NOAA/NWS/NCEP, Camp Springs, MD 20746

1. INTRODUCTION

After two years of highly successful real-time storm-scale ensemble forecast experiment in 2007 and 2008, performed by the Center for Analysis and Prediction of Storms (CAPS), University of Oklahoma in collaboration with the Storm Prediction Center (SPC) and the National Severe Storm Laboratory (NSSL), to support the NOAA Hazardous Weather Testbed (HWT) Spring Experiment Program (Xue et al. 2007, 2008; Kong et al. 2007, 2008), significant changes are made to the ensemble system in the still on-going 2009 Spring Experiment that features a doubling of ensemble members from 10 to 20 and two more numerical weather prediction model systems, the Nonhydrostatic Mesoscale Model WRF dynamical core (WRF-NMM) and the Advanced Regional Prediction System (ARPS) in addition to the WRF-ARW modeling system already been used in previous years. As the real-time ensemble forecast Experiment is currently underway, this extended abstract presents some examples of the real-time ensemble forecast product and preliminary assessment of the multimodel storm-scale ensemble system.

2. EXPERIMENT HIGHLIGHT

The CAPS 2009 Spring Program started on 20 April 2009 and will end on 5 June, encompassing the NOAA HWT 2009 Spring Experiment that is officially between 4 May and 5 June. Three numerical weather models are used to produce a 20 member 30 h ensemble forecast during weekdays, initialized at 0000 UTC, covering a near-CONUS domain at 4 km horizontal grid spacing. Ten members are produced using the Weather Research and Forecast (WRF) Advanced Research WRF core (ARW), eight members are produced using the WRF Nonhydrostatic Mesoscale Model core (NMM), and two members are produced using the Advanced Regional Prediction System (ARPS). Both WRF cores are V3.0.1.1 release. A companion paper to this conference (Xue et al. 2009) has detail description for the entire 2009 Spring Experiment design. This extended abstract only highlights key configurations for the ensemble system.

Corresponding Author Address: Dr. Fanyou Kong, Center for Analysis and Prediction of Storms, Univ. of Oklahoma, Norman, OK 73072; e-mail: *fkong@ou.edu*

As in 2008 Experiment (Kong et al. 2008; Xue et al. 2008), daily 30 h forecasts were initiated at 0000 UTC, using NAM 12 km (218 grid) 00Z analyses as background for initialization with the initial condition perturbations for the ensemble members coming from the NCEP Short-Range Ensemble Forecast (SREF). Doppler radar radial wind and reflectivity data from 120 available WSR-88D stations within the domain are assimilated through ARPS 3DVAR and cloud analysis package into all but three members (one from each model group).

The daily 30 h ensemble forecasts, for the weekdays from Monday through Friday, started at 0000 UTC and ended at 0600 UTC of the next day. Special weekend runs are arranged if it is requested by SPC based on the severe weather outlook. All ARW and NMM forecasts are produced on Bigben, a Cray XT3 supercomputing system, at the Pittsburgh Super-computing Center (PSC), while ARPS forecasts are produced on Kraken, a Cray XT5 system, at the National Institute of Computational Sciences (NICS). Hourly model outputs are archived on Mass storage facilities at PSC and NICS. Selected Figure 1 shows the coverage area of the model domains.

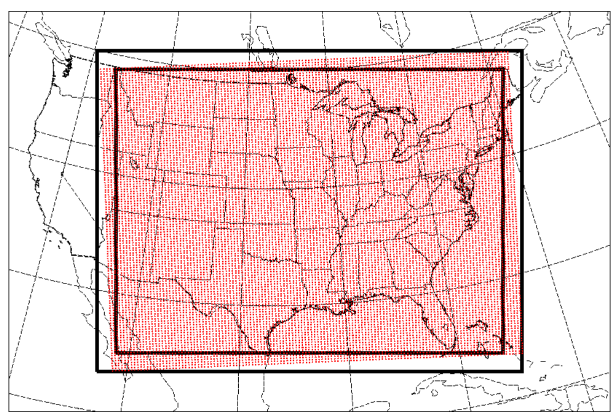

Figure 1. Computational domains for the 2009 Spring Experiment. The outer thick rectangular box represents the domain for performing 3DVAR/Cloud Analysis (1000×760). The red dot area represents the WRF-NMM forecast domain (650×979). The inner thick box is the domain for WRF-ARW and ARPS forecast and also for common verification (900×672), which is the same as the domain used in 2008 Spring Experiment.

Since NMM uses rotated E-grid while both ARW and ARPS use C-grid, special software codes were developed at CAPS to convert between the two grids in order to utilize a single 3DVAR/Cloud Analysis over a larger outer domain that encompasses both forecast domains (Figure 1) by converting the analysis to the forecast domains, and to convert NMM forecast to a common verification domain that is the same as the ARW and ARPS forecast domain.

