12.1 SEVERE-WEATHER FORECAST GUIDANCE FROM THE FIRST GENERATION OF LARGE DOMAIN CONVECTION-ALLOWING MODELS: CHALLENGES AND OPPORTUNITIES

John S. Kain*[,@], S. J. Weiss[@@], S. R. Dembek[†,‡], J. J. Levit[@@], D. R. Bright[@@], J. L. Case[††], M. C. Coniglio[@], A. R. Dean[‡,**], R. A. Sobash[‡‡], and C. S. Schwartz[‡‡]

[@]NOAA/OAR/National Severe Storms Laboratory, Norman, OK
[@@]NOAA/NWS/NCEP/Storm Prediction Center, Norman, OK
[†]Universities Space Research Association/SPoRT Center, Huntsville, AL
[††]ENSCO Inc./Short-term Prediction Research and Transition (SPoRT) Center, Huntsville, AL
[‡]Cooperative Institute for Mesoscale Meteorological Studies, Norman, OK
[‡‡]School of Meteorology, University of Oklahoma, Norman, OK

1. INTRODUCTION

Since the inception of operational numerical weather prediction (NWP), convective parameterization has been an integral component of prediction systems because computational constraints have precluded the routine use of models with resolution high enough to resolve convective processes explicitly. In recent years, this has begun to change and operational NWP centers around the world have started implementing high-resolution regional models with either no convective parameterization (e.g., Siefert et al. 2008; Weiss et al. 2008) or a relaxed form of parameterization that is much less heavy-handed than traditional approaches (e.g., Narita and Ohmori 2007; and Lean et al. 2008). This trend represents a significant transition in the evolution of operational NWP.

At the NOAA[1] Hazardous Weather Testbed (HWT) in Norman, OK, as part of a collaborative initiative of the NOAA/OAR[2]/National Severe Storms Laboratory (NSSL) and the NOAA/NWS[3]/NCEP[4]/Storm Prediction Center (SPC), we have been actively involved in evaluating NWP models without parameterized convection (hereafter collectively referred to as CAMs - convection-allowing models) since 2003. For example, we have conducted annual 6-8 week experiments known as the HWT Spring

Experiment in which evaluation of CAMs has been a primary focus. These experiments benefit from active participation by forecasters, model developers, research scientists, and university faculty who have a passion for operationally relevant meteorological challenges (see Kain et al. 2006 and http://www.nssl.noaa.gov/hwt). Moreover, we have worked closely with severe-convection forecasters at the SPC, not only during spring experiments, but throughout the year. The SPC has been receiving experimental CAM forecasts from NCEP/Environmental Modeling Center) on a daily basis since 2004 (hereafter WRF-EMC, where WRF refers to the Weather Research and Forecasting model), and this output has been closely interrogated every day as part of the decision making process at the SPC. Finally, at NSSL, we have been running a near CONUS[5]-scale CAM on a daily basis since late 2006 (hereafter WRF-NSSL). Output from these forecasts has been ingested into the SPC operational data stream and has become another important piece of information in the operational forecast-preparation process. The output has been evaluated internally at the NSSL as well.

These HWT activities have been coordinated with additional testing and evaluation efforts in Huntsville, AL, where scientists from NASA's Short-term Prediction Research and Transition (SPoRT) Center seek to accelerate the infusion of modeling research into operational forecasting, much like their HWT counterparts. SPoRT scientists are experimenting with a configuration of the WRF-NSSL modeling domain in which NASA physical parameterizations and input data sources are tested for selected cases from the 2007 and 2008 Spring Experiments (Case et al. 2008a). The physics schemes being tested within the WRF-NSSL include a revised Goddard short-wave and new long-wave radiation parameterizations

*Corresponding Author Address:
Jack Kain
NSSL
120 David L. Boren Blvd.
Norman, OK 73072

[1] National Oceanic and Atmospheric Administration
[2] Office of Atmospheric Research
[3] National Weather Service
[4] National Centers for Environmental Prediction

[5] CONtinental U. S.

not currently available in the public release of WRF (Matsui et al. 2007; Chou and Suarez 1999; Chou et al. 2001), and the new Goddard 3-ice microphysics scheme available in WRF version 3 (Tao et al. 2003). In addition, the impacts of high-resolution lower boundary data from the NASA Land Information System (LIS; Kumar et al. 2006, 2007; Case et al. 2008b) and a SPoRT Center high-resolution sea surface temperature composite derived from the Moderate Resolution Imaging Spectroradiometer instruments aboard the Aqua and Terra satellites (Haines et al. 2007; LaCasse et al. 2008) are being evaluated.

These collaborative efforts provide us with the basis for a thorough, if somewhat subjective, assessment of the utility of CAMs as guidance for severe weather forecasters. However, although we draw from an assessment period of several years and a diverse range of expertise (i.e., forecasters, model developers, etc.), it is important to acknowledge that, from a global perspective, the context of this assessment might be considered rather limited. For example, it is based on:

1) The WRF (Weather Research and Forecasting) modeling framework: Our experiments involve different dynamic cores within this framework (e.g., see Skamarock et al. 2005 for a description of the ARW dynamic core and Janjic 2005 for a description of the NMM dynamic core), as well as multiple physical parameterizations, but it is likely that this modeling system has characteristic behaviors that are quite different from those used at various operational and research centers around the globe.

2) "Cold-start" initialization: The CAM forecast systems that we have analyzed have typically been initialized by simply interpolating data from coarse resolution operational models.

	WRF-EMC	WRF-NSSL
Horiz. Grid (km)	4.0	4.0
Vertical Levels	35	35
PBL/Turb. Param.	MYJ	MYJ
Microphys. Param	Ferrier	WSM6
Radiation (SW/LW)	GFDL/GFDL	Dudhia/RRTM
Initial Conditions	32 km NAM	40 km NAM
Dynamic Core	NMM	ARW

Table 1. Model configurations for the daily WRF-EMC and WRF-NSSL forecasts. NAM is the operational North American Mesoscale model run by NCEP/EMC. Descriptions of the different WRF physical parameterizations can be found at URL http://www.mmm.ucar.edu/wrf/users/docs/user_guide_V3/users_guide_chap5.htm#Phys

With this approach, the first 6-12h of the forecast are considered the "spin up" period, during which smaller scale structures develop as a function of meso and larger scale features in the initial conditions. Most of the evaluation presented herein involves these cold-start forecasts, but in 2008 we evaluated a relatively limited number of forecasts that were initialized with radar data, as described later in this paper.

3) Next-day model forecasts: For the most part, we have examined model predictions for the 18-36h time period (i.e., next afternoon and night) from models initialized at 0000 UTC. This approach implicitly acknowledges that the solution will be strongly modulated by initial and lateral-boundary conditions from which the model is initialized (Weisman et al. 2008).

4) Forecasts for the eastern two-thirds of the CONUS: Computational constraints continue to limit the size of high-resolution domains, so this region has been selected because of its high frequency of severe convection and other types of
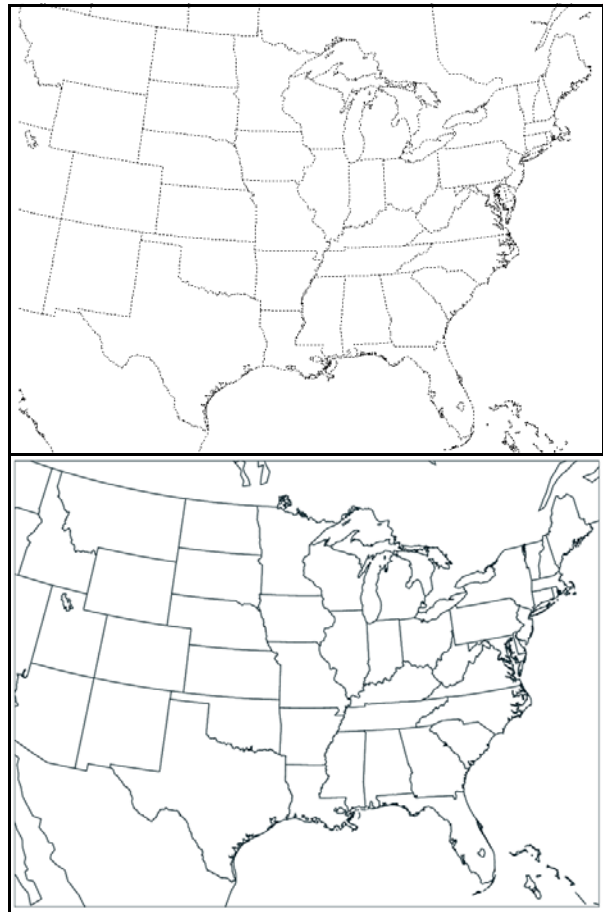

Fig. 1. Computational domains for the spring 2008 WRF-EMC and WRF-NSSL daily forecasts.

hazardous weather. However, environmental fields in this part of the world are climatologically different from other areas on the globe, so some important characteristics of model behavior over this region may not be as relevant to model performance in other areas. Also, localized orographic and land-sea contrasts are less important over this region than in many parts of the world.

These considerations are necessary caveats to keep in mind as the results are considered. But, on a positive note, they also limit the degrees for freedom, which can make it easier to untangle cause and effect relationships in evaluation of models.

The purpose of this paper is to draw upon our experiences at the HWT, in addition to coordinated efforts with NASA/SPoRT, to present some challenges and opportunities as we wade deeper into the operational usage of CAMS. It draws heavily from HWT Spring Experiments in 2008, 2007, and 2005 (hereafter SE2008, SE2007, and SE2005, respectively). This paper is presented as
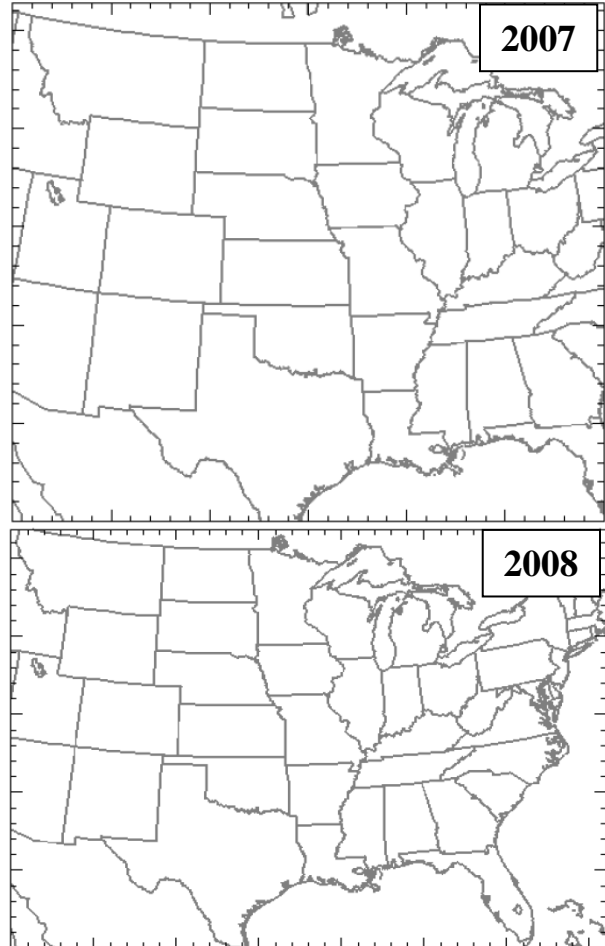

Fig. 2. Computational domains for the CAPS WRF-ARW forecasts from 2007 and 2008

a less formal complement to other publications on this topic (e.g., Weisman et al. 2008; Kain et al. 2006, and others)

2. MODELING SYSTEMS

2.1 *WRF-EMC and WRF-NSSL*

As described in the introduction, hourly output from 0000 UTC initializations of the WRF-EMC and WRF-NSSL configurations are used by SPC forecasters every day. The two forecasts cover nearly the same domain (Fig. 1), but they differ significantly in their physical parameterizations and dynamic core (Table 1.). Both forecasts integrate out to 36h.

