A FEASIBILITY STUDY FOR PROBABILISTIC CONVECTION INITIATION FORECASTS BASED ON EXPLICIT NUMERICAL GUIDANCE

BY JOHN S. KAIN, MICHAEL C. CONIGLIO, JAMES CORREIA, ADAM J. CLARK, PATRICK T. MARSH, CONRAD L. ZIEGLER, VALLIAPPA LAKSHMANAN, STUART D. MILLER JR., SCOTT R. DEMBEK, STEVEN J. WEISS, FANYOU KONG, MING XUE, RYAN A. SOBASH, ANDREW R. DEAN, ISRAEL L. JIRAK, AND CHRISTOPHER J. MELICK

For practical purposes, the convection initiation forecasting challenge should be framed in terms of the initiation of mesoscale convective events rather than the formation and growth of individual cumulonimbus clouds.

The annual Spring Forecasting Experiment (SFE) is conducted by the National Severe Storms Laboratory (NSSL) and Storm Prediction Center (SPC) in the National Oceanic and Atmospheric Administration (NOAA) Hazardous Weather Testbed (HWT) during the climatological peak of severe convective weather in the United States (Clark et al. 2012a). This experiment is uniquely designed to bring together meteorological scientists who conduct applied research and practitioners who focus on forecasting to work on emerging problems of mutual interest. In 2011, a major component of the SFE (hereafter SFE2011) was a pilot study on the initiation of thunderstorms, also known as convection initiation (CI).[1] The focus on CI was motivated by two primary factors: 1) CI remains a scientifically challenging problem (e.g., see Markowski and Richardson 2010, chapter 7), and 2) there is a growing awareness that CI prediction is one of the weak links in forecasts of thunderstorm activity, presenting a significant challenge for forecasters, including those who specialize in prediction of severe convective weather, flash flooding, aviation hazards, and other specific threats.

A CI focus was also inspired by the encouraging performance over the contiguous United States (CONUS) of emerging convection-allowing models (CAMs) with 3–4-km grid spacing, as documented by numerous recent studies (e.g., Weisman et al. 2008; Kain et al. 2008; Coniglio et al. 2010; Clark et al. 2009, 2010a,b, 2012a). These models have sufficiently fine resolution to allow deep moist convection to develop and evolve explicitly on the model grid and they have shown skill in the prediction of multiple aspects of severe weather (e.g., Sobash et al. 2011; Marsh et al. 2012; Clark et al. 2012b), but the skill of their CI forecasts has never been evaluated objectively and systematically over the CONUS. A primary goal of SFE2011 was to provide

[1] The alternate term "convective initiation" has been applied to CI in some previous studies.

a preliminary quantitative assessment of the skill of CAMs in predicting CI.

Such an evaluation requires specific definitions, but to our knowledge there is no generally accepted definition of CI. From a process-oriented perspective, CI can be thought of as the sequence of events in which air parcels accelerate upward beyond their level of free convection, resulting in a clearly visible growth of cloud top, a rapid increase in cloud depth, and possibly the formation of precipitation and development of lightning (e.g., Weckwerth and Wakimoto 1992; Crook 1996; Ziegler and Rasmussen 1998; Doswell 2001; Houston and Niyogi 2007). For practical purposes (i.e., societal impacts), this sequence of events is not necessarily noteworthy. For example, sometimes this process results in nothing more than "turkey towers" (e.g., Corfidi et al. 2008) that have little impact at ground level. Other times, it leads to short-lived isolated thunderstorms that bring lightning and brief heavy downpours but affect only a very small and localized segment of the population. In other cases, however, CI episodes are clustered in association with specific mesoscale dynamic forcing and/or focusing mechanisms (e.g., Wilson and Roberts 2006). In these cases, CI can mark the beginning of a major disruptive event such as a squall line, an outbreak of supercells, or a derecho that interferes with human activities over a large area. Thus, if societal impact is to be considered in defining CI, it might be appropriate to include criteria for the intensity and longevity of individual (or representative) storms, changes in frequency of nearby CI episodes, tendencies for consolidation of storms, upscale growth rates, overall coverage, and perhaps other factors. A second goal of SFE2011 was to explore the use of different criteria for defining CI, considering both scientific and practical priorities.

This exploration was conducted in the context of experimental forecasting exercises, which have become a cornerstone of interactions between forecasters and researchers in the HWT (e.g., Clark et al. 2012a). Strongly forced convective events were the primary focus of these exercises, partly because of the predominance of these events east of the Rockies during the spring but also because this focus complemented simultaneous experimental forecasts for severe convection that were conducted as part of another component of SFE2011. During forecasting exercises CAM output was examined for evidence of CI itself and for the presence of meteorological features and physical processes that are often precursors to CI. For example, CI appears to be modulated by low-level airmass boundaries and convergence zones (e.g., Wilson and Schreiber 1986), horizontal convective rolls (HCRs) (e.g., Weckwerth 2000), the depth and intensity of cloud-scale lifting along boundaries and HCR-like circulations (Ziegler and Rasmussen 1998; Ziegler et al. 2007), orography (e.g., Groenemeijer et al. 2009; Barthlott et al. 2011; Behrendt et al. 2011), mesoscale and larger-scale dynamic effects aloft (e.g., Griffiths et al. 2000), gravity waves and bores (e.g., Koch and O'Handley 1997), and many other factors (see Browning et al. 2007; Wilson and Roberts 2006; Weckwerth and Parsons 2006, and references therein). A third goal of SFE2011 was to examine how well CAMs with 3–4-km grid spacing can represent some of these important features and processes.

SFE2011 was conducted from 9 May through 10 June 2011, Monday through Friday. Experimental activities and preliminary results from the CI component are highlighted in this paper, beginning with a description of relevant data and methods, followed by a presentation of key results, and finishing with concluding remarks.

NUMERICAL MODELS, DIAGNOSTIC TOOLS, AND EXPERIMENTAL FORECASTS.

Models. The primary guidance tool for CI forecasting and diagnostics was the Advanced

AFFILIATIONS: Kain, Coniglio, and Ziegler—NOAA/National Severe Storms Laboratory, Norman, Oklahoma; Correia and Melick—NOAA/Storm Prediction Center, and Cooperative Institute for Mesoscale Meteorological Studies, Norman, Oklahoma; Clark and Lakshmanan—NOAA/National Severe Storms Laboratory, and Cooperative Institute for Mesoscale Meteorological Studies, Norman, Oklahoma; Marsh—NOAA/National Severe Storms Laboratory, and Cooperative Institute for Mesoscale Meteorological Studies, and School of Meteorology, University of Oklahoma, Norman, Oklahoma; Miller—School of Meteorology, University of Oklahoma, Norman, Oklahoma, and Civilian Institution Programs, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio; Dembek—Cooperative Institute for Mesoscale Meteorological Studies, Norman, Oklahoma; Weiss, Jirak, and Dean—NOAA/Storm Prediction Center, Norman, Oklahoma; Kong and Xue—School of Meteorology, University of Oklahoma, and Center for Analysis and Prediction of Storms, Norman, Oklahoma; Sobash—NOAA/National Severe Storms Laboratory, and School of Meteorology, University of Oklahoma, Norman, Oklahoma
CORRESPONDING AUTHOR: Dr. John S. Kain, National Severe Storms Laboratory, 120 David L. Boren Blvd., Norman, OK 73072
E-mail: jack.kain@noaa.gov

The abstract for this article can be found in this issue, following the table of contents.
DOI:10.1175/BAMS-D-11-00264.1

In final form 5 December 2012

Research Weather Research and Forecasting model (ARW-WRF) modeling framework (Skamarock et al. 2008). The configurations of two forecasting systems used during the experiment are highlighted below.