Table 1-3 outline the basic configuration for each individual members of each model group. *cn* refers to the control member, with radar data analysis, *c0* is the same as *cn* except no radar data. *n1-n4* and *p1-p4* are members with initial perturbation added on top of *cn* initial condition, NAMa and NAMf refer to 12 km NAM analysis and forecast, respectively. ARPSa refers to ARPS 3DVAR and Cloud Analysis using NAMa as background. For the perturbed members, the ensemble initial conditions consist of a mixture of bred perturbations coming from the 21Z SREF perturbed members (one pair each from WRF-em (ARW), WRF-nmm (NMM), ETA-KF, and ETA-BMJ) and physics variations (grid-scale microphysics, radiation, land-surface model and PBL physics). The lateral boundary conditions come from the corresponding 21Z SREF forecasts directly for those perturbed members and from the 00Z 12 km NAM forecast for the non-perturbed members (*cn* and *c0*).

For the ARPS model group, the only members are cn and c0.

Table 1. Configurations for each individual member with WRF-ARW core. NAMa and NAMf refer to the 12 km NAM analysis and forecast, respectively. ARPSa refers to ARPS 3DVAR and cloud analysis

member	IC	BC	Radar data	mp_phy	sw-phy	sf_phy	pbl_phy
arw_cn	00Z ARPSa	00Z NAMf	yes	Thompson	Goddard	Noah	MYJ
arw_c0	00Z NAMa	00Z NAMf	no	Thompson	Goddard	Noah	MYJ
arw_n1	arw_cn – em_pert	21Z SREF em-n1	yes	Ferrier	Goddard	Noah	YSU
arw_p1	arw_cn + em_pert	21Z SREF em-p1	yes	WSM 6-class	Dudhia	Noah	MYJ
arw_n2	arw_cn – nmm_pert	21Z SREF nmm-n1	yes	Thompson	Dudhia	RUC	MYJ
arw_p2	arw_cn + nmm_pert	21Z SREF nmm-p1	yes	WSM 6-class	Dudhia	Noah	YSU
arw_n3	arw_cn – etaKF_pert	21Z SREF etaKF-n1	yes	Thompson	Dudhia	Noah	YSU
arw_p3	arw_cn + etaKF_pert	21Z SREF etaKF-p1	yes	Ferrier	Dudhia	Noah	MYJ
arw_n4	arw_cn – etaBMJ_pert	21Z SREF etaBMJ-n1	yes	WSM 6-class	Goddard	Noah	MYJ
arw_p4	arw_cn + etaBMJ_pert	21Z SREF etaBMJ-p1	yes	Thompson	Goddard	RUC	YSU

* For all members: ra_lw_physics= RRTM; cu_physics= NONE

Table 2. Configurations for each individual member with WRF-NMM core

member	IC	BC	Radar data	mp_phy	lw_phy	sw-phy	sf_phy	pbl_phy
nmm_cn	00Z ARPSa	00Z NAMf	yes	Ferrier	GFDL	GFDL	Noah	MYJ
nmm_c0	00Z NAMa	00Z NAMf	no	Ferrier	GFDL	GFDL	Noah	MYJ
nmm_n1	nmm_cn – em_pert	21Z SREF em-n1	yes	Thompson	RRTM	Dudhia	Noah	MYJ
nmm_p1	nmm_cn + em_pert	21Z SREF em-p1	yes	WSM 6-class	GFDL	GFDL	RUC	MYJ
nmm_n2	nmm_cn – nmm_pert	21Z SREF nmm-n1	yes	Ferrier	RRTM	Dudhia	Noah	YSU
nmm_p2	nmm_cn + nmm_pert	21Z SREF nmm-p1	yes	Thompson	GFDL	GFDL	RUC	YSU
nmm_n3	nmm_cn – etaKF_pert	21Z SREF etaKF-n1	yes	WSM 6-class	RRTM	Dudhia	Noah	YSU
nmm_p3	nmm_cn + etaKF_pert	21Z SREF etaKF-p1	yes	Thompson	RRTM	Dudhia	RUC	MYJ
nmm_n4	nmm_cn – etaBMJ_pert	21Z SREF etaBMJ-n1	yes	WSM 6-class	RRTM	Dudhia	RUC	MYJ
nmm_p4	nmm_cn + etaBMJ_pert	21Z SREF etaBMJ-p1	yes	Ferrier	RRTM	Dudhia	RUC	YSU

* For all members: cu_physics= NONE. The two grayed out rows are removed in final real-time forecast system due to computation constrains at PSC, leaving total eight NMM contributing members.

Table 3. Configurations for each individual member with ARPS

member	IC	BC	Radar data	Microphysics	radiation	PBL	Turb	sf_phy
arps_cn	00Z ARPSa	00Z NAMf	yes	Lin	Goddard	TKE	3D TKE	Force-restore
arps_c0	00Z NAMa	00Z NAMf	no	Lin	Goddard	TKE	3D TKE	Force-restore

* For all members: no cumulus parameterization

Selected 2D weather fields from each ensemble member are written in GEMPAK format and are directly transferred into SPC's N-AWIPS system to be reviewed by participants to the HWT Spring Experiment at SPC's daily weather briefing. In addition, CAPS also makes available a webpage showing the Spring Experiment products (http://www.caps.ou.edu/wx/spc), and a supplemental webpage [1] in demonstrating ensemble products and high frequency (5 min interval) reflectivity movies from the real-time forecast.