2.2 *WRF-ARW 10-Member Ensemble – 2007*

In 2007, the 10 member ensemble produced by CAPS (The University of Oklahoma's Center for

member	IC	BC	mp_phy	pbl_phy
Cntl	21Z NAMa	18Z NAMf	WSM 6-class	MYJ
n1	Cntl – em_pert	21Z SREF em-n1	Ferrier	MYJ
p1	Cntl + em_pert	21Z SREF em-p1	Thompson	MYJ
n2	Cntl – nmm_pert	21Z SREF nmm-n1	Thompson	YSU
p2	Cntl + nmm_pert	21Z SREF nmm-p1	WSM 6-class	YSU
Ph1	21Z NAMa	18Z NAMf	Thompson	MYJ
Ph2	21Z NAMa	18Z NAMf	Ferrier	MYJ
Ph3	21Z NAMa	18Z NAMf	WSM 6-class	YSU
Ph4	21Z NAMa	18Z NAMf	Thompson	YSU
Ph5	21Z NAMa	18Z NAMf	Ferrier	YSU

Table 2. Variations in initial conditions (IC), lateral boundary conditions (BC), microphysics (mp_phy), and planetary boundary layer physics (pbl_phy) for the 2007 CAPS WRF-ARW ensemble. NAMa – 12km NAM analysis; NAMf – 12km NAM forecast. All members used the RRTM longwave radiation scheme, the Goddard short-wave radiation scheme, and the Noah land-surface scheme.

3

Analysis and Prediction of Storms) and run at the Pittsburgh Supercomputing Center (PSC) used physics diversity throughout and initial and lateral-boundary condition (IC and LBC) diversity in 4 members. All members were initialized at 2100 UTC and forecasts were 33h in duration. The domain was focused on the Great Plains (Fig. 2, top) and grid spacing was 4 km, with 51 vertical levels. In addition, CAPS ran a single forecast using 2 km grid spacing over the same domain, with physics configuration identical to the ensemble control member. The matrix of configurations for the individual ensemble members is summarized in Table 2.

2.2   WRF-ARW   10-member Ensemble – 2008

In 2008, the 10 member ensemble produced by CAPS and run at PSC used physics and IC/LBC diversity in 9 out of 10 members (one "control" member and eight perturbed members). Furthermore, radar data (reflectivity and radial velocity) were assimilated into all 9 of these members, using the CAPS 3DVAR assimilation system (Hu et al. 2006a, b) as a last step in the initialization process. The tenth member was configured identically to the control member, but it was not subjected to the final radar-data assimilation step, allowing for a systematic assessment of the impact of the radar data. All members were initialized at 0000 UTC and forecasts were 30h in duration. The domain was expanded from the one used in 2007 to cover nearly three-fourths of the CONUS (Fig. 2, bottom) while grid spacing remained at 4 km, with 51 vertical levels. The configuration matrix for the SE2008 ensemble is shown in Table 3.

3. CHALLENGES

Fundamentally, the biggest impediments to forecast improvements with CAMs appear to be the same as those for coarser resolution models: Uncertanties in IC/LBCs and weaknesses in model formulations - especially model physical parameterizations. During SE2008, participants were asked to spend ~1.5 h each day analyzing guidance from daily 18-30 h WRF-NSSL and WRF-EMC forecasts and identifying potential

member	IC	LBC	Radar data	mp_phy	sw-phy	pbl_phy
Cntl	00Z NAMa	00Z NAMf	yes	Thompson	Goddard	MYJ
C0	00Z NAMa	00Z NAMf	no	Thompson	Goddard	MYJ
N1	Cntl – em_pert	21Z SREF em_n1	yes	Ferrier	Goddard	YSU
P1	Cntl + em_pert	21Z SREF em_p1	yes	WSM 6-class	Dudhia	MYJ
N2	Cntl – nmm_pert	21Z SREF nmm_n1	yes	Thompson	Goddard	MYJ
P2	Cntl + nmm_pert	21Z SREF nmm_p1	yes	WSM 6-class	Dudhia	YSU
N3	Cntl – etaKF_pert	21Z SREF etaKF_n1	yes	Thompson	Dudhia	YSU
P3	Cntl + etaKF_pert	21Z SREF etaKF_p1	yes	Ferrier	Goddard	MYJ
N4	Cntl – etaBMJ_pert	21Z SREF etaBMJ_n1	yes	WSM 6-class	Goddard	MYJ
P4	Cntl + etaBMJ_pert	21Z SREF etaBMJ_p1	yes	Thompson	Dudhia	YSU

Table 3. Variations in IC, LBC, microphysics (mp_phy), shortwave radiation (sw_phy), and planetary boundary layer physics (pbl_phy) for the 2008 CAPS WRF-ARW ensemble. NAMa – 12km NAM analysis; NAMf – 12km NAM forecast. All members used the RRTM longwave radiation scheme, and the Noah land-surface scheme. Additional details about the IC and LBC perturbations can be found at URL http://www.emc.ncep.noaa.gov/mmb/SREF/SREF.html and in Xue et al. (2008)

problems with ICs and/or physical parameterizations. These subjective analyses were thoroughly documented and they provide a valuable reference for more detailed ongoing investigations. They form the basis for much of the discussion in this section.

In regard to the impact of ICs/LBCs, SE2008 participants were asked to compare 18-30 h model predictions of large-scale fields to verifying analyses of the same fields. On some days, it was possible to identify errors in phase and amplitude within these fields that clearly had a negative impact on CAM forecasts of convective initiation and evolution. The errors apparently originated
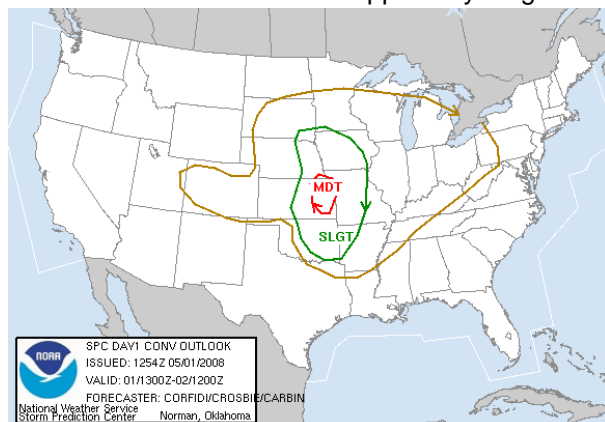

Fig. 3. SPC categorical Day 1 Convective Outlook for 1 May 2008, issued at 1300 UTC.

from ICs and/or LBCs provided by the NAM (North American Mesoscale model), and on some days it was noted that NAM forecasts of convective precipitation showed biases similar to the CAMs. Obviously, the skill of model forecasts is limited by the accuracy of ICs/LBCs (see Weisman et al. 2008 for a discussion of the modulating effect of source-model ICs/LBCs). In the first case shown below, participants agreed that significant errors in the forecast resulted from a poor representation of meso and larger-scale fields, apparently emanating from the initial state provided by the NAM.

On other days, poor model performance

Observed Reflectivity   WRF-NSSL Simulated Reflectivity
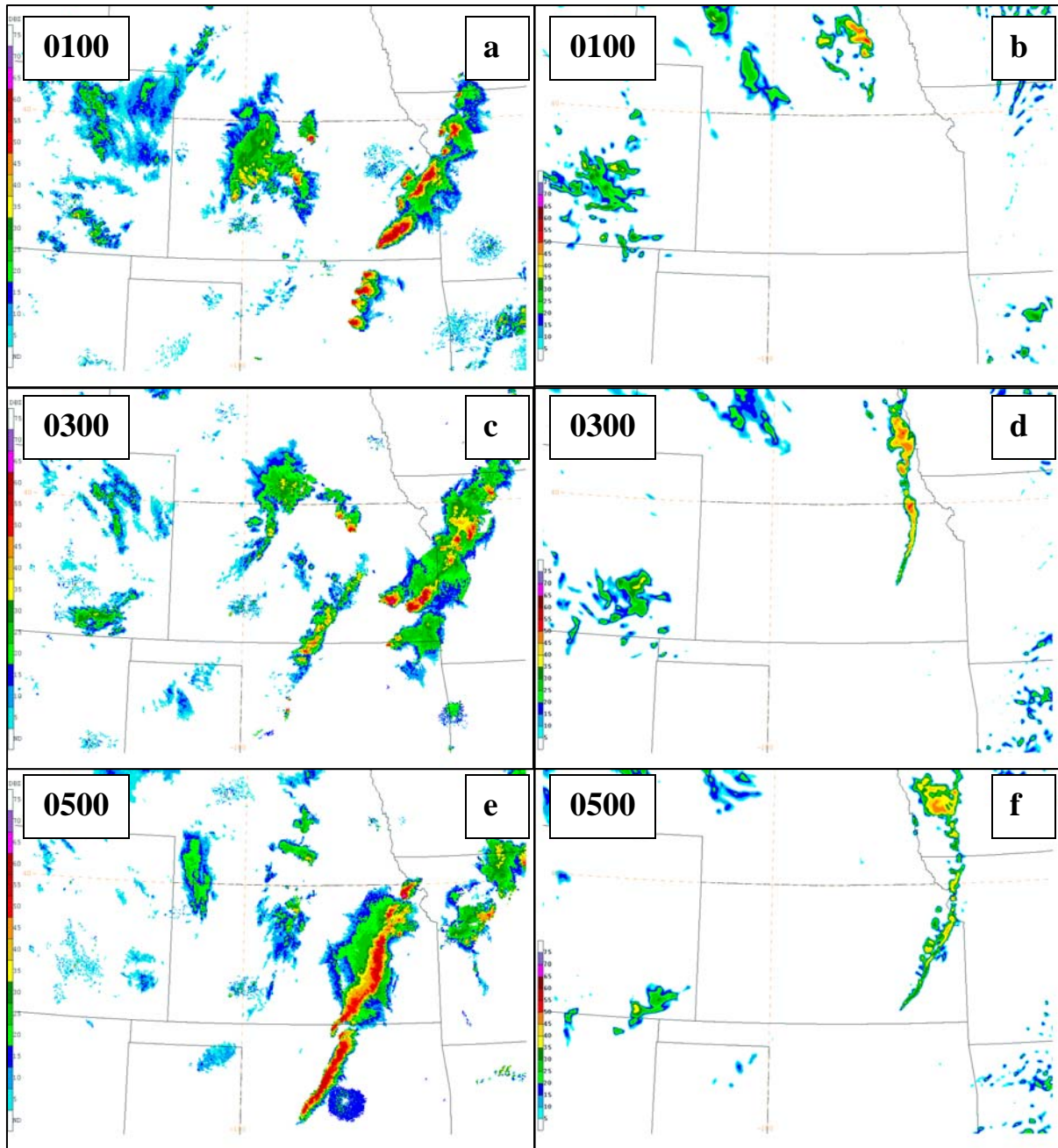
Fig. 4. Observed mosaic of lowest-elevation-angle radar reflectivity (left) and simulated 1-km-AGL reflectivity from the WRF-NSSL 4 km forecast (right), valid 0100 (a, b), 0300 (c, d), and 0500 (e, f) UTC 2 May 2008. Forecast times in right panels are 25, 27, and 29 h from top to bottom.

seemed more closely linked to biases in physical parameterizations. There were many days on which participants noted that the models seemed slow to dissipate clouds at the top of the boundary layer (BL), resulting in anomalously cool and moist low-level thermodynamic fields, and excessive convective inhibition. There was speculation that this tendency could be due to problems with parameterizations of BL, radiation, and/or microphysical (MP) processes, or a combination thereof. Of course, it could also be related to problems with ICs, especially if the ICs contributed to spurious convection over the focus region early in the model integration, prior to the period of primary interest. In the second case selected below, erroneous ICs may have played a role, but physical parameterizations appeared to exacerbate the problem and lead to serious deficiencies in next-day forecasts of convection.