1) *NSSL-WRF*: SPC forecasters have used output from this single-member experimental modeling system, configured with 4-km grid spacing, since the fall of 2006. This system is run by NSSL once daily at 0000 UTC throughout the year over a full CONUS domain with forecasts to 36 h. Initial and lateral-boundary conditions (IC/LBCs) are provided by the operational North American Mesoscale model (NAM; Rogers et al. 2009). Forecast graphics are available online at www.nssl.noaa.gov/wrf/. This modeling system is used as the "alpha" testing framework for many diagnostic tools that are used in SFEs, as was the case for the CI diagnostics used in SFE2011. The model configuration is described in Kain et al. (2010).
2) *CAPS ensemble*: The University of Oklahoma Center for Analysis and Prediction of Storms (CAPS) generated numerical guidance using a 50-member storm-scale ensemble forecast (SSEF) system with grid spacing of 4 km, multiple dynamic cores, and forecasts to 36 h (0000 UTC initialization) covering a CONUS domain during SFE2011 (Kong et al. 2011). Only 29 members of the ensemble were used for the CI component of the experiment in order to focus on the ARW-WRF dynamic core, systematic variations in planetary boundary layer (PBL) parameterizations, and a core subensemble with perturbations in both IC/LBCs and model physics. Specifically, of these 29 members, 12 used identical IC/LBCs but different parameterizations of the PBL (hereafter the PBL members), while 18 were initialized with both IC/LBCs and physics perturbations (hereafter the core members). Note that the "control" member was included in both subsets.

Diagnostic tools. Unique model diagnostic tools were developed to identify convectively active (CA) grid points and to provide insight into simulated physical processes related to CI.

1) Identifying CA points. Prior to the start of SFE2011, three distinct definitions were developed to identify CA points objectively, at 5-min intervals, during integration of the models. One was based on simulated total lightning (McCaul et al. 2009), another was based on explicit kinematic properties and hydrometeor content of simulated updrafts, and a third was based on simulated reflectivity. All three definitions were used during SFE2011 and were found to be equally useful for identifying significant deep convective features, but only the simulated reflectivity approach was readily verifiable with directly analogous CONUS-scale observational datasets [Zhang et al. (2011) describe the observed-reflectivity dataset], so the other two are not discussed further here.

   For the reflectivity-based definition, a threshold value of 35 dB*Z* was used as the discriminator between convective and nonconvective points, following Roberts and Rutledge (2003), Mecikalski and Bedka (2006), and others. This threshold check was applied at the −10°C level to avoid brightbanding effects (see Gremillion and Orville 1999). The height of this temperature level was found by searching downward from the model top and interpolating within individual layers, using hourly temperature analyses from the Rapid Update Cycle (RUC) model (Benjamin et al. 2004) for observed reflectivity and individual CAM temperature fields for simulated reflectivity. Simulated reflectivity was computed as in Kain et al. (2008).
2) Model forecast soundings. Hourly soundings were generated for over 1100 locations across the CONUS for the PBL members of the CAPS ensemble. Each of these soundings was then processed through a custom-built version of the SPC's National Skew-*T*/Hodograph Analysis and Research Program (NSHARP; Hart and Korotky 1991) and output for each forecast location was postprocessed into an ensemble-data format readable by the National Weather Service Warning Decision Training Branch's Buffalo Toolkit for Lake Effect Snow (BUFKIT) sounding analysis tool (Mahoney and Niziol 1997). This tool enabled very efficient visualization of large quantities of sounding output with powerful user controls. For example, it allowed users to view forecast soundings from multiple members simultaneously (overlaid, color coded) at any of the preset model output times and locations.
3) Convective boundary layer kinematics and thermodynamics. Interactive analysis of full three-dimensional output grids from CAPS ensemble members was not possible during the real-time experiment, but convective boundary layer processes were sampled effectively in many situations by saving data from just one specific lower-tropospheric model level, located approximately 1.1 km above ground level. The

water vapor mixing ratio, temperature, pressure, and u, v, and w wind components from all ensemble members were saved at this level for every forecast hour. In addition, following Kain et al. (2010), diagnostic code was implemented to extract unique individual gridpoint data at this same vertical level, including the maximum vertical velocity in the last hour and the number of model time steps that boundary layer vertical velocity exceeded 0.25 m s^{-1} in the last hour. These data were designed to measure the maximum magnitude and persistence of boundary layer updrafts, respectively. The threshold of 0.25 m s^{-1} was chosen because empirical evidence in pre-experiment testing showed that it was useful for identifying the stronger, deeper mesoscale updrafts that were associated with surface-based CI in the CAMs. This threshold was effective but not necessarily optimal. Standard surface variables were also extracted and used to identify mesoscale features such as drylines, fronts, and cold storm-outflow boundaries.

Human forecasts for CI. During SFE2011, forecasters worked together with researchers and other participants, using model guidance, diagnostics, and all available observations to make experimental team forecasts for CI. These forecasts consisted of spatial categorical outlooks over limited regional domains for 3-h periods. The spatiotemporal domain was relocated each day with the intention of isolating "clean slate" environments (i.e., environments with no preexisting deep moist convection) in which specific anticipated CI events would occur. Within the selected spatial domain, any areas in which the likelihood of CI in the 3-h period was deemed to be greater than 10%[2] were outlined as categorical "slight" probability areas. Contours of moderate and high probability were added on some days to indicate areas of enhanced likelihood of CI, but these higher categories were not assigned specific quantitative probability levels for this preliminary effort. Each participant was also asked to predict the specific location where the first CI episode was most likely. Clusters of these points were highlighted with hatching in the forecast graphic.

The forecasts were typically issued in the morning, focusing on anticipated afternoon CI. Disciplined preparation of these experimental forecasts was a very valuable exercise because it catalyzed an open exchange of ideas among scientists, forecasters, and other participants. Furthermore, it effectively promoted a critical examination of model guidance and key considerations for defining the CI problem.

WHAT MAKES A SUCCESSFUL CAM FORECAST FOR CI? CAM output was used as guidance for daily experimental CI forecasts and subjective comparisons with observations were made following each event. The motivation for this effort was to develop metrics for identifying and measuring those characteristics of model forecasts that made them useful as guidance for CI. The ultimate goal is to develop ways to objectively quantify these characteristics so that benchmark measures of skill for CI prediction by CAMs can be established. Many of these characteristics are evident in the two example events highlighted below. The first is a clean slate event in which clearly discernible CI leads directly to a mesoscale convective event. A significant event occurs in the second case as well, but its evolution is more complex.

24 May 2011, Oklahoma. A significant severe weather outbreak was predicted for the eastern two-thirds of Oklahoma and surrounding areas on this day as exceptionally strong upper-level forcing for large-scale ascent, vertical wind shear, and convective instability moved over the area. The first storms formed along a dryline (not shown) in western and southwestern Oklahoma around 1900 UTC (Fig. 1a) and quickly evolved into discrete supercells. The dryline moved eastward and it focused meridional expansion of active CI episodes over the next two hours (Figs. 1d,g). Some cells consolidated and new cells continued to form for several hours thereafter and a quasi-linear broken line with embedded supercells was evident from south-central Kansas into north-central Texas by 2300 UTC (Figs. 1j,m) This sequence of observed radar images provides an excellent example of a "clean slate" convection initiation event with no preexisting storms that began with a singular instance of initiation followed in rapid succession by multiple additional initiation episodes along a common boundary, consolidation and upscale growth, and eventual emergence of a clearly identifiable mesoscale convective system (MCS).