[1] http://www.caps.ou.edu/~fkong/sub_atm/spring09.html

3. ENSEMBLE PRODUCT EXAMPLES

Figure 2-4 show example ensemble forecast reflectivity products generated during the 2009 real-time Spring Experiment. Showing are 30 h forecast initialized at 0000 UTC on 5 May, 7 May, and 8 May 2009, including Probability Matching (Ebert 2001; Kong et al 2008) composite reflectivity, un-calibrated probability of composite reflectivity exceeding 35 dBZ, and spaghetti chart of composite reflectivity equal to 35 dBZ, validated against the observed composite reflectivity mosaic valid at the same times (0600 UTC on 6 May, 8 May, and 9 May 2009, respectively).

Owing to very high spatial and temporal variance associated with precipitation, especially when produced with very high-resolution model runs like in this case, ensemble mean of precipitation field tends to be excessively broad in area coverage and too weak in magnitude, and thus is not a useful QPF product Probability Matching (PM, hereafter) (Ebert 2001; Clark et al. 2008; Kong et al. 2008) has been proven to be a more useful deterministic QPF variable derived from ensemble forecasts by assuming that the best spatial representation of rainfall (or reflectivity) is given by the ensemble mean and that the best frequency distribution of rainfall (reflectivity) is given by the ensemble member QPFs. PM products for QPF are produced by first pooling QPF amounts of all ensemble members and over all grid points for a given forecast lead time and sorted from the highest to the lowest to obtain a QPF distribution. The ensemble mean QPF amounts are also sorted from the highest to the lowest. Then the QPF values from the ensemble mean are reassigned using values from the corresponding ranks of the QPF distribution. Given N ensemble members and M total grid numbers of model domain, there are MN elements in QPF distribution versus N in ensemble mean. Kong et al (2008) modified Ebert's method by averaging the N elements and assigning the mean to the corresponding rank of ensemble mean. This new approach reduces extremely high peak values and preserves the element of averaging in the resulting PM field. Still there is drawback. The assumption of ensemble mean's spatial distribution makes PM unable to reflect storm detail structures that can be produced from convection-resolving or convection-permitting high-resolution ensemble forecasts, such as in this project.

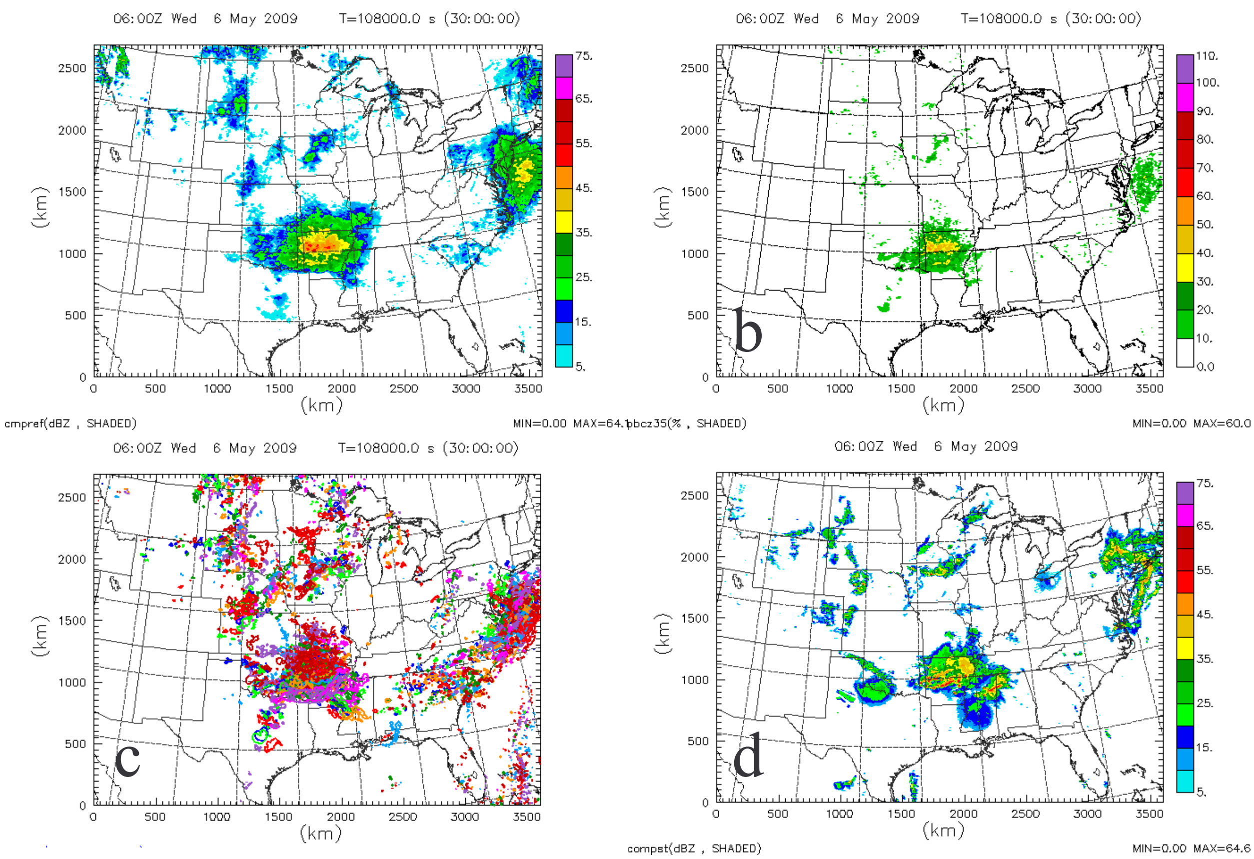

Figure 2. 30 h forecast of composite reflectivity products, valid at 0600 UTC 6 May 2009. (a) Probability Matching), (b) probability of composite reflectivity ≥35 dBZ, and (c) spaghetti chart of composite reflectivity = 35 dBZ, and (d) observed composite reflectivity mosaic.

