COLLABORATIVE EFFORTS BETWEEN THE UNITED STATES AND UNITED KINGDOM TO ADVANCE PREDICTION OF HIGH-IMPACT WEATHER

John S. Kain, Steve Willington, Adam J. Clark, Steven J. Weiss, Mark Weeks, Israel L. Jirak, Michael C. Coniglio, Nigel M. Roberts, Christopher D. Karstens, Jonathan M. Wilkinson, Kent H. Knopfmeier, Humphrey W. Lean, Laura Ellam, Kirsty Hanley, Rachel North, and Dan Suri

The international partnership between NSSL, SPC, and the Met Office has been enthusiastic and productive, driven by strong mutual interests at a grassroots level and generous institutional support from parent government agencies.

The National Oceanic and Atmospheric Administration (NOAA) conducts atmospheric research and development (R&D) primarily within the Office of Oceanic and Atmospheric Research (OAR), with much of the R&D effort aimed at improving weather forecasts issued from a separate NOAA line office, the National Weather Service (NWS). Although they fall in different NOAA branches, the designated national centers for both research and forecasting of *severe* thunderstorms—OAR's National Severe Storms Laboratory (NSSL) and NWS's Storm Prediction Center (SPC)—are located side by side within the National Weather Center in Norman, Oklahoma. Scientists and forecasters from NSSL and SPC regularly engage in collaborative R&D as well as experimental forecasting activities to make continuous improvements in the products and services provided to the public and the larger weather enterprise (e.g., other government agencies, emergency managers, and broadcast media).

The Met Office functions much like the atmospheric component of NOAA, conducting R&D in the atmospheric sciences and providing operational forecasting services for citizens of the United Kingdom. As

AFFILIATIONS: Kain, Clark, and Coniglio—NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma; Willington, Weeks, Roberts, Wilkinson, Lean, Ellam, Hanley, North, and Suri—Met Office, Exeter, United Kingdom; Weiss and Jirak—NOAA/NWS/NCEP/Storm Prediction Center, Norman, Oklahoma; Karstens and Knopfmeier—Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, and NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma
CORRESPONDING AUTHOR: John S. Kain, 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-15-00199.1

In final form 28 July 2016
©2017 American Meteorological Society

Brought to you by NOAA Central Library | Unauthenticated | Downloaded 09/24/21 07:35 PM UTC

with NSSL and SPC, organized collaborations between the research and forecasting divisions of the Met Office have been critically important for informing researchers where new R&D is needed and for moving the fruits of that R&D into forecasting operations. As with their American counterparts, this collaborative strategy results in considerable benefits for the United Kingdom.

In recent years, severe weather experts from NOAA and the Met Office have found compelling reasons to combine efforts in improving forecasts of high-impact weather. The impetus for this enhanced collaboration has been a growing recognition on both sides of the Atlantic of the potential value of high-resolution numerical prediction systems as providers of guidance for forecasting high-impact weather events. Over the last decade the Met Office has very successfully introduced progressively higher-resolution operational numerical weather prediction (NWP) systems into forecasting operations, focusing on a unified modeling framework, while NOAA has used somewhat coarser resolution but simultaneously developed and implemented multiple high-resolution NWP dynamic cores. Both organizations have developed innovative postprocessing strategies and unique tools for visualization and verification of model output. Both have also worked with frontline forecasters to tailor model output for specialized forecast problems. Significantly, much of the development in the two organizations has been complementary rather than duplicative, providing extra incentive to combine efforts for mutual benefit.

The primary venue for recent NSSL, SPC, and Met Office interactions has been the annual Spring Forecasting Experiment (SFE), conducted in the NOAA Hazardous Weather Testbed (HWT). Specific objectives of the SFE change each year, but the general framework for this experiment is to conduct real-time forecasting exercises in a simulated operational forecasting environment, with experimental forecast products prepared jointly by operational forecasters and "bench" scientists. These exercises differ from forecasting operations in that the normal set of observations and model guidance used to inform key decisions is supplemented with new scientific concepts and guidance products that are under development within the research community (including SPC's Science Support Branch). The utility of emerging techniques and guidance products is documented during this process and further assessment of these products is conducted during postforecasting evaluation exercises. This process accelerates the transfer of promising new ideas and technologies into operations and offers key insights for further development of others.

NSSL and SPC have been conducting the annual SFE since 2000 (Kain et al. 2003b; Kain et al. 2006; Clark et al. 2012), and the Met Office became a contributing partner beginning with the 2012 experiment. The benefits of this partnership have been substantial for all organizations and for the SFE. These benefits are discussed herein, beginning with a brief historical overview, followed by a summary of positive outcomes from assessing high-resolution model guidance products from NSSL and the Met Office, and a summary of the benefits of the strategic partnership between NSSL, the Met Office, and SPC.

A BRIEF HISTORY. Collaborations between NSSL and Met Office scientists on operationally relevant research topics date back to at least the early 1990s, when short-range ensemble forecasting was a topic of mutual interest. In 1994, NSSL and the U.S. National Meteorological Center cohosted a workshop on this topic (Brooks et al. 1995) at which Met Office scientists David Richardson and Mike Harrison participated. Gil Ross from the Met Office also visited NSSL around this time to collaborate on research interests related to model verification and ensemble forecast systems (Harold Brooks, NSSL, 2015, personal communication).

NSSL also had ongoing collaborations with SPC in the late 1980s and early 1990s (e.g., Lewis et al. 1989; Johns and Doswell 1992), and the breadth of these interactions increased dramatically in 1997 when SPC relocated from Kansas City, Missouri, to a specifically designed workspace within NSSL facilities in Norman, Oklahoma. The proximity of SPC operations and support scientists invigorated NSSL research and that burst of activity included studies evaluating the use of ensemble forecast systems to improve the prediction of severe weather (Stensrud et al. 1999, 2000; Stensrud 2001; Stensrud and Weiss 2002; Bright et al. 2004), which provided a scientific foundation for the implementation of an operational short-range ensemble forecasting (SREF) system (Tracton et al. 1998; Du and Tracton 2001) at the Environmental Modeling Center (EMC), which, like SPC, is part of the NWS's National Centers for Environmental Prediction (NCEP).

The Met Office started developing a nonhydrostatic version of the Unified Model (MetUM) in the mid-1990s (Davies et al. 2005), with the first operational forecasts from this model produced in 2003. In 2005, the first convection-allowing (also called convection permitting) operational forecast was made using a 4-km configuration of the MetUM over the United Kingdom. This was followed in 2009 by a U.K.-wide model with variable resolution (UKV; Tang et al. 2013), having 1.5-km grid length in the interior stretching to 4 km

Brought to you by NOAA Central Library | Unauthenticated | Downloaded 09/24/21 07:35 PM UTC

at the boundaries; the UKV currently operates with a 3-h three-dimensional variational data assimilation (3D-Var) analysis cycle, which at the time of writing is about to include a 1-h four-dimensional variational data assimilation (4D-Var) analysis cycle for short-range forecast (out to $T + 12$ h). Along with the deterministic UKV, the Met Office also uses the Met Office Global and Regional Ensemble Prediction System (MOGREPS; Bowler et al. 2008) to drive a 2.2-km downscaled ensemble over the United Kingdom called MOGREPS-UK (Golding et al. 2016). A more detailed review of the development and use of convection-allowing models (CAMs) within the Met Office can be found in Clark et al. (2016).