This event was predicted fairly well by the CAMs used in the CAPS ensemble. Output fields from two of the PBL members were selected to highlight the comparison. The primary CI event in the member

[2] Specifically, the probability was defined as CI within 20 km of any point, to be consistent with the SPC's Thunderstorm Outlooks.

using the Yonsei University (YSU) PBL parameterization was about an hour late and too far east by several counties (cf. Figs. 1b,e and 1a,d), but the early evolution of the overall predicted convective event was quite similar to observations, showing a rapid formation of discrete storms both to the north and south along the dryline (cf. Figs. 1h,k,n and 1g,j,m). The member using the Mellor–Yamada–Janjić (MYJ) PBL also initiated the first storms a bit late. Furthermore, it suffered from a northeastward location bias (Figs. 1c,f) and differed in its prediction of the early evolution of the event. Specifically, the rapid expansion of storm coverage in the first two hours after initiation was conspicuously absent (Figs. 1i,l). However, by 2300 UTC this member did "fill in" along the dryline, suggesting the formation of a line of strong discrete storms, similar to observations and the forecast from the YSU member. Thus, the MYJ member also provided a strong signal for a significant convective event, but there was some ambiguity about when and where the "event" started in the forecast from this member because of the relatively discontinuous evolution between the first deep convective cells and subsequent widespread convective development.

The real-time experimental forecast for this event, valid for the 1900–2200 UTC time period, is shown in Fig. 2. It included a large area of high probability for CI, indicating high confidence that initiation would occur within the specified time–space window. The forecast team hedged westward and earlier in time compared to guidance from the CAPS ensemble, in response to model and observational data updates from the morning of the event.[3] This

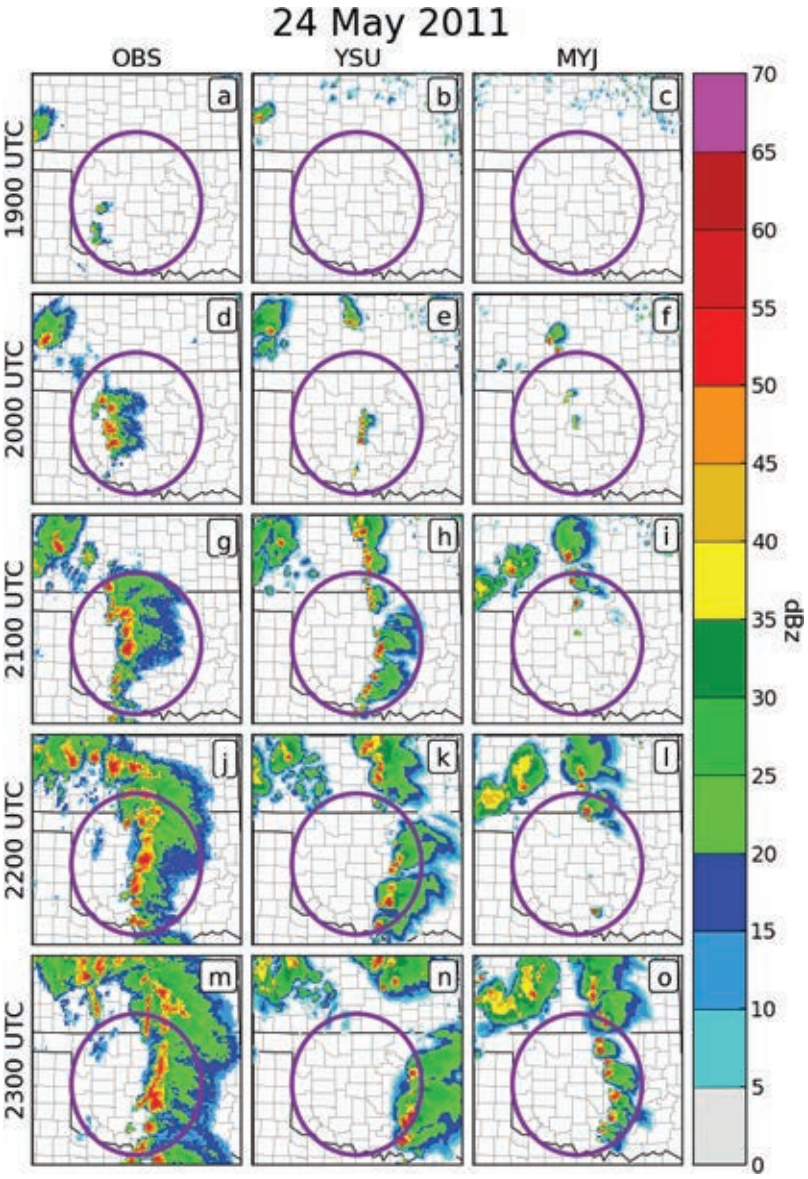

Fig. 1. Hourly composite reflectivity plots from (left) observations and (middle),(right) two PBL members from the CAPS ensemble for a CI event from 24 May 2011. The purple ellipse outlines a hypothetical "isolation domain" over which event-specific diagnostic analyses could be performed. The two PBL members used the YSU (Hong 2010) and MYJ (Janjić 1990) PBL parameterizations, respectively, but were otherwise configured identically.

[3] The forecasted dryline location from the 0000 UTC 24 May NAM (which provided time-dependent lateral boundary conditions for the CAMs) was considerably farther east than in the subsequent 1200 UTC 24 May NAM forecast and much farther east than the observed position (not shown). The dryline was the focusing mechanism for CI on this day; the CAMs were inherently biased toward the east in their CI guidance. This highlights an important challenge of forecasting CI probability from CAM guidance given the possible occurrence of forecast errors of dryline or other synoptic-scale boundary locations (e.g., Ziegler et al. 1997).

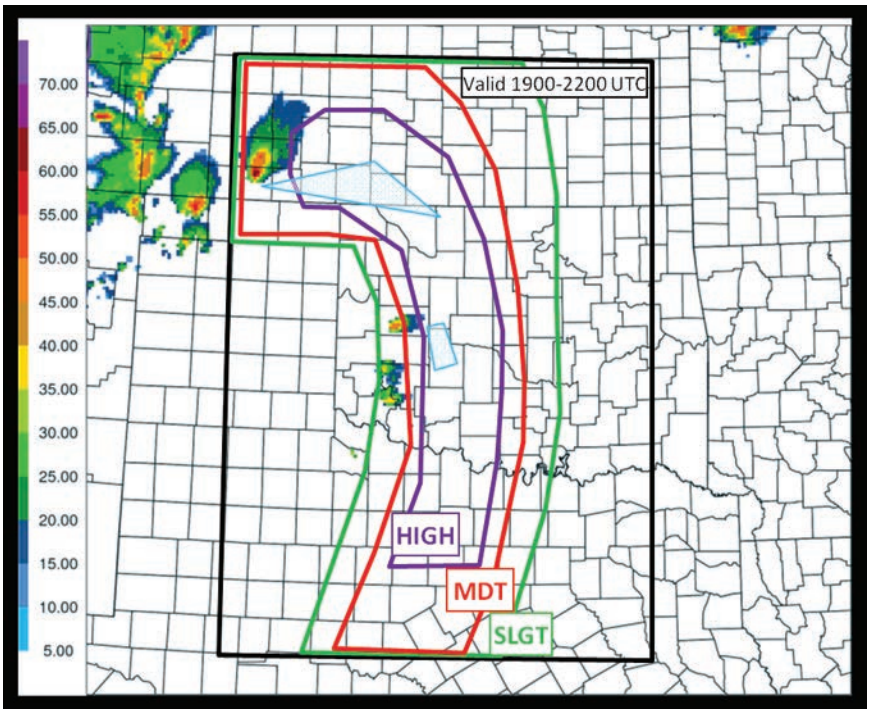

Fig. 2. Observed reflectivity at the –10°C level at 1901 UTC 24 May 2011 with the experimental CI forecast for the period 1900–2200 UTC overlaid. The thick black line outlines the intended boundaries of the forecast domain, and the green, red, and purple lines outline the slight, moderate, and high probability areas for CI, respectively. The hatched areas represent clusters of first CI locations predicted by individual members of the forecast team.