3.1 *Meso- and Larger-Scale Initial/Lateral-boundary conditions*

On 1 May 2008 upper-air charts revealed an
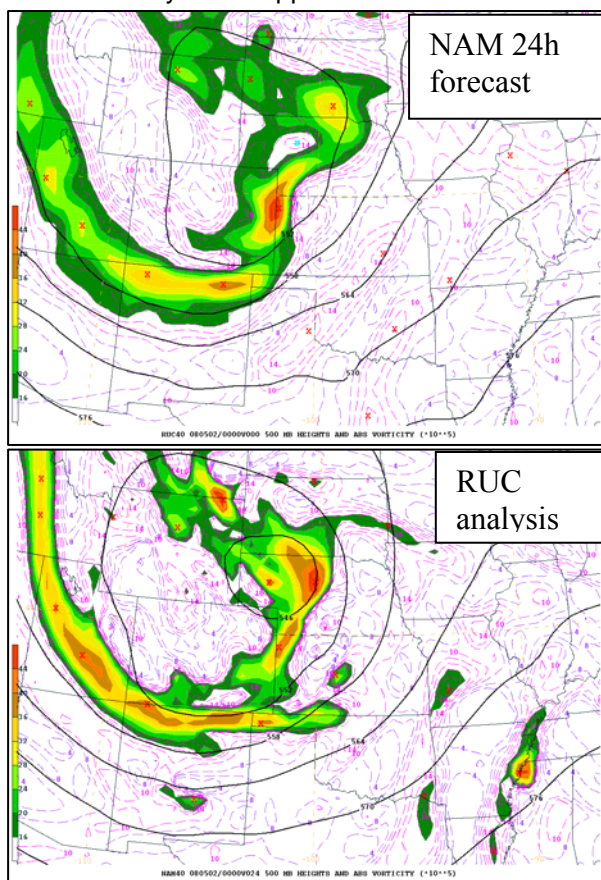
Fig. 5. 500 hPa heights and vorticity valid 0000 UTC 2 May 2008, with 24h NAM forecast on top and RUC analysis on bottom.

upper low migrating eastward across the High Plains and a strong mid-level jet streak rounding the base of the trough over the Four Corners region (not shown). CAPE (Convective Available Potential Energy) and shear fields were expected to be favorable for severe convection over the Southern and Central Plains. SPC forecasters issued a Moderate Risk for severe weather in eastern Kansas and a Slight Risk over a relatively broad area stretching from the eastern half of Oklahoma and western Arkansas as far northward as southeastern South Dakota and southwestern Minnesota (Fig. 3).

Isolated convective cells developed along a dryline between 2200 and 2300 UTC in eastern Kansas and farther south into central Oklahoma over the next couple of hours (Fig. 4a). Later in the evening, a strong squall line formed along a cold front as it swept through the same region (Figs. 4c, e). The WRF-EMC (not shown) and WRF-NSSL (Figs. 4b, d, and f) forecasts were severely deficient on this day, failing to give a strong signal for the development of either phase of convective activity. SE2008 participants wrote that the "models failed to capture the large scale forcing", citing errors in the phasing and configuration of the upper-level trough and low-level boundaries. When a 24 h forecast from the 0000 UTC initialization of the NAM (the same initial conditions used for the CAMs) is compared to the RUC (Rapid Update Cycle; Benjamin et al. 2004) analysis valid at the same time, some of these errors are apparent (Fig. 5). Furthermore, it is worth noting that the NAM precipitation forecast for this region was also poor (not shown), providing further evidence that there were fundamental problems with the initial fields in the model.

Oftentimes, larger-scale errors are quite subtle and hard to discern from visual inspection. Linking forecast errors to specific deficiencies in either initial conditions or model physics is one of the most challenging tasks for modelers and a serious impediment to improvement of the models. For example, in the face of poor model performance, it can be tempting to adjust physical parameterizations to achieve a more desirable result, but if deficient forecasts are in fact due to errors in ICs and LBCs, adjustments to physical parameterizations will be misguided and may ultimately do more harm than good. Recent efforts to identify and quantify large-scale errors and their propagation through the forecast (i.e., Elmore et al. 2006; Coniglio et al. 2008) should allow us to focus efforts in this area more clearly.
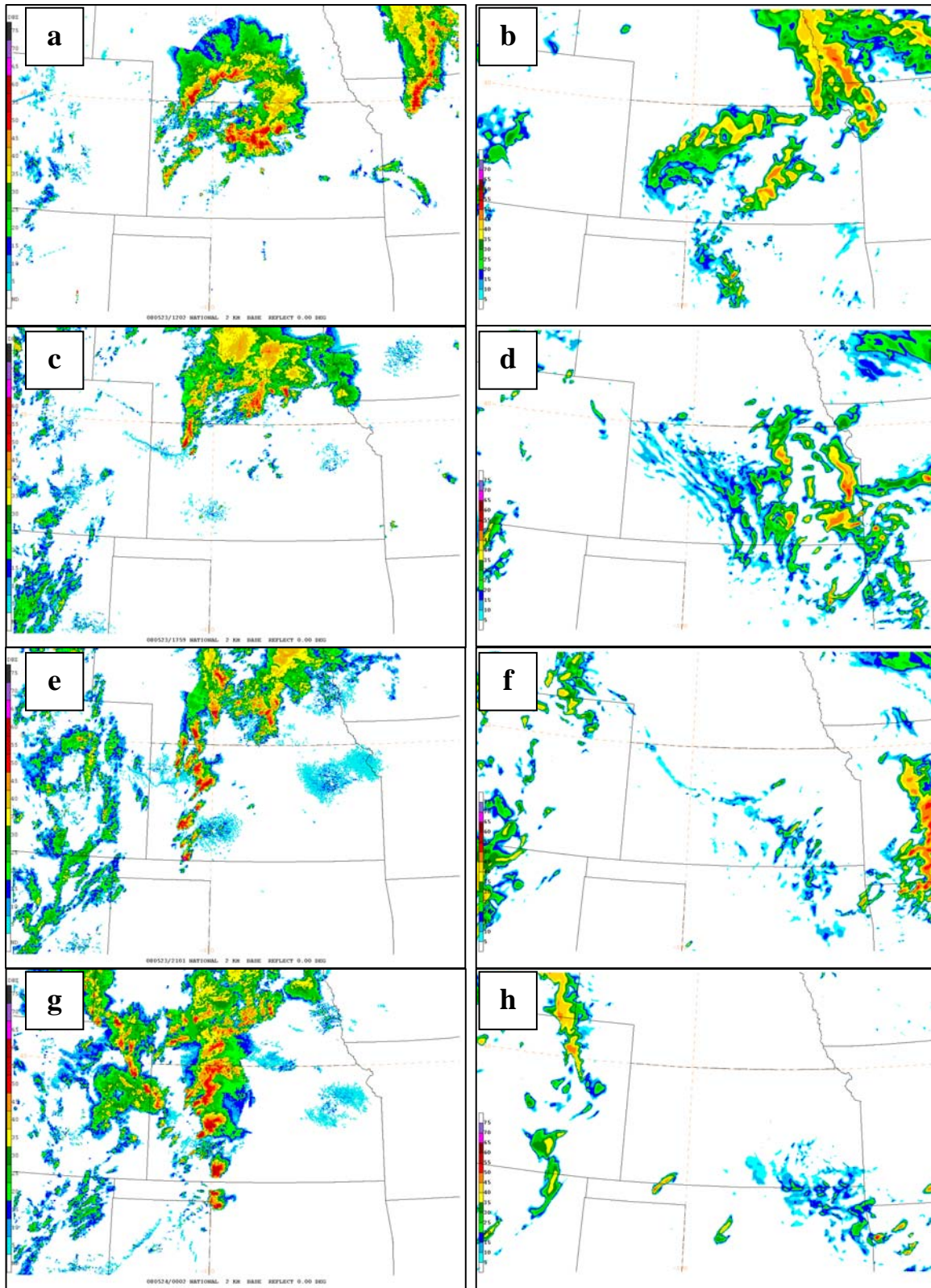
6

Fig. 6. Observed mosaic of lowest-elevation-angle radar reflectivity (left) and simulated 1-km-AGL reflectivity from the WRF-EMC 4 km forecast (right) valid 1200 (a, b), 1800 (c, d), 2100 (e,f) UTC 23 May 2008 and 0000 UTC 24 May 2008 (g, h). Forecast times in right panels are 12, 18, 21, and 24 h from top to bottom.

3.2 *Physical Parameterizations*

On the morning of 23 May 2008, a large upper cyclone covered the western third of the CONUS and a short wave trough was ejecting out of this system across the Northern Plains (not shown). A second mid-level disturbance was forecast to round the base of the larger system and propagate into the southern and central High Plains during the afternoon and evening hours. Elevated convection associated with the leading system covered much of the area from west-central Kansas into south-central Nebraska, and a separate MCS was evident over Iowa and northern Missouri at 1200 UTC (Fig. 6a). The latter system moved rapidly eastward, while the Kansas-Nebraska activity moved relatively slowly to the northeast as primarily light to moderate "stratiform" precipitation. But by 1800 UTC, stronger convective cells were beginning to develop along a boundary in northwestern Kansas, on the southwestern periphery of this activity (Fig. 6c), and by 2100 UTC isolated intense cells extended as far south as the Oklahoma Panhandle (Fig. 6e). Several of these storms grew into mature supercells over the next three hours and numerous tornadoes were reported in western Kansas (Fig. 6g). A separate area of tornadic supercells developed in northeastern Colorado.

Both the WRF-NSSL (not shown) and WRF-EMC configurations struggled to capture this complex evolution of events. At 1200 UTC, the 12h WRF-EMC forecast showed reasonable agreement with observations in that it predicted two distinct systems in the region, both with reflectivity patterns similar to observed reflectivity (Fig. 6b). However, the predicted systems appeared to be displaced upstream from the observed locations, and the primary system over Kansas had an additional branch of weak model reflectivities extending southward into Oklahoma. Furthermore, while the observed system moved northeastward over the next six hours, the simulated system drifted slowly on more of an easterly trajectory (Fig. 6d). No deep convection developed in the model over western Kansas through 0000 UTC (Figs. 6f and h), providing woefully inadequate guidance to forecasters.

During SE2008, this failure was attributed to a poor representation of overnight and early morning convection in the CAMs. As noted above, precipitation during this period in the WRF-EMC was too far south and too slow to dissipate. SE2008 analysts suggested that this spurious activity generated a cold pool that was too strong,

propagated too far south, and was too resistant to dissipation. Late-afternoon fields from the model are consistent with this assessment. For example, low-level temperature and wind fields show relatively cold air and easterly flow much farther south and west than a verifying RUC analysis at 2100 UTC (compare Figs. 7a and b). Furthermore, the 0000 UTC model -forecast sounding for Dodge City reveals a deep surface-
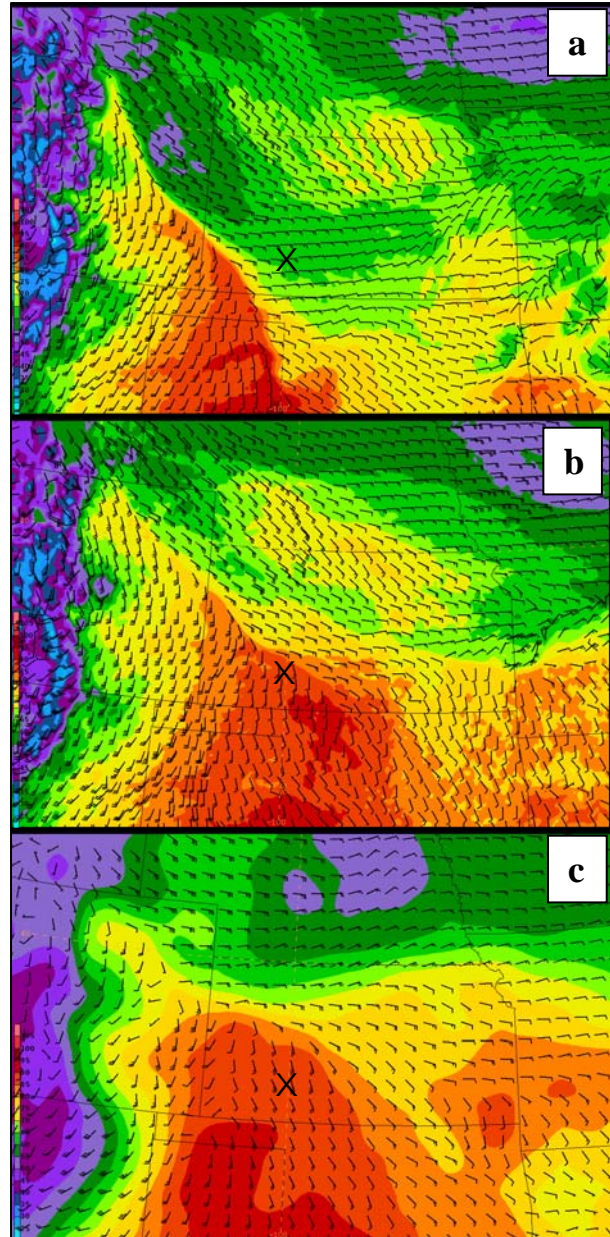

