115 VERIFICATION OF 10-METER WIND FORECASTS FROM NSSL-WRF IN PREDICTING SEVERE WIND-PRODUCING MCSS

Jeffrey M. Milne*[1,2,3], Harold E. Brooks[4], Israel L. Jirak[3], Robert M. Hepper[2]

[1]University of Oklahoma, [2]Cooperative Institute for Mesoscale Meteorological Studies, [3]NOAA/NWS/NCEP Storm Prediction Center, [4]NOAA/OAR/NSSL

1. INTRODUCTION

Forecasting for severe convective wind is a challenge for forecasters. Additionally, severe wind reports account for approximately half of all severe storm reports received between 2012-14, and severe wind events are likely to increase as the Earth's climate changes (Brooks 2013). As convection-allowing models (CAMs) have become more prevalent, forecasters receive explicit details regarding thunderstorm intensity from model forecasts. Though CAMs can provide information about convective mode (i.e. linear, cellular, or clustered) (Weisman et al. 2008; Done et al. 2004), there has been no study done to determine CAMs' ability to forecast severe wind events. To develop improved guidance in forecasting severe wind events, it is necessary to determine how well CAMs currently forecast severe wind events. To do this, forecasts of 10-meter wind speeds will be verified for the 2012-14 period to determine CAMs' ability to forecast severe wind-producing mesoscale convective systems (MCSs). As part of the verification effort, a spatial and temporal climatology of severe wind-producing MCSs will be developed. Object-based verification will be explored to determine its utility in verifying severe wind-producing MCSs.

Severe wind-producing MCSs have been examined on an observational basis using soundings that sample the environment of mature MCSs. Cohen et al. (2007) found that the best discriminators for distinguishing severe wind-producing MCSs from non-severe MCSs were deep-layer wind shear and upper level winds. However, they only examined observed soundings taken ahead of or within MCSs and not model forecasts of MCSs. No studies have examined model forecasts in an attempt at assessing model skill in forecasting severe wind-producing MCSs, as this study will do.

Smith et al. (2013) found that measured severe wind gusts associated with quasi-linear convective systems (QLCS) were most common, relative to gusts associated with supercells and disorganized convection, between November and April. The most measured gusts associated with QLCSs occurred in June. Measured gusts from QLCSs made up 42% of severe measured gusts associated with deep convection. QLCS gusts occurred most often east of the Rockies between the plains and the Ohio River Valley.

2. DATA AND METHODS

Object-based verification is used in this study for verifying a field forecast with sporadic point observations. While there are other methods to verify a field forecast with point observations, exploring object-based verification was felt to be worthwhile since no convection-allowing model verification has been done with object-based verification for severe convective wind events. Additionally, object-based verification will allow for the development of a severe wind-producing MCS climatology.

The Method for Object-based Diagnostic Evaluation (MODE) is an object-based method for evaluating forecasts developed by the Developmental Testbed Center and National Center for Atmospheric Research as part of their

* *Corresponding author address*: Jeffrey M. Milne, University of Oklahoma, School of Meteorology, 120 David L. Boren Blvd., Suite 5900, Norman, OK 73072; email: jeffrey.milne@noaa.gov

Model Evaluation Tools (Davis et al. 2006). MODE identifies objects by applying a Gaussian filter with a user-defined spatial smoothing radius, σ. A user-defined intensity threshold is then applied. After the thresholding, a user-defined minimum area may also be applied, with objects smaller than a certain area being excluded. The resulting objects are then used to mask the original, unsmoothed data.

Before matching is done between forecast and observed objects, grid-based verification metrics are computed for the objects. That is, every grid square within both a forecast and observation objects is counted as a hit; every grid square within a forecast object but not an observed object is counted as a false alarm; every grid square within an observed object but not a forecast object is counted as a miss; and every grid square in neither a forecast or observed object is counted as a correct null. This tabulation allows for the computation of grid-based verification metrics using the forecast and observed objects.

To match objects between the forecast and observed field, MODE uses a fuzzy logic engine to generate an interest score between pairs of objects. The fuzzy logic attempts to determine whether the forecast object and the observed object are the same event. This method attempts to remove the double-penalty for spatial errors in which one missed forecast counts as both a miss and as a false alarm as described in Ahijevych et al. (2009).