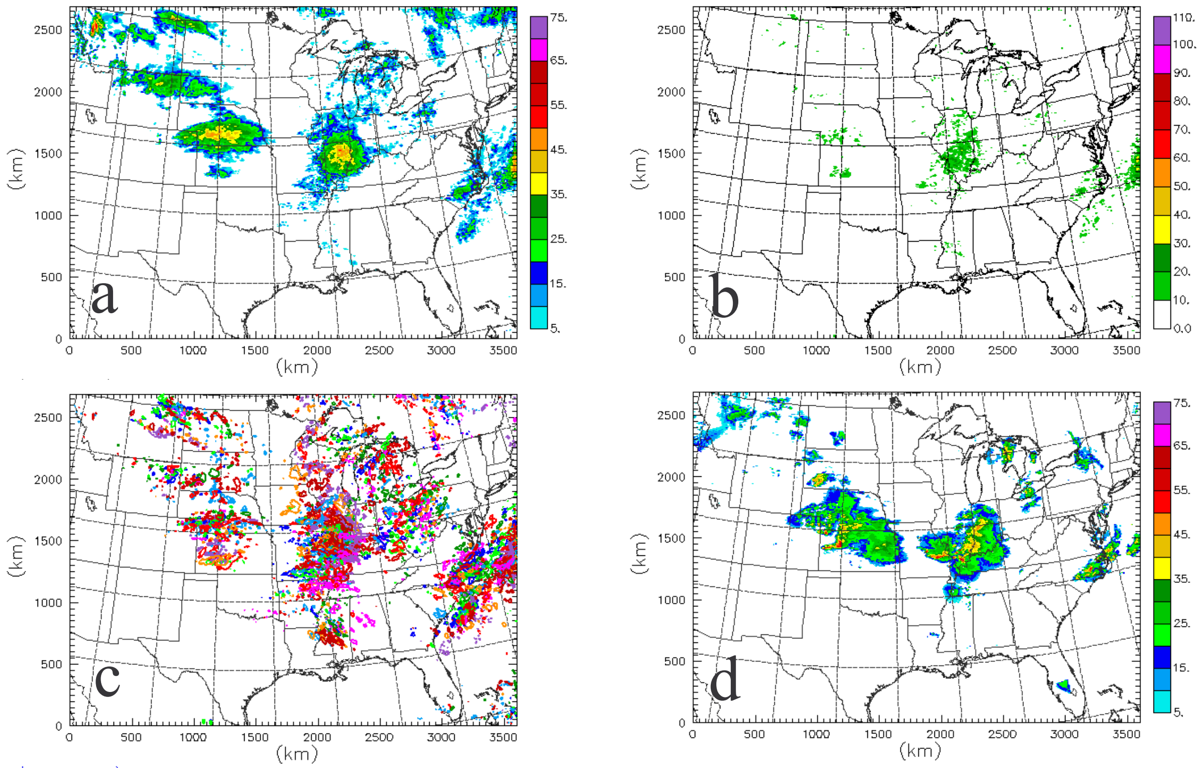

Figure 3. 30 h forecast of composite reflectivity products, valid at 0600 UTC 8 May 2009. (a) Probability Matching), (b) probability of composite reflectivity ≥35 dBZ, and (c) spaghetti chart of composite reflectivity = 35 dBZ, and (d) observed composite reflectivity mosaic.