adjustment was in the right direction, but perhaps not bold enough—early CI episodes occurred to the west of clusters of predicted first CI points (see caption) in both the northern and central parts of the outlook area.

26 May 2011, mid-Atlantic region. The upper-level system that triggered severe weather over the southern plains on 24 May played a role in focusing storm development over the mid-Atlantic region two days later. However, the convective forcing and focusing mechanisms were not so sharply defined on this day. This assessment was reflected in the SPC 1300 UTC day 1 convective outlook (www.spc.noaa.gov/products/outlook/archive/2011/day1otlk_20110526_1300.html), which stated that, despite weak upper-level forcing, storms would likely form as daytime heating proceeded within a broad warm sector extending from the central Appalachians into upstate New York. Specific forcing mechanisms were expected to include low-level convergence along residual outflow boundaries from overnight convection and local terrain effects, but there was no discernible dryline and the main synoptic-scale surface cold front was well to the west of this area (not shown).

The CI focus area for this day extended from southeastern West Virginia northeastward to north-central Pennsylvania. CI was observed at multiple locations within this area between 1800 and 1900 UTC (Fig. 3a). Storms generally remained discrete and moved to the northeast for the next 1–2 h (Figs. 3d,g), but by 2100 UTC storms began to consolidate, especially from south-central Pennsylvania into southeastern West Virginia (Figs. 3g,j,m). By 2300 UTC 27 May, an intense line segment had taken shape in south-central Pennsylvania and northern Maryland, posing a substantial threat for severe weather and disruption to aviation over this region.

Output from the same two ensemble members examined in the first event reveals that both members initiated convection close to the high terrain near northeastern West Virginia and also over north-central Pennsylvania by 1900 UTC (Figs. 3b,c). Furthermore, they both initiated and maintained discrete storms over the northern half of the focus area through 2200 UTC (Figs. 3e,f,h,i,k,l). However, they failed to predict the discrete CI episodes over southeastern West Virginia and consequently failed to predict the convective evolution that led to the most significant feature observed in this time–space window—the intense line segment over south-central Pennsylvania and northern Maryland (Figs. 3n,o). Interestingly, both members eventually (after 0000 UTC; not shown) produced intense convective lines that moved from Pennsylvania into New York, but these lines appeared to emanate from convection directly linked to the synoptic-scale boundary rather than the discrete initiation events that actually occurred well into the warm sector. These differences are significant not only from a timing and evolution perspective, but also because the model solutions imply a lower threat of supercellular convection owing to their apparent association with linear (i.e., frontal) forcing.

In this case the models appeared to do well in predicting the timing of the first CI episodes in the northern half of the region—and the isolated character of these storms. However, they missed important initiation episodes farther south, and it was the activity associated with these episodes that eventually grew upscale and became the dominant convection feature. Clearly, there were significant errors in the details of the model convection forecasts, but it is not clear what specific metrics might help to quantify the relevant errors.

Lessons learned. These two cases highlight some of the important considerations for objectively measuring the success of numerical forecasts of CI. Specifically, they demonstrate that accurate predictions of the timing and location of specific CI episodes are certainly desirable, but the timing and location of the first convective cells tell only part of the story. Perhaps even more important is the subsequent evolution of activity, including changes in areal coverage, intensity, and convective mode, particularly if these changes involve upscale growth of activity into organized MCSs.

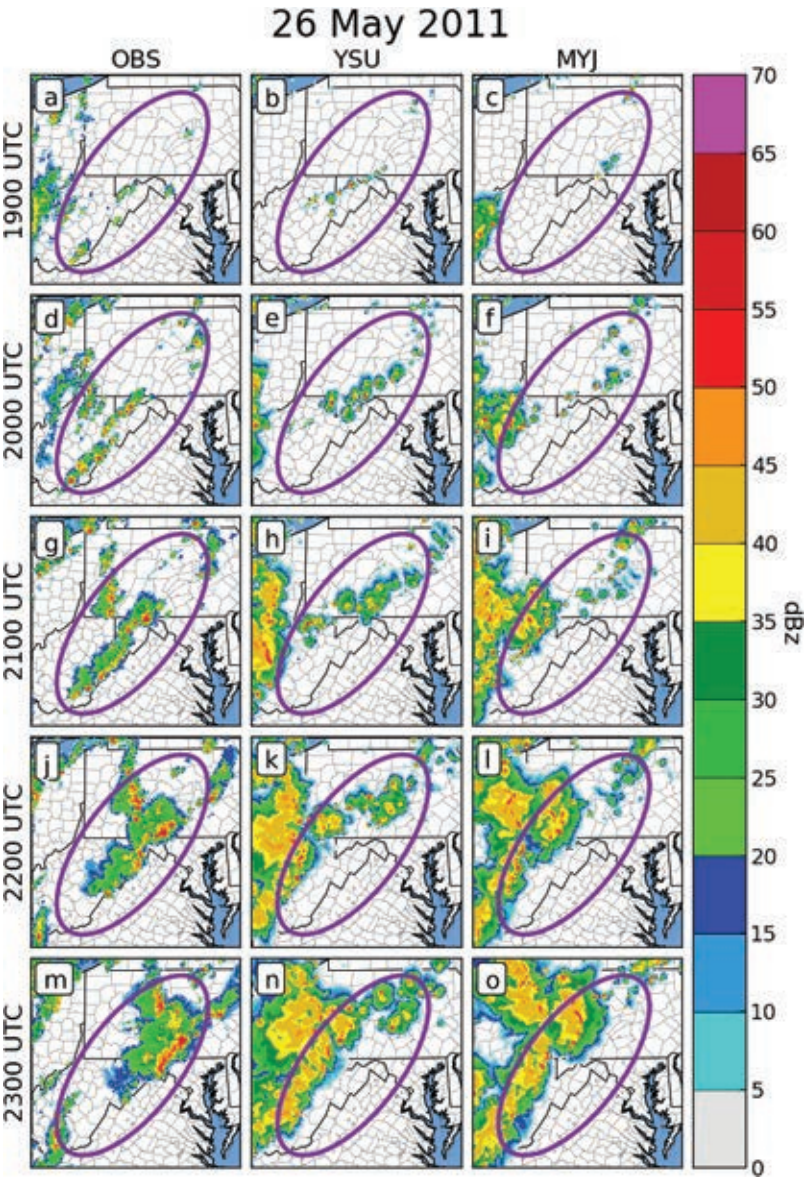

Fig. 3. As in Fig. 1, but for 26 May 2011.