The National Weather Service defines severe wind as wind gusts greater than or equal to 50 knots (25.7ms^{-1}, 58mph). Severe wind reports accounted for 65% of all severe local storm reports (i.e., hail with diameter > 1", wind speed > 50kts, or a tornado) from 2012 through 2014. Local storm reports are collected by NWS Weather Forecast Offices (WFOs) and then grouped from 12Z on one day through 1159Z on the next day (referred to as a convective day, hereafter all references to days are referring to convective days). Depending on the time of day and region of the country, this roughly corresponds to the 24-hour period beginning at or just before sunrise. Sunrise is the approximate daily minimum of severe storm reports (Kelly et al. 1985), so the number of events occurring across two convective days is minimized.

In addition to the magnitude overestimation issues shown in Doswell et al. (2005), there are other issues with severe reports overall. As described in Doswell and Burgess (1988), for a report to show up in the local storm report or *Storm Data* databases, three things must happen: someone must observe the event, they must recognize that the event meets the severe criteria, and they must report the event to the relevant authority. This leads to potential population biases, since it is less likely someone will observe the event if there is low population density except for near roads (Weiss et al. 2002). Weiss also mentions diurnal biases, since it is more difficult to observe events at night. Additionally, the use of wind damage reports can introduce biases. As Trapp et al. (2006) note, a report of "trees were downed" could be "a few bent-over saplings, a large grove of snapped hardwood trees with ~0.5 m diameters, or something in between." Trapp et al. (2006) also note that report (either local storm report or *Storm Data*) concentration or counts do not necessarily correlate with wind speed magnitude. An event that had only three local storm reports resulted in $1 million in damages from 70mph wind gusts, while an event with 55 reports resulted in only $0.3 million in damages with less significant damage than the other case (Trapp et al. 2006). The lack of correlation between reports and monetary damages is not an issue for this study.

The 10-m wind forecasts were generated from a 4-km grid spacing configuration of the Weather Research and Forecasting model (WRF) run by the National Severe Storms Laboratory (NSSL, the model is hereafter referred to as the NSSL-WRF) initialized at 00Z and run to 36 hours (until 12Z the next day). The NSSL-WRF uses Mellor-Yamada-Janjić boundary layer and turbulence parameterization (Mellor and Yamada 1982), WRF Single-Moment 6-Class microphysics

(Hong and Lim 2006), Rapid Radiative Transfer Model longwave radiation (Mlawer et al. 1997), Dudhia (1989) shortwave radiation, and Noah land surface model (Chen and Dudhia 2001). The NSSL-WRF has 35 vertical levels, a 24 second time step, and uses the NAM model interpolated to a 40-km grid for initial and boundary conditions.

To use MODE to objectively verify model forecasts, it was first necessary to figure out how to best identify severe wind-producing MCSs that occurred during 2012-14. Initially, this was done manually using severe wind reports and radar reflectivity, then the manually identified severe wind-producing MCSs were used to determine the best objective approach to identify severe wind-producing MCSs using MODE.

To identify a coherent damaging wind swath from severe wind-producing MCSs, a spatial Gaussian kernel density estimation was applied to the severe wind reports. This was done using the practically perfect method described in Hitchens et al. (2013) modified to be used on a 4-km grid. Hitchens et al. (2013) used an 80-km grid with a 120km Gaussian smoother. All grid squares containing a report were assigned a value of 1 and the Gaussian smoother was applied to produce a probabilistic field that should match what a Storm Prediction Center forecaster would forecast given perfect foreknowledge of the day's severe storm reports. The modification to the practically perfect method in this study was to put all severe wind reports on a 4-km grid and assign a value of 1 to all grid squares within 10 grid squares (40km) of a wind report. The same 120km Gaussian smoother was applied. All storm reports for 29 June 2012 are shown in Fig. 1, and the practically perfect hindcast for severe wind is shown in Fig. 2.

