SPRING FORECASTING EXPERIMENT 2016

Conducted by the

EXPERIMENTAL FORECAST PROGRAM

of the

NOAA/HAZARDOUS WEATHER TESTBED

http://hwt.nssl.noaa.gov/Spring_2016/

HWT Facility – National Weather Center
2 May - 3 June 2016

Program Overview and Operations Plan

4/29/2016

Adam Clark[2,3], Israel Jirak[1], Chris Melick[1,3], James Correia[1,3], Steven J. Weiss[1], Jack Kain[2], Andy Dean[1,3], Kent Knopfmeier[2,3], Chris Karstens[2,3] and Burkely Twiest[4]

(1) NOAA/NWS/NCEP Storm Prediction Center, Norman, Oklahoma
(2) NOAA/OAR National Severe Storms Laboratory, Norman, Oklahoma
(3) Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma
(4) School of Meteorology, University of Oklahoma, Norman, Oklahoma

1. Introduction

Each spring, the Experimental Forecast Program (EFP) of the NOAA/Hazardous Weather Testbed (HWT), organized by the Storm Prediction Center (SPC) and National Severe Storms Laboratory (NSSL), conducts a collaborative experiment to test emerging concepts and technologies designed to improve the prediction of hazardous convective weather. The primary goals of the HWT are to accelerate the transfer of promising new tools from research to operations, to inspire new initiatives for operationally relevant research, and to identify and document sensitivities and the performance of state-of-the art convection-allowing (1 to 4 km grid-spacing) experimental modeling systems (CAMs).

The 2016 Spring Forecasting Experiment (SFE2016), a cornerstone of the EFP, will be conducted 2 May – 3 June with participation expected from more than 80 forecasters, researchers, and model developers from around the world. Building upon successful experiments of previous years, a main emphasis of SFE2016 will be the generation of probabilistic forecasts of severe weather valid over shorter time periods than current SPC operational products. This will be an important step toward addressing a strategy within the National Weather Service of providing nearly continuous probabilistic hazard forecasts on increasingly fine spatial and temporal scales, consistent with the NOAA Forecasting a Continuum of Environmental Threats (FACETs) vision. As in previous experiments, a suite of new and improved experimental CAM guidance contributed by our large group of collaborators will be central to the generation of these forecasts. However, this year a major effort has been made to coordinate CAM-based ensemble configurations much more closely than in previous years. Specifically, instead of each group providing a separate, independently designed CAM-based ensemble, all groups have agreed on a set of model specifications (e.g., grid-spacing, vertical levels, domain size, physics) so that the simulations contributed by each group can be used in carefully designed controlled experiments. This design will allow us to conduct several experiments geared toward identifying optimal configuration strategies for CAM-based ensembles, and is especially well timed to help inform the design of the first operational CAM-based ensemble for the US, which is planned for implementation by NOAA's NCEP/Environmental Modeling Center (EMC) in the upcoming years. This large number of CAM members has been termed the Community Leveraged Unified Ensemble, or CLUE, and includes 65 members using 3-km grid-spacing that will allow a set of eight unique experiments.

This operations plan summarizes the core interests of SFE2016 and provides information on the operations of the experiment. Detailed information on the organizational structure of the HWT and information on various forecast tools and diagnostics can also be found in this document. The remainder of the operations plan is organized as follows: Section 2 provides details on a number of new products being introduced during SFE2016 and Section 3 describes the core interests and new concepts being introduced for SFE2016. A list of daily participants, details on the SFE forecasting, and more general information on the HWT are found in appendices.