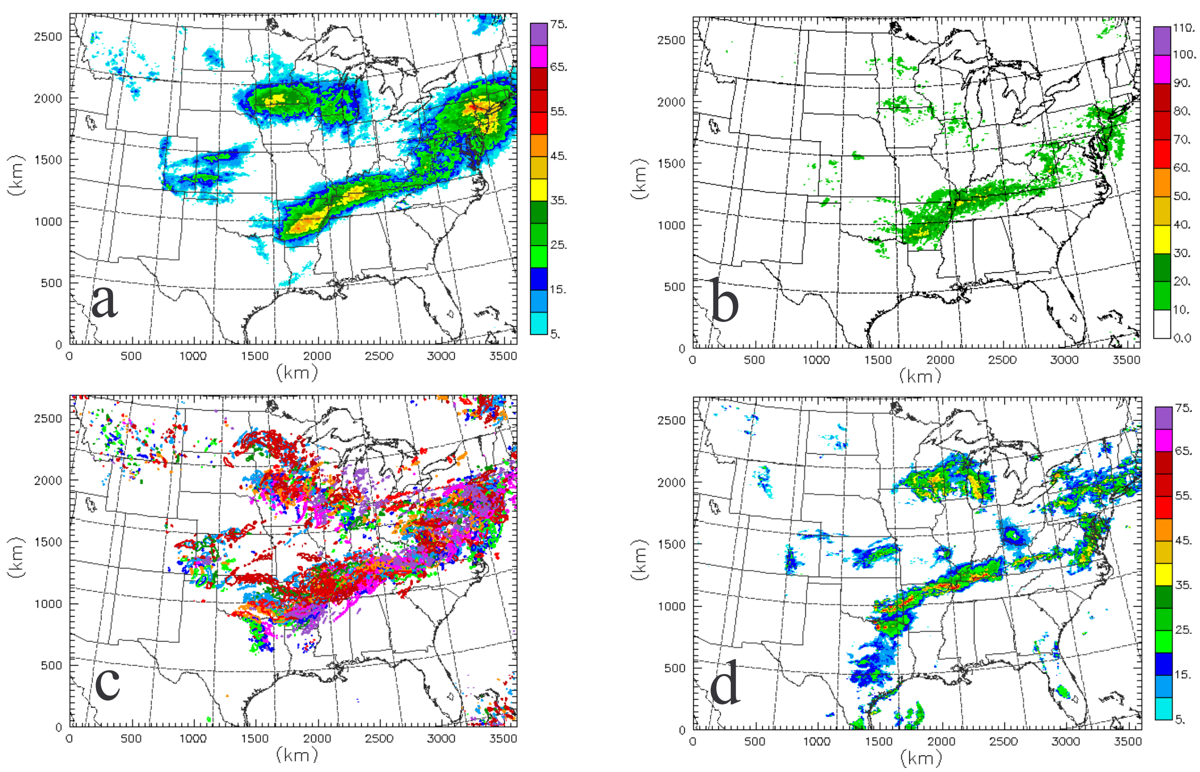

Figure 4. Same as Fig 3, but valid at 0600 UTC 9 May 2009.

4. ASSESSMENT OF THE ENSEMBLE SYSTEM

As the real-time storm-scale ensemble forecast Experiment is still underway at the completion of this extended abstract, analysis on a small dataset has been performed to have a preliminary assessment on the statistic feature and performance of the expanded ensemble system. Results presented in this section are calculated from ten dates that have complete dataset available for all 20 ensemble members. As in 2007 and 2008 Experiment seasons, the experimental fine grid (1 km) national radar mosaic and QPE products generated by the NSSL/NMQ project[2] are first interpolated to the 4 km verification domain and used as verification dataset to verify the predicted QPF quantities (1 h accumulated precipitation and composite reflectivity). Verification scores are calculated using data from eight dates among the ten, due to a NSSL/NWQ system outage that spanned two days (18-19 May 2009) and resulted incomplete radar mosaic data. More thorough post-season evaluation study will be performed in the summer and fall following the completion of the 2009 Spring Experiment.

In order to examine multimodel impact on the storm-scale ensemble system and its performance in QPF, all analyses are performed over three sub-groups of ensembles: The 10 member ARW member group; The 8 member NMM member group; And the sum of all 20 members (ALL).

[2] http://www.nmq.nssl.noaa.gov/

4.1 Ensemble spread

Figure 5 shows the domain-mean ensemble spread (defined as standard deviation against ensemble mean) of selected weather fields, averaged over 10 complete forecast dates. For the mean sea level pressure and 500 hPa geopotential height, both ARW and NMM ensembles have comparable spread level though ARW has slightly more dispersion. However, the spread from the full ensemble (ALL) are significantly higher from the start of the forecast (Fig. 5a,b), indicating difference between the two WRF cores in initial mass-related fields. The addition of 2 ARPS members in the full ensemble is not the contributor to the large ALL spread since removing ARPS members from ALL does not change the result (not shown). The full ensemble also shows significantly higher spread than individual sub-ensembles for 2 m temperature (Fig. 5c). The spread for the hourly accumulated precipitation (Fig. 5d) exhibit highest values from NMM ensemble and lowest values for ARW ensemble, while the full ensemble lies in between. With a small sample base, the diurnal pattern for the hourly accumulated precipitation is still evident but the morning low is less clear than in previous years (Kong et al. 2007, 2008) .