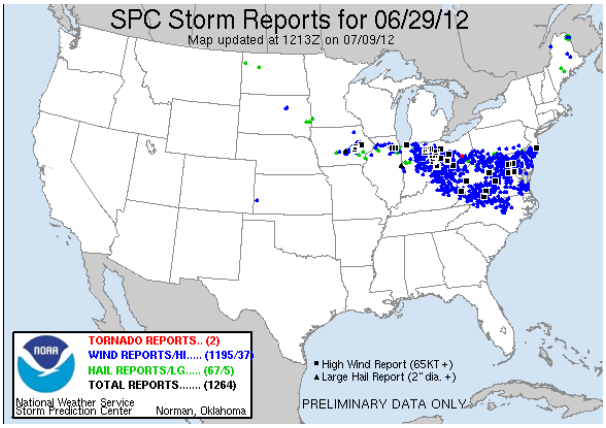

Figure 1. All local storm reports for 29 June 2012

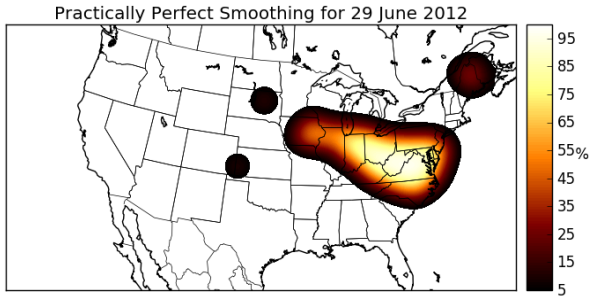

Figure 2. Practically perfect smoothed hindcast of all local storm wind reports for 29 June 2012.

Due to the potential issues with severe wind reports and to eliminate severe wind reports not associated with MCSs, observed radar reflectivity was used to eliminate reports not associated with large areas of organized convection. Reports were discarded if they were not within 40km of a contiguous area of radar reflectivity greater than 35dBZ covering 500 grid square (8000km^2 on a 4km grid). The 35dBZ radar reflectivity threshold was chosen to capture convective radar echoes as in Mecikalski and Bedka (2006). The 40km radius of influence was chosen to account for wind reports caused by outflow boundaries and gust fronts and to match Hitchens et al. (2013). The 8000km^2 minimum area represents a circle with a ~100km diameter. This minimum area ensures that the area of radar reflectivity is larger than 100km in at least one direction to match Parker and Johnson (2000). An example of the practically perfect hindcast for severe wind reports filtered by radar reflectivity is shown in Fig. 3. Notice removal of severe wind areas in Maine and the plains by filtering out reports not associated with MCSs.

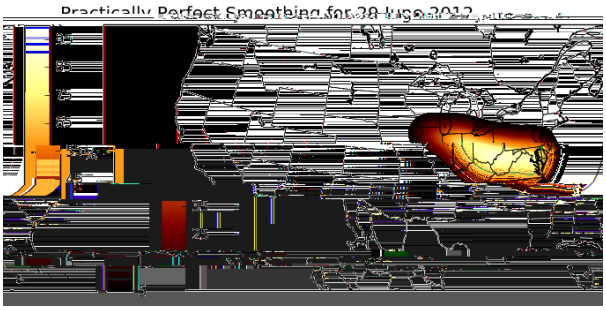

Figure 3. Practically perfect hindcast of radar-filtered local storm wind reports for 29 June 2012.

A spatial climatology is developed using MODE identified objects. A temporal climatology is developed using the formula:

$$C_j = \frac{\sum_{i=j-M}^{i=j+M} N_i}{2M+1} \quad (1)$$

Where C_j is the number of severe wind producing-MCSs expected on day j. M determines the size of the window for the moving average. Results are presented for M = 15 days (for a 31 day window) and M = 45 days (for a 91 day window). N_i is the number of severe wind-producing MCSs observed on day i. When i reaches above 365 or below 0, 365 is added or subtracted as appropriate so that i stays between 0 and 365.

For model forecasts, a 24-hour maximum 10-m wind field was created by taking the maximum of the 24 individual hourly maximum 10-m wind fields from forecast hours 12 through 36 (i.e. from 12Z to 12Z for a 00Z forecast, a convective day). A simulated radar filter was applied that matches the observed reflectivity filter that was applied to wind reports. Areas of model forecast wind were discarded if they were not within 40km of an area of 500 contiguous grid squares of simulated hourly maximum reflectivity higher than 35dBZ. A filtered forecast for 29 June 2012 is shown in Fig.4.