2. Overview of Experimental Products and Models

A primary goal of the SFE2016 forecasting activities will be to test methods for generating probabilistic forecasts of severe weather that are valid over shorter time windows than current SPC operational products. Two separate groups, led by SPC and NSSL staff, will issue slightly different sets of convective outlooks for this testing. One group will issue Day 1 and 2 full-period outlooks (1600 to 1200 UTC for Day 1 and 1200 to 1200 UTC for Day 2) for individual severe weather hazards (tornado,

wind, and hail), along with two 4-h period outlooks within the Day 1 period for each hazard covering the periods 1800 to 2200 UTC and 2200 to 0200 UTC. The other group will be issuing Day 1, and/or Day 2 and Day 3 full-period outlooks for total severe (i.e., outlook for combined hazards of severe hail, wind, or tornadoes), along with five 4-h period outlooks within the Day 1 period for total severe valid 1800-2200 UTC, 2000-0000 UTC, 2200-0200 UTC, 0000-0400 UTC, and 0200-0600 UTC. Additionally, isochrones of severe weather will be drawn at two-hour intervals on top of the full period Day 1 total severe probabilities to delineate the start-time of the 4-h time window with the highest total severe probabilities. For example, an area encompassed by the 2000 UTC and 0000 UTC isochrones would expect total severe probabilities to be highest during the 2000-0000 UTC period. We are exploring the feasibility of drawing isochrones of severe weather to add more detailed timing information to outlooks as an alternative (or supplement) to issuing more frequent outlooks valid for shorter periods of time. These products are slightly different from those of SFE2014 and SFE2015, but the goals are the same – namely, to explore different ways of introducing probabilistic severe weather forecasts on time/space scales that are not currently addressed with categorical forecast products (e.g., SPC Mesoscale Discussions and Severe Thunderstorm/Tornado Watches). This is an effort to begin exploring ways of seamlessly merging probabilistic severe weather outlooks with probabilistic severe weather warnings as part of NOAA's Warn-on-Forecast (WoF; Stensrud et al. 2009) and Forecasting a Continuum of Environmental Threats (FACETs; http://www.nssl.noaa.gov/projects/facets/) initiatives.

Generating the forecasts described above will be intensive and will thus rely on deterministic and ensemble CAM output for guidance and to generate first guesses for the severe weather probabilities. Most of the CAMs will be based on recent versions of the Advanced Research Weather Research and Forecasting (WRF-ARW) model, as well as the Nonhydrostatic Multi-Scale Model on the B grid (NMMB). In addition to the ensemble subsets contained within the 65-member CLUE system, the United Kingdom Meteorological (UKMET) Office will provide three CAM forecasts (two with 2.2 km grid spacing and one with 1.1 km grid spacing) that are based on their Unified Modeling System, and 3-km grid-spacing forecasts over the CONUS will be provided from the National Center for Atmospheric Research (NCAR) Model for Prediction Across Scales – a global model with a variable-resolution mesh and "scale-aware" physics. Also, SPC's 7-member Storm Scale Ensemble of Opportunity (SSEO) comprised of CAM output available to SPC operationally (or near operationally) since 2011 will be used, along with NSSL's 10-member, 4-km grid-spacing ensemble that began running year-round in 2014.

For the generation of first-guess guidance forecasts from the CAM ensembles, it is important to extract explicit and proxy variables in the forecasts that track the potential of severe weather in the models. Previous SFEs and operational experience has shown that fields like hourly-maximum updraft helicity (UH) and hourly-maximum wind speed at the near the surface can be effective for highlighting the likelihood of severe weather in the model (Sobash et al. 2010, Kain et al. 2010, Clark et al. 2013). To support the goal of the SFE2016 to generate forecasts of individual hazards, there will be further efforts to explore the ability of new model fields to delineate individual hazards, particularly for the size of hail and threat for tornadoes. Based on ongoing research and findings from previous SFEs, updates have been made to the HAILCAST algorithm (Adams-Selin 2012; Adams-Selin and Ziegler 2016), which predicts maximum hail size using a hail growth model coupled to WRF, so this will be tested again in SFE2016. Also, new hail size diagnostics will be available in the CLUE members from CAPS and NSSL that use WRF-ARW, which are based directly on information in the microphysical parameterizations from Greg Thompson of NCAR. Finally, probabilistic