# 13B.3 Evaluation of the Performance and Distribution of Hourly Maximum Fields from Storm-Scale Ensemble Forecasts

Israel L. Jirak[1*], Steven J. Weiss[1], Christopher J. Melick[1], Patrick T. Marsh[2,3], John S. Kain[2], Adam J. Clark[2], Ming Xue[4], Fanyou Kong[4], and Kevin W. Thomas[4]

[1]NOAA/NWS/NCEP/Storm Prediction Center
[2]NOAA/OAR/National Severe Storms Laboratory
[3]School of Meteorology, University of Oklahoma
[4]Center for Analysis and Prediction of Storms, University of Oklahoma
Norman, OK

## 1. INTRODUCTION

Since 2007, the NOAA Hazardous Weather Testbed Spring Experiment has featured data from an ensemble of convection-allowing numerical models (Xue et al. 2007; Xue et al. 2008; Xue et al. 2009). The 2010 storm-scale ensemble forecast (SSEF) system is a multi-model, multi-analysis, and multi-physics 26-member ensemble produced by the University of Oklahoma (OU) Center for Analysis and Prediction of Storms (CAPS). While a large number of members with varied configurations and initializations is a desirable trait of an ensemble system, the abundance of data makes the interpretation and understanding of ensemble characteristics quite challenging.

The objective of this study is to examine the characteristics and performance of individual SSEF members in simulating convective storm attribute fields related to storm intensity and potential severe storm occurrence. The attribute fields are investigated to assess the relative role of each member in contributing to SSEF statistical results, and to assess the relationship between member-generated fields and observed severe weather. Ultimately, a better understanding of the distribution and performance characteristics of the 2010 SSEF system will aid in interpretation and configuration of future convection-allowing ensemble systems.

## 2. SSEF OVERVIEW

The 2010 Spring Experiment took place over a five-week period from 17 May – 18 June (see Weiss et al. 2010 for more details). From Monday through Friday of every week during this period, each SSEF member generated a 30-h forecast initialized at 00 UTC for a total of 25 runs. Each model run covered the CONUS with 4-km horizontal grid spacing. Owing to file size considerations and a general paucity of severe weather over the Western U.S. during this time of year, a subset of the full domain that included the eastern two-thirds of the U.S. was extracted for analysis of severe weather.

* *Corresponding author address*: Israel L. Jirak, NOAA/NWS/NCEP/Storm Prediction Center, 120 David L. Boren Blvd., Norman, OK 73072; e-mail: Israel.Jirak@noaa.gov

During the Spring Experiment, CAPS produced post-processed ensemble fields (e.g., means, exceedance probabilities, maxima, etc.) from a subset of 15 members with the most varied perturbations. These members (minus the ARPS control member) were the focus for this study. The members examined in this study come from two dynamic cores (i.e., 4 WRF-NMM and 10 WRF-ARW members) and have a variety of initial condition and physics perturbations (Table 1).

| Member | IC | MP | LSM | PBL |
|---|---|---|---|---|
| **nmm_cn** | 00Z ARPSa | Ferrier | Noah | MYJ |
| **nmm_m3** | nmm_cn + nmm-n1 | Thompson | Noah | MYJ |
| **nmm_m4** | nmm_cn + nmm-n2 | WSM6 | RUC | MYJ |
| **nmm_m5** | nmm_cn + em-n1 | Ferrier | RUC | MYJ |
| **arw_cn** | 00Z ARPSa | Thompson | Noah | MYJ |
| **arw_m6** | arw_cn + em-p1 | Morrison | RUC | YSU |
| **arw_m7** | arw_cn + em-p2 | Thompson | Noah | QNSE |
| **arw_m8** | arw_cn – nmm-p1 | WSM6 | RUC | QNSE |
| **arw_m9** | arw_cn + nmm-p2 | WDM6 | Noah | MYNN |
| **arw_m10** | arw_cn + rsmSAS-n1 | Ferrier | RUC | YSU |
| **arw_m11** | arw_cn – etaKF-n1 | Ferrier | Noah | YSU |
| **arw_m12** | arw_cn + etaKF-p1 | WDM6 | RUC | QNSE |
| **arw_m13** | arw_cn – etaBMJ-n1 | WSM6 | Noah | MYNN |
| **arw_m14** | arw_cn + etaBMJ-p1 | Thompson | RUC | MYNN |

**Table 1. Configurations of initial conditions (IC), microphysics schemes (MP), land surface models (LSM), and planetary boundary layer (PBL) schemes for SSEF members examined.**

Given the explicit representation of convective processes in convection-allowing models, traditional model output every 3-h or even every hour is not sufficient to sample the rapidly evolving model-generated storms. A method of obtaining storm attribute information at higher temporal resolution is to track the maximum value of a certain field at each grid point and every model time step and then output these "hourly maximum fields" (HMFs; Kain et al. 2010). Six different HMFs are examined for their association with severe convective weather: updraft helicity (Kain et al. 2008) which represents a rotating updraft (i.e., supercell thunderstorm) in a simulated storm, updraft speed, downdraft speed, 10-m AGL wind speed to identify convective wind gusts, 1-km AGL simulated reflectivity, and vertically integrated graupel, as a proxy for hail in a model storm.

Previous studies (e.g., Done et al. 2004; Weisman et al. 2008) have shown the ability of convection-allowing models to provide some information about convective mode. On occasion, HMFs can offer a unique perspective on convective mode and simulated storm tracks. For example, discrete supercells formed in the Texas and Oklahoma panhandles on the afternoon/evening of 18 May 2010, producing several reports of tornadoes and large hail. Inspection of the hourly maximum updraft helicity and simulated reflectivity for that day (Fig. 1) suggests that storms would likely be rotating and remain discrete, as evidenced by the long, isolated swaths of updraft helicity co-located with swaths of simulated reflectivity. Although the exact location of the storms may not be correctly predicted, the HMFs highlighted the occurrence of discrete simulated storms. It is recommended, however, that the HMF fields be viewed together with instantaneous simulated reflectivity to more completely ascertain aspects of convective mode.