Fig. 7. 2m temperature field (color fill) and 10m wind barbs from the a) 21h WRF-EMC forecast, b) 21h WRF-NSSL forecast, and c) the corresponding RUC analysis, all valid 2100 UTC 23 May 2008. The "X" marks the location of Dodge City, where the sounding shown in Fig. 8 was taken.
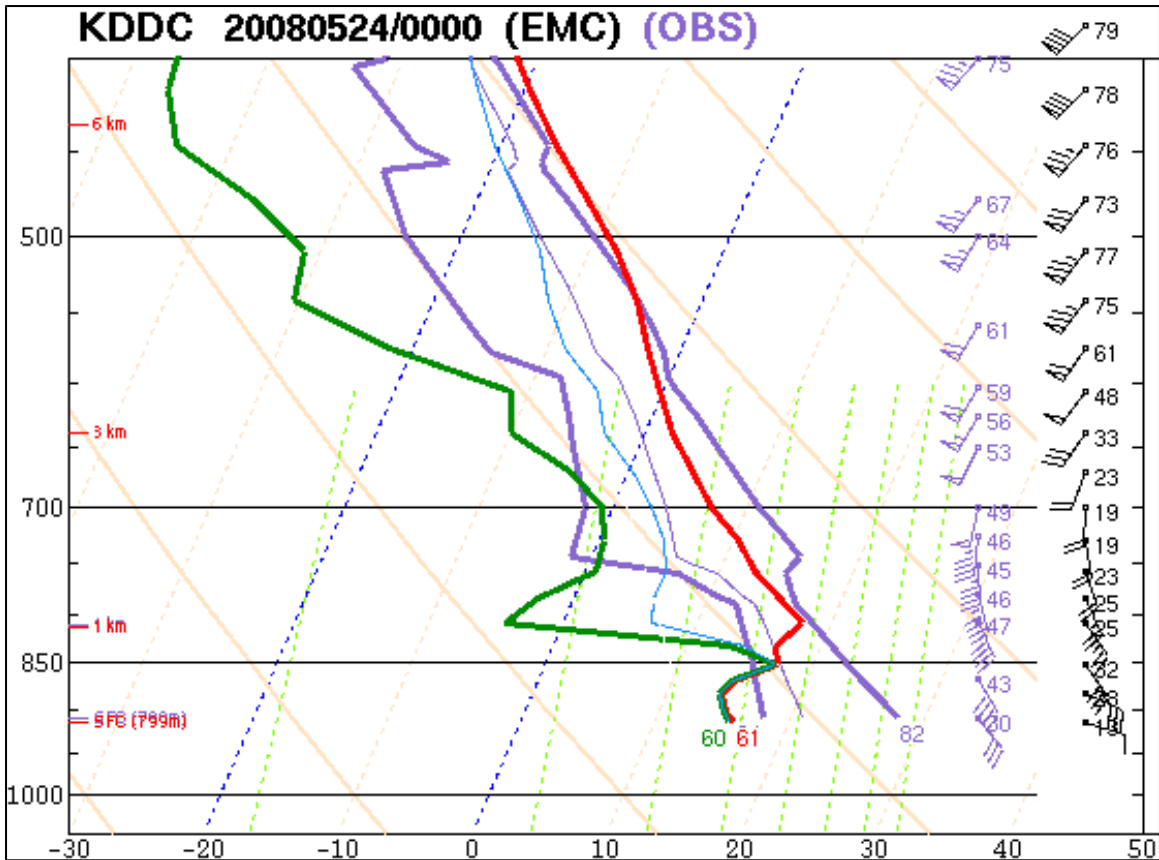
Fig. 8. 24h WRF-EMC forecast sounding (Temperature –red; Dew point – green, wind-barbs – black) for Dodge City, KS, overlaid on the observed sounding (purple), valid 0000 UTC 24 May 2008.

based cold pool, in sharp contrast to observations (Fig. 8).

During SE2008 analysts speculated that biases in the WRF-EMC's physical parameterizations may have caused the model to over-estimate the intensity of cold outflow from the early convection, leading to the later forecast failures. Indeed, the low-level temperature forecast from the WRF-NSSL run, which uses some different physical parameterizations (see Table 1), was in somewhat better agreement with observations (compare Fig. 7b to Fig. 7a and 7c). Yet, the WRF-NSSL run also failed to initiate any convection in western Kansas during the period when the severe activity was observed. While this may imply that both configurations were handicapped by deficient initial conditions, the NAM correctly produced a band of precipitation over this region (Fig. 9). This suggests that there was a dynamic forcing mechanism in larger-scale fields, but that the over-stabilized environment in the CAMs was unable to respond appropriately. Clearly, this case would be a good candidate for a more detailed investigation of the relative

importance of model physics versus initial conditions.

The potential impact of ICs notwithstanding, precipitation forecasts with the WRF model are clearly sensitive to the choice of model physics. During SE2007 CAPS made daily forecasts with a
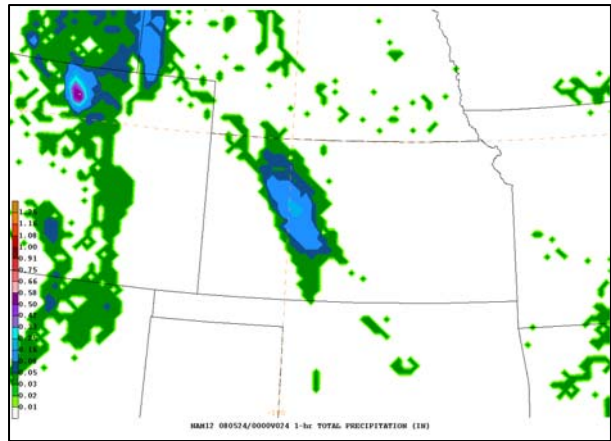

Fig. 9. Accumulated precipitation (inches) for the 1 hour ending 0000 UTC 24 May 2008 (24 h forecast), from the operational NAM.

9

10 member CAM ensemble, with physics diversity a key ingredient in the ensemble. Specifically, 6 of the 10 members started with identical initial conditions, but different combinations of microphysical and boundary-layer parameterizations; the other 4 members used SREF[6]-based variations in IC and LBC, but also different combinations of the same physical parameterizations. A simple plot of the areal coverage of precipitation as a function of forecast time (averaged over all days during SE2007) reveals the strong sensitivity to formulations for both microphysics and BL processes (Fig. 10). Although there were not enough variations to perform a complete separation of contributions from the different parameterizations, one can gain a qualitative sense of the individual biases by grouping together all forecasts produced using a given parameterization and computing the deviation of each group from the mean. When this is done, it suggests that the MYJ BL scheme tends to be wetter than YSU, while Ferrier is the wettest microphysical parameterization, followed by WSM6 and Thompson. The sensitivities are quite large, but it is unclear at this stage which particular schemes are "best" for a given application. Development and refinement of physical parameterizations remains one of the greatest
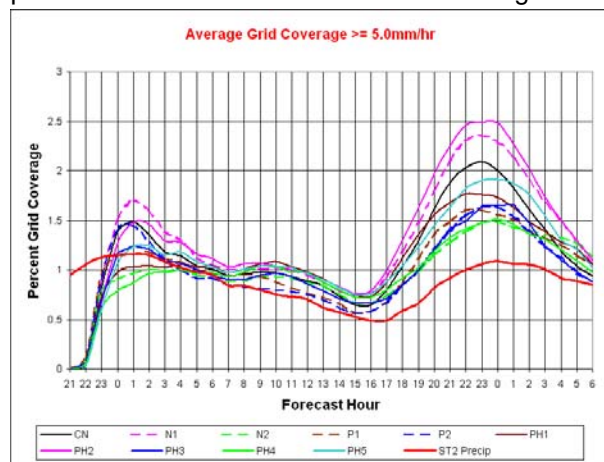

Fig. 10. Percent coverage of precipitation rate $\geq$ 5 mm h[-1] as a function of forecast hour for all members of the CAPS ensemble, averaged over all days during SE2007. Microphysical and planetary-boundary-layer parameterizations for the individual members listed in the legend can be found in Table 2. Stage II precipitation observations are represented by the red curve.

challenges for the modeling community (e.g., Mass 2006)

4. OPPORTUNITIES

As a new class of NWP models, CAMs provide both researchers and forecasters with numerous opportunities. Some of these opportunities represent "low hanging fruit", while others will require years of development.

4.1 *The Low Hanging Fruit: Postprocessing Opportunities.*

By definition, CAM configurations allow a whole new set of phenomena to develop compared to coarser resolution models that rely on parameterized convection. Thus, it is incumbent upon numerical modelers to develop post-processing strategies that extract information about these phenomena to use as guidance for forecasters.

4.1.1 *Identification of a new set of simulated phenomena*

The mesoscale organizational mode of convection is an important concern for severe-weather forecasters, such as those at the SPC, because the type of severe convection that occurs is strongly linked to mode. For example, tornadoes and large hail are often associated with supercells, while damaging straight-line winds are the primary threat from bowing line segments. Thus, efforts to identify and characterize convective mode in CAM output are providing important contributions to new and unique forecaster-guidance products.

Simulated reflectivity, computed directly from model hydrometeor fields, has proven to be very useful for subjective assessments of convective mode. In particular, simulated reflectivity fields can provide important clues about structures and circulations associated with precipitation features (see Koch et al. 2005) and they are very valuable for comparison with observed reflectivity patterns (as demonstrated in the previous section).

Subjective assessments can become even more useful when they are combined with objective routines that identify and highlight specific features in CAM output. For example, NSSL scientists and SPC forecasters developed several diagnostic routines to identify lower to middle tropospheric rotating updrafts in preparation for SE2005. One of these routines specifically identifies updraft helicity and it has

10

proven to be a useful proxy for supercells in the CAM output (Kain et al. 2008). This is significant because supercellular convection produces a disproportionate share of severe convective weather (i.e., tornadoes, damaging winds, and/or large hail). Fig. 11a zooms in on a 23 h forecast of simulated reflectivity over northern Kansas and southern Nebraska, with purple hatching indicating high UH values. Note that the reflectivity pattern of the dominant storm resembles that of a prototypical supercell, and the UH diagnostic provides additional confidence that this is indeed a supercellular structure, as a UH maximum is located above the inflow notch where one would expect to find a mesocyclone in a storm with this configuration. Although it is unreasonable to expect the model to routinely have skillful, timely, and reliable predictions of individual storms like this, it is worth noting that a storm having supercellular characteristics was observed in this same county within about 2 hours of the predicted

Fig. 11. a) Simulated 1 km AGL reflectivity from the WRF-NSSL 4 km forecast, with area of UH ≥ 25 m^2 s^{-2} indicated by purple hatching and b) observed mosaic of lowest-elevation-angle radar reflectivity. The valid time is 2300 UTC for the model forecast and 2054 UTC for the observations, 31 May 2007, and the indicated county outlines are in northern Kansas and southern Nebraska.

storm (Fig. 11b)

As discussed in Kain et al. (2008), the UH diagnostic has proven to be quite useful for highlighting areas and conditions that favor supercell formation. SPC forecasters/scientists have also developed routines that identify other manifestations of convective mode, such as quasi-linear and bowing reflectivity structures. Explicit forecasts of convective mode have proven to be useful for SPC forecasters. Ongoing efforts are aimed at determining how such information compares to inferences made from environmental parameters.

4.1.2 *Tracking severe phenomena every timestep.*

Traditionally, NWP guidance is presented as a series of snapshots in time (an exception is accumulated precipitation). This approach has been adequate for most applications because common features of interest evolve slowly compared to the time between outputs. In CAMS, output frequency is typically no higher than hourly, partly because higher frequency output could overwhelm operational dataflow and processing systems due to the size of the output files. Yet convective features of interest can develop, move, and vary in intensity on time scales much shorter than one hour. Thus, we have modified the WRF code to track certain fields and features every time step (24 s) while outputting the average or extreme values of the data at normal output intervals. This approach allows us to monitor phenomena of interest every time step without compromising computational efficiency or producing an overflow of data.