Within their jointly occupied facility, NSSL and SPC conducted the inaugural SFE in 2000. Mesoscale deterministic NWP provided the foundation for the annual experiment through 2002 (e.g., Kain et al. 2003a), but SFE leaders shifted gears to focus on mesoscale ensemble forecast systems, such as EMC's SREF in 2003 (Bright et al. 2004; Homar et al. 2006). A key participant in the 2003 SFE was Ken Mylne, who had worked for 6 years as an operational forecaster in the Met Office before assuming responsibility for the Met Office's ensemble forecasting team in 1999. Ken contributed his insights for a full week during the SFE (Fig. 1).

The SFE continued during 2004 and 2005, with a primary focus on experimental CAM forecasts that were generated by collaborators at EMC and the National Center for Atmospheric Research. After witnessing firsthand the potential impact of CAMs in the 2005 SFE, EMC scientists began generating daily contiguous United States (CONUS)-scale CAM forecasts for the SPC on an experimental basis: the first such effort in the United States and a harbinger of things to come. There was a 1-yr hiatus of the SFE in 2006 owing to the relocation of both NSSL and SPC to the National Weather Center building. During this period there was diminished formal interaction between NSSL/SPC and the Met Office, but the organizations conducted parallel research initiatives focusing on applications of CAMs for the prediction of localized high-impact weather. Innovative Met Office concepts in representing uncertainty and in verifying high-resolution model output (e.g., Roberts 2005; Roberts and Lean 2008) informed ongoing research at NSSL (e.g., Schwartz et al. 2010). Meanwhile, seemingly contradictory results related to resolution sensitivities (Kain et al. 2008; Schwartz et al. 2009; Roberts and Lean 2008; Lean et al. 2008) begged further investigation on both sides.

The 2007 SFE featured the first demonstration and evaluation of CAM-based ensembles, generated for the experiment by the University of Oklahoma Center for Analysis and Prediction of Storms (CAPS; see Xue et al. 2007). Ken Mylne returned for another week of participation to exchange knowledge about ensemble forecasting, both the convection-allowing and convection-parameterizing varieties, sharing his perspective on how ensemble systems should be configured, interpreted, and evaluated. His immersion in the SFE's application of convection-allowing ensembles for severe convection forecasting influenced his decision to give added weight to the preexisting plans for the development of CAM-based ensemble systems for the Met Office. In 2009, another Met

Fig. 1. A forecast team composed of research scientists and operational forecasters examines experimental model guidance during the 2003 SFE. Team members included (left to right) John Gaynor (NWS/U.S. Weather Research Program), Dave Sills (research forecaster, Meteorological Services of Canada), Jim Hansen (postdoctoral fellow, Massachusetts Institute of Technology), Ken Mylne (then manager, ensemble forecasting team, Met Office), Greg Mann (science and operations officer, NWS forecast office), and Harold Brooks (research scientist, NSSL).

Brought to you by NOAA Central Library | Unauthenticated | Downloaded 09/24/21 07:35 PM UTC

Office scientist, Nigel Roberts, visited the SFE for 1 week to explore further development of both deterministic and ensemble configurations of CAMs.

Over the next couple of years, Met Office, NSSL, CAPS, SPC, and EMC scientists continued to explore new forecast applications using their CAMs and they worked on the development and optimization of CAM-based ensemble systems. This resulted in the first use of a real-time year-round CAM ensemble by SPC forecasters in 2011 (Jirak et al. 2012), with membership derived from operational forecasts at EMC and semi-operational runs at NSSL. In addition, NSSL and SPC collaborators developed new diagnostic strategies and postprocessing routines that were specifically designed for CAMs and the severe weather forecasting challenges faced by the SPC (e.g., Sobash et al. 2011; Kain et al. 2010; Clark et al. 2012). These tools were migrated to SPC operations and were soon embraced by the SPC forecast staff.

Complementary knowledge bases and skill sets were being developed by NSSL/SPC and the Met Office during 2010/11, yet formal interactions between NOAA and the Met Office again waned. However, in 2012 the Met Office made a bold move to provide real synergy to the collaboration by assigning two top severe weather forecasters—Steve Willington and Dan Suri—to participate for a week in that year's SFE and for the full 5 weeks of the experiment in 2013 (Fig. 2). In addition, beginning in 2013 the Met Office increased its investment by generating real-time CAM forecasts for the SFE, using operational versions of the MetUM, and executing the model forecasts on high-priority, continuously supported operational computing systems. Furthermore, they began assigning a broader group of scientists and forecasters to participate in the SFE, contributing personnel for every day of the experiment in 2014 and 2015. The MetUM CAMs have provided an important dataset for comparison with CAMs that have been run by NOAA for the SFE, and Met Office personnel have provided unique and valuable perspectives on forecasting and high-resolution NWP, no doubt elevating the positive impact of the experiment on the broader meteorological community.

POSITIVE EARLY OUTCOMES. *New diagnostic tools, visualization strategies, and related initiatives.* In their efforts to help operational forecasters extract the most useful severe weather guidance from high-resolution NWP models, NSSL and SPC have developed numerous visualization and postprocessing tools (e.g., Karstens et al. 2015). One positive outcome of Met Office interactions with NSSL and SPC has been the migration of diagnostic tools and visualization strategies into the MetUM diagnostic toolkit. For example, simulated radar reflectivity was one of the first—and most revealing—diagnostic output fields to emerge from early CAM testing in the United States (e.g., Koch et al. 2005), and it was quickly adopted by the operational forecasting community in this country. However, when Met Office participation in the SFE ramped up in 2013, simulated reflectivity was not available in the MetUM's operational diagnostic package (although offline code existed so it could be generated from other MetUM output fields; e.g., see Stein et al. 2014). The pervasive use and obvious value of simulated reflectivity as a diagnostic tool in the SFE provided motivation for the Met Office to incorporate this field as a standard diagnostic in the MetUM, facilitating direct comparison with the U.S. models. These comparisons revealed previously undetected, higher-than-expected graupel concentrations associated with the operational MetUM microphysical parameterization. As a result, additional testing and research was

Fig. 2. NSSL Laboratory Director Steve Koch (left) and Met Office Chief Forecaster Steve Willington (right) discuss collaborative activities in the HWT, with SPC operations room visible in the background.