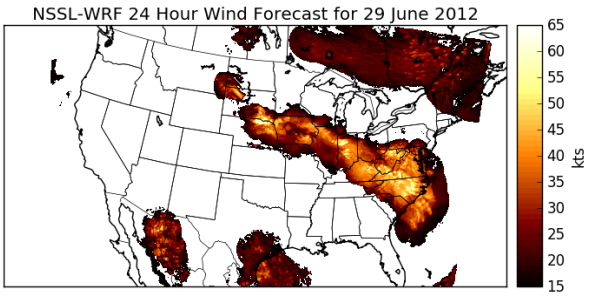

Figure 4. 24-hour maximum forecast 10-m wind field filtered by simulated reflectivity.

Practically perfect smoothing was applied to the forecast field so that forecast and observed objects would have similar characteristics and could be matched using MODE. To generate practically perfect fields, thresholds were applied to the forecast field at 5-knot wind speed intervals between 15 and 60kts (7.7 - 30.9ms^{-1}). Areas with forecast wind speeds higher than the threshold were assigned a value of 1 and all other areas assigned a value of 0. The same practically perfect smoothing was applied as with the reports: 1) All grid squares within 40km of an area of wind speeds above the threshold were given a value of 1 and 2) a 120km Gaussian smoother was then applied to create a probabilistic field. A selection of practically perfectly smoothed

forecasts based on the filtered wind field are shown for several wind speed thresholds in Fig. 5.

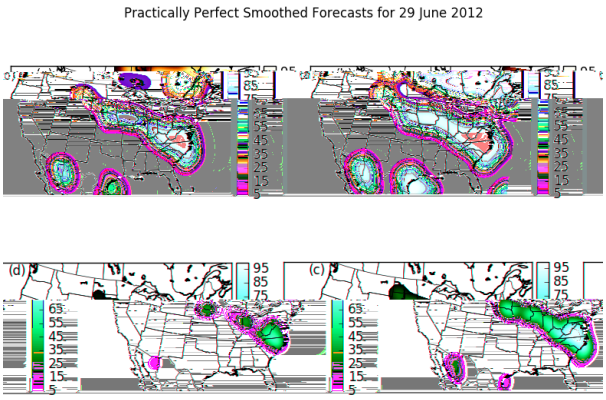

Figure 5. Practically perfect smoothed forecasts for 29 June 2012. The forecasts were thresholded at (a) 20kt, (b) 30kt, (c) 40kt, and (d) 50kt.

Verification metrics were computed in two ways: grid-based verification in which objects are identified, then every grid point is counted as a hit, miss, false alarm, or correct null based on whether it is in both a forecast and observed object, an observed object only, a forecast object only, or neither, respectively; and object-based verification, in which matched forecast and observed objects are counted as a hit, unmatched observed objects as a miss, and unmatched forecast objects as a false alarm. Verification results are presented in a performance diagram (Roebber 2009).

3. SEVERE WIND-PRODUCING MCS CLIMATOLOGY

The locations of all MODE-identified severe wind-producing MCSs are shown in Figure 6. The distribution of severe wind-producing MCSs is similar to that shown in (Smith et al. 2013). There are two maxima with more than 30 severe wind-producing MCSs occurring over the three year period: one in Kentucky and another in southwestern Georgia. There is also a relative maximum in eastern Pennsylvania. The prevalence of severe wind-producing MCSs decreases towards the western Great Plains based on this approach. The reason for the decrease over the Great Plains may not be meteorological, but may be explained by a couple of other factors: population is lower in the western Great Plains, so there are fewer possibilities for severe winds to be reported, and there are fewer trees and structures to be damaged, so there are fewer instances of reported wind damage.