This concept is currently being used to extract sub-output-time information about convective storms in WRF-NSSL forecasts, initially focusing on five different output fields. The first two are updraft and downdraft velocities below the 400 hPa level. These two fields have obvious implications regarding the intensity of convective overturning. The third field is simulated reflectivity computed at the lowest model level. Without much additional computational effort, other levels (or a composite reflectivity field) could be used as well. As with updraft and downdraft velocities, this field is related to the intensity of convection and may provide some guidance for forecasting large hail, even though hail is not explicitly represented in the current microphysical scheme (WSM6). The fourth field is 10-m wind speed, which may be helpful in predicting the magnitude of convectively generated surface wind gusts. The last field is
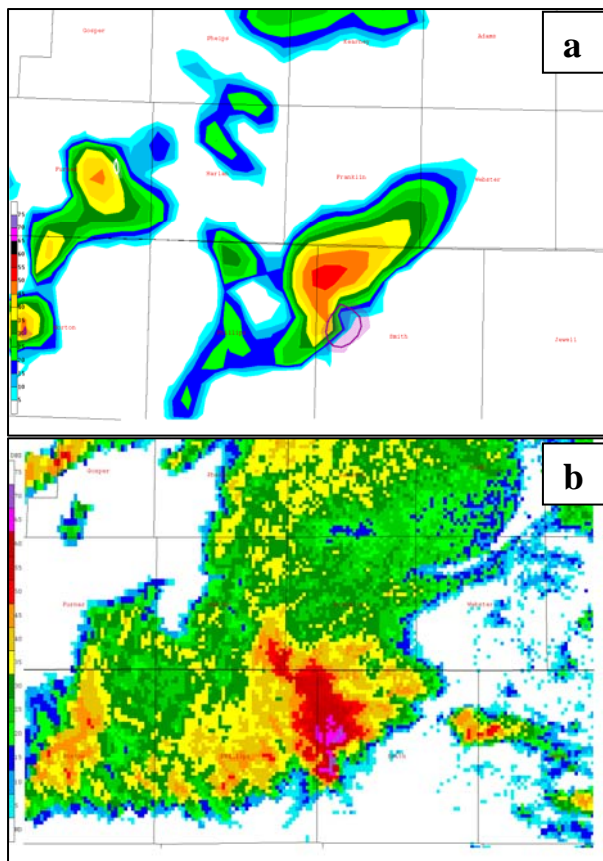
UH. Each of these fields is tracked every time step at every grid point and 2-D fields representing the maximum value over the hour (again, at each grid point) are saved along with normal output fields at the top of the hour.

This approach not only provides information about between-output-time maximum values, it yields features such as wind swaths and supercell tracks in the model. For example, Fig. 12 highlights the evolution of convective activity that began in the western Oklahoma Panhandle late in the afternoon of 3 June 2008 as isolated convective cells. One of these cells emerged as a dominant, large-hail-producing supercell, then the cluster consolidated, went through a process of upscale growth, and eventually developed into a meso- -scale bow echo that produced a swath of damaging winds. The WRF-NSSL model
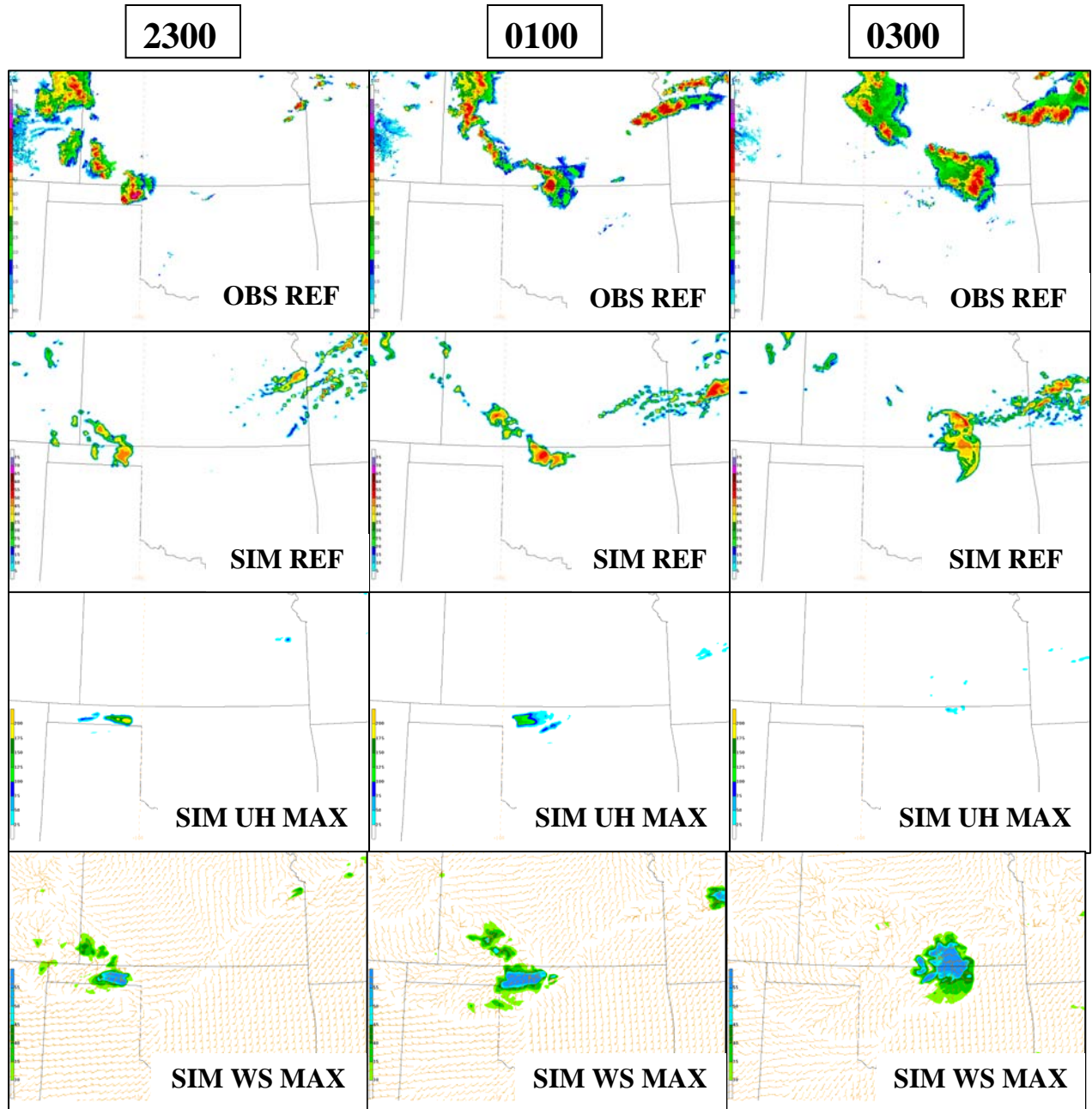
Fig. 12. Observed and simulated evolution of a convective complex from 2300 UTC 3 June 2008 through 0300 UTC 4 June 2008, showing observed lowest-elevation-angle reflectivity (top row), simulated 1 km AGL reflectivity (second row), and simulated hourly maximum values of UH (third row) and 10 m wind speed (bottom row).

performed exceptionally well in predicting this convective system, capturing the entire evolution in a realistic manner (compare top two rows in Fig. 12). Furthermore, the hourly-maximum UH and 10-m wind speed fields reveal the development of a narrow path of high winds associated with a strong mesocyclone in the early stages, followed by upscale growth, a bowing convective system, and a relatively broad swath of high winds in later stages (bottom two rows in Fig. 12). Although the basic character of this evolution can be inferred from the hourly simulated-reflectivity snapshots, the supplementary UH and 10 m wind-speed output provides valuable additional insight into the dynamic evolution of the convective activity within the model.

4.1.3 *Customizing output fields to provide specific guidance similar to operational products*

SPC forecasters issue a wide range of forecast products, but numerical model guidance is used more for Convective Outlooks than any other product. The Day 1, Day 2, and Day 3 outlooks include probabilistic forecasts for severe convection over the entire CONUS. For example, the Day 2 Convective Outlook issued 1730 UTC 7 May 2008, valid for the period 1200 UTC 8 May through 1200 UTC 9 May is shown in Fig. 13a. It identifies three areas of enhanced probability for severe convection, one centered over western Kansas, another over northern Alabama, and the third over the piedmont of the Carolinas.

The distribution of severe weather reports for this day (Fig. 13b) shows that the areas of concentrated activity were predicted quite well. This subjective comparison is useful, but a more quantitative verification requires additional processing of the data. One approach is to compare the probabilistic forecast to the smoothed spatial density of observed reports using a spatial smoother to create fields of report density for direct comparison with probabilistic forecasts. Brooks et al. (1996) used this approach with the goal of generating retrospectively a "practically perfect" probabilistic forecast – the best probabilistic forecast that a human could be
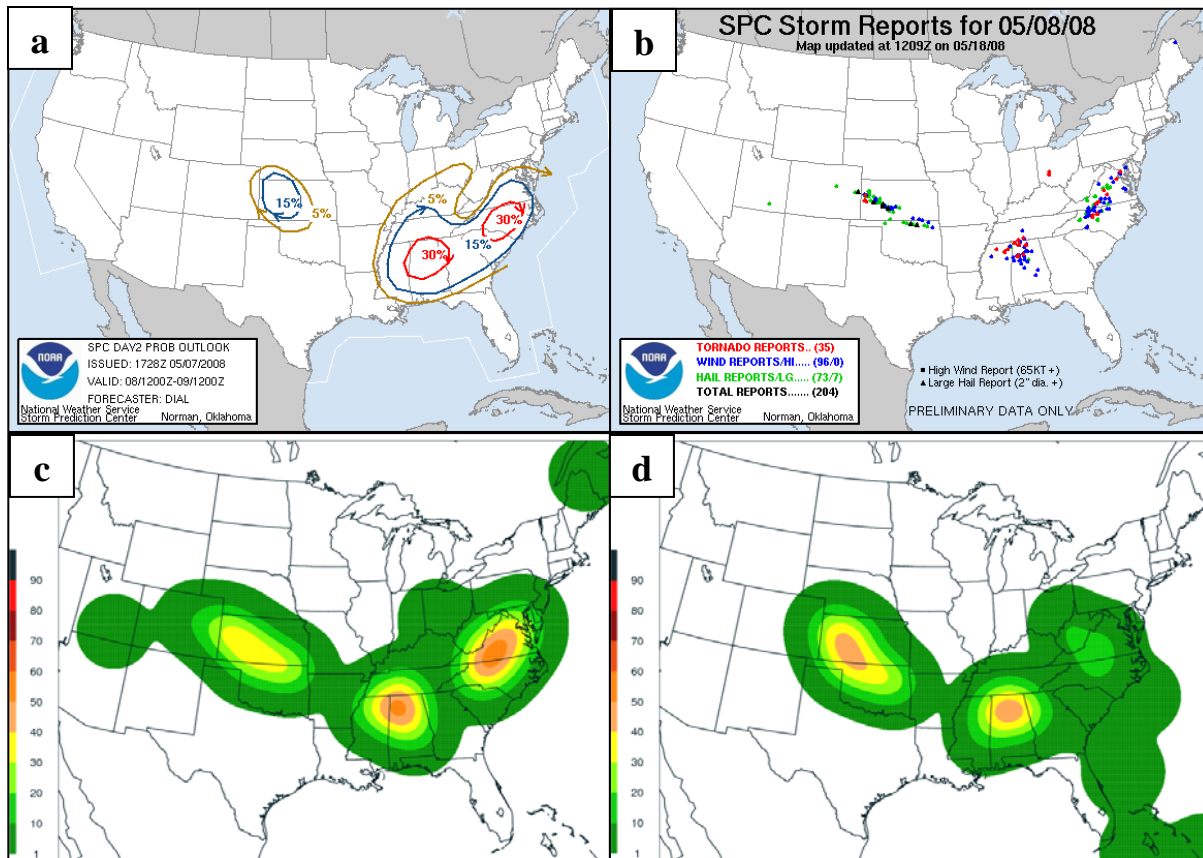

Fig. 13. SPC Day 2 Outlook for the period 1200 UTC 8 May – 1200 UTC 9 May 2008 (a) and the locations of corresponding observed severe-weather reports plotted as individual points (b) and as a density/probability field (c), along with a density/probability field of simulated "surrogate" reports from the WRF-NSSL forecast for the same time period (12-36 h forecast period).

13

expected to make given the observed distribution of reports. Figure 13c shows the "practically perfect" forecast corresponding to the reports shown in Fig 13b.

If this concept can be used to make single-point observations of severe weather equitable with probabilistic forecasts, why not apply it to single-grid-point "observations" from CAM output? During SE2008, we exploited this idea to produce an experimental guidance product for the likelihood of severe convection, based on CAM output. Specifically, we developed algorithms to identify severe phenomena in CAM forecasts and treat each occurrence of these phenomena as a surrogate severe weather report. Then we applied the "practically perfect" smoother to the field of surrogate reports. Results from this initial effort have been quite encouraging. For example, when this concept was applied to the WRF-NSSL output from 8 May, using extreme values of the five hourly-maximum fields described in the previous section, the resulting probabilistic forecast highlighted the three main areas of severe activity quite well (Fig. 13d). This approach, described in detail by Sobash et al. (2008), shows promise in providing useful guidance for preparation of SPC Outlooks and perhaps other products. Similar efforts to produce probabilistic guidance from deterministic model forecasts have also been documented, e.g., by Theis et al. (2005), Roberts (2005), and Schwartz et al. (2008b).

