16B.2 ASSESSING THE SKILL OF CONVECTION-ALLOWING ENSEMBLE FORECASTS OF SEVERE MCS WINDS FROM THE SSEO

Robert M. Hepper[1,2*], Israel L. Jirak[1], and Jeffrey M. Milne[1,2]

[1]NOAA/NWS/NCEP/Storm Prediction Center, Norman, OK

[2]CIMMS, University of Oklahoma, Norman, OK

1. INTRODUCTION

The development and usage of convection-allowing models (CAMs) over the past decade has resulted in a general improvement in convective forecasts, due to their ability to reasonably represent the location, mode, and evolution of convection (Done et al. 2004; Weisman et al. 2008). Significant questions remain, however, in the ability to diagnose the explicit threat of individual severe hazards (tornadoes, large hail, and damaging winds) from CAMs. Model diagnostics, such as updraft helicity (Kain et al. 2008) and various hail proxies (Adams-Selin and Ziegler 2016; Gagne et al. 2015), have been developed to assess the potential for tornadoes and large hail. Little work, however, has been done in developing techniques to diagnose severe winds from mesoscale convective systems (MCSs). As a result, forecasting severe MCSs currently relies on an examination of a combination of traditional environmental conditions favorable for MCSs, such as instability and vertical wind shear (Johns and Doswell 1992), alongside simulated radar and probabilistic 10-meter wind fields produced by CAM ensemble guidance. Verification of these CAM ensemble probabilistic 10-meter wind fields is therefore crucial.

This study will examine severe wind forecasts from the Storm Prediction Center (SPC) Storm-Scale Ensemble of Opportunity (SSEO; Jirak et al. 2012), a seven-member ensemble comprised of individual deterministic CAMs, including three WRF-ARW members and four WRF-NMM/NEMS-NMMB members (Table 1). Additionally, two of the seven members are the 12-hour time-lagged members of the NCEP HiRes Window runs. The objectives of this study include developing a technique for verifying probabilistic ensemble forecasts of severe MCS winds, providing guidance for the design of future CAM ensembles by examining the relative importance of ARW, NMM, NMMB, and time-lagged members to the ensemble through a stepwise removal of members and groups of members from the ensemble, and suggesting future paths for the improvement of ensemble forecasts with regard to severe MCSs. The following section will describe the verification methodology used in this study. Section 3 will show overall results and examine differences in ensemble performance based on the member or groups of members removed from the ensemble. Section 4 will provide two cases exemplifying current weakness of ensemble MCS forecasts and suggest paths for improvement. The final section will summarize the findings.

Updated 12 Aug 2014	Grid Spacing	Vert Levels	Fcst Length	ICs/ LBCs	PBL	Micro
NSSL WRF-ARW	4 km	35	36 h	NAM/ NAM	MYJ	WSM6
EMC HRW WRF-ARW	4.2 km	40	48 h	RAP/ GFS	YSU	WSM6
EMC HRW WRF-ARW; 12-h time lag	4.2 km	40	48 h	RAP/ GFS	YSU	WSM6
EMC HRW NMMB	3.6 km	40	48 h	RAP/ GFS	MYJ	Ferrier updated
EMC HRW NMMB; 12-h time lag	3.6 km	40	48 h	RAP/ GFS	MYJ	Ferrier updated
EMC CONUS WRF-NMM	4 km	35	36 h	NAM/ NAM	MYJ	Ferrier
EMC CONUS NAM NEST	4 km	60	60 h	NAM/ NAM	MYJ	Ferrier-Aligo

Table 1. Membership configuration of the SPC SSEO

2. VERIFICATION METHODOLOGY

Verification of severe wind forecasts is heavily complicated by the quality and coverage of observations available for verification, requiring the use of innovative techniques. Additionally, as this study is principally interested in examining forecasts of severe wind associated with organized MCSs, rigorous filtering of both the

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

forecast and observations fields is required to ensure association with an organized MCS.