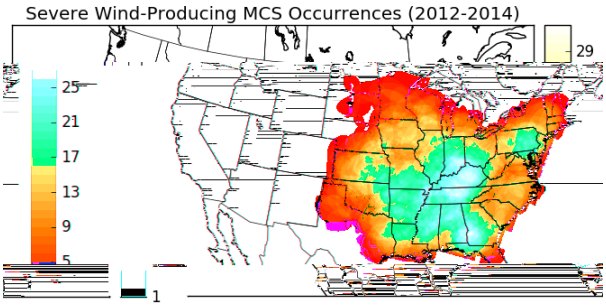

Figure 6. Severe wind-producing MCS climatology for 2012-14

The number of severe wind-producing MCSs across the CONUS expected per day is shown in Fig. 7 with a maximum occurrence in the summer. The 31-day window has an absolute maximum on 13 June, when 0.70 severe wind-producing MCSs occurred per day. The 91-day window has an absolute maximum on 14 June, when 0.48 severe wind-producing MCSs occurred per day. There are several relative maxima with the 31-day window, though none are apparent in the 91-day window. This matches Smith et al. (2013), which found that wind gusts associated with QLCSs occurred most often in June. Derechos tend to have a maximum in May (Bentley and Sparks 2003). Jirak et al. (2003) found that MCSs in general (not restricted to severe wind-producing MCSs) have a maximum in July, but that May and June also have a similarly large number of MCSs.

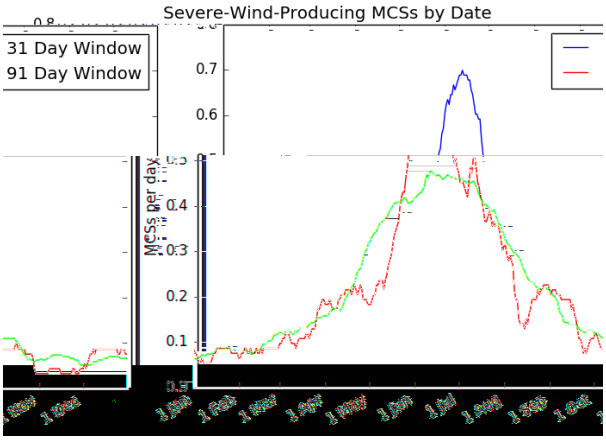

Figure 7. Severe wind-producing MCSs by date.

4. VERIFICATION OF NSSL-WRF FORECASTS

When using MODE and considering a grid-based verification, CSI is below 0.10 for all forecast thresholds (Fig.8). The maximum CSI of 0.075 occurs at a forecast threshold of 45kts. For forecast thresholds between 30kts and 50kts, the CSI stays relatively close to 0.070 (highlighted in Fig. 8). For low thresholds, the forecast approaches "always yes." That is, forecast objects cover a large portion of the domain for every forecast. As the forecast threshold increases, POD decreases dramatically, while FAR only decreases slightly. At the 45kt threshold, the bias is near 1, and as the forecast threshold increases above 45kts, POD still decreases, though less dramatically and FAR starts to decrease more rapidly. For the highest threshold, 60kts, POD is slightly above 0.

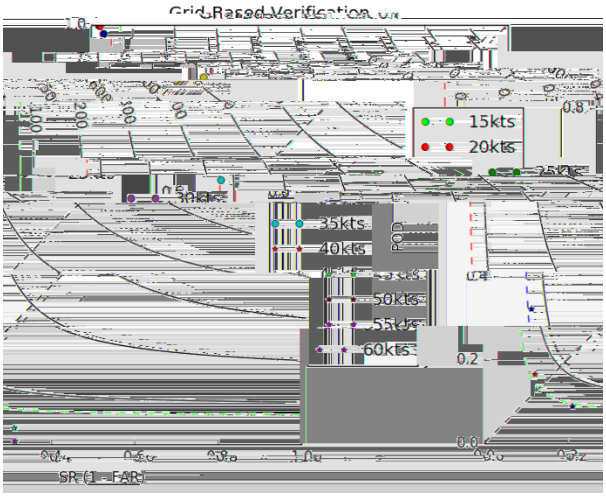

Figure 8. A performance diagram for grid-based verification using MODE to identify 10-m wind objects from the NSSL-WRF from 15-60 knots. The 0.070 CSI contour is highlighted. A perfect forecast would be in the upper right.