A PRELIMINARY ESTIMATE OF SKILL IN ENSEMBLE FORECASTS FOR CI.

Most meteorologists would say they know CI when they see it, but as suggested above, objective quantification of skill in CI forecasts is challenging. A relatively simple verification framework was established based on lessons learned during SFE2011 to provide a preliminary estimate of timing skill in CAM-based ensemble forecasts. In this framework, an event-specific spatiotemporal window was specified to isolate a particular observed CI event from other convective activity. Spatially, this involved subjective drawing of an isolation domain, based on the location and early evolution of a specific observed convective event. Representative examples of these isolation domains are shown in Figs. 1 and 3.

Comparably sized mesoscale domains were drawn to isolate 14 other events from the SFE2011 time period, for a total sample size of 16 events. These events were scattered over the eastern two-thirds of the CONUS and all occurred during the late morning/afternoon time period (15–24-h forecasts for the CAPS ensemble). The specific measured quantity in this evaluation was the time of first CI in each model forecast (within the limited domain) relative to observed CI. The search encompassed a time window covering ±5 h of each observed CI event.

CI points were identified as a subset of CA points using an object-based time-domain methodology similar to the one described in Clark et al. (2012b). Specifically, in postanalysis of individual events,

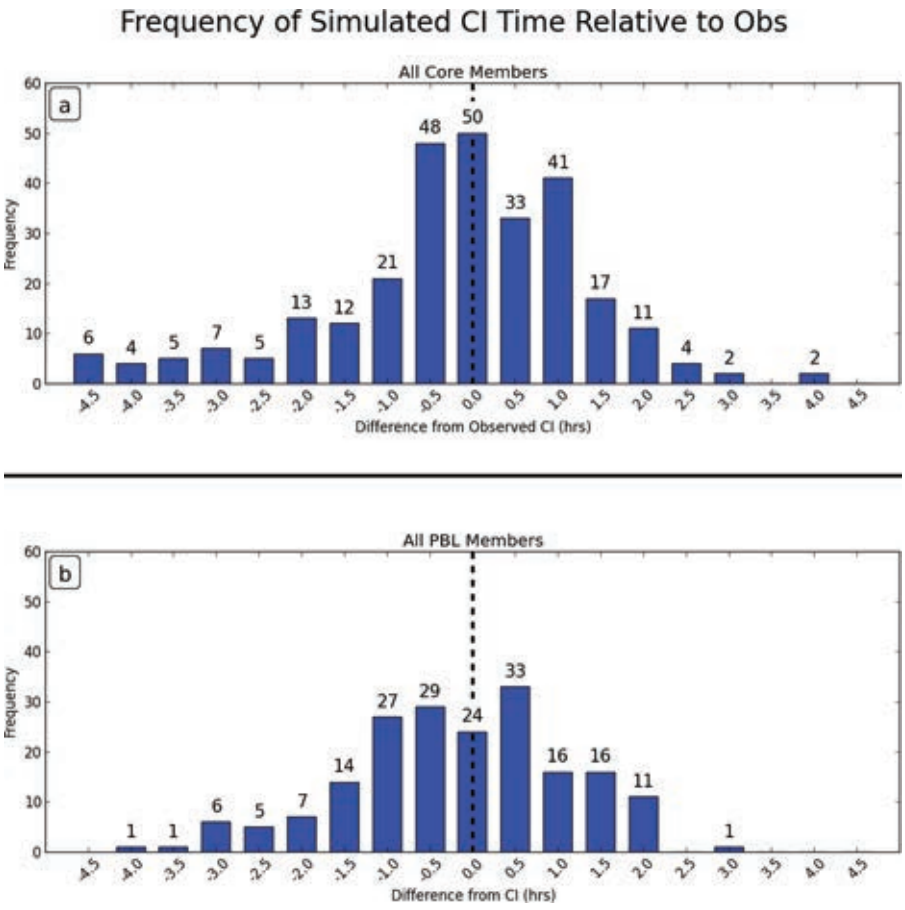

Fig. 4. Frequency histograms of the difference between the CI time predicted by CAPS ensemble members and the observed time, for (a) the core 18 ARW-WRF members and (b) the PBL members. The data are based on 16 selected events from SFE2011.

time-domain CA objects were defined as groups of CA points contiguous in space and time. For example, adjacent grid points became part of a single CA object if they were classified as CA points at the same sample time, immediately preceding the sample time, or immediately following the sample time. Objects that did not span a time period greater than 30 min were not considered. Within each CA object, CI points were identified as local time minima (i.e., CA points without adjacent CA points at earlier times). This algorithm appeared to successfully identify CI points associated with any storms lasting greater than 30 min. It was applied to model output on the native grid and observed reflectivity interpolated to the same grid.

The probability of detection (POD) for CI was very high for these 16 events. Specifically, CI was detected in 456 out of 464 model forecasts (16 events, 29 members), within the specified time–space window, for a POD of 0.98.

The difference in first CI time between each model forecast and the corresponding observation was determined for all events, and a frequency histogram of the differences was constructed. For the 18 core ensemble members, CI times were clustered within about ±2 h of the observed time, with a few outliers, especially on the early side of the histogram (Fig. 4a). The 12 PBL members appeared to be clustered more tightly within the same time window and with fewer outliers (Fig. 4b).

These results are encouraging, suggesting that *when CI occurred in both models and observations* there was very little systematic bias in CI timing for the 4-km CAPS ensemble. The variance in timing was substantial, but the temporal histogram likely provides at least an approximate reflection of true temporal uncertainty in CI and this concept may prove to be useful for developing probabilistic guidance for the timing of CI, and perhaps of other processes as well.

ENVIRONMENTAL PRECURSORS TO CI IN CAMS. Surface-based[4] CI is strongly modulated by a multitude of PBL processes and constrained by thermodynamic profiles within and near the top of the PBL. As described previously, numerous diagnostic tools were developed for SFE2011 to help visualize relevant processes in model output. One set of diagnostic fields was based on model-predicted mass and momentum fields on a fixed sigma level corresponding to about 1.1 km AGL. This level was

[4] A storm is often referred to as "surface based" (as opposed to "elevated") if the air parcels that feed its main updraft originate in a convective planetary boundary layer.

typically within or near the top of the PBL during the afternoon heating cycle within the focus areas for experimental forecasts. Examination of these fields often allowed one to infer important information about PBL processes related to model CI. For example, judicious presentation of vertical velocity and moisture fields can indicate the presence of horizontal convective rolls within the boundary layer, transverse rolls near the top of the boundary layer, and the correlation between the upward-motion branch of these rolls and the moisture field (Fig. 5). A causal relationship between these features and specific CI events has been suggested by observational studies (e.g., Ziegler and Rasmussen 1998; Wilson et al. 1992; Ziegler et al. 2007) and an analogous association was often discernible in the model output during the experiment.

Examination of model-predicted sounding structures revealed that temperature and moisture profiles often varied significantly within the simulated late-day convective boundary layer (e.g., Fig. 6). Over the course of SFE2011, there was a clear trend for certain PBL members of the ensemble to produce relatively cool and moist boundary layer structures, while others were systematically warmer and drier. The process of quantifying these biases and explaining their association with specific PBL parameterizations is underway (Coniglio et al. 2013). CI is undoubtedly sensitive to these details in sounding structure.