Brought to you by NOAA Central Library | Unauthenticated | Downloaded 09/24/21 07:35 PM UTC

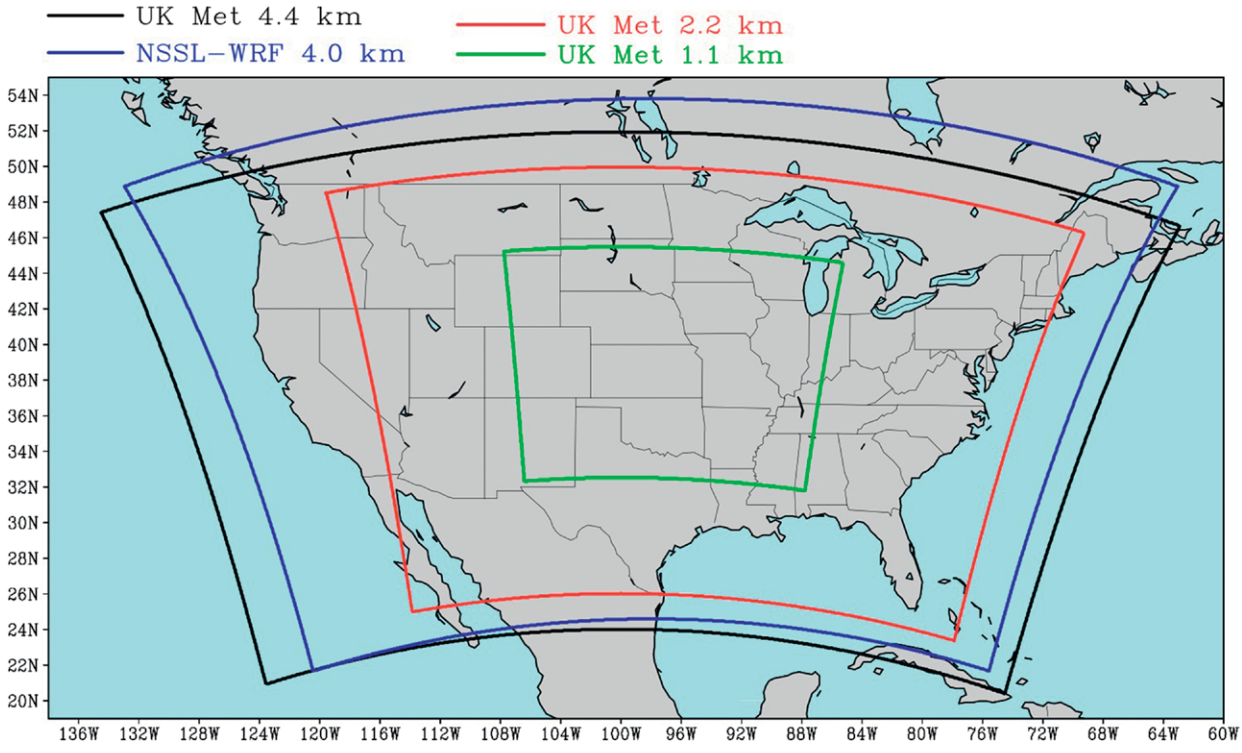

Fig. 3. Forecast domains used for CAMs discussed herein. The blue border encloses the NSSL-WRF domain, which uses 4.0-km grid spacing; the black border denotes the 4.4-km MetUM domain, which was used for the 2013 and 2014 SFEs; the red border marks the outer limits of the 2.2-km MetUM domain, which was used in 2013–15; and the green border outlines the 1.1-km grid mesh used by the MetUM in 2015. Note that NSSL runs a full 10-member ensemble over NSSL-WRF domain, but only the control member of this ensemble is discussed here.

conducted, eventually leading to important adjustments and improvements to the representation of microphysics in the MetUM. The Met Office is now routinely outputting the reflectivity fields from its U.K. NWP suite. Several MetUM users outside of the United Kingdom have also started to make use of the simulated reflectivity for verification and forecasting purposes.

Another widely used CAM diagnostic field, updraft helicity (Kain et al. 2008), has also become part of the output diagnostic suite in the MetUM. This field has proven to be valuable for identifying and tracking supercell-like features in high-resolution model forecasts, which is significant because, among all thunderstorms, supercells are associated with a disproportionate share of severe weather reports (e.g., Duda and Gallus 2010; Smith et al. 2012).

The HWT experience has also inspired complementary research initiatives within the Met Office. For example, Met Office scientists have simulated several high-profile tornado events from the 2013 season using a domain with 100-m grid spacing nested within the 2.2-km grid that they used in real time during the 2013 SFE. At 100-m grid length the MetUM is able to produce realistic supercells with tornado-like vortices (Hanley et al. 2016).

Comparison of model performance 2014–15. CAM forecasts from numerous contributors were considered in the annual SFE during 2014 and 2015 (e.g., Jirak et al. 2014), but the focus here is on comparisons of those generated by NSSL and the Met Office. Scientists from both organizations initialized their CAMs daily at 0000 UTC, deriving initial and lateral boundary conditions by simply downscaling from coarser-resolution operational models. Specifically, the NSSL CAM was initialized using EMC's 12-km North American Mesoscale Model (NAM) (Rogers et al. 2009), while initial conditions for MetUM CAM configurations were derived from operational global versions of this model, which had 17-km grid spacing during this time period.

The NSSL CAM was a 4-km configuration of the Advanced Research version of Weather Research and Forecasting (WRF) Model (WRF-ARW; Skamarock et al. 2008) with 35 vertical levels (NSSL-WRF). It has been run year-round by NSSL in support of SPC operations since late 2006 (Kain et al. 2010), and the configuration was held constant during 2014 and 2015. The Met Office used multiple CAM configurations over these 2 years, with each configuration being a high-resolution, slightly modified version of the operational MetUM, including 70 vertical levels. Two

Brought to you by NOAA Central Library | Unauthenticated | Downloaded 09/24/21 07:35 PM UTC

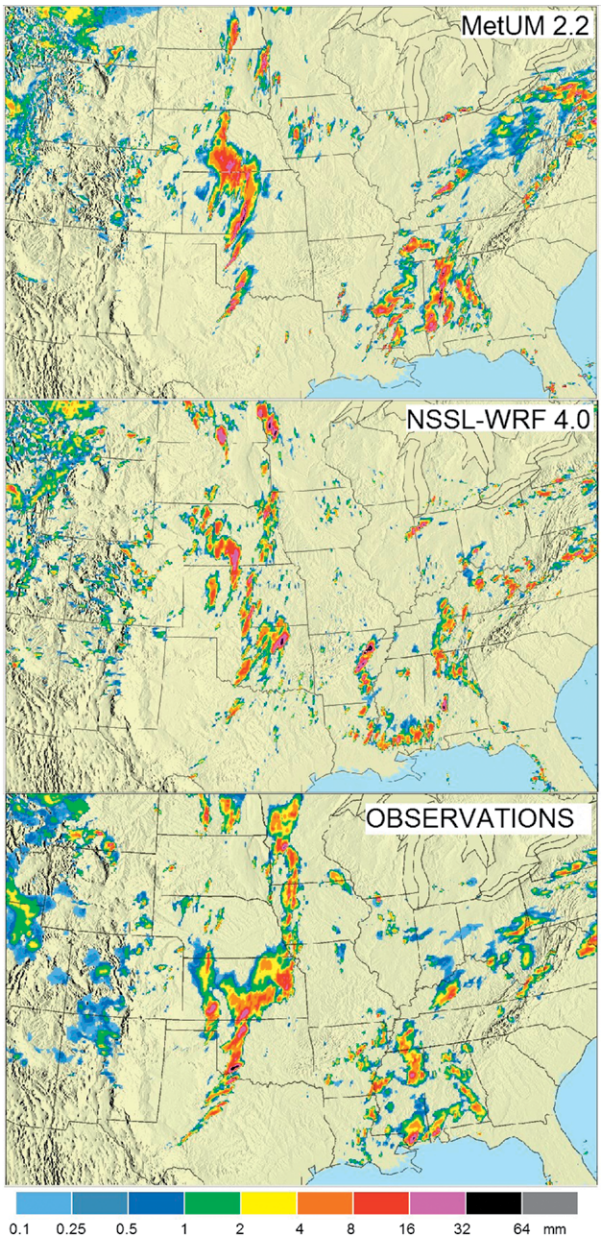

Fig. 4. An example of 24-h forecasts of 1-h accumulated precipitation from the (top) MetUM 2.2 km, (middle) NSSL-WRF 4 km, and the (bottom) National Centers for Environmental Prediction hourly accumulation radar composite valid at 0000 UTC 17 May 2015.