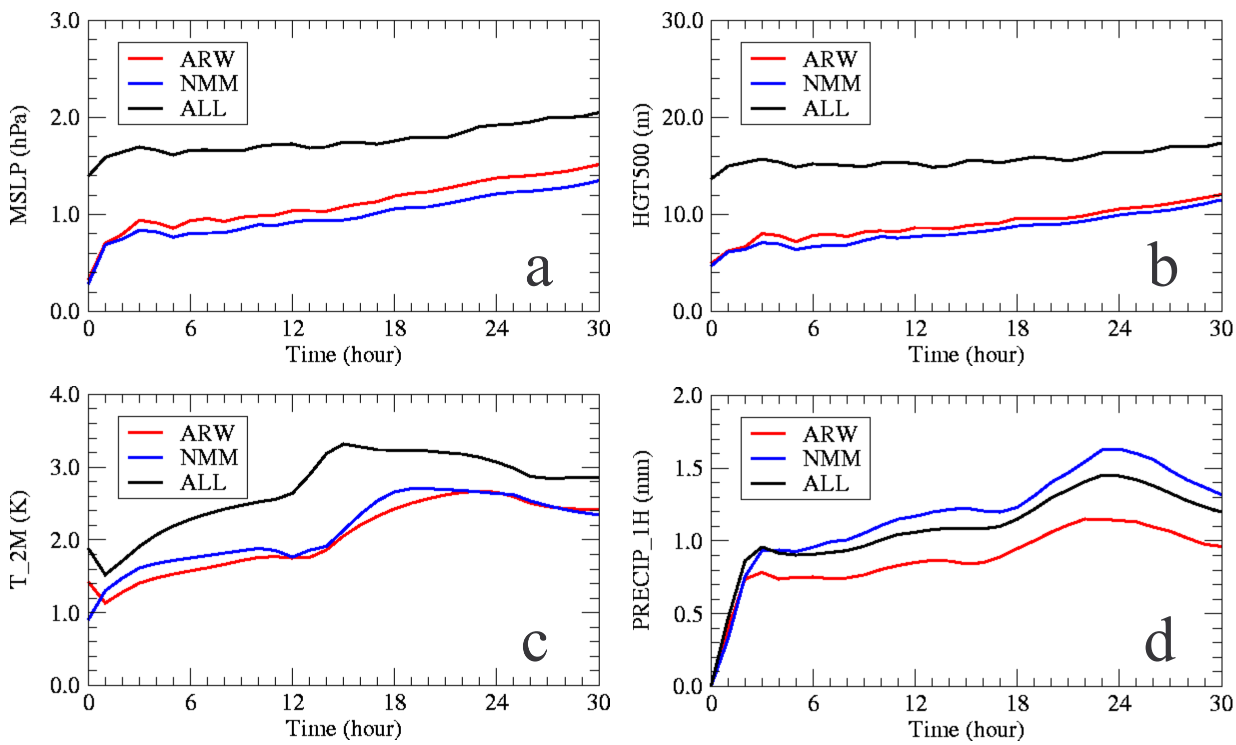

Figure 5. Domain-mean ensemble spread of (a) mean sea level pressure, (b) 500 hPa geopotential height, (c) 2 m temperature, and (d) 1 h accumulated precipitation, from three ensemble sub-groups, averaged over ten case dates.

4.2 BIAS score

Figure 6 shows the BIAS scores of 1 h accumulated precipitation exceeding 0.01 and 0.1 in (0.254 and 2.54 mm, respectively) for individual members of the full ensemble system, differentiated by contour styles into sub-ensemble groups. BIAS scores of ensemble mean and PM from the full ensemble are also shown. For higher thresholds, the small sample dataset makes interpretation of BIAS scores less meaningful (not shown). Model-wise, the NMM ensemble members have higher BIAS scores than ARW members, while ARPS members have lower values. The ensemble mean has remarkably larger BIAS than any individual member for the light rain threshold. The ensemble PM just lie in the middle. The three initially low value curves are from the three no-radar members (c0), one from each model group.

The BIAS scores for the forecasted composite reflectivity ≥ 20 and 30 dBZ are shown in Figure 7. While ARPS members are clearly exhibiting better scores (close to 1) than the rest, ARW and NMM members are more mixed. Again, ensemble PM has BIAS scores in the middle of all individual members. For the given thresholds (and higher, not shown), however, ensemble mean shows under-prediction to the composite reflectivity field.

Figure 6 and 7 also indicate that the ensemble system over-forecast more in composite reflectivity than in 1 h accumulated precipitation with respect to BIAS score.

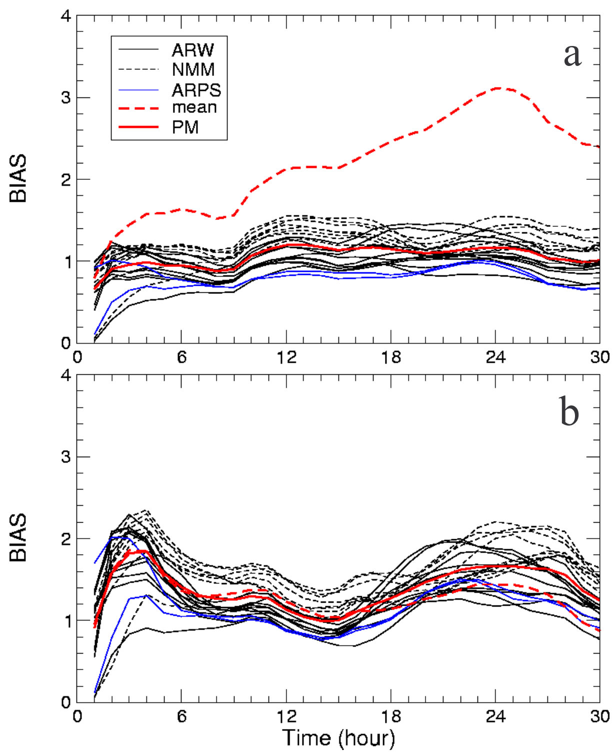

Figure 6. BIAS scores of 1 h accumulated precipitation ≥ 0.01 in (a) and 0.1 in (b) averaged over eight dates with complete forecast and verification dataset. The BIAS of ensemble mean and PM are calculated from the full ensemble system.