## 3. METHODOLOGY

The HMFs were compared among the 14 SSEF members listed in Table 1 from forecast hour 12 (i.e., 12 UTC on Day 1) to forecast hour 30 (i.e., 06 UTC on Day 2). The first 12 hours were excluded from this analysis to focus on new convective development in the afternoon. The data were analyzed using three primary methods: distribution of domain-wide statistics, contribution of each member to the ensemble maximum, and verification near severe storm reports.

In calculating the domain-wide statistics, the maximum value of each hourly maximum field was determined for every grid point in each model run (i.e., 12Z through 06Z) for every member. Thus, the member average is simply the sum of the grid-point maxima for a given model run divided by the total number of grid points while the maximum is the largest value found across the domain. Since the HMFs are primarily descriptors of storm attributes, most of the grid points have undefined or very small values, hence the focus on grid-point maxima rather than averages.

The ensemble maximum is simply the largest value at a given grid point among all analyzed members. This ensemble statistic was one of the most commonly viewed plots of the HMFs during the Spring Experiment. This type of plot provides useful information about the areal spread of the ensemble in the placement of storms (Fig. 2). However, this plot alone does not provide information about which members are contributing to the ensemble maximum. The contribution (in terms of number of grid points) by each member to the ensemble maximum was tallied for every forecast hour and summed over the entire period to determine the overall frequency by member.

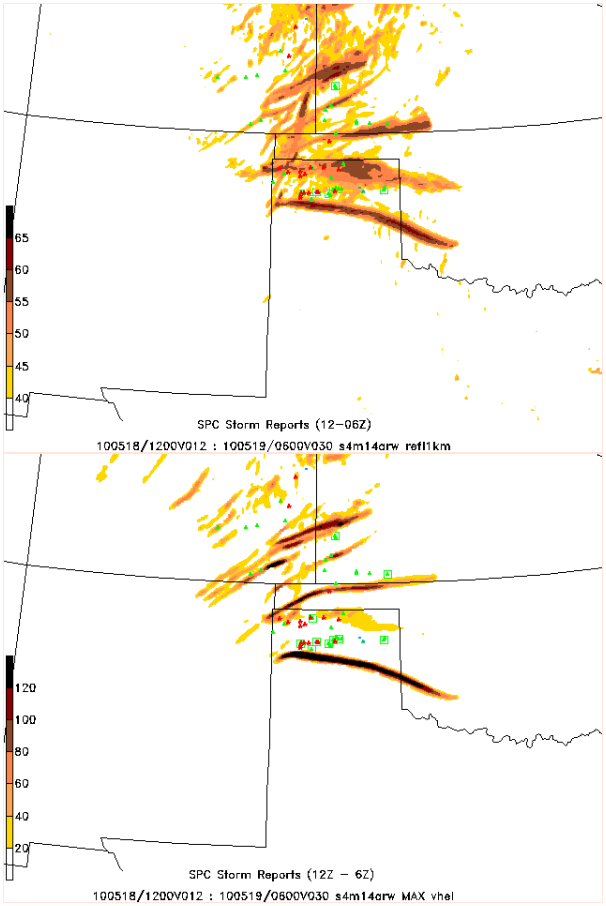



**Fig. 1: Maximum hourly maximum simulated reflectivity (dBZ; top) and updraft helicity ($m^2s^{-2}$; bottom) from 12 UTC on 18 May to 06 UTC on 19 May for SSEF member, s4m14, overlaid with preliminary storm reports (tornado – red, hail – green, box – significant severe).**

Finally, the members are evaluated as a predictor of severe weather by inspecting the distribution of the HMFs surrounding severe local storm reports in time and space. The SPC preliminary storm reports of tornadoes, severe wind, and large hail were used in this study. During the period studied, there were 307 tornado reports, 1392 hail reports [145 of which were significant (>2" diameter)], and 2441 wind reports [47 of which were significant (≥65 kts)]. For every severe storm report, the maximum value of each field was extracted within a ~25 mile box surrounding that report.

To allow for some model error in the timing of storm initiation, the maximum value within that box was also extracted from the hour before and after the hour during which the report occurred, resulting in a three-hour verification window. The distribution of these maximum values is compared among all members for each type of storm report.

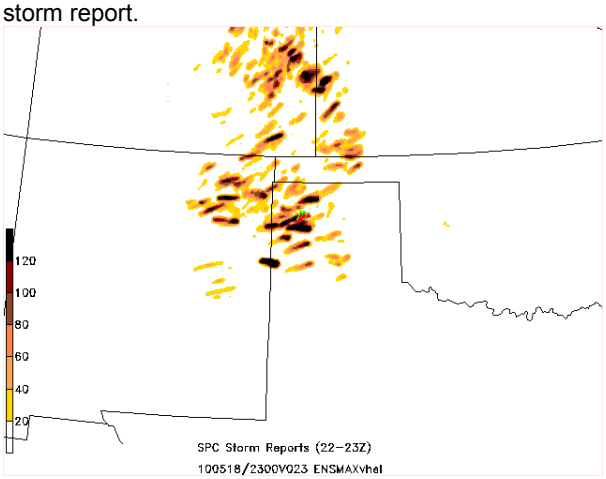



**Fig. 2: Ensemble maximum hourly maximum updraft helicity ($m^2s^{-2}$) at 23 UTC on 18 May overlaid with preliminary storm reports (tornado – red, hail – green, box – significant severe).**