different configurations were used in 2014. The first had 4.4-km grid spacing and was downscaled directly from the global MetUM, while the second had 2.2-km spacing and was nested within the coarser-resolution (4.4 km) domain. In 2015, three different configurations were used. Specifically, there were two 2.2-km versions with identical forecast domains and identical initial conditions derived from the global model, but one differed by using an experimental parameterization of partial cloudiness (labeled PC2 in graphics shown here). The third configuration used in 2015 had 1.1-km grid spacing, non-PC2 physics, and was nested within the non-PC2 2.2-km grid, from which it drew its initial and lateral boundary conditions forecasts. The forecast domains are shown in Fig. 3. A pair of 24-h forecasts from the 2.2-km MetUM and 4-km NSSL-WRF models is shown in Fig. 4, indicating that both CAMs are capable of generating rainfall fields with a great deal of realism.

Subjective evaluations of radar reflectivity forecasts. A staple of the SFE is the systematic subjective assessment of weather forecasts, which is a forecast from both NWP models and human forecast teams (Kain et al. 2003a). In the 2014 SFE, this included a direct comparison of daily NSSL-WRF and 4.4-km MetUM forecasts of high-impact weather. Specifically, SFE forecast teams were asked to indicate (retrospectively) which CAM provided better guidance for the high-impact convective weather that occurred, that is, which model forecasts of convective storms corresponded better with observed radar. As shown in Fig. 5, the MetUM forecasts were judged to be superior 50% of the time, compared to 30% for the NSSL-WRF. Neither one was considered better than the other on 20% of the days. In addition to this simple either/or assessment, the teams were asked to note distinguishing characteristics of and differences in CAM forecasts. This information has proven to be very useful over the years for identifying and documenting systematic biases, strengths, and weaknesses of different modeling systems.

Objective verification of quantitative precipitation forecasts. Precipitation forecasts from the CAMs were examined during the SFE on a daily basis, and after the experiment they were objectively verified in an aggregate sense, using precipitation observations from the Multi-Radar Multi-Sensor system developed by NSSL (Zhang et al. 2016; Smith et al. 2016). Results from three different metrics are shown here.

The first metric is the equitable threat score (ETS; also called the Gilbert skill score), which has been used in the United States as a bellwether metric for precipitation forecasts for decades (e.g., Olson et al. 1995; Baldwin and Kain 2006). It provides a measure of the degree of spatial overlap between forecasted and observed precipitation events. It is commonly agreed that, for CAM forecasts, it is useful to compute this score on a "neighborhood" basis, where exact gridpoint correspondence is relaxed, owing to inherently low predictability on the small scales of individual forecast grid points (Roberts and Lean 2008; Ebert 2009). Here, ETSs are measured by

Brought to you by NOAA Central Library | Unauthenticated | Downloaded 09/24/21 07:35 PM UTC

comparing CAM forecasts and observations at the grid point [radius of influence (ROI) of zero] and within a radius of 24 km, focusing on the 2014 results. Not surprisingly, the scores are consistently higher with the nonzero radius (Figs. 6a–c), regardless of precipitation threshold. Another obvious systematic difference is higher ETSs for the MetUM during the first 12 h of the 36-h forecasts, both for the 4.4- and 2.2-km configurations. The ETSs are quite similar for all CAMs during the 12–36-h period of the forecasts.

The second metric is frequency bias (commonly referred to as simply the bias), which is often used in combination with ETS. The bias is a ratio of total areal forecast coverage to observed coverage, without consideration of location. With a perfect bias being 1, it is evident that both of the MetUM CAMs tend to overestimate the areal coverage of precipitation for the entire forecast period, but the high bias appears to be more pronounced at higher precipitation thresholds and particularly during the first 12 h (Figs. 6d–f). These scores are from 2014, but the same tendencies were noted in 2015 (bias scores from 2015 are not shown).

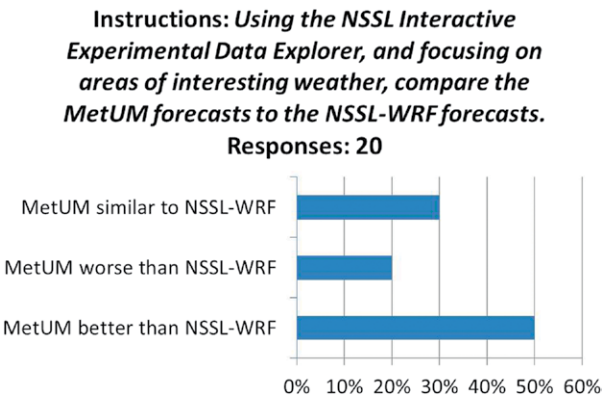

Fig. 5. Subjective comparison of daily MetUM and NSSL-WRF forecasts during the 2014 SFE.

The third metric used here is the fractions skill score (FSS; Roberts and Lean 2008), which provides information about the degree of correspondence between forecast and observed features as a function of spatial scale. Focusing now on 2015 results, Fig. 7 (left) shows that FSSs measured using a 16 mm $(24\ h)^{-1}$ accumulation threshold improve when assessed over increasingly larger scales. The MetUM scores are somewhat