Perhaps the most intriguing aspect of these results is that, in individual CAM simulations during SFE2011, the combination of sounding evolution and dry dynamic processes—including subcloud-layer updrafts with a seemingly appropriate combination of persistence, penetration depth, and intensity—allowed some parcels to reach their level of free convection and start the CI process at approximately the right time and place. Like the timing statistics, these results are encouraging. Nonetheless, they only scratch the surface of our efforts to understand the CI process in CAMs more generally because, for example, they are derived

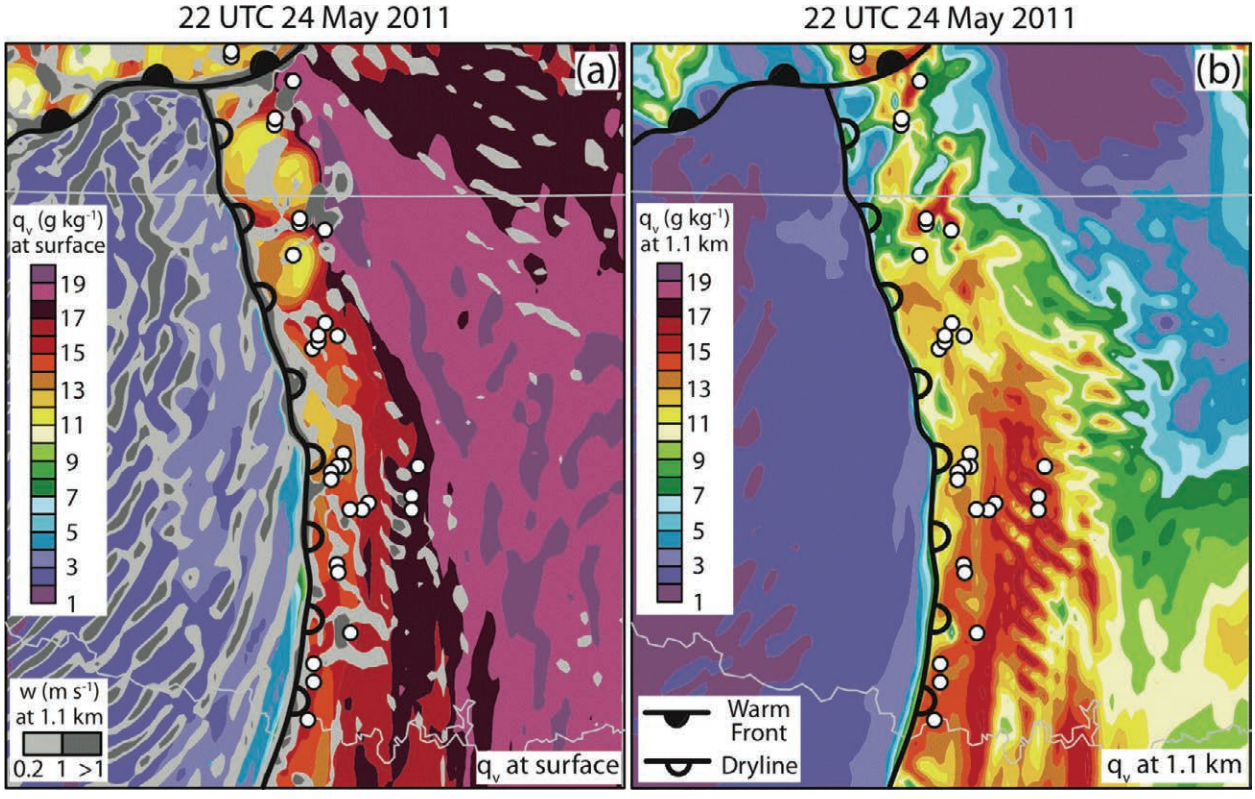

Fig. 5. Sample diagnostic plot from the CAPS ARW-WRF ensemble control member valid at 2200 UTC 24 May 2011 showing diagnosed CI points within ±30 min of the plot time (white-filled black circles) and water vapor mixing ratio (color fill) at (a) the surface and (b) a model level at approximately 1.1 km AGL. In (a), updraft at 1.1 km AGL is indicated by gray fill. Note the HCR-like features in the dry air west of the dryline, the pronounced (>1 m s^{-1}) updrafts along the dryline and warm front, and the elevated transverse rolls in stable air at the top of the PBL to the east of the dryline in central Oklahoma. The CI points are associated with the combination of dryline- and transverse-roll updrafts with deeply lifted, large PBL water vapor mixing ratios.

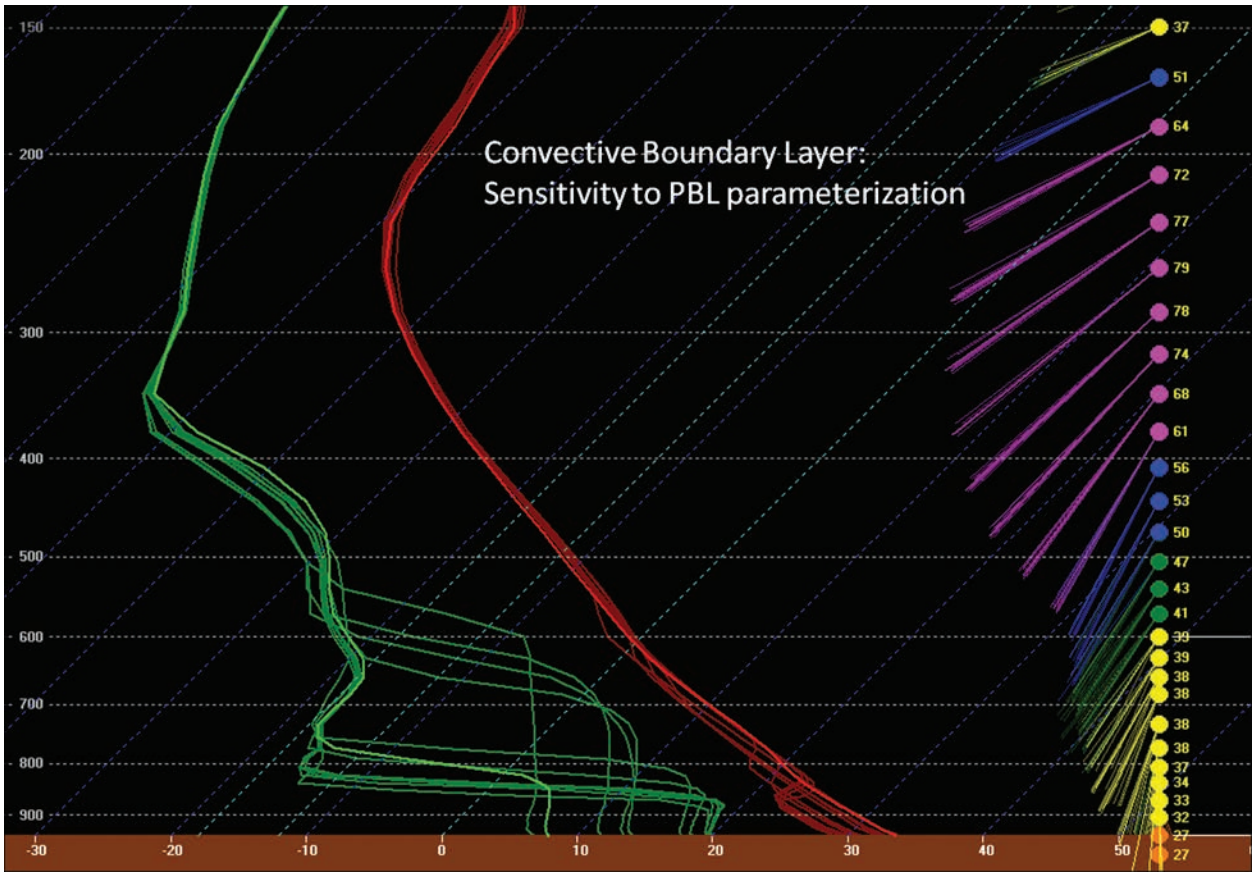

Fig. 6. Forecast soundings valid at a single time and location from each of the PBL members from the CAPS ensemble.

from a limited set of environmental conditions that are commonly observed over the eastern two-thirds of the U.S. landmass in the springtime; most of the sampled events were strongly forced and all of them were surface-based; and the results are more qualitative than quantitative. Much more work is needed to gain complete insight into how sounding structure and dynamic PBL processes modulate the CI process in CAMs.