4.2 *Grid Resolution*

NWP's transition from ~10 km grid spacing (with parameterized convection) to sub-5-km spacing with explicit convection is a positive and exciting development. Yet there is still considerable debate about how much resolution is necessary for reliable and skillful predictions involving convective storms. Is 4 km grid spacing adequate, or would forecasts be much more skillful with 2 km or 1 km grid lengths? These issues are discussed at some length in Kain et al. (2008) and in a follow-up study (Schwartz et al. 2008a). Results shown here are extracted from the latter study.

During SE2007, scientists from CAPS provided output from WRF-ARW forecasts with 2 and 4 km grid spacing every day. Aside from horizontal resolution, the model configurations used to produce these forecasts were identical. Furthermore, ICs and LBCs were generated from the same NAM grids. This experimental set-up provided a unique framework in which the added value of 2 versus 4 km could be assessed.

Comparison of forecasts focused on daily subjective assessments during SE2007, followed by a thorough objective verification after the formal experiment.

In general, results showed that the value added by doubling the horizontal resolution was marginal, and almost certainly not worth the ~ten-fold increase in computational expense. As first shown in Kain et al. (2008) and later substantiated in Schwartz et al. (2008a), the two model configurations seem equally skillful in predicting the mesoscale organizational mode of convection as well as the formation (though not exact placement) of significant smaller-scale phenomena such as mesocyclones. In terms of quantitative precipitation forecasting (QPF), the CAM forecasts have skill comparable to that of the NAM for low precipitation thresholds, and considerably greater than the NAM for higher thresholds (Fig. 14). In other words, they are much better than the NAM in predicting extreme precipitation events that can be a threat to life and property. [Note that Fig. 14 shows the fractions skill score (Roberts and Lean 2008), a "neighborhood" verification metric that does not require exact grid-point-by-grid-point correspondence between observational and forecast fields.]

These results notwithstanding, it is likely that certain applications will benefit substantially from higher resolution, (e.g., forecasts in regions of complex terrain, or perhaps assimilation of high-resolution data such as radar data). But the relative insensitivity revealed by these HWT experiments is a very favorable result because it implies that there need not be a rush to higher and higher resolution for NWP. Instead, operationally-minded modelers can concentrate on improving the models, adapting guidance products to the new capabilities of CAMs, and developing new strategies, such as ensemble approaches. The idea of using CAM output as guidance for operational forecasting is still in its infancy and there is much that can be learned even if convective-scale circulations are only coarsely resolved.

4.3 *Land Surface Impacts*

The land surface state, particularly soil moisture distribution, has a substantial impact on the surface heat fluxes and the development of mesoscale circulations, which in turn can affect convective initiation and precipitation distribution in CAMs (Trier et al. 2004). Heterogeneity in surface fluxes lead to strong lower-tropospheric,

mesoscale "solenoidal" circulations, behaving like an inland sea-breeze circulation. These circulations are not significantly impacted by the prevailing large-scale flow, and depending on stability, the strong vertical motions associated with these circulations can initiate deep, moist convection (Pielke 2001; Weaver 2004).

Collaborative studies at the SPoRT Center are using the NASA LIS to initialize land surface variables in WRF at CAM resolutions. This approach produced improvements in the diurnal cycle of 2-m temperature predictions over Florida

for a month of daily WRF simulations (Case et al. 2008b). LIS provides the capability to conduct long-term offline integrations or "spinups" to allow the surface and soil profiles to reach thermodynamic equilibrium, using bias-adjusted meteorological inputs from sources such as the Global Data Assimilation System and Stage IV precipitation analyses.

Producing high-resolution spinups is not currently possible using the standard WRF version, and therefore most applications initialize surface and soil fields by interpolating from a
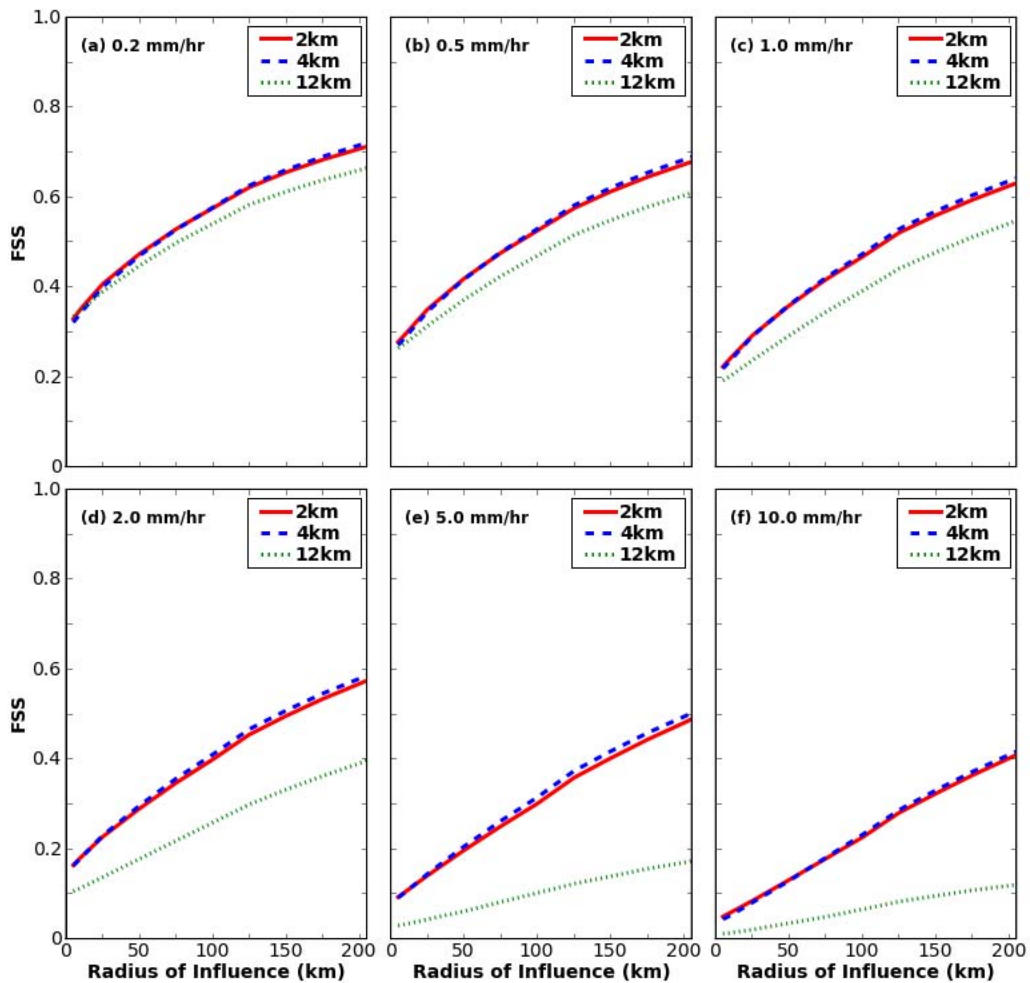
Fig. 14. Fractions skill scores as a function of radius of influence (see Roberts and Lean 2008) for different precipitation thresholds. The legend indicates the grid spacing used in the individual forecasts: 2 and 4 km refers to the WRF-ARW CAM forecasts provided by CAPS and 12 km denotes the operational NAM forecasts. These aggregate scores include hourly output from all models valid between 1800 and 0600 UTC (21-33 h forecast for the CAMs, 18-30 h forecast period period for the NAM) for all days during SE2007. Stage II hourly precipitation analyses (Seo 1998) were used as the verifying field. See Schwartz et al. (2008) for additional details.

coarser-resolution analysis/forecast system such as the NAM, potentially leading to unrepresentative initial soil moisture and temperature fields. Offline LIS output is generated at the same resolution as the local/regional grids, and is then used directly as input to the WRF simulation, eliminating the need for horizontal spatial interpolation (Case et al. 2008b). Such a methodology produces an initial soil moisture field consistent with the high-resolution land surface

properties of the CAMs. SPoRT is conducting sensitivity tests using LIS land surface initialization for selected severe weather cases from SE2007 and SE2008 to examine the potential improvements to the NSSL-WRF resulting from the LIS input.

4.4 Ensemble Applications

Given the uncertainties in both initial

| Obs Comp Refl | Sim Comp Refl: W/Rad Assim. | Sim Comp Refl: No Rad Assim. |
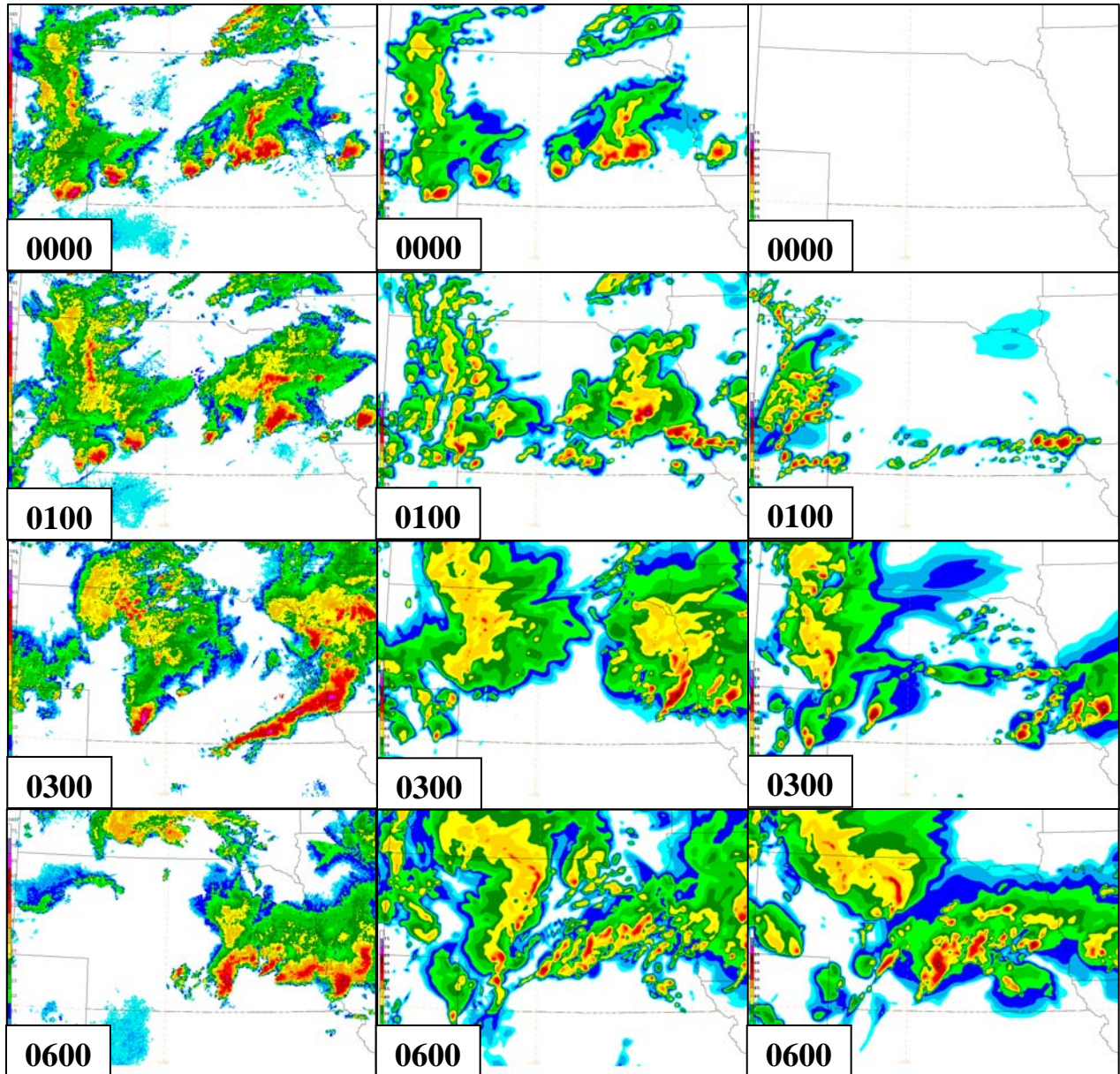
Fig. 15. Comparison of observed and simulated composite reflectivity for forecast hours 0, 1, 3, and 6, beginning at 0000 UTC 5 June 2008. Observations are in the left-hand column, the WRF-ARW forecast that includes assimilation of radar data is in the middle column, and the identically configured "cold-start" forecast is in the right column.