The forecast field examined in this study is the 0000 UTC SSEO 24-hour neighborhood probability (i.e., forecast hours 12-36; valid 1200-1200 UTC) of 10-meter wind greater than 30 knots over a three year period from 2012 to 2014. The forecast field is generated by first objectively filtering individual deterministic member 10-meter wind fields through removal of any wind values which are not within 40 km of a simulated radar reflectivity object with a major axis length of at least 100 km. Ensemble neighborhood probabilities are then calculated with a radius of influence of 40 km and smoothed using a 2-D Gaussian kernel density estimation (Brooks et al. 1998) with a smoothing parameter of 120 km. An example of such a filtered forecast field from 24 June 2013 is shown in Fig. 1, with a general large swath of ≥30% probabilities across Nebraska, Iowa, and Illinois. Additional smaller areas of probabilities can be seen over the Oklahoma and Texas panhandles, as well as Montana.

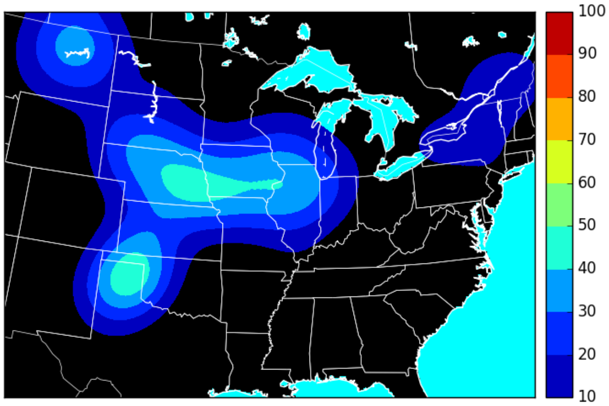

Figure 1. 24 Jun 2013 filtered SSEO 24-hour neighborhood probability (%) of 10-meter wind ≥ 30 knots.

Ensemble forecasts are verified with local storm reports of wind gusts or damage from the SPC storm report archive. The usage of storm reports in verification presents significant challenges, as the wind storm report database has been shown to have significant limitations (Weiss et al. 2002; Trapp et al. 2006). In order to address these challenges, wind reports are filtered in a manner similar to that of the forecast field. Specifically, if a report does not occur within 40 km of an observed radar object with a major axis length of at least 100 km, it is filtered out. The filtering process ensures that reports used in verification are associated with organized MCS activity. After filtering reports are re-gridded to the SSEO 4-km model grid.

An example of the results of this filtering technique can be seen in Fig. 2 from 24 June 2013. The unfiltered report field (Fig. 2a) highlights two main swaths of concentrated reports in which MCSs could conceivably have occurred. The first swatch stretches across the Corn Belt in Iowa, Illinois, and Indiana, and the second covers the Mid-Atlantic and Northeast regions. The observed radar valid at 2116 UTC 24 June 2013 (Fig. 2b) shows a well-organized bowing MCS over western Illinois, and scattered, cellular convection over the Northeast. The final filtered report field (Fig. 2c) shows that nearly all of the reports associated with the MCS over the corn belt were kept, while all but a couple of the reports associated with the scattered convection over the northeast have been removed.