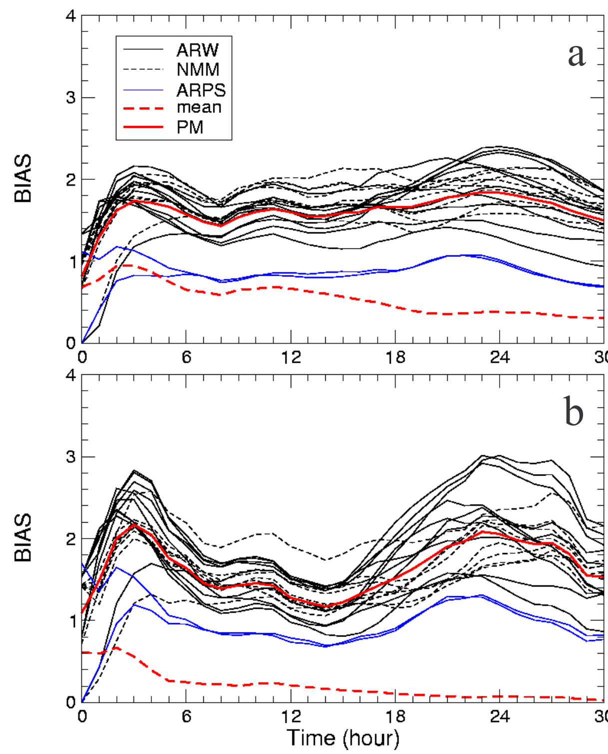

Figure 7. BIAS scores of composite reflectivity ≥ 20 dBZ (a) and 30 dBZ (b) averaged over eight dates with complete forecast and verification dataset. The BIAS of ensemble mean and PM are calculated from the full ensemble system.

4.3 ETS scores

The ETS scores are presented for the thresholds of 1 h accumulated precipitation ≥ 0.01 in and 0.1 in for all members, the ensemble mean and PM (Figure 8). As demonstrated in 2008 season (Kong et al. 2008), Figure 8 shows that the inclusion of radar data helps boost the ETS scores for the initial hours. Among model groups, ARPS member *cn* outscores any other model members over the first 18 h for the light rain threshold (Fig. 8a). ARW and NMM groups do not show which is better in ETS scores.

Both the ensemble mean and PM in Figure 8 generally outscore all individual members. Given the fact that PM has better forecast value than ensemble mean with respect to high-resolution QPF, the higher ETS scores add to PM's advantage.

ETS scores of composite reflectivity have similar results (not shown).

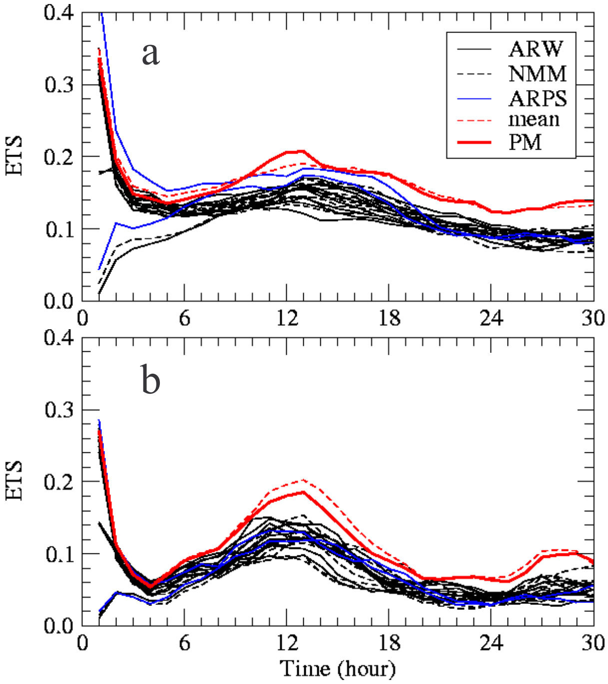

Figure 8. ETS of 1 h accumulated precipitation ≥ 0.01 in (a) and 0.1 in (b) averaged over eight forecast dates with complete dataset. The ETS of ensemble mean and PM are calculated from the full ensemble.

4.4 Reliability of PQPF

Figure 9 shows a reliability diagram for the 12 h probabilistic QPF (un-calibrated PQPF) of 1 h accumulated precipitation ≥ 0.1 in. Reliability curves from three sub-ensemble groups are presented. The ARW ensemble has more reliable PQPF than NMM ensemble. The full ensemble system improves the reliability to some degree.

5. DISCUSSION AND FUTURE WORKS

This extended abstract only presents some very preliminary analysis/assessment of the expanded storm-scale ensemble forecast system using a small sample dataset from the 2009 real-time Spring Experiment. Results may change with a larger sample base. More thorough analysis will be performed in the summer and fall following the completion of the Experiment.

As a future focus, we also plan to further examine various post-processing techniques including developing and applying proper bias correction approaches, and apply to 3 h accumulation or longer period.

6. ACKNOWLEDGMENTS

This work is mainly funded by a grant to CAPS from the NOAA CSTAR program. Supplementary support is also provided by the NSF ITR project LEAD (ATM-0331594), and by NSSL. The numerical forecasting of ARW and NMM ensembles is performed at the Pittsburgh Supercomputing Center (PSC). David O'Neal of PSC provided enormous technical and logistic support for the experiment. ARPS forecasting members are produced at the National Institute of Computational Sciences (NICS) at the Oakridge National Laboratory /University of Tennessee (ORNL/UT). The authors also thank Drs. Jimy Dudhia, Morris Weisman, Greg Thompson and Wei Wang for their very helpful suggestions and input on the WRF model configurations.