16

conditions and model physics, and the potential for ensemble-based techniques to provide quantitative estimates of uncertainty, exploration of CAM-based ensembles has become an important initiative. During SE2007 and SE2008 CAPS generated CAM-based ensemble output every day. Specifically, they provided forecasts from 10 member ensembles of 4 km WRF-ARW configurations. The ensembles had diversity in both model physics and IC/LBC. Given the sensitivity to initial land surface states, future CAM-based ensembles could consist of members that perturb the initial soil moisture. Sutton et al. (2006) found that 5-km WRF forecast differences in explicit precipitation due to soil moisture variations were as large as convective precipitation variations resulting from the use of different convective parameterization schemes on a 20-km simulation grid.

Numerous applications of the high-resolution ensemble output were explored during SE2007 and SE2008. Because convective mode is such

an important concern for severe-weather forecasters, one of the more exciting efforts focused on computing probabilistic guidance for different convective modes, with a particular emphasis on supercells and linear/bowing structures. Much of this work is described elsewhere in this conference volume (e.g., Xue et al. 2008; Schwartz et al. 2008a; Kong et al. 2008; Bright et al. 2008).

4.5 *Data Assimilation*

During SE2008 participants examined and compared the first 12 forecast hours from the two CAPS ensemble control members – one with and one without assimilation of radar data – every day. The point of this exercise was to provide a subjective assessment of the impact of the radar-data assimilation for the 0-12 h period. Note that aside from the radar-data assimilation, model configurations and initialization procedures were identical. An example is provided below.
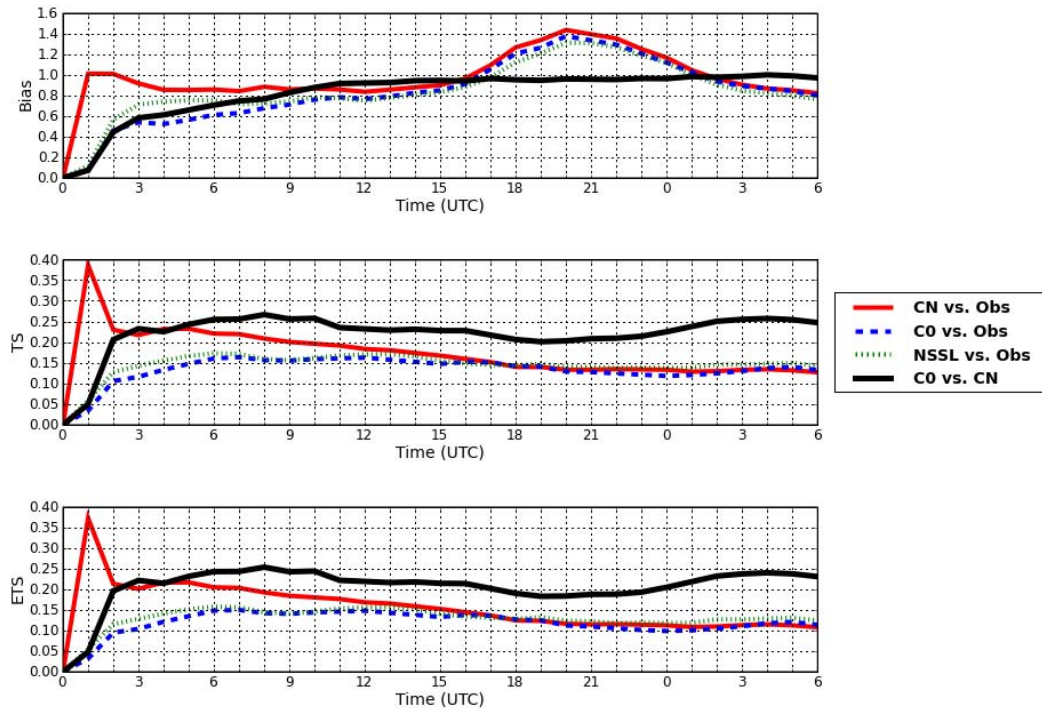

Fig. 16. Coverage bias, threat score (TS), and equitable threat score (ETS) as a function of time for all forecast hours and all days during SE2008, for the 1 mm hr^{-1} precipitation-rate threshold. The different datasets indicated in the legend correspond to the CAPS WRF-ARW forecasts with (CN) and without (C0) assimilation of radar data and the WRF-NSSL runs that also do not assimilate radar data. Note that the black curve in each image compares the C0 and CN fields, while the other curves compare the respective model forecasts to Stage II observed precipitation data.

Comparison of the first two images in the top row of Fig. 15 suggests that the CAPS procedures for assimilating hydrometeors from radar data are quite skillful (note that the CAPS 3DVAR system assimilated radar from individual sites – so-called nids2 data - whereas the national mosaic fields shown in the left-most column of Fig. 15 have been processed using proprietary Unisys algorithms, so an exact agreement should not be expected). The correspondence between observed and assimilated mesoscale structures is remarkably good and even individual convective cells appear to be initialized quite well. Within the first hour, the model seems to lose the smaller scale details and the simulated reflectivity pattern suggests that small-scale adjustments are ubiquitous (noisy pattern), but a good semblance of two distinct convective complexes remains after the first hour of integration (compare first two images in the second row). Meanwhile, the "cold-start" forecast develops some isolated storms in southeastern Nebraska during this first hour and begins to form a cohesive convective system in the western part of the state (rightmost image in second row). As integration time progresses, the two forecasts begin to look more and more alike, and perhaps less like observations, but the run with data assimilation clearly appears to get a
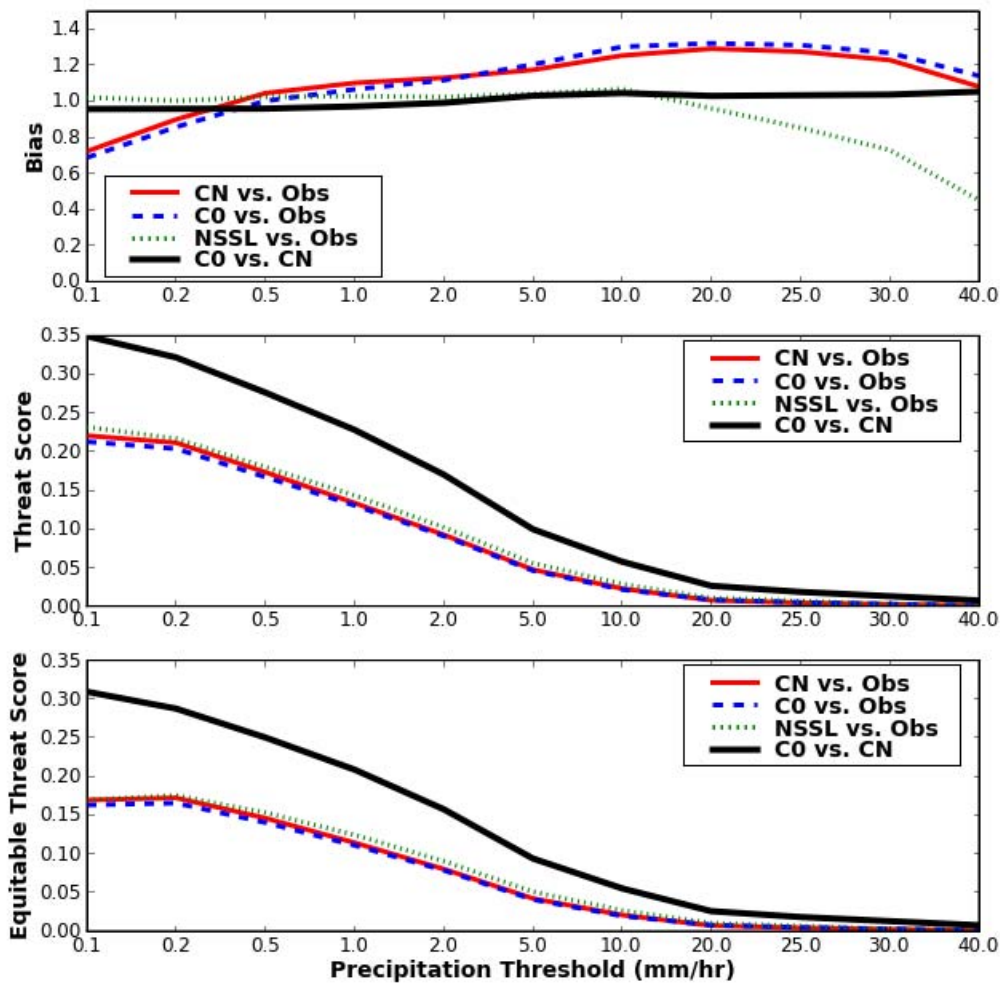

Fig. 17. Aggregate coverage bias, threat score (TS), and equitable threat score (ETS) as a function of precipitation-rate threshold for the 18-30 h forecast period (1800-0600 UTC) on all days during SE2008. Definitions in the legend are as in Fig. 16.

"head start" in its representation of two coherent convective systems.

During much of the SE2008 period, CAPS scientists were modifying strategies and adjusting parameters in the radar assimilation system. Also, the input radar dataset was incomplete on many days. Thus, the case shown here is one of the "best" examples from the experiment. Nonetheless, and in spite of some inconsistencies in performance from day to day, the net impact of the radar-data assimilation was positive. Aggregate verification of hourly precipitation data from all days suggests that the assimilation procedures led to higher precipitation skill scores through at least the first 12h of the forecast (Fig. 16), perhaps due to a higher bias during this period. At the same time, objective metrics confirm our perception that the two forecasts begin to look more like each other than like observations within the first few hours (CN vs. C0 TS and ETS scores are higher than either CN or C0 vs. Obs in Fig. 16). During the next-afternoon and evening time period, the two forecasts appear to be about equally skillful and continue to look more similar to each other than to observations (Fig. 17).

This result is encouraging, but assimilation of radar data into CAMs remains one of the most challenging problems in NWP. Scientists who work more closely on this effort could provide a much more nuanced and insightful assessment, but our observations during SE2008 suggested that assimilated convective scale information was lost very quickly after the model integration started. Indeed, SE2008 participants often noted that reflectivity structures looked "non-meteorological" between hour 0 and, say, the 3-6 hour period when realistic looking features began to spin up in both the run with radar-data assimilation and the "cold start" run. So the value added by assimilation of radar data appeared to come mainly in the category of QPF rather than convective mode, even though mode is generally the attribute of greatest added value for CAMs (compared to models using parameterized convection) for the later forecast periods.

We appear to be a long way from the point at which assimilation procedures can consistently insert convective storms that "hit the ground running", with surrounding circulations and mass fields that enable a natural evolution of individual updrafts, downdrafts, and near-convective scale circulations. But progress is being made on several fronts. For this long-term effort, approaches based on 4DVAR and ensemble-Kalman-filter (EnKF) appear to be just as viable as the 3DVAR technique used by CAPS during SE2008 (D. Dowell, NCAR, personal communication, 2008), and some of the most talented individuals in the modeling community are developing and assessing each of these different approaches.

5. SUMMARY

The emergence of CAMs in operational NWP presents us with exciting opportunities as well as scientific challenges that will keep us occupied for many years to come. In terms of opportunities, the "low-hanging fruit" appears to be in the realm of post-processing – a matter of extracting new output fields and/or computing new diagnostics that were not available from coarser resolution models that parameterized convection. This effort appears to be most effective when it is done collaboratively by model developers and practitioners (forecasters). Such collaboration is a primary focus of annual NOAA HWT Spring Experiments.

Another opportunity looms in the area of CAM-based ensembles. As with coarser resolution models, there is considerable uncertainty in both IC/LBCs and model physics. Ensembles provide us with a conceptual framework for addressing the magnitude and character of this uncertainty, but at this stage very little is known about appropriate perturbation strategies for CAMs, and high resolution ensembles are computationally expensive. Fortunately, 4 km grid spacing seems to be sufficient to obtain significant added value from CAMs, compared to traditional approaches with parameterized convection. Output from 4 km CAM-based ensembles, provided by CAPS and PSC, was evaluated during SE2007 and SE2008. Ensemble forecasts were certainly intriguing and they show promise in providing useful probabilistic guidance for convective storms, but many questions remain about how ensembles might be applied most effectively on these scales.