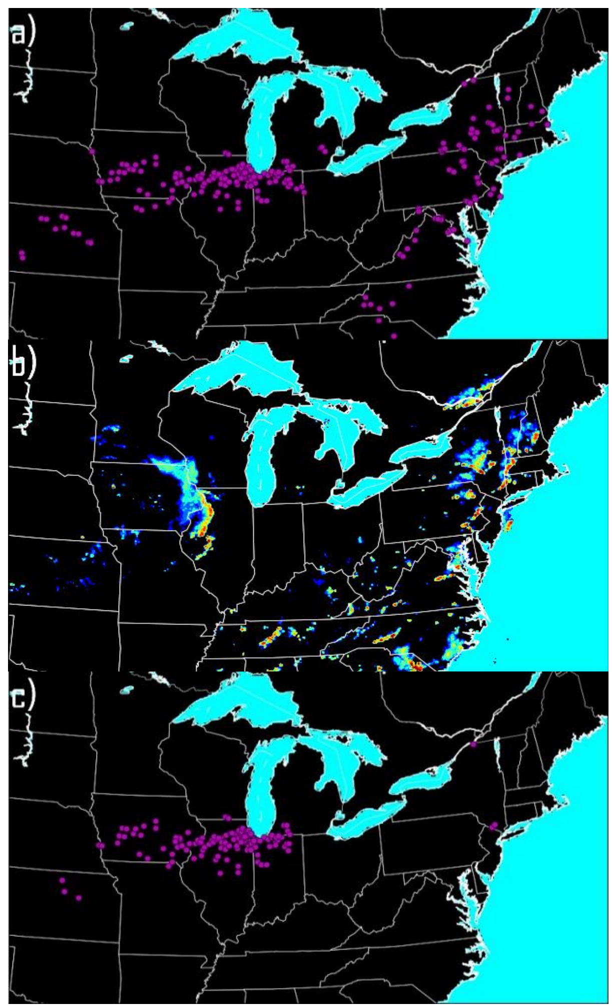

Figure 2. Example of report filtering by observed radar reflectivity from 24 June 2013. a) 24 June 2013 unfiltered wind reports b) Observed radar reflectivity valid 2116 UTC 24 June 2013 c) 24 June 2013 filtered wind reports based on observed radar reflectivity.

The final step in processing the report field for verification is to generate a probabilistic hindcast using the Practically Perfect technique (Hitchens et al. 2013), which applies a binary neighborhood around reports with a radius of influence of 40 km, and then smooths with a 2-D Gaussian kernel with a smoothing parameter of 120 km (Fig. 3).

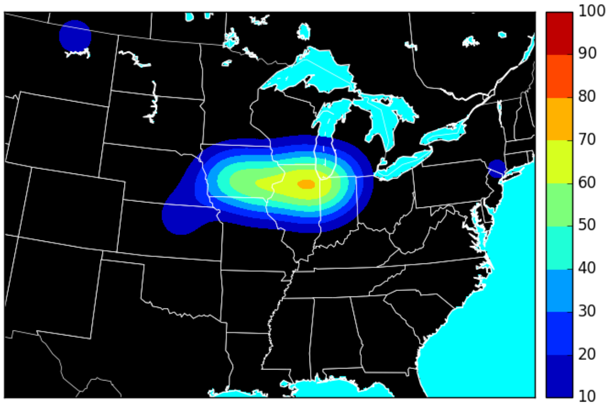

Figure 3. 24 Jun 2013 filtered Practically Perfect hindcast (%) generated from wind reports.

Thresholds can then be applied to the individual forecast and observed probabilistic swaths, allowing them to be treated as individual objects for object-based verification using the Method for Object-based Diagnostic Evaluation (MODE; Davis et al. 2006a,b). MODE has primarily been used to perform verification of precipitation forecast fields and has, until now, not been applied to the verification of ensemble wind probabilistic fields. MODE compares object attributes of forecast and observation objects and generates an interest score for each pair of objects between the fields. First, individual attribute interest scores are generated based on comparing the following object attributes for an object pair:

- Distance between object centroids
- Minimum separation between object boundaries
- Difference in orientation angle
- Area ratio
- Intersection area

These individual object attributes are then combined in an overall interest score for the object pair. If that score is above 0.7, the object pair is determined to be a match. Contingency table statistics can then be generated from the MODE output, where a matched forecast and observed object is considered a hit, an unmatched observed object is considered a miss, and an unmatched forecast object is considered a false alarm.

In order to define objects, MODE requires user-defined area and probability thresholds. In order to determine these values for this study, initial MODE tests using a range of area and probability thresholds over the three-year period were performed, matching MODE identified observation objects to manually identified MCS swaths. As seen in the performance diagram (Roebber 2009) shown in Fig. 4, the best-performing pair of thresholds for manually identified MCSs was the 25% probability threshold and 3000 grid space (48,000 km^2) area threshold pair, which had the combination of highest CSI score and most neutral bias. Based on the performance of this threshold pair for this initial verification, it was selected as the primary thresholds moving forward for forecast verification.