SUMMARY. A primary objective of the CI component of SFE2011 was to examine the potential utility of CAMs with ~4-km grid spacing in providing guidance for CI forecasts. The models proved to be quite useful in spite of the fact that many convective and preconvective processes were coarsely represented with this grid spacing. For example, in the model forecasts, the pattern and orientation of updrafts rooted in the convective boundary layer often implied the presence of HCR-like features and local airmass boundaries that are known to play a role in the CI process. Because of the relatively coarse model resolution, the scale of these features was likely too large, but their mere presence and association with CI inspired confidence among SFE2011 participants that the CAMs were able to represent physical processes known to be important for CI.

The models also showed promise in predicting the timing of CI over targeted mesoscale areas, particularly when a probabilistic, ensemble-based approach was considered. Specifically, for those events in which convective cells were both observed by radar and predicted by a 4-km ARW-WRF ensemble, there appeared to be no systematic ensemble bias in the timing of first cell initiation. Frequency histograms of timing differences between observations and model predictions, aggregated over many events, had a quasi-Gaussian distribution, suggesting that the ensemble could provide useful probabilistic forecast guidance on the timing of CI with minimal calibration.

Before the start of SFE2011, a considerable amount of effort was spent developing prototype algorithms that first identified convectively active grid points and then found the subset of these points that corresponded to CI. Only one of these algorithms is described here, but they all proved to be very useful for automating the identification of "new" deep

convective cells (i.e., cells that are not directly associated with ongoing deep convection) in either model output or radar-based observations.

Identification of these cells is important, but the consensus that emerged from the experiment was that the algorithms were often inadequate indicators of impending hazardous or disruptive weather. In the springtime when the SFE is conducted, patterns of deep convection can be very complex, with merging and splitting of convective activity on multiple scales and convective instability emanating from multiple layers in the atmosphere. During SFE2011 CI algorithms often flagged initiation episodes related to weak, elevated convection in the warm sector of larger-scale systems, but these flags were false alarms for the initiation of the stronger, surface-based convective events that forecast teams were targeting. When the algorithms were applied to CAM output, interpretation was further complicated by the fact that, while CAMs are known to have skill in predicting the occurrence of some significant convective events (e.g., Done et al. 2004), they commonly suffer from substantial errors in timing, location, orientation, and other system attributes in the prediction of such events—and, of course, they totally miss some events.

Since algorithms that identify local CI processes are useful but not sufficient for predicting the disruptive potential of convective storms and larger-scale systems, perhaps future CI algorithm development should be couched in terms of feature-specific prediction (Carley et al. 2011). In this context the timing and location of initiation are only two of many relevant attributes of convective features. These features could be individual convective cells or they could be scale-selective envelopes of convective activity. For example, other *convective cell* attributes might include size, intensity, rotation characteristics, longevity, and movement, while relevant *convective system* attributes could include initiation, porosity, expansion rate, orientation, movement, etc. It is anticipated that development and testing of feature-specific prediction algorithms will be an important part of future SFEs in the NOAA HWT.

ACKNOWLEDGMENTS. Many individuals contributed to the success of the CI component of SFE2011, including Gregg Grosshans, Jay Liang, Joe Byerly, Linda Crank, and Peggy Stogsdill (SPC); and Brett Morrow, Jeff Horn, Steve Fletcher, Brad Sagowitz, Greg Trotter, and Linda McGuckin (NSSL).

Authors Clark, Correia, Dembek, Lakshmanan, Marsh, Melick, Miller, and Sobash were partially funded by the NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma Cooperative Agreement NA17RJ1227, U.S. Department of Commerce.

CAPS ensemble forecasts were supported by the NOAA Collaborative Science, Technology, and Applied Research (CSTAR) Program with supplementary support from NSF Grant AGS-0802888. M. Xue was also supported by NSF Grants OCI-0905040, AGS-0941491, AGS-1046171, and AGS-1046081. CAPS forecasts were supported National Science Foundation XSEDE grant allocation, and computations were performed at the National Institute for Computational Sciences. Many scientists at CAPS contributed to the design and production of the ensemble forecasts, including Kevin Thomas, Yunheng Wang, Keith Brewster, Jidong Gao, and Xuguang Wang.

REFERENCES

Barthlott, C., and Coauthors, 2011: Initiation of deep convection at marginal instability in an ensemble of mesoscale models: A case-study from COPS. *Quart. J. Roy. Meteor. Soc.,* **137,** 118–136, doi:10.1002/qj.707.

Behrendt, A., and Coauthors, 2011: Observation of convection initiation processes with a suite of state-of-the-art research instruments during COPS IOP 8b. *Quart. J. Roy. Meteor. Soc.,* **137,** 81–100, doi:10.1002/qj.758.

Benjamin, S. G., and Coauthors, 2004: An hourly assimilation–forecast cycle: The RUC. *Mon. Wea. Rev.,* **132,** 495–518.

Browning, K. A., and Coauthors, 2007: The Convective Storm Initiation Project. *Bull. Amer. Meteor. Soc.,* **88,** 1939–1955.

Carley, J. R., B. R. J. Schwedler, M. E. Baldwin, R. J. Trapp, J. Kwiatkowski, J. Logsdon, and S. J. Weiss, 2011: A proposed model-based methodology for feature-specific prediction for high-impact weather. *Wea. Forecasting,* **26,** 243–249.

Clark, A. J., W. A. Gallus, M. Xue, and F. Kong, 2009: A comparison of precipitation forecast skill between small convection-allowing and large convection-parameterizing ensembles. *Wea. Forecasting,* **24,** 1121–1140.

——, ——, ——, and ——, 2010a: Growth of spread in convection-allowing and convection-parameterizing ensembles. *Wea. Forecasting,* **25,** 594–612.

——, ——, ——, and ——, 2010b: Convection-allowing and convection-parameterizing ensemble forecasts of a mesoscale convective vortex and associated severe weather environment. *Wea. Forecasting,* **25,** 1052–1081.

——, and Coauthors, 2012a: An overview of the 2010 Hazardous Weather Testbed Experimental Forecast Program Spring Experiment. *Bull. Amer. Meteor. Soc.,* **93,** 55–74.

——, J. S. Kain, P. T. Marsh, J. Correia Jr., M. Xue, and F. Kong, 2012b: Forecasting tornado path lengths using a three-dimensional object identification algorithm applied to convection-allowing forecasts. *Wea. Forecasting,* **27,** 1090–1113.