When considering object-based verification (Fig. 9), the general trend is similar to that observed in the MODE grid-based verification, though CSI is higher for object-based verification. The maximum CSI of 0.16 occurs at a forecast threshold of 50kts, which is the severe wind threshold set by the NWS. The 50kt forecast threshold also produces the bias nearest to 1. Between the 35kt and 50kt thresholds, CSI stays very close to 0.15 (highlighted in the figure). As was seen with grid-based verification, forecasts at very low thresholds have a large frequency bias (i.e., more forecast objects than observed objects), though there is a lower FAR for object-based verification than for grid-based verification. For the 3 lowest thresholds, POD is equal to 1. As the forecast threshold increases, POD decreases dramatically, and FAR decreases slightly until the 45kt threshold. For thresholds above 45kts, FAR decreases nearly as fast as POD. As in grid-based verification, the highest threshold, 60kts, has a POD just above 0. The dramatic decrease in POD without much decrease in FAR for thresholds between 30kts and 45kts suggests that 30kts may be a more useful threshold to forecasters. The slight decrease in FAR may not be worth the dramatic decrease in POD when using a higher threshold. If the costs of a missed event are

greater than the costs of a false alarm, then choosing a lower threshold may be worthwhile.

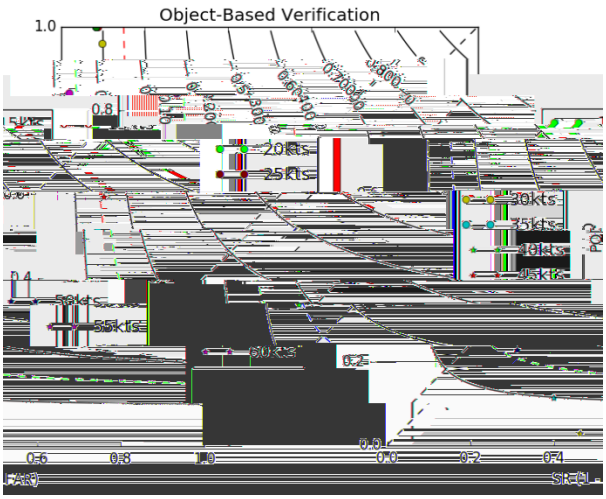

Figure 9. As in Fig.8, but for object-based verification. The 0.15 CSI contour is highlighted.

5. CONCLUSIONS

A climatology of severe wind-producing MCSs from 2012 - 2014 was developed using an object-based approach using severe wind reports and observed radar reflectivity. To develop the climatology, MODE was used to identify severe wind objects that were based on a practically perfect hindcast of severe wind reports filtered by radar reflectivity. Severe wind-producing MCSs occurred most often in the Ohio River Valley, with a secondary maximum in southwestern Georgia and southeastern Alabama. Temporally, severe wind-producing MCSs occurred most often in the warm season with a peak in June. This climatology is generally consistent with other climatologies of severe wind reports and MCSs.

To verify the NSSL-WRF, a 24-hour maximum 10-m wind field was generated from forecast hours 12 - 36 from the 0000 UTC model run and filtered by simulated reflectivity using the same parameters as were used to filter the wind reports. Thresholds between 15kt and 60kt were examined in the forecast field, then the forecast field was smoothed using the same parameters that were applied to create the hindcast of the severe wind reports. MODE was again used to identify forecast objects that were matched to observed objects using a fuzzy logic engine. Verification metrics were computed two ways: grid-based verification using MODE and object-based verification using MODE. Grid-based and object-based verification using MODE both showed that the model had a relatively constant critical success index across a range of forecast wind speed thresholds. Object-based verification using MODE yielded higher values of CSI than grid-based verification using MODE. Additionally, though wind-speed forecasts at higher thresholds yielded biases nearest to one and slightly higher CSI values, wind-speed forecasts at lower thresholds with slightly lower CSIs might be more useful for forecasters since POD is dramatically higher at lower thresholds without a large penalty in FAR. Using a lower forecast wind-speed threshold captures more events (e.g. a 35kt threshold captures more events than a 50kt threshold).