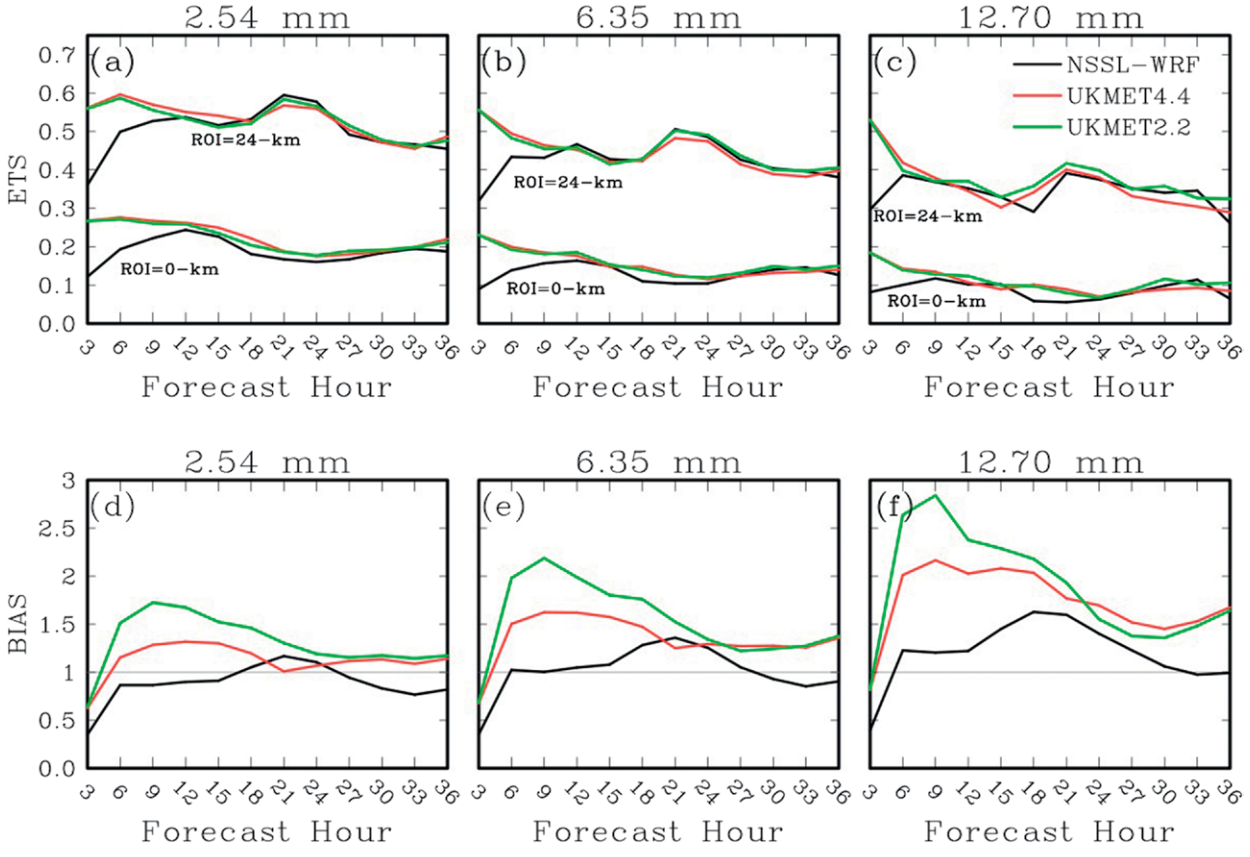

Fig. 6. (a)–(c) Aggregate neighborhood equitable threat scores and (d)–(f) frequency bias scores from NSSL-WRF and MetUM (Met Office) model forecasts during the 2014 HWT SFE as a function of forecast hour for different hourly accumulation thresholds.

Brought to you by NOAA Central Library | Unauthenticated | Downloaded 09/24/21 07:35 PM UTC

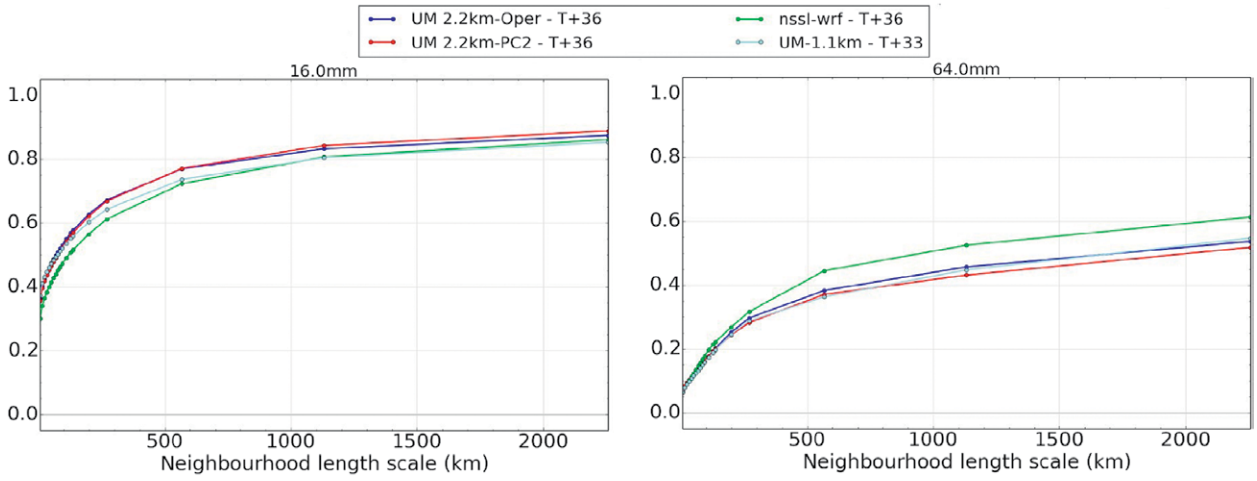

Fig. 7. Fractions skill score for 24-h precipitation accumulation using a (left) 16-mm threshold and (right) 64-mm threshold for the NSSL-WRF and MetUM model forecasts during the 2015 SFE.

higher than the NSSL-WRF on all scales. For a higher threshold of 64 mm (24 h)$^{-1}$, the NSSL-WRF outscores the MetUM at all but the smallest scales. The relatively low FSS from the MetUM for larger scales and higher accumulation thresholds may reflect the model's high bias at these thresholds, as high bias tends to have a negative impact on FSSs, especially on larger scales (Mittermaier and Roberts 2010).

In combination, these ETS, FSS, and bias scores reveal important systematic differences in model performance. For example, both of the MetUM configurations appear to "spin up" precipitation at the start of the forecast period much more quickly than the NSSL-WRF does. Yet, this more timely adjustment appears to come at a cost: overprediction of precipitation coverage. It seems likely that these systematic biases result from some combination of deficiencies in model physics and initial conditions. Whatever the source, it is important to understand the physical reasons for these systematic differences and find an optimal balance between rapid spinup and appropriate coverage/intensity.

Further objective information can be obtained by examining the characteristics of the rainfall cells (Fig. 8). The MetUM 2.2- and 1.1-km CAMs and NSSL-WRF 4-km CAM all produce too many small cells, and the NSSL-WRF tends to produce too few larger cells (diameter > 250 km). The NSSL-WRF 4-km model produces too many cells of all intensities, whereas the MetUM models are closer in number for the very light showers but have noticeably too many very heavy rain cells even beyond values measured by radar. Once again this provides more insight into the differing behavior of the models, tying in with the findings

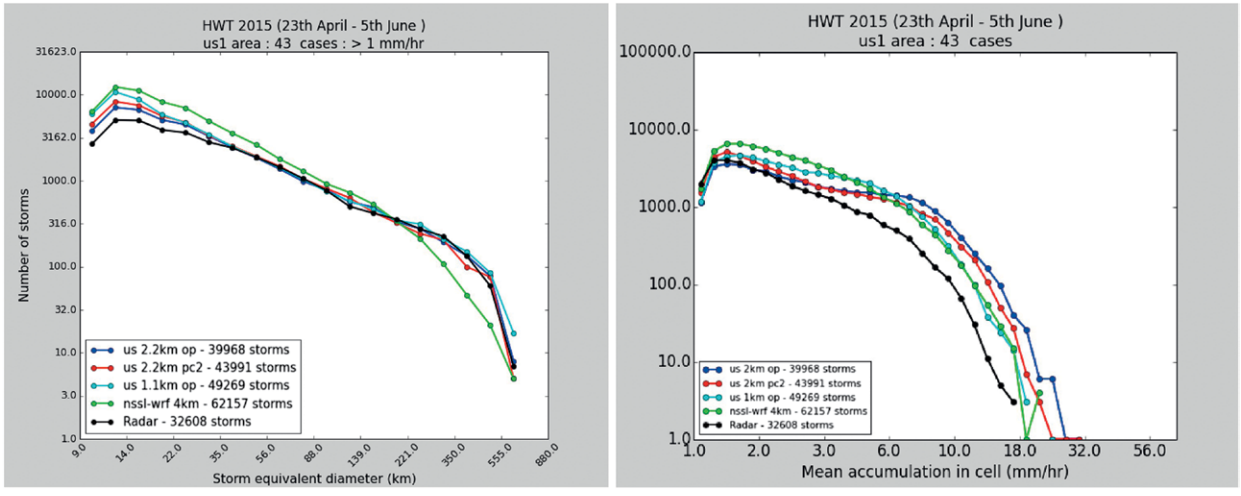

Fig. 8. Cell statistics showing (left) the numbers of cells with different diameters for rainfall exceeding 1 mm h^{-1} and (right) the numbers of cells with different mean accumulations over an hour from the MetUM 1.1- and 2.2-km, NSSL-WRF 4-km, and the NCEP composite radar. All the processing was done on a common 4-km grid.