Coniglio, M. C., K. L. Elmore, J. S. Kain, S. J. Weiss, M. Xue, and M. L. Weisman, 2010: Evaluation of WRF model output for severe weather forecasting from the 2008 NOAA Hazardous Weather Testbed Spring Experiment. *Wea. Forecasting,* **25,** 408–427.

——, J. Correia Jr., P. T. Marsh, and F. Kong, 2013: Verification of convection-allowing WRF model forecasts of the planetary boundary layer using sounding observations. *Wea. Forecasting,* in press.

Corfidi, S. F., S. J. Corfidi, and D. M. Schultz, 2008: Elevated convection and castellanus: Ambiguities, significance, and questions. *Wea. Forecasting,* **23,** 1280–1303.

Crook, N. A., 1996: Sensitivity of moist convection forced by boundary layer processes to low-level thermodynamic fields. *Mon. Wea. Rev.,* **124,** 1767–1785.

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,** 110–117.

Doswell, C. A., III, 2001: Severe convective storms—An overview. *Severe Convective Storms, Meteor. Monogr.,* No. 50, Amer. Meteor. Soc., 1–26.

Gremillion, M. S., and R. E. Orville, 1999: Thunderstorm characteristics of cloud-to-ground lightning at the Kennedy Space Center, Florida: A study of lightning initiation signatures as indicated by the WSR-88D. *Wea. Forecasting,* **14,** 640–649.

Griffiths, M., A. J. Thorpe, and K. A. Browning, 2000: Convective destabilization by a tropopause fold diagnosed using potential-vorticity inversion. *Quart. J. Roy. Meteor. Soc.,* **126,** 125–144.

Groenemeijer, P., and Coauthors, 2009: Observations of kinematics and thermodynamic structure surrounding a convective storm cluster over a low mountain range. *Mon. Wea. Rev.,* **137,** 585–602.

Hart, J. A., and W. D. Korotky, 1991: The SHARP Workstation v. 1.50: A skew *T*/hodograph analysis and research program for the IBM and compatible PC: User's manual. NOAA/NWS Forecast Office, 62 pp.

Hong, S.-Y., 2010: A new stable boundary-layer mixing scheme and its impact on the simulated East Asian summer monsoon. *Quart. J. Roy. Meteor. Soc.,* **136,** 1481–1496.

Houston, A. L., and D. Niyogi, 2007: The sensitivity of convective initiation to the lapse rate of the active cloud-bearing layer. *Mon. Wea. Rev.,* **135,** 3013–3032.

Janjić, Z. A., 1990: The step-mountain coordinate: Physical package. *Mon. Wea. Rev.,* **118,** 1429–1443.

Kain, J. S., and Coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting,* **23,** 931–952.

——, 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,* **25,** 1536–1542.

Koch, S. E., and C. O'Handley, 1997: Operational forecasting and detection of mesoscale gravity waves. *Wea. Forecasting,* **12,** 253–281.

Kong, F., and Coauthors, 2011: CAPS storm-scale ensemble forecasts in NOAA HWT 2011 Spring Experiment: Sensitivities of WRF physics on QPF. *Extended Abstracts, 12th WRF Users' Workshop*, Boulder, CO, NCAR. [Available online at www.mmm.ucar.edu/wrf/users/workshops/WS2011/Power%20Points%202011/6_3_Kong_WRFworkshop_11.pdf.]

Mahoney, E. A., and T. A. Niziol, 1997: BUFKIT: A software application tool kit for predicting lake-effect snow. Preprints, *13th Int. Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology,* Long Beach, CA, Amer. Meteor. Soc., 388–391.

Markowski, P., and Y. Richardson, 2010: *Mesoscale Meteorology in Midlatitudes*. Wiley-Blackwell, 407 pp.

Marsh, P. T., J. S. Kain, V. Lakshmanan, A. J. Clark, N. M. Hitchens, and J. Hardy, 2012: A method for calibrating deterministic forecasts of rare events. *Wea. Forecasting,* **27,** 531–538.

McCaul, E. W., S. J. Goodman, K. M. LaCasse, and D. J. Cecil, 2009: Forecasting lightning threat using cloud-resolving model simulations. *Wea. Forecasting,* **24,** 709–729.

Mecikalski, J. R., and K. M. Bedka, 2006: Forecasting convective initiation by monitoring the evolution of moving cumulus in daytime GOES imagery. *Mon. Wea. Rev.,* **134,** 49–78.

Roberts, R. D., and S. Rutledge, 2003: Nowcasting storm initiation and growth using *GOES-8* and WSR-88D data. *Wea. Forecasting,* **18,** 562–584.

Rogers, E., and Coauthors, 2009: The NCEP North American Mesoscale modeling system: Recent changes and future plans. Preprints, *23rd Conf. on Weather Analysis and Forecasting/19th Conf. on Numerical Weather Prediction*, Omaha, NE, Amer. Meteor. Soc., 2A.4. [Available online at https://ams.confex.com/ams/pdfpapers/154114.pdf.]

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech

Note NCAR/TN-475+STR, 113 pp. [Available online at www.mmm.ucar.edu/wrf/users/docs/arw_v3.pdf.]

Sobash, R A., J. S. Kain, D. R. Bright, A. R. Dean, M. C. Coniglio, and S. J. Weiss, 2011: Probabilistic forecast guidance for severe thunderstorms based on the identification of extreme phenomena in convection-allowing model forecasts. *Wea. Forecasting,* **26,** 714–728.

Weckwerth, T. M., 2000: The effect of small-scale moisture variability on thunderstorm initiation. *Mon. Wea. Rev.,* **128,** 4017–4030.

——, and R. M. Wakimoto, 1992: The initiation and organization of convective cells atop a cold-air outflow boundary. *Mon. Wea. Rev.,* **120,** 2169–2187.

——, and D. B. Parsons, 2006: A review of convection initiation and motivation for IHOP_2002. *Mon. Wea. Rev.,* **134,** 5–22.

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.

Wilson, J. W., and W. E. Schreiber, 1986: Initiation of convective storms at radar-observed boundary-layer convergence lines. *Mon. Wea. Rev.,* **114,** 2516–2536.

——, and R. D. Roberts, 2006: Summary of convective storm initiation and evolution during IHOP: Observational and modeling perspective. *Mon. Wea. Rev.,* **134,** 23–47.

——, G. B. Foote, N. A. Crook, J. C. Fankhauser, C. G. Wade, J. D. Tuttle, C. K. Mueller, and S. K. Krueger, 1992: The role of boundary-layer convergence zones and horizontal rolls in the initiation of thunderstorms: A case study. *Mon. Wea. Rev.,* **120,** 1785–1815.

Zhang, J., and Coauthors, 2011: National Mosaic and Multi-Sensor QPE (NMQ) system: Description, results, and future plans. *Bull. Amer. Meteor. Soc.,* **92,** 1321–1338.

Ziegler, C. L., and E. N. Rasmussen, 1998: The initiation of moist convection at the dryline: Forecasting issues from a case study perspective. *Wea. Forecasting,* **13,** 1106–1131.

——, T. Lee, and R. Pielke, 1997: Convective initiation at the dryline: A modeling study. *Mon. Wea. Rev.,* **125,** 1001–1026.

——, E. Rasmussen, M. Buban, Y. Richardson, L. J. Miller, and R. Rabin, 2007: The "triple point" on 24 May 2002 during IHOP. Part II: Ground-radar and in situ boundary layer analysis of cumulus development and convection initiation. *Mon. Wea. Rev.,* **135,** 2443–2472.