Perhaps the most exciting possibilities are in the arena of data assimilation, particularly assimilation of radar data. Over the past 5 years or so, the value of next-day CAM forecasts has been demonstrated. The success of these forecasts appears to confirm that even CAMS with grid spacing as large as 4 km can develop near-convective scale structures and skillful predictions of the mesoscale organizational mode of convection when larger-scale forcing mechanisms are represented accurately. However, we appear to be a long way from being able to "insert" convective scale features in the initial conditions without seeing their character disappear rapidly

once the model integration begins. This is the most pressing data assimilation problem (and the most captivating possibility) for CAMS. This work is a high priority for several research groups, focusing on 3DVAR, 4DVAR, and EnKF approaches.

A perpetual challenge that should not be overlooked is improvement of forecast models themselves, especially their physical parameterizations. Progress in the development of physical parameterizations has been slow historically, and it is likely to continue that way. But any NWP system, whether it involves advanced data assimilation, ensembles, and/or other sophisticated approaches, is only as good as the forecast model at its core. It is imperative that scientists make efforts to improve forecast models a top priority and that funding agencies find ways to support this fundamental need.

Acknowledgements

REFERENCES

Benjamin, S.G., D. Dévényi, S.S. Weygandt, K.J. Brundage, J.M. Brown, G.A. Grell, D. Kim, B.E. Schwartz, T.G. Smirnova, T.L. Smith, and G.S. Manikin, 2004: An hourly assimilation–forecast cycle: The RUC. *Mon. Wea. Rev.*, **132**, 495–518.

Bright, D. R., S. J. Weiss, J. J. Levit, and R. S. Schneider, 2008: The evolution of multi-scale ensemble guidance in the prediction of convective and severe convective storms at the Storm Prediction Center. *Preprints, 24th Conference on Severe Local Storms*, Amer. Meteor. Soc., Savannah, GA. CD-ROM P10.7

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

Case, J. L., S. R. Dembek, J. S. Kain, S. V. Kumar, T. Matsui, J. J. Shi, W. M. Lapenta, and W-K. Tao, 2008a: A sensitivity study of the operational NSSL WRF using unique NASA assets. *Preprints, 9th Annual WRF Users' Workshop,* Boulder, CO, National Center for Atmospheric Research, P9.4. [Available online at http://www.mmm.ucar.edu/wrf/users/workshops/WS2008/abstracts/P9-04.pdf]

Case, J. L., W. L. Crosson, S. V. Kumar, W. M. Lapenta, and C. D. Peters-Lidard, 2008b:

Impacts of high-resolution land surface initialization on regional sensible weather forecasts from the WRF model. *J. Hydrometeor., In Press*.

Chou, M.-D., and M. J. Suarez, 1999: A solar radiation parameterization for atmospheric studies. NASA Tech. Pre. NASA/TM-1999-10460, vol. 15, 38 pp.

Chou, M.-D., M. J. Suarez, X-Z. Liang, and M. M.-H. Yan, 2001: A thermal infrared radiation parameterization for atmospheric studies. NASA/TM-2001-10406, vol. 19, 55 pp.

Coniglio, M. C., J. S. Kain, S. J. Weiss, D. R. Bright, J. J. Levit, G. W. Carbin, K. W. Thomas, F. Kong, M. Xue, M. L. Weisman, and M. E. Pyle, 2008: Evaluation of WRF model output for severe-weather forecasting from the 2008 NOAA Hazardous Weather Testbed Spring Experiment. *Preprints, 24th Conference on Severe Local Storms*, Amer. Meteor. Soc., Savannah, GA. CD-ROM 12.4.

Du, J. J., and Coauthors, 2004: The NOAA/NWS/NCEP short-range ensemble forecast (SREF) system: Evaluation of an initial condition vs multiple model physics ensemble approach. *Preprints, 16th Conf. on Numerical Weather Prediction*, Seattle, WA, Amer. Meteor. Soc., CD-ROM, 21.3.

Elmore, K. L., M. E. Baldwin, and D. M. Schultz, 2006: Field significance revisited: spatial bias errors in forecasts as applied to the Eta model, *Mon. Wea. Rev.,* **134**, 519-531.

Haines, S. L., G. J. Jedlovec, and S. M. Lazarus, 2007: A MODIS sea surface temperature composite for regional applications. *IEEE Trans. Geosci. Remote Sens.,* **45**, 2919-2927.

Hu, M., M. Xue, and K. Brewster, 2006a: 3DVAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part I: Cloud analysis and its impact. *Mon. Wea. Rev.*, **134**, 675–698.

Hu, M., M. Xue, J. Gao, and K. Brewster, 2006b: 3DVAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part II: Impact of radial velocity analysis via 3DVAR. *Mon. Wea. Rev.*, **134**, 699–721.

Janjic, Z. I., T. L. Black, M. Pyle, H.-Y. Chuang, E. Rogers and G. DiMego, 2005: An evolutionary approach to nonhydrostatic modeling. [Available at URL: http://www.wrf-model.org/wrfadmin/publications/Chuang_Janjic_NWP50yearsfinalshort.pdf]

Kain, J. S., S. J. Weiss, J. J. Levit, M. E. Baldwin, and D. R. Bright, 2006: Examination of convection-allowing configurations of the WRF model for the prediction of severe convective weather: The SPC/NSSL Spring Program 2004. *Wea. Forecasting,* **21**, 167-181

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.

Koch, S. E., B. Ferrier, M. Stolinga, E. Szoke, S. J. Weiss, and J. S. Kain, 2005: The use of simulated radar reflectivity fields in the diagnosis of mesoscale phenomena from high-resolution WRF model forecasts. *Preprints, 11th Conference on Mesoscale Processes*, Albuquerque, NM, Amer. Meteor. Soc., CD-ROM, J4J.7

Kong, F, M. Xue, K. K. Droegemeier, K. W. Thomas, Y. Wang, J. S. Kain, S. J. Weiss, D. R. Bright, and J. Du, 2008: Real-time storm-scale ensemble forecast 2008 Spring Experiment. *Preprints, 24th Conference on Severe Local Storms*, Amer. Meteor. Soc., Savannah, GA. CD-ROM 12.3.

Kumar, S. V., and Coauthors, 2006. Land Information System − An Interoperable Framework for High Resolution Land Surface Modeling. *Environmental Modeling & Software,* **21 (10),** 1402-1415, doi:10.1016/j.envsoft.2005.07.004.

Kumar, S. V., C. D. Peters-Lidard, J. L. Eastman, and W.-K. Tao, 2007: An integrated high-resolution hydrometeorological modeling testbed using LIS and WRF. *Environmental Modeling & Software,* **23 (2),** 169-181, doi: 10.1016/j.envsoft.2007.05.012.

LaCasse, K. M., M. E. Splitt, S. M. Lazarus, and W. M. Lapenta, 2008: The impact of high-resolution sea surface temperatures on the simulated nocturnal Florida marine boundary layer. *Mon. Wea. Rev.,* **136,** 1349-1372.

Lean, H. W., P. A. Clark, M. Dixon, N. M. Roberts, A. Fitch, R. Forbes, and C. Halliwell, 2008: Characteristics of high resolution versions of the Met Office Unified Model for forecasting convection over the UK. *Mon. Wea. Rev.*, **136**, 3408-3424.

Mass, C. F., 2006: The uncoordinated giant: Why U.S. weather research and prediction are not

achieving their potential. *Bull. Amer. Meteor. Soc.*, **87**, 573–584

Matsui, T., W.-K. Tao, and J.-J. Shi, 2007: Goddard radiation and aerosol direct effect in Goddard WRF. NASA/UMD WRF meeting, College Park, MD, Sep 14, 2007.

Narita, M., and Ohmore, S. 2007: Improving precipitation forecasts by the operational nonhydrostatic mesoscale model with the Kain-Fritsch convective parameterization and cloud microphysics. *Preprints, 12th Conference on Mesoscale Processes*, Watervillle Valley, NH, CDROM, 3.7

Pielke, R. A., 2001: Influence of the spatial distribution of vegetation and soils on the prediction of cumulus convective rainfall. *Rev. Geophys.,* **39,** 151-177.

Roberts, N. M., 2005: An investigation of the ability of a storm scale configuration of the Met Office NWP model to predict flood-producing rainfall. UK Met Office Technical Report No. 455. (Available from http://www.metoffice.gov.uk/research/nwp/publications/papers/technical_reports/2005/FRTR455/FRTR455.pdf)

Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78-97.

Schwartz, C. S., J. S. Kain, S. J. Weiss, D. R. Bright, M. Xue, F. Kong, K. W. Thomas, J. J. Levit, and M. C. Coniglio, 2008a: Next-day convection-allowing WRF model guidance: A second look at 2- vs. 4-km grid spacing. *Preprints, 24th Conference on Severe Local Storms*, Amer. Meteor. Soc., Savannah, GA. CD-ROM P10.3. ...

Schwartz, C. S., J. S. Kain, D. R. Bright, S. J. Weiss, M. Xue, F. Kong, J. J. Levit, M. C. Coniglio, and M. S. Wandishin, 2008b: Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Preprints, 24th Conference on Severe Local Storms*, Amer. Meteor. Soc., Savannah, GA. CD-ROM 13A.6

Seifert, A., M. Baldauf, K. Stephan, U. Blahak, and K. Beheng, 2008: The Challenge of convective scale quantitative precipitation forecasting. *Preprints, 5th International Conference on Clouds and Precipitation*, Cancun, Mexico, CD-ROM 3.12 [available from http://conventioncenter.net/iccp2008/abstracts/Program_on_line/Oral_03/SeifertAxelEetal_extended_3.pdf]

Seo, D. J., 1998: Real-time estimation of rainfall fields using radar rainfall and rain gauge data. *J. Hydrol.*, **208**, 37-52.

Skamarock, W.C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, J. G. Powers, 2005: A Description of the Advanced Research WRF Version 2. NCAR Tech Note, NCAR/TN-468+STR, 88 pp. [Available from UCAR Communications, P. O. Box 3000, Boulder, CO 80307].

Sobash, R. A., D. R. Bright, A. R. Dean, J. S. Kain, M. Coniglio, S. J. Weiss, and J. J. Levit, 2008: Severe storm forecast guidance based on explicit identification of convective phenomena in WRF-model forecasts. *Preprints, 24th Conference on Severe Local Storms*, Amer. Meteor. Soc., Savannah, GA. CD-ROM 11.3.

Sutton, C., T. M. Hamill, and T. T. Warner, 2006: Will perturbing soil moisture improve warm-season ensemble forecasts? A proof of concept. *Mon. Wea. Rev.,* **134,** 3174-3189.

Tao, W.-K., J. Simpson, D. Baker, S. Braun, M.-D. Chou, B. Ferrier, D. Johnson, A. Khain, S. Lang, B. Lynn, C.-L. Shie, D. Starr, C.-H. Sui, Y. Wang and P. Wetzel, 2003: Microphysics, radiation and surface processes in a non-hydrostatic model, *Meteorology and Atmospheric Physics,* **82**, 97-137.

Theis S. E., A. Hense, and U. Damrath, 2005: Probabilistic precipitation forecasts from a deterministic model: A pragmatic approach. *Meteor. Appl.*, **12**, 257–268.

Trier, S. B., F. Chen, and K. W. Manning, 2004: A study of convective initiation in a mesoscale model using high-resolution land surface initial conditions. *Mon. Wea. Rev.,* **132,** 2954-2976.

Weaver, C. P., 2004: Coupling between large-scale atmospheric processes and mesoscale land-atmosphere interactions in the U.S. Southern Great Plains during summer. Part I: Case studies. *J. Hydrometeor.,* **5,** 1223-1246.

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., M. E. Pyle, Z. Janjic, D. R. Bright, and G. J. DiMego, 2008: The operational High Resolution Window WRF model runs at NCEP: Advantages of multiple model runs for severe convective weather forecasting. *Preprints, 24th Conference on Severe Local Storms*, Amer. Meteor. Soc., Savannah, GA. CD-ROM P10.8

Xue, M., F. Kong, K. W. Thomas, J. Gao, Y. Wang, K. Brewster, K. K. Droegemeier, J. S. Kain, S. J. Weiss, D. R. Bright, M. C. Coniglio, and J. Du, 2008: CAPS realtime storm-scale ensemble and high-resolution forecasts as part of the NOAA Hazardous Weather Testbed 2008 Spring Experiment. *Preprints, 24th Conference on Severe Local Storms*, Amer. Meteor. Soc., Savannah, GA. CD-ROM 12.2