## 4. RESULTS

### *4.1 Distribution of Domain-Wide Statistics*

The distribution of daily grid maximum updraft helicity (Fig. 3a) reveals that the majority of NMM and ARW members produce similar maximum values of updraft helicity centered around 150 $m^2s^{-2}$, with a few ARW members (i.e., s4m8, s4m12, and s4m13) producing larger values and NMM member, s4m3, generating lower values. The similarity in distributions between most NMM and ARW members, however, does not hold true for the other kinematic HMFs (Figs. 3b-3d). The NMM members clearly produce weaker maximum updraft speeds than the ARW members. The median updraft speed grid maximum (Fig. 3b) is ~7 $ms^{-1}$ larger for the ARW members compared to the NMM members. While the overall magnitude of downdrafts is less than updrafts, the ARW members also generally produce stronger downdrafts than the NMM members on the order of 2 $ms^{-1}$ (Fig. 3c). In terms of horizontal wind speed, the ARW members also generate stronger 10-m AGL winds than the NMM members (Fig. 3d). These results are consistent with anecdotal operational experience of SPC forecasters in using 4-km WRF-NMM and WRF-ARW deterministic model guidance, with NMM storms tending to exhibit broader, less well-defined convective structures and weaker attribute fields.

In comparing just the ARW members for these kinematic fields, a few trends stand out. Members s4m8 and s4m12 tend to produce larger magnitudes of these fields while members s4m6, s4m10, and s4m11 tend to produce smaller magnitudes. This difference in magnitudes appears to be linked to the PBL scheme utilized. Members at the upper end employ the QNSE scheme while members at the lower end use the YSU scheme.

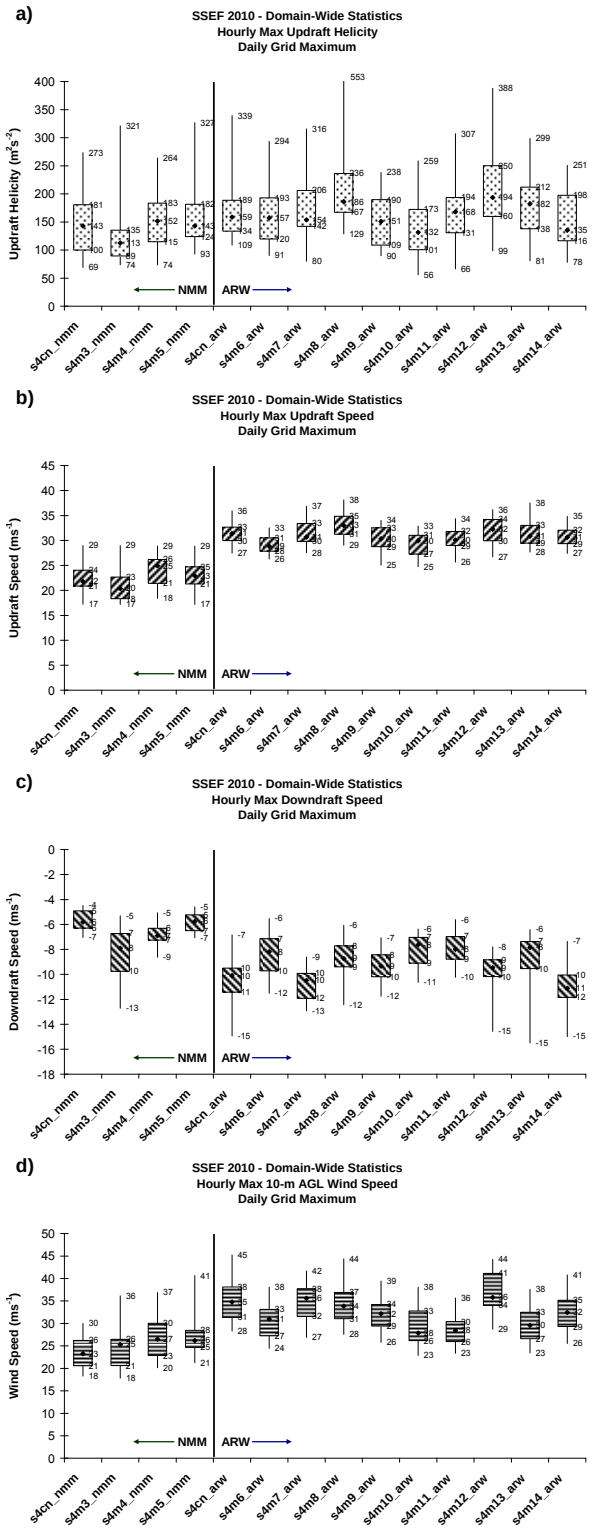



**Fig. 3: Box plots of domain-wide daily grid maxima of each member over the 25-day Spring Experiment for a) updraft helicity, b) updraft speed, c) downdraft speed, and d) 10-m AGL wind speed.**

The remaining HMFs (i.e., simulated reflectivity and vertically integrated graupel) are strongly influenced by microphysical properties. For simulated reflectivity (Fig. 4a), there is not only a stratification between dynamic cores (i.e., NMM and ARW), but also a strong stratification among microphysics schemes. Two of the ARW members with double-moment schemes [Thompson (s4cn, s4m7, s4m14) and WDM6 (s4m9, s4m12)] produce much higher maximum reflectivity than the other members. There is less difference between NMM and ARW members for vertically integrated graupel (Fig. 4b); however, the Thompson and WDM6 ARW members generally show higher concentrations of graupel than the other members.

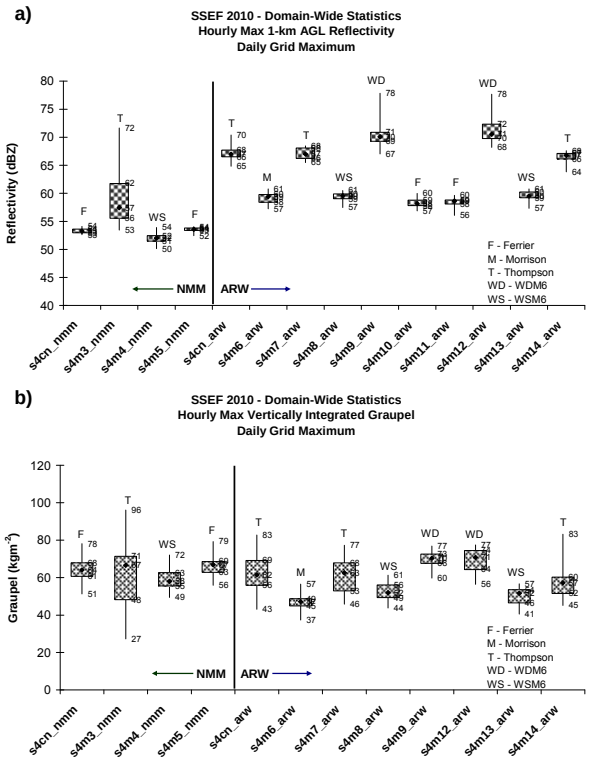



**Fig. 4: Box plots of domain-wide daily grid maxima of each member over the 25-day Spring Experiment for a) 1-km AGL simulated reflectivity, b) vertically integrated graupel. Data (large ice) not available for Ferrier members, s4m10 and s4m11.**

### *4.2 Contribution to Ensemble Maximum*

Based on the domain-wide statistics, the ARW members would generally be expected to make the largest contributions to the ensemble maximum, which is confirmed in Fig. 5. For all fields except updraft helicity (Fig. 5a) and vertically integrated graupel (Fig. 5f), the NMM members are the smallest contributors to the ensemble maximum. Even the larger contributions to vertically integrated graupel from NMM members s4cn and s4m5 may be questionable, as large ice from the Ferrier microphysics scheme was substituted for graupel. Updraft helicity does appear to be evenly distributed among all 14 members examined even though the NMM members generate weaker updrafts.

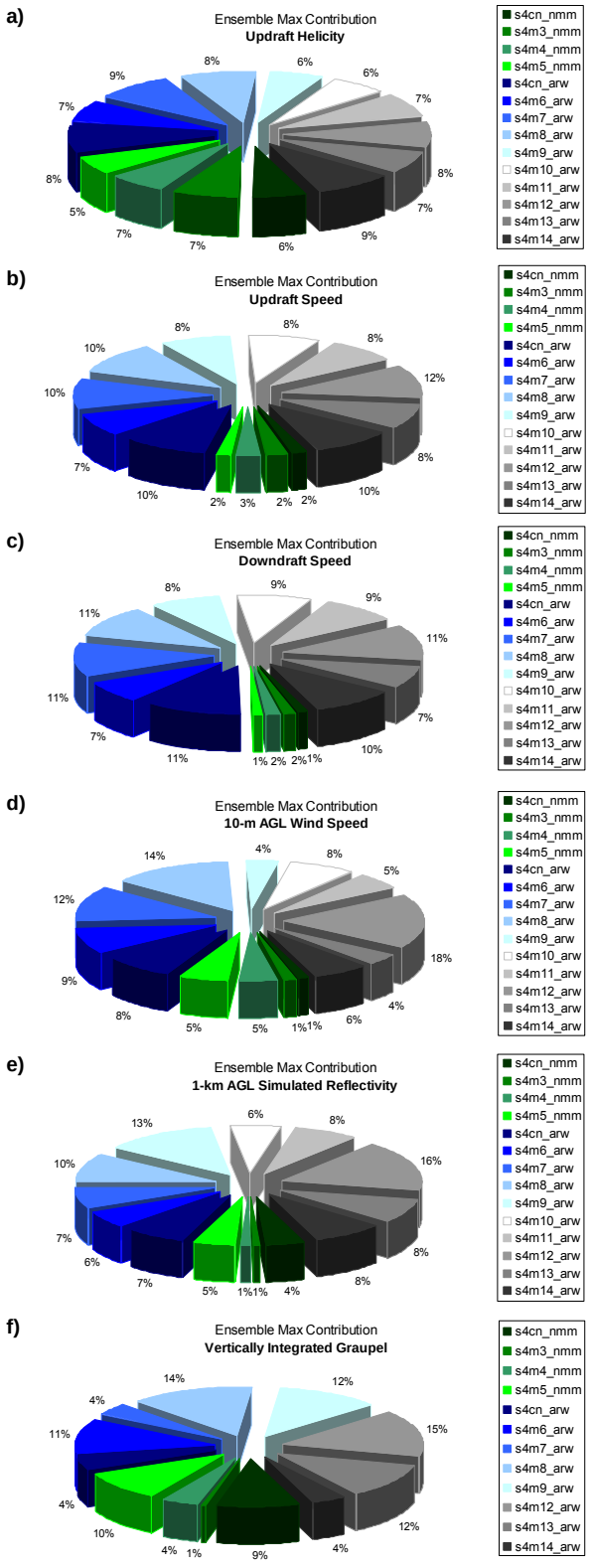



**Figure 5: Pie charts showing the contribution of each member to the ensemble maximum throughout the Spring Experiment for a) updraft helicity, b) updraft speed, c) downdraft speed, d) 10-m wind speed, e) 1-km AGL simulated reflectivity, and f) vertically integrated graupel.**

Other results from the domain-wide statistics are also confirmed. The ARW QNSE members (s4m7, s4m8, and s4m12) are the largest contributors to the ensemble maximum for updraft, downdraft and 10-m wind speed. For simulated reflectivity (Fig. 5e), the WDM6 members (s4m9 and s4m12) dominate with a combined contribution of ~30% of the ensemble maximum grid points. Finally, the members utilizing the Thompson microphysics scheme (NMM_s4m3, ARW_s4cn, ARW_s4m7, and ARW_s4m14) tend to contribute less than expected to the ensemble maximum of vertically integrated graupel (Fig. 5f).

### *4.3 Verification near Severe Storm Reports*

To evaluate the ability of each member to generate storms with severe characteristics, the HMF values are verified against observed severe storm reports. In the vicinity of tornado (Figs. 6 and 7), hail (not shown), and wind (not shown) reports, the ARW members produce larger values of the HMFs than the NMM members. As the most evenly distributed field among NMM and ARW members, updraft helicity (Fig. 6a) provides the most meaningful comparison. Although the maximum and upper quartile values of updraft helicity are similar among the members, the median and lower quartile values of all 10 ARW members are larger (means larger at 99% confidence level, except for s4m6, s4m10, & s4m11) than those for any NMM member. As expected from previous results for the other kinematic fields (Figs. 6b-d), the ARW member generate stronger vertical and horizontal flow fields in the neighborhoods surrounding storm reports.

For the microphysical hourly max fields (Fig. 7), the results are similar to those found previously. The NMM members generate lower reflectivity fields than the ARW members in the vicinity of storm reports (Fig. 7a). In particular, NMM members s4m3 and s4m4 are not skillful in producing storms with high reflectivity, given a third quartile value of 0. An obvious stratification by microphysics scheme is again visible among the ARW members. For vertically integrated graupel (Fig. 7b), the WDM6 ARW members generate the largest values.

With performance data for each HMF and each type of storm report, comparisons can also be made to determine whether an hourly maximum field provides indication of a specific threat of severe weather. For the ARW control member, updraft helicity (Fig. 8a) tends to be slightly higher for more extreme events (i.e., tornadoes, significant hail, and significant wind) than for non-significant severe hail and wind events. The lower quartile of updraft speed (Fig. 8b) is larger for tornado and significant hail events compared to other severe events. Downdraft speed (Fig. 8c) and reflectivity (Fig. 9a) do not show any significant delineation among severe type. The 10-m AGL wind (Fig. 8d) reveals higher values for significant severe wind events compared to non-significant severe wind events. Finally, vertically integrated graupel (Fig. 9b), which was examined as a predictor for hail, does show the largest values for significant severe hail reports. Overall, however, there is considerable overlap among the HMF distributions for different severe storm types, which suggests that individual HMF fields by themselves cannot be used to effectively discriminate between event types.

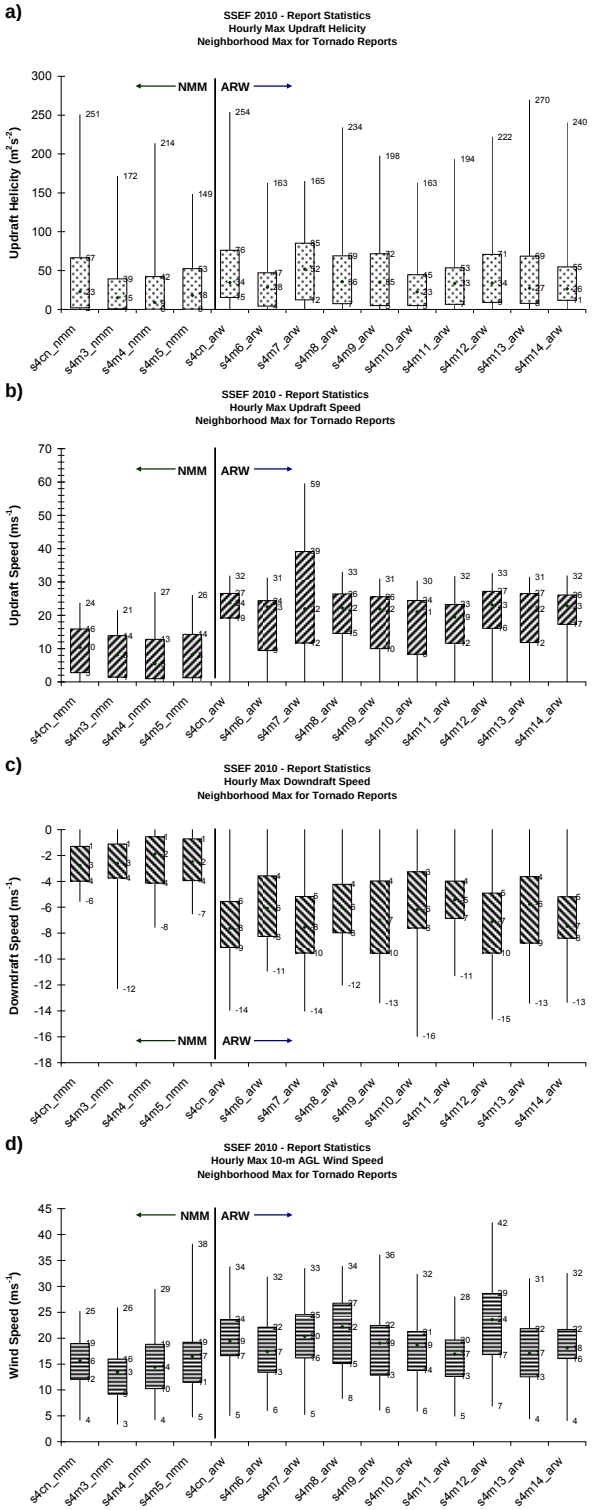



**Fig. 6: Box plots for each member of maxima within 25-mile and +/- 1-hour neighborhood of tornado reports during the Spring Experiment for a) updraft helicity, b) updraft speed, c) downdraft speed, and d) 10-m AGL wind speed.**

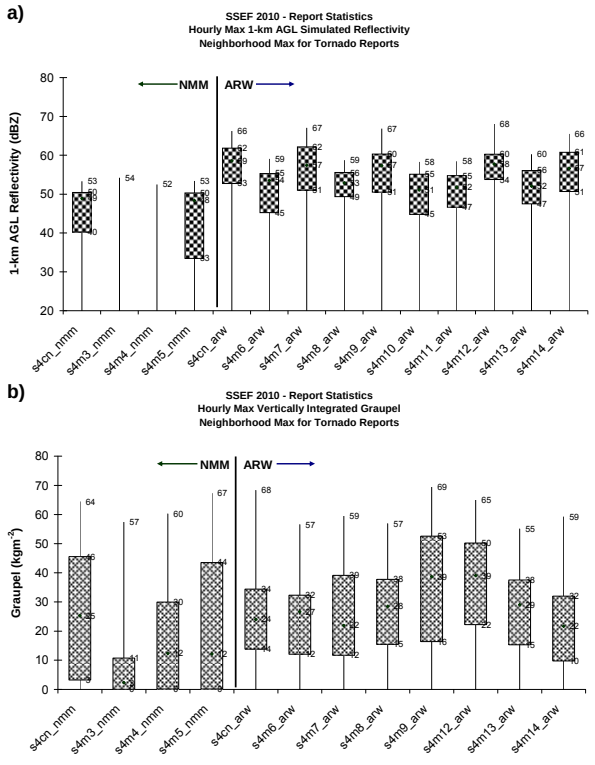
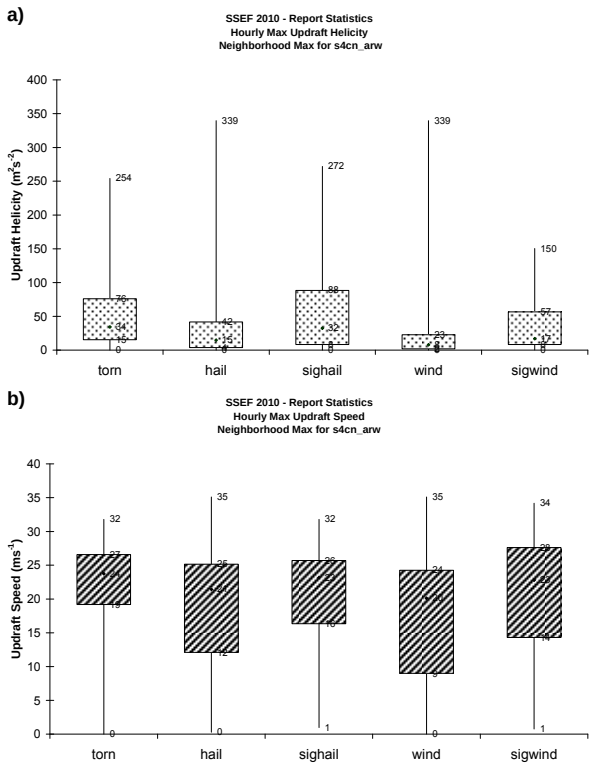



**Fig. 7: Box plots for each member of maxima within 25-mile and +/- 1-hour neighborhood of tornado reports during the Spring Experiment for a) 1-km AGL simulated reflectivity, b) vertically integrated graupel.**



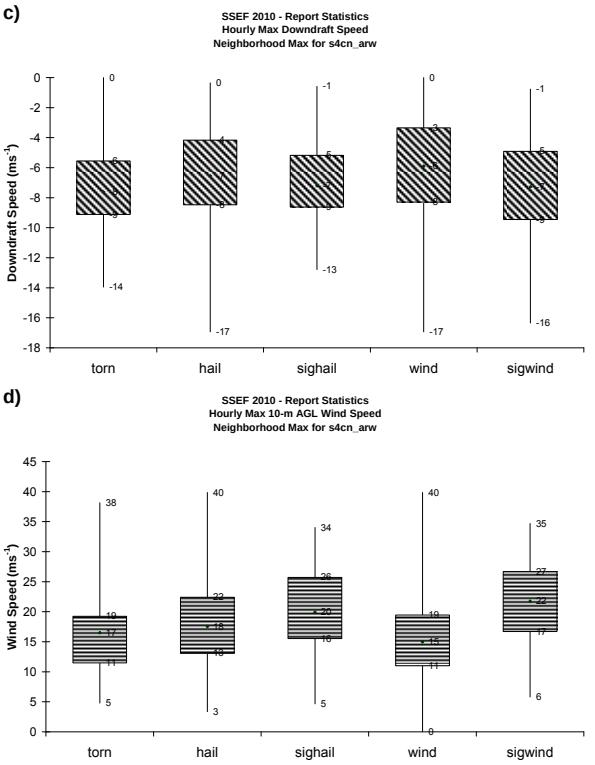




**Fig. 8: Box plots for the ARW control member, s4cn, of maxima within 25-mile and +/- 1-hour neighborhood of the different storm reports during the Spring Experiment for a) updraft helicity, b) updraft speed, c) downdraft speed, and d) 10-m AGL wind speed.**

### *4.4 Example*

An example is provided in Fig. 10 to compare HMFs from two different SSEF members. The members chosen for this comparison include one of the configurations, ARW_s4cn, that tended to produce higher HMF values and one of the members that produced lower HMF values (NMM_s4m3). These members have different dynamic cores, but utilize the same microphysics, PBL, long-wave radiation, and land surface schemes.

The updraft helicity fields (Figs. 10a-b) look qualitatively similar in regions of simulated storms. A close examination reveals that the ARW member produces swaths of updraft helicity in eastern North Dakota near several tornado reports while the NMM member does not generate rotating storms in that area. The NMM member does appear to produce higher values of updraft helicity near the tornadoes reported in southern Minnesota than the ARW member.

While the updraft helicity fields appear similar between the models, the updraft (Figs. 10c-d) and downdraft (Figs. 10e-f) fields look very different, as might be suspected from results presented previously. The ARW member produces much stronger updrafts and downdrafts over a much larger area than the NMM member. The ARW member even generates very strong updrafts for the isolated severe convection in the Texas panhandle.

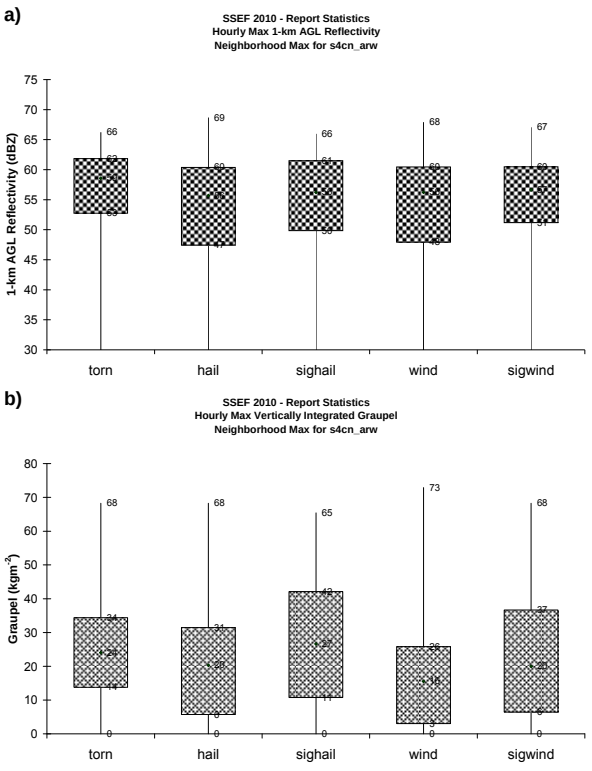



**Fig. 9: Box plots for the ARW control member, s4cn, of maxima within 25-mile and +/- 1-hour neighborhood of the different storm reports during the Spring Experiment for a) 1-km AGL simulated reflectivity, b) vertically integrated graupel.**

## 5. SUMMARY AND CONCLUSIONS

The 26-member SSEF system produced for the 2010 Spring Experiment provided an abundance of high-resolution model data. A unique method of examining the data was to output hourly maximum values of selected storm-attribute fields. These HMFs were examined for 14 members with mixed IC-physics perturbations to determine their statistical distribution and performance characteristics in simulating storm attributes.

Examination of the HMFs for individual members revealed noticeable differences between members with different dynamic cores. The ARW members typically generated larger values for the HMFs than the NMM members. The only notable exception was updraft helicity; however, the ARW members still showed a higher median value of updraft helicity in the vicinity of severe storm reports. Such a systematic difference in field values based on the WRF dynamic core can impact the interpretation of ensemble output. For example, the NMM members were underrepresented for most of the HMFs with regard to the ensemble maximum, which was one of the most commonly viewed ensemble plots during the Spring Experiment. Although not explored directly here, the probability of exceedance would also be impacted by a systematic difference in values. Choosing a static exceedance threshold (e.g., updraft >15 $ms^{-1}$) for all members will impact the result if a subset of members is less likely to exceed that value (NMM members for this example).

Among the ARW members, the differences were more subtle. For the kinematic HMFs (i.e., updraft, downdraft, and 10-m wind speed), the differences often appeared to be linked to the PBL scheme utilized. The members employing the YSU scheme tended to generate weaker kinematic fields, resulting in lower HMF values near storm reports. For the microphysical HMFs (i.e., simulated reflectivity and vertically integrated graupel), the differences were more pronounced and strongly stratified by the microphysics scheme utilized.

With many members comprised of various physics schemes, storm-scale ensembles require considerable analysis to understand their statistical performance. Attempting to understand some of the characteristics of individual members has been shown to better clarify some of the unknowns. Hopefully, the results presented in this study will aid in the interpretation and configuration of future convection-allowing ensemble systems.

*Acknowledgements.* This research was supported by an allocation of advanced computing resources provided by the National Science Foundation. The computations were performed on Athena (a Cray XT4) at the National Institute for Computational Science (NICS; http://www.nics.tennessee.edu/).

## REFERENCES

Done, J., C. Davis, and M. Weisman, 2004: The next generation of NWP: Explicit forecasts of convection using the Weather Research and Forecasting (WRF) model. *Atmos. Sci. Lett.*, **5** (6), 110–117.

Kain, J. S., S. J. Weiss, D. R. Bright, M. E. Baldwin, J. J. Levit, G. W. Carbin, C. S. Schwartz, M. L. Weisman, K. K. Droegemeier, D. B. Weber, K. W. Thomas, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting*, **23**, 931-952.

Kain, J. S., S. R. Dembek, S. J. Weiss, J. L. Case, J. J. Levit, and R. A. Sobash, 2010: Extracting unique information from high resolution forecast models: Monitoring selected fields and phenomena every time step. *Wea. Forecasting*, in press.

Weisman, M. L., C. Davis, W. Wang, K. W. Manning, and J. B. Klemp, 2008: Experiences with 0–36-h explicit convective forecasts with the WRF-ARW model. *Wea. Forecasting*, **23**, 407–437.

Weiss, S. J., A. J. Clark, I. L. Jirak, C. J. Melick, C. W. Siewert, R. Sobash, P. T. Marsh, A. R. Dean, M. Xue, F. Kong, K. W. Thomas, J. Du, D. R. Novak, F. E. Barthold, M. J. Bodner, J. J. Levit, C. B. Entwistle, T. Jensen, J. S. Kain, M. C. Coniglio, and R. S. Schneider, 2010: An overview of the 2010 NOAA Hazardous Weather Testbed spring forecasting experiment. *25th Conf. Severe Local Storms*, Denver, CO, Ameri. Meteor. Soc., Paper 7B.1.

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. Pred.*, Amer. Meteor. Soc., 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. *24th Conf. Severe Local Storms*, Savannah, GA, Ameri. Meteor. Soc., 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 4-km multi-model convection-allowing ensemble and 1-km convection-resolving forecasts for the NOAA Hazardous Weather Testbed 2009 Spring Experiment. *23rd Wea. Analysis and Forecasting*, Omaha, NE, Ameri. Meteor. Soc., Paper 16A.2.

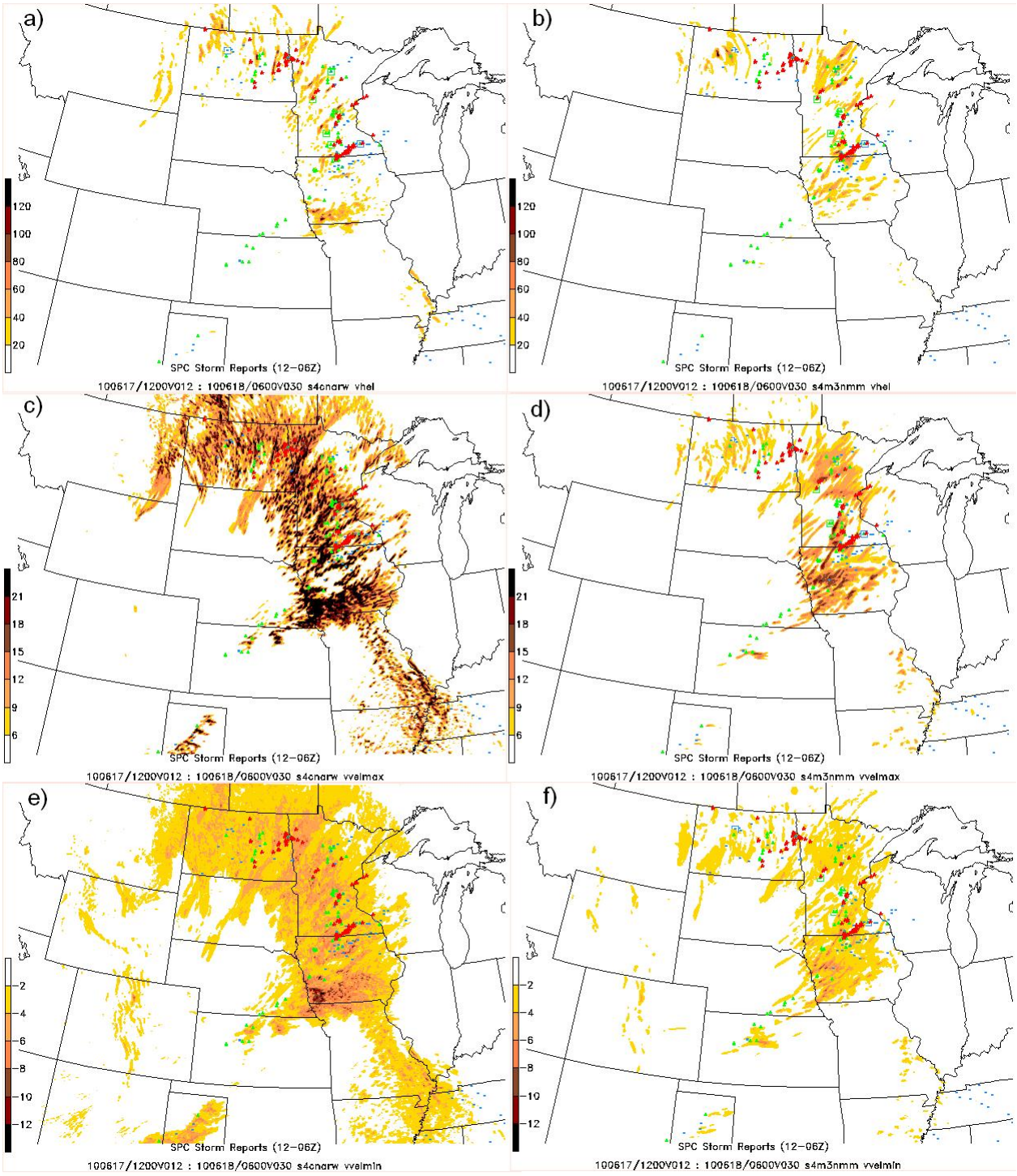



**Fig. 10: Maximum hourly maximum updraft helicity ($m^2s^{-2}$; a,b), updraft speed ($ms^{-1}$; c,d), and downdraft speed ($ms^{-1}$; e,f) from 12 UTC on 17 June to 06 UTC on 18 June 2010 for ARW_s4cn (left panels) and NMM_s4m3 (right panels) overlaid with preliminary storm reports (tornado – red, hail – green, wind – blue, box – significant severe) .**