Brought to you by NOAA Central Library | Unauthenticated | Downloaded 09/24/21 07:35 PM UTC

from the verification scores, and is useful when delving deeper into the physical processes responsible.

Vertical profiles. Severe weather forecasters rely heavily on soundings from raobs and from model forecast soundings [point forecast soundings (PFCs)] to help them diagnose and predict the likelihood of initiation, potential intensity, mode of organization, and likelihood of hazards such as tornadoes, large hail, high winds, and/or heavy rain associated with deep convection (e.g., Baldwin et al. 2002; Coniglio et al. 2013). Consequently, corresponding PFCs from the NSSL-WRF and MetUM were compared during the 2014 SFE. A glaring difference was revealed in this comparison, particularly in environments with a convective boundary layer topped by a sharp, stable transition to an elevated mixed layer (commonly referred to as a capping inversion). The NSSL-WRF tended to provide an overly smooth depiction of this capping inversion or convective inhibition (CIN) layer (e.g., Fig. 9a), while the MetUM often represented this with noticeably greater fidelity (e.g., Fig. 9b). This difference is especially important because the structure of the CIN layer is believed to play an important role in modulating the timing, location, and incidence of convective initiation. The undesirable tendency for the NSSL-WRF to smooth out the vertical structure is the subject of ongoing investigation. Early testing suggests that this tendency cannot be avoided by simply increasing the vertical resolution of the model.

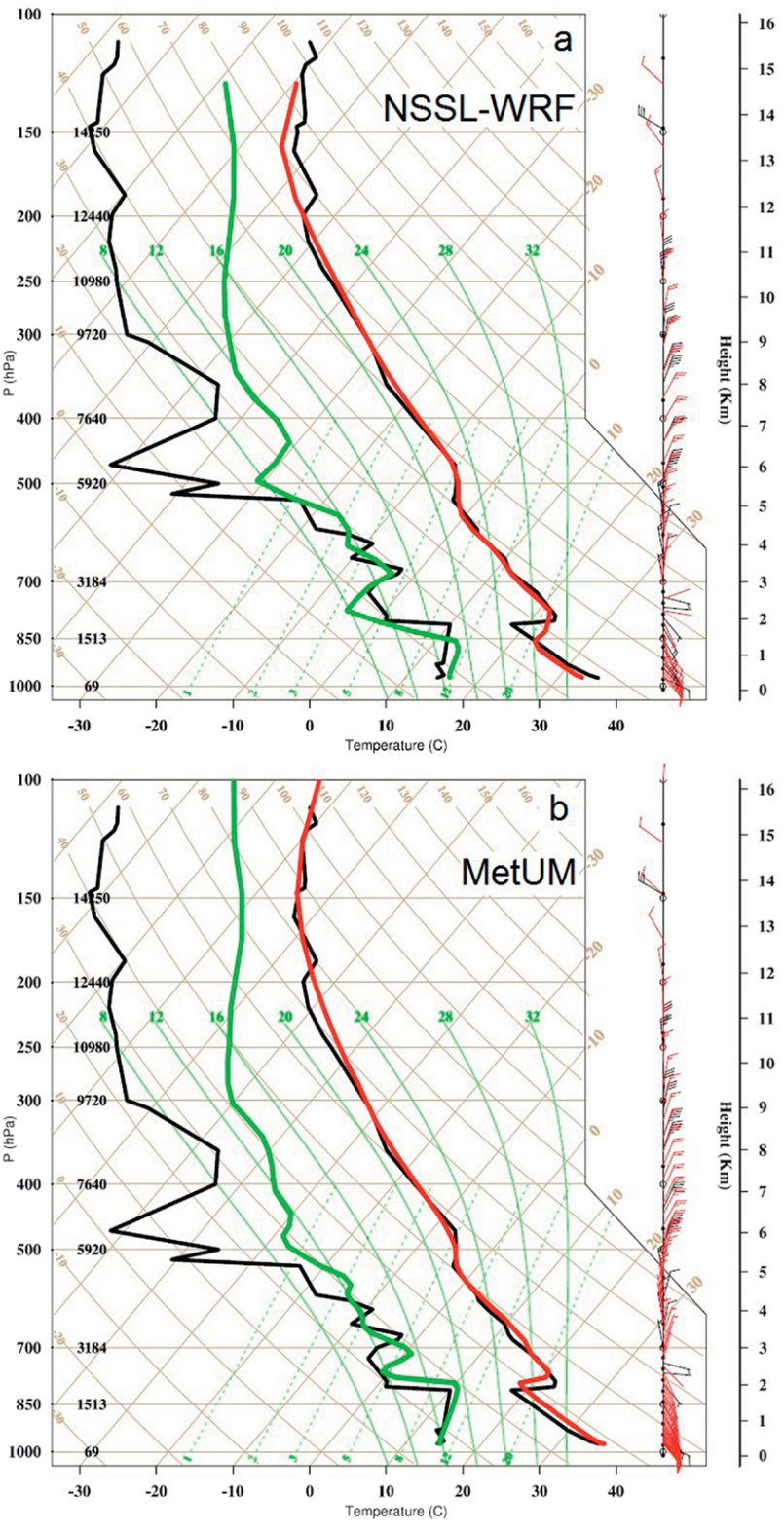

Fig. 9. An example showing corresponding PFCs (24-h forecasts) for the (a) NSSL-WRF and (b) MetUM, overlaid on observations (thick black profiles).

SUMMARY. The NOAA HWT provides a stimulating environment that cultivates collaborative activities between meteorological researchers and practitioners. For more than a decade this test bed has provided an environment in which individuals from these groups of professionals have learned to appreciate the insights, skills, and efficacy of the other, enhancing collaborative relationships and resulting in operational forecasting practices that more quickly

Brought to you by NOAA Central Library | Unauthenticated | Downloaded 09/24/21 07:35 PM UTC

adapt to emerging technologies and scientific advances while better informing operationally relevant research initiatives. The enhancement of complementary research-to-operations (R2O) and operations-to-research (O2R) interactions is a core mission of NOAA test beds (Ralph et al. 2013).

The underlying enthusiasm for activities in the HWT is fuelled by a compelling mutual interest in applied research related to prediction of high-impact weather. In the HWT's annual SFE, this grassroots enthusiasm has its foundation in interactions between NOAA's national centers for expertise in severe weather research (NSSL) and forecasting (SPC), with additional contributions from other domestic agencies. In recent years, the SFE has been invigorated by contributions from the Met Office, which in 2013 initiated a strong commitment to the SFE that significantly expanded international participation and provided complementary insight and technical skills in both forecasting and research. Specifically, the Met Office brought expertise gained from its efforts using CAMs to better represent the convective storms that bring flash flooding in the United Kingdom. The infusion of Met Office models and perspectives dovetailed exceptionally well with the rapidly growing NSSL and SPC proficiency in using CAMs to help predict tornadoes, large hail, and damaging winds.

The challenges associated with prediction of these and other high-impact weather phenomena are too great to be addressed entirely by any single institution. The successful collaborative efforts of the HWT, NSSL, SPC, and Met Office are demonstrating that international collaboration can provide synergy, efficiency, and important scientific advances when it is strongly supported at both grassroots and institutional levels.

ACKNOWLEDGMENTS. The annual SFE, conducted in the NOAA HWT, has benefited for many years from the generous support of the NSSL, the SPC, and the University of Oklahoma Cooperative Institute for Mesoscale Meteorological Studies (CIMMS), including the technical support and administrative support staffs from these organizations. In addition, the NOAA/OAR/Office of Weather and Air Quality has dramatically increased its support of the HWT in recent years, with considerable benefit to the organization, facilities, and experimental activities. Authors Knopfmeier and Karstens were funded by NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma Cooperative Agreement NA11OAR4320072, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of NOAA or the U.S. Department of Commerce.

Since 2012 the SFE has received additional benefits from the generous support of the following Met Office teams: Forecasting and Service Delivery, with particular thanks to Iain Forsyth (executive head), Paul Davies (executive head forecasting), and Martin Cumper (strategic resources manager); Public Weather Service, with particular thanks to Derrick Ryall (executive head) and Becky Moore (international defense); Weather Science, with particular thanks to Dale Barker (deputy director), Ken Mylne (head, numerical modelling), Clive Wilson (manager mesoscale model development), Adrian Semple (forecast monitoring project leader), and Tom Blackmore (satellite applications research scientist); Foundation Science, with particular thanks to Douglas Boyd (senior scientist, science partnerships); Simon Vosper and Adrian Lock (atmospheric processes and parameterization); and Technology and Information Services, with particular thanks to Dave Robinson (operational NWP suite team leader) and Siân Collins (project manager).

Collaborations between the Met Office and NSSL are facilitated by a formal memorandum of agreement between these two organizations, dated 28 February 2014.

Many thanks to Harold Brooks of NSSL and four anonymous reviewers for their helpful and constructive comments on earlier versions of this manuscript.

REFERENCES

Baldwin, M. E., and J. S. Kain, 2006: Sensitivity of several performance measures to displacement error, bias, and event frequency. *Wea. Forecasting*, **21**, 636–648, doi:10.1175/WAF933.1.

——, ——, and M. P. Kay, 2002: Properties of the convection scheme in NCEP's Eta Model that affect forecast sounding interpretation. *Wea. Forecasting*, **17**, 1063–1079, doi:10.1175/1520-0434(2002)017<1063:POTCSI>2.0.CO;2.

Bowler, N. E., A. Arribas, K. R. Mylne, K. B. Robertson, and S. E. Beare, 2008: The MOGREPS short-range ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **134**, 703–722, doi:10.1002/qj.234.

Bright, D. R., S. J. Weiss, J. J. Levit, M. S. Wandishin, J. S. Kain, and D. J. Stensrud, 2004: Evaluation of short-range ensemble forecasts during the 2003 SPC/NSSL Spring Program. 22nd Conf. on Severe Local Storms, Hyannis, MA, Amer. Meteor. Soc., P15.5. [Available online at https://ams.confex.com/ams/11aram22sls/techprogram/paper_83403.htm.]

Brooks, H. E., M. S. Tracton, D. J. Stensrud, G. DiMego, and Z. Toth, 1995: Short-range ensemble forecasting: Report from a workshop (25-27 July 1994). *Bull. Amer. Meteor. Soc.*, **76**, 1617–1624.

Clark, A. J., and Coauthors, 2012: An overview of the 2010 Hazardous Weather Testbed Experimental Forecast

Brought to you by NOAA Central Library | Unauthenticated | Downloaded 09/24/21 07:35 PM UTC

Program Spring Experiment. *Bull. Amer. Meteor. Soc.*, **93**, 55–74, doi:10.1175/BAMS-D-11-00040.1.

Clark, P., N. Roberts, H. Lean, S. P. Ballard, and C. Charlton-Perez, 2016: Convection-permitting models: A step-change in rainfall forecasting. *Meteor. Appl.*, **23**, 165–181, doi:10.1002/met.1538.

Coniglio, M. C., 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*, **28**, 842–862, doi:10.1175/WAF-D-12-00103.1.

Davies, T., M. J. P. Cullen, A. J. Malcolm, M. H. Mawson, A. Staniforth, A. A. White, and N. Wood, 2005: A new dynamical core for the Met Office's global and regional modelling of the atmosphere. *Quart. J. Roy. Meteor. Soc.*, **131**, 1759–1782, doi:10.1256/qj.04.101.

Du, J., and M. S. Tracton, 2001: Implementation of a real-time short-range ensemble forecasting system at NCEP: An update. Preprints, *Ninth Conf. on Mesoscale Processes*, Ft. Lauderdale, FL, Amer. Meteor. Soc., 355–356.

Duda, J. D., and W. A. Gallus Jr., 2010: Spring and summer midwestern severe weather reports in supercells compared to other morphologies. *Wea. Forecasting*, **25**, 190–206, doi:10.1175/2009WAF2222338.1.

Ebert, E. E., 2009: Neighborhood verification: A strategy for rewarding close forecasts. *Wea. Forecasting*, **24**, 1498–1510, doi:10.1175/2009WAF2222251.1.

Golding, B., N. Roberts, G. Leoncini, K. Mylne, and R. Swinbank, 2016: MOGREPS-UK convection-permitting ensemble products for surface water flood forecasting: Rationale and first results. *J. Hydrometeor.*, **17**, 1383–1406, doi:10.1175/JHM-D-15-0083.1.

Hanley, K. E., A. I. Barrett, and H. W. Lean, 2016: Simulating the 20 May 2013 Moore, Oklahoma tornado with a 100-metre grid-length NWP model. *Atmos. Sci. Lett.*, **17**, 453–461, doi:10.1002/asl.678.

Homar, V., D. J. Stensrud, J. J. Levit, and D. R. Bright, 2006: Value of human-generated perturbations in short-range ensemble forecasts of severe weather. *Wea. Forecasting*, **21**, 347–363, doi:10.1175/WAF920.1.

Jirak, I. L., S. J. Weiss, and C. J. Melick, 2012: The SPC storm-scale ensemble of opportunity: Overview and results from the 2012 Hazardous Weather Testbed Spring Forecasting Experiment. *26th Conf. on Severe Local Storms*, Nashville, TN, Amer. Meteor. Soc., P9.137. [Available online at https://ams.confex.com/ams/26SLS/webprogram/Paper211729.html.]

——, and Coauthors, 2014: An overview of the 2014 NOAA Hazardous Weather Testbed Spring Forecasting Experiment. *27th Conf. on Severe Local Storms*, Madison, WI, Amer. Meteor. Soc., 46. [Available online at https://ams.confex.com/ams/27SLS/webprogram/Paper254650.html.]

Johns, R. H., and C. A. Doswell III, 1992: Severe local storms forecasting. *Wea. Forecasting*, **7**, 588–612, doi:10.1175/1520-0434(1992)007<0588:SLSF>2.0.CO;2.

Kain, J. S., M. E. Baldwin, P. R. Janish, S. J. Weiss, M. P. Kay, and G. W. Carbin, 2003a: Subjective verification of numerical models as a component of a broader interaction between research and operations. *Wea. Forecasting*, **18**, 847–860, doi:10.1175/1520-0434(2003)018<0847:SVONMA>2.0.CO;2.

——, P. R. Janish, S. J. Weiss, M. E. Baldwin, R. S. Schneider, and H. E. Brooks, 2003b: Collaboration between forecasters and research scientists at the NSSL and SPC: The Spring Program. *Bull. Amer. Meteor. Soc.*, **84**, 1797–1806, doi:10.1175/BAMS-84-12-1797.

——, 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, doi:10.1175/WAF906.1.

——, and Coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting*, **23**, 931–952, doi:10.1175/WAF2007106.1.

——, 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, doi:10.1175/2010WAF2222430.1.

Karstens, C. D., and Coauthors, 2015: Evaluation of a probabilistic forecasting methodology for severe convective weather in the 2014 Hazardous Weather Testbed. *Wea. Forecasting*, **30**, 1551–1570, doi:10.1175/WAF-D-14-00163.1.

Koch, S. E., B. Ferrier, M. Stoelinga, 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. *11th Conf. on Mesoscale Processes*, Albuquerque, NM, Amer. Meteor. Soc., J4J.7. [Available online at https://ams.confex.com/ams/32Rad11Meso/techprogram/paper_97032.htm.]

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 United Kingdom. *Mon. Wea. Rev.*, **136**, 3408–3424, doi:10.1175/2008MWR2332.1.

Lewis, J. M., C. M. Hayden, R. T. Merrill, and J. M. Schneider, 1989: GUFMEX: A study of return flow in the Gulf of Mexico. *Bull. Amer. Meteor. Soc.*, **70**,

Brought to you by NOAA Central Library | Unauthenticated | Downloaded 09/24/21 07:35 PM UTC

24–29, doi:10.1175/1520-0477(1989)070<0024:GASORF>2.0.CO;2.

Mittermaier, M., and N. Roberts, 2010: Intercomparison of spatial forecast verification methods: Identifying skillful spatial scales using the fractions skill score. *Wea. Forecasting*, **25**, 343–354, doi:10.1175/2009WAF2222260.1.

Olson, D. A., N. W. Junker, and B. Korty, 1995: Evaluation of 33 years of quantitative precipitation forecasting at the NMC. *Wea. Forecasting*, **10**, 498–511, doi:10.1175/1520-0434(1995)010<0498:EOYOQP>2.0.CO;2.

Ralph, F. M., and Coauthors, 2013: The emergence of weather-related test beds linking research and forecasting operations. *Bull. Amer. Meteor. Soc.*, **94**, 1187–1211, doi:10.1175/BAMS-D-12-00080.1.

Roberts, N., 2005: An investigation of the ability of a storm scale configuration of the Met Office NWP model to predict flood-producing rainfall. Met Office Forecasting Research Tech. Rep. 455, 80 pp. [Available online at www.metoffice.gov.uk/media/pdf/m/g/FRTR455.pdf.]

——, and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, doi:10.1175/2007MWR2123.1.

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

Schwartz, C. S., and Coauthors, 2009: Next-day convection-allowing WRF Model guidance: A second look at 2-km versus 4-km grid spacing. *Mon. Wea. Rev.*, **137**, 3351–3372, doi:10.1175/2009MWR2924.1.

——, and Coauthors, 2010: Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Wea. Forecasting*, **25**, 263–280, doi:10.1175/2009WAF2222267.1.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., doi:10.5065/D68S4MVH.

Smith, B. T., R. L. Thompson, J. S. Grams, C. Broyles, and H. E. Brooks, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part I: Storm classification and climatology. *Wea. Forecasting*, **27**, 1114–1135, doi:10.1175/WAF-D-11-00115.1.

Smith, T. M., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) severe weather and aviation products: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 1617–1630, doi:10.1175/BAMS-D-14-00173.1.

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, doi:10.1175/WAF-D-10-05046.1.

Stein, T. H. M., R. J. Hogan, K. E. Hanley, J. C. Nicol, H. W. Lean, R. S. Plant, P. A. Clark, and C. E. Halliwell, 2014: The three-dimensional morphology of simulated and observed convective storms over southern England. *Mon. Wea. Rev.*, **142**, 3264–3283, doi:10.1175/MWR-D-13-00372.1.

Stensrud, D. J., 2001: Using short-range ensemble forecasts for predicting severe weather events. *Atmos. Res.*, **56**, 3–17, doi:10.1016/S0169-8095(00)00079-X.

——, and S. J. Weiss, 2002: Mesoscale model ensemble forecasts of the 3 May 1999 tornado outbreak. *Wea. Forecasting*, **17**, 526–543, doi:10.1175/1520-0434(2002)017<0526:MMEFOT>2.0.CO;2.

——, H. E. Brooks, J. Du, M. S. Tracton, and E. Rogers, 1999: Using ensembles for short-range forecasting. *Mon. Wea. Rev.*, **127**, 433–446, doi:10.1175/1520-0493(1999)127<0433:UEFSRF>2.0.CO;2.

——, J.-W. Bao, and T. T. Warner, 2000: Using initial condition and model physics perturbations in short-range ensembles of mesoscale convective systems. *Mon. Wea. Rev.*, **128**, 2077–2107, doi:10.1175/1520-0493(2000)128<2077:UICAMP>2.0.CO;2.

Tang, Y., H. Lean, and J. Bornemann, 2013: The benefits of the Met Office variable resolution NWP model for forecasting convection. *Meteor. Appl.*, **20**, 417–426, doi:10.1002/met.1300.

Tracton, S., J. Du, Z. Toth, and H. Juang, 1998: Short-range ensemble forecasting (SREF) at NCEP/EMC. Preprints, *12th Conf. on Numerical Weather Prediction*, Phoenix, AZ, Amer. Meteor. Soc., 269–272.

Xue, M., and Coauthors, 2007: CAPS realtime storm-scale ensemble and high-resolution forecasts as part of the NOAA Hazardous Weather Testbed 2007 Spring Experiment. *22th Conf. on Weather Analysis and Forecasting/18th Conf. on Numerical Weather Prediction*, Park City, UT, Amer. Meteor. Soc., 3B.1. [Available online at https://ams.confex.com/ams/22WAF18NWP/techprogram/paper_124587.htm.]

Zhang, J. K., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 621–638, doi:10.1175/BAMS-D-14-00174.1..

Brought to you by NOAA Central Library | Unauthenticated | Downloaded 09/24/21 07:35 PM UTC