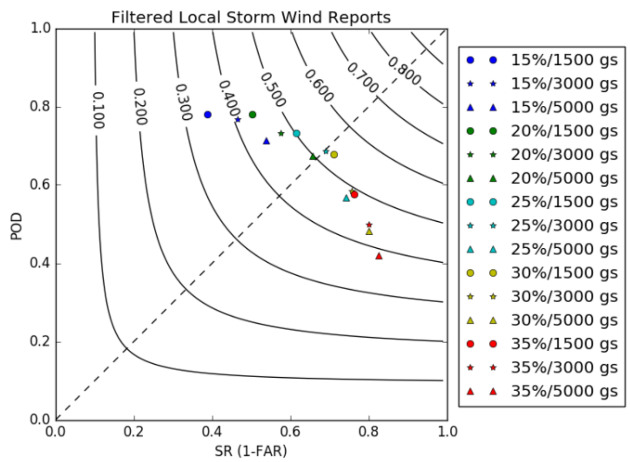

Figure 4. Performance diagram showing results of verification for MODE identified observed MCSs to manually identified MCS tracks.

3. RESULTS

MODE verification was performed over the three-year period from 2012 to 2014 for the full SSEO ensemble, as well as each individual deterministic model. These results are shown in the performance diagram in Fig. 5. The most noticeable result from this initial verification is the strong performance in terms of POD of the NSSL WRF-ARW. Additionally, both of the 12-hour time-lagged members from the HiRes Window had significantly lower PODs when compared to the more current HiRes Window runs. Finally, the SSEO ensemble outperformed all of the deterministic members in terms of CSI, confirming that an ensemble will generally outperform its deterministic members.

In order to examine the relative importance of individual members and groups of members to the ensemble, a systematic stepwise removal of members from the ensemble membership was performed. MODE verification was performed over the same three-year period for seven different versions of SSEO ensemble probabilities in which each single member was individually removed from the ensemble membership. Additionally, verification was performed for versions in which all three WRF-ARW members, all four NMM/NMMB members, and the two time-lagged members were removed from the ensemble membership.

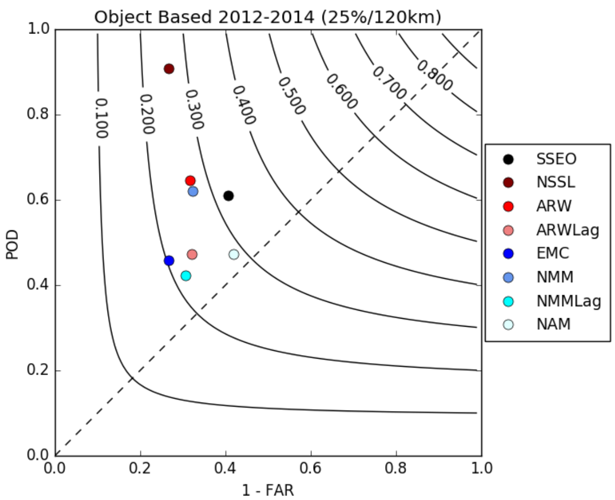

Figure 5. Performance diagram comparing deterministic member performance to the full SSEO.

The results for this member removal verification are shown in the performance diagram in Fig. 6. Six of the seven single member removals do not generally have significant impacts on the forecasts, as shown by the cluster of points around the full SSEO in black. The major exception occurs when removing the NSSL WRF-ARW from the SSEO, as its removal resulted in a significant drop in POD. This is not surprising given the NSSL-WRF had a significantly higher POD than any other deterministic member in Fig.5, and shows the relative importance of the NSSL-WRF to the SSEO when compared to the other deterministic members.