When considering severe wind-producing MCSs, the NSSL-WRF overforecasts at low wind speed thresholds and underforecasts at higher wind speed thresholds. By filtering the model winds with simulated radar reflectivity, model performance for severe MCSs was improved. Even though object-based verification of severe MCS winds, as highlighted by the utility of CAM forecasts over traditional grid-point verification approaches, there may room for developing improved severe wind proxies from CAMs. Further research investigating fields in addition to 10-m winds as potential severe wind proxies is ongoing. The current work provides baseline verification metrics for any potential new proxy. That is, a new proxy would have to generate better verification scores than 10-meter wind forecasts to show utility in forecasting severe wind-producing MCSs.

6. REFERENCES

Ahijevych, D., E. Gilleland, B. G. Brown, and E. E.

Ebert, 2009: Application of Spatial Verification Methods to Idealized and NWP-Gridded Precipitation Forecasts. *Weather Forecast.*, **24**, 1485–1497.

Bentley, M. L., and J. a Sparks, 2003: A 15 yr climatology of derecho-producing mesoscale convective systems over the central and eastern United States. *Clim. Res.*, **24**, 129–139.

Brooks, H. E., 2013: Severe thunderstorms and climate change. *Atmos. Res.*, **123**, 129–138.

Chen, F., and J. Dudhia, 2001: Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity. *Mon. Weather Rev.*, **129**, 569–585.

Cohen, A. E., M. C. Coniglio, S. F. Corfidi, and S. J. Corfidi, 2007: Discrimination of Mesoscale Convective System Environments Using Sounding Observations. *Weather Forecast.*, **22**, 1045–1062.

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

Done, J., C. A. 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**, 110–117.

Doswell, C. A., and D. W. Burgess, 1988: On Some Issues of United States Tornado Climatology. *Mon. Weather Rev.*, **116**, 495–501.

——, H. E. Brooks, and M. P. Kay, 2005: Climatological Estimates of Daily Local Nontornadic Severe Thunderstorm Probability for the United States. *Weather Forecast.*, **20**, 577–595.

Dudhia, J., 1989: Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model. *J. Atmos. Sci.*, **46**, 3077–3107.

Hitchens, N. M., H. E. Brooks, and M. P. Kay, 2013: Objective Limits on Forecasting Skill of Rare Events. *Weather Forecast.*, **28**, 525–534.

Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF Single-Moment 6-Class Microphysics Scheme (WSM6). *J. Korean Meteorol. Soc.*, **42**, 129–151.

Jirak, I. L., W. R. Cotton, and R. L. McAnelly, 2003: Satellite and Radar Survey of Mesoscale Convective System Development. *Mon. Weather Rev.*, **131**, 2428–2449.

Kelly, D. L., J. T. Schaefer, and C. A. Doswell, 1985: Climatology of Nontornadic Severe Thunderstorm Events in the United States. *Mon. Weather Rev.*, **113**, 1997–2014.

Mecikalski, J. R., and K. M. Bedka, 2006: Forecasting Convective Initiation by Monitoring the Evolution of Moving Cumulus in Daytime GOES Imagery. *Mon. Weather Rev.*, **134**, 49–78.

Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys.*, **20**, 851.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res. Atmos.*, **102**, 16663–16682.

Parker, M. D., and R. H. Johnson, 2000: Organizational Modes of Midlatitude Mesoscale Convective Systems. *Mon. Weather Rev.*, **128**, 3413–3436.

Roebber, P. J., 2009: Visualizing Multiple Measures of Forecast Quality. *Weather Forecast.*, **24**, 601–608.

Smith, B. T., T. E. Castellanos, A. C. Winters, C. M. Mead, A. R. Dean, and R. L. Thompson, 2013:

Measured severe convective wind climatology and associated convective modes of thunderstorms in the contiguous United States, 2003-2009. *Weather Forecast.*, **228**, 229–236.

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. *Weather Forecast.*, **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. *Weather Forecast.*, **23**, 407–437.

Weiss, S., J. Hart, and P. Janish, 2002: An examination of severe thunderstorm wind report climatology: 1970–1999. *Prepr. 21st Conf. Sev. Local Storms, San …*, 2–5.