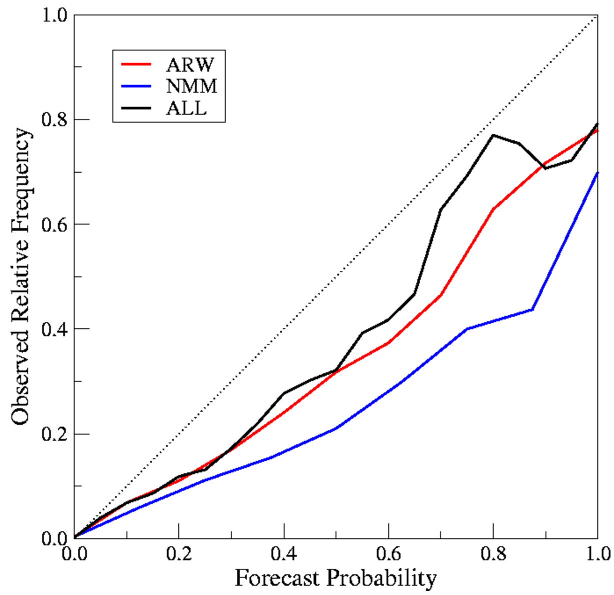

Figure 9. Reliability diagrams for the 1 h accumulated precipitation ≥ 0.1 in at 12 h lead time.

REFERENCES

Clark, A. J., W. A. Gallus, M. Xue, and F. Kong, 2009: A comparison of precipitation forecast skill between small near-convection-permitting and large convection-parameterizing ensembles. *Weather and Forecasting, accepted*.

Ebert, E. E., 2001: Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Mon. Wea. Rev.*, **129**, 2461-2480.

Kong, F., M. Xue, K. K. Droegemeier, D. Bright, M. C. Coniglio, K. W. Thomas, Y. Wang, D. Weber, J. S. Kain, S. J. Weiss, and J. Du,, 2007: Preliminary analysis on the real-time storm-scale ensemble forecasts produced as a part of the NOAA Hazardous Weather Testbed 2007 Spring Experiment. *Preprints, 22th Conf. on Weather Analysis and Forecasting and 18th Conf. on Numerical Weather Prediction* Amer. Meteor. Soc., Park City, UT, 3B.2.

Kong, F., M. Xue, K. K. Droegemeier, K. W. Thomas, Y. Wang, J. S. Kain, S. J. Weiss, D. Bright, and J. Du , 2008: Real-time storm-scale ensemble forecast experiment. *9th WRF User's Workshop*, NCAR

Center Green Campus, 23-27 June 2008, Paper 7.3.

Kong, F., M. Xue, M. Xue, K. K. Droegemeier, K. W. Thomas, Y. Wang, J. S. Kain, S. J. Weiss, D. Bright, and J. Du, 2008: Real-time storm-scale ensemble forecast experiment - Analysis of 2008 spring experiment data. *24th Conf. Several Local Storms*, Savannah, GA, Ameri. Meteor. Soc., Paper 12.3.

Schwartz, C. S., J. S. Kain, S. J. Weiss, M. Xue, D. R. Bright, F. Kong, K. W.Thomas, J. J. Levit, M. C. Coniglio, and M. S. Wandishin, 2009: Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Wea Forcasting*, Accepted

Xue, M., F. Kong, D. Weber, K. W. Thomas, Y. Wang, K. Brewster, K. K. Droegemeier, J. S. K. S. J. Weiss, D. R. Bright, M. S. Wandishin, M. C. Coniglio, and J. Du , 2007: CAPS realtime storm-scale ensemble and high-resolution forecasts as part of the NOAA hazardous weather testbed 2007 spring experiment. *22nd Conf. Wea. Anal. Forecasting/18th Conf. Num. Wea. Prediction*., Amer. Meteor. Soc., Park City, UT, Paper 3B.1.

Xue, M., F. Kong, K. W. Thomas, J. Gao, Y. Wang, K. Brewster, K. K. Droegemeier, J. Kain, S. Weiss, D. Bright, M. Coniglio, and J. Du, 2008: CAPS Realtime Storm-scale Ensemble and High-resolution Forecasts as Part of the NOAA Hazardous Weather Testbed 2008 Spring Experiment. *Preprints*, *24th Conf. on Severe Local Storm*, Amer. Meteor. Soc., Savannah, GA, Paper 12.2.

Xue, M., F. Kong, K. W. Thomas, J. Gao, Y. Wang, K. Brewster, K. K. Droegemeier, X. Wang J. Kain, S. Weiss, D. Bright, M. Coniglio, and J. Du, 2009: CAPS realtime multi-model convection-allowing ensemble and 1-km convection-resolving forecasts for the NOAA Hazardous Weather Testbed 2009 Spring Experiment. *Preprints*, *23th Conf. on Weather Analysis and Forecast/1 th Conf. on Numerical Weather Prediction*, Amer. Meteor. Soc., Savannah, Omeha, NE, Paper 16A.2.