Removing groups of members based on model core reveals significant differences in the performances of WRF-ARW and NMM/NMMB comprised ensembles. First, removing all NMM/NMMB members results in a significant increase in both POD and frequency bias. Meanwhile, removing all WRF-ARW members results in a significant decrease in POD. This generally indicates that WRF-NMM/NMMB members are responsible for lowering 30-knot convective wind probabilities in the SSEO, while WRF-ARW members are responsible for raising convective wind probabilities. This shows the importance of having WRF-ARW members in a CAM ensemble for detecting severe MCS wind events.

Removing the two 12-hr time-lagged members of the SSEO results in a small increase in POD and frequency bias, which is to be expected based on the reduced POD of the deterministic time-lagged members in Fig. 5. This result generally indicates that including time-lagged members in a convection allowing ensemble will result in slightly lower convective wind probabilities generated by the ensemble (i.e., increased diversity/spread).

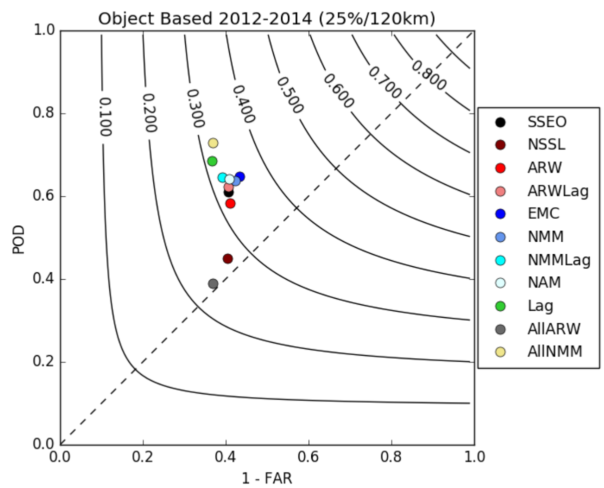

Figure 6. Performance diagram showing results of removing members from the SSEO membership.

4. IMPROVING ENSEMBLE FORECASTS OF SEVERE MCS WIND

While ensemble forecasts of severe wind from the SSEO have shown skill in outlining general severe MCS wind threat areas, forecasts are limited by the current ~4 km resolution of convection-allowing guidance. It has been show that at this relatively coarse resolution, CAMs are not able to fully resolve convective process necessary for the correct representation of convective wind gusts at the 10-meter level (Bryan et al. 2003). For this reason, the 30-knot wind threshold is used in the SSEO as a proxy for severe convective wind. This low threshold, however, introduces a potential limitation in which SSEO ensemble probabilities struggle to delineate between a high impact derecho-type event producing widespread significant wind reports greater than 65 knots, and a lower end MCS event producing mostly non-severe winds throughout its lifecycle.

An example of this limitation is presented in Figs. 7 and 8, showing two MCS cases which were forecast hits based on MODE verification of SSEO 30 knot wind probabilities. In the first case (Fig. 7a), 16 June 2014, a long-lived bowing MCS moved across northern and eastern Iowa, producing several 65 to 85 knot wind reports along its path. Comparatively, the second case (Fig. 8a), 14 June 2013, featured a much less impactful MCS moving across eastern Nebraska and western Iowa and producing several damage and 50 knot measured reports. The SSEO 30 knot wind probabilities for 16 June 2014 and 14 June 2013 are shown in Fig. 7b and Fig. 8b, respectively. While both forecasts perform well in covering the general area impacted by the MCS wind threat, there is little to no information which would provide guidance as to the expected magnitude of the MCS threat. Maximum forecast probabilities in the 16 June 2014 case were in the low 50s, while maximum forecast probabilities for the 14 June 2013 case were in the 40s.

This example shows the limitations of using a 10-meter wind variable to generate CAM ensemble probabilities. The 30-knot threshold is too low to generate meaningful guidance for MCS magnitude and potential impacts. Due to their current resolution, however, CAMs do not generate 50 knot winds with enough frequency to make useful ensemble probabilities at that threshold or above. With that in mind it becomes necessary to investigate the development of proxy guidance from CAMs for severe MCS wind. This proxy guidance would need to provide information on both the general threat areas and potential magnitude of MCS events, which would provide forecasters with better information for producing high impact MCS forecasts.

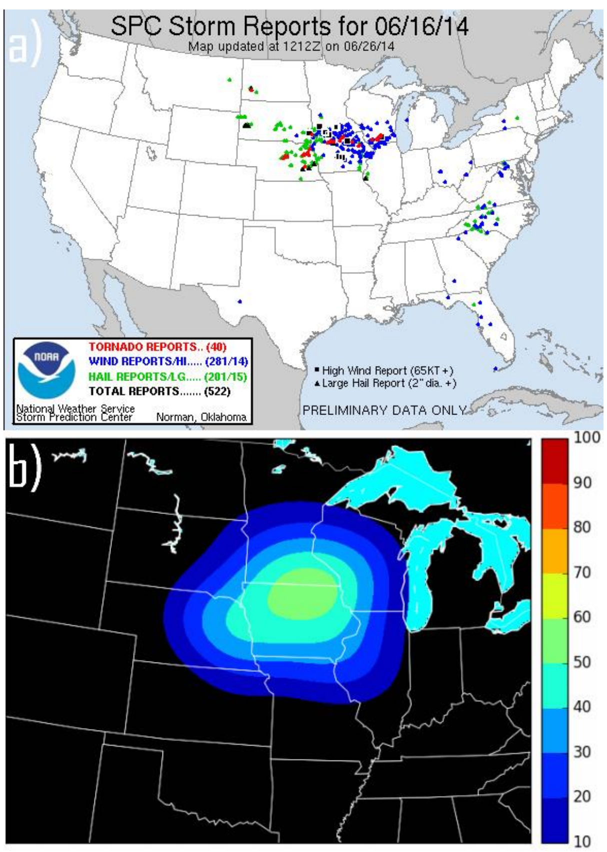

Figure 7. 16 June 2014 a) SPC local storm reports and b) SSEO 30 knot wind ensemble probabilities.

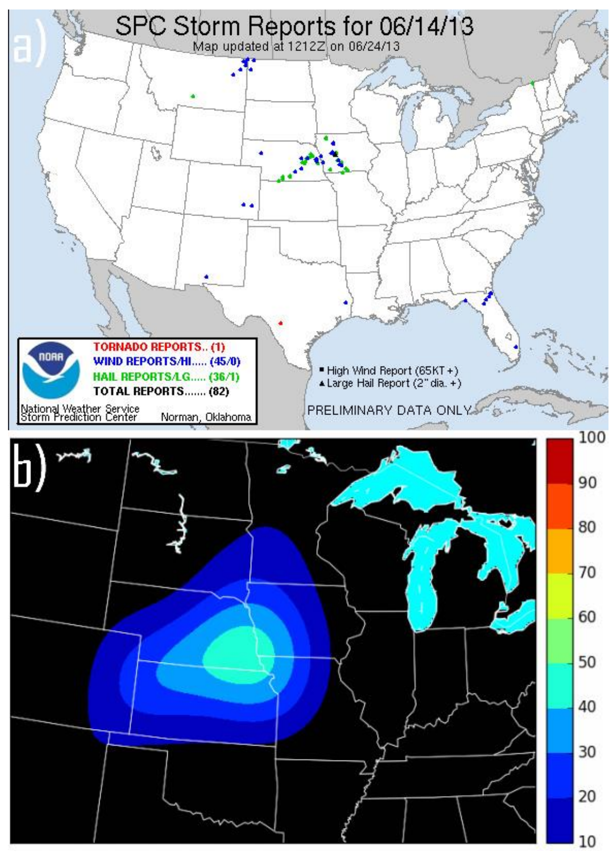

Figure 8. 14 June 2013 a) SPC local storm reports and b) SSEO 30 knot wind ensemble probabilities.

5. SUMMARY AND CONCLUSIONS

A technique for the verification of ensemble probabilistic forecasts of severe MCS wind in an object-based framework using MODE has been presented. While the usage of storm reports of damaging wind gusts presents serious challenges and limitations for verification, filtering of the reports by observed radar reflectivity ensures that the storm reports used in verification are associated with organized MCS activity. MODE verification was performed for SSEO forecasts of 24-hour probabilities of 10-meter wind speeds ≥30 knots over a three-year period from 2012 to 2014. Results show that the SSEO ensemble outperformed its individual determinstic members in terms of CSI over that period.

Additional MODE verification was performed over the same time period in which single members and groups of members were removed from the SSEO membership. Results show that removing the NSSL WRF-ARW from the SSEO membership had a significant negative impact on the performance of the ensemble. Additionally, removing all NMM/NMMB members significantly increased POD, while removing all WRF-ARW members from the membership significantly decreased POD. Removing the two 12-hour time-lagged members from the membership saw a small increase in POD. These results highlight the importance of WRF-ARW members in a CAM ensemble for the prediction of severe MCS wind.

Finally, two example cases were presented which illustrate the limitations of current CAM ensembles to provide guidance on the magnitude of expected MCS threats. This limitation is related to resolution limitations of current CAMs, and their inability to fully resolve physical processes resulting in convective gusts. In order to provide more meaningful ensemble guidance for MCS magnitude prediction, the development of proxy variables for severe wind is necessary and is currently ongoing.

REFERENCES

Adams-Selin, R. and C. L. Ziegler, 2016: Forecasting hail using a one-dimensional hail growth model within WRF. *Mon. Wea. Rev.*, Early Online Release.

Brooks, H.E., M. Kay, and J.A. Hart, 1998: Objective limits on forecasting skill of rare events. Preprints, *19th Conf. on Severe Local Storms*, Minneapolis, MN, Amer. Meteor. Soc., 552–555.

Bryan, G. H., J. C. Wyngaard, and J. M. Fritsch, 2003: Resolution requirements for the simulation of deep moist convection. *Mon. Wea. Rev.,* **131**, 2394–2416.

Davis, C. A., B. Brown, and R. Bullock, 2006a: Object-Based Verification of Precipitation Forecasts. Part I: Methodology and Application to Mesoscale Rain Areas. *Mon. Wea. Rev.*, **134,** 1772–1784.

Davis, C. A., B. Brown, and R. Bullock, 2006b: Object-Based Verification of Precipitation Forecasts. Part II: Application to Convective Rain Systems. *Mon. Wea. Rev.,* **134**, 1785–1795.

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.

Gagne II, D.J., A. McGovern, J. Brotzge, M. Coniglio, J. Correia Jr., and M. Xue, 2015: Day-Ahead Hail Prediction Integrating Machine Learning with Storm-Scale Numerical Weather Models. *IAAI.*

Hitchens, N.M., H.E. Brooks, and M.P. Kay, 2013: Objective limits on forecasting skill of rare events. *Wea. Forecasting*, **28**, 525–534.

Johns, R. H. and C. A. Doswell III, 1992: Severe local storms forecasting. *Wea. Forecasting*, **7**, 588–612.

Jirak, I. L., S. J. Weiss, and C. J. Melick, 2012: The SPC storm-scale ensemble of opportunity: Overview and results from the 2012 Hazardous Weather Testbed Spring Forecasting Experiment. *Preprints*, 26th Conf. Severe Local Storms, Nashville, TN. Amer. Meteor. Soc., P9.137.

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.

Roebber, P. J., 2009: Visualizing multiple measures of forecast quality. *Wea. Forecasting*, **24**, 601–608.

Trapp, R. J., D. M. Wheatley, N. T. Atkins, R. W. Przybylinski, and R. Wolf, 2006: Buyer beware: Some words of caution on the use of severe wind reports in postevent assessment and research. *Wea. Forecasting*, **21**, 408–415.

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., J. A. Hart, and P. R. Janish, 2002: An examination of severe thunderstorm wind report climatology: 1970–1999. Preprints, *21st Conf. on Severe Local Storms*, San Antonio, TX, Amer. Meteor. Soc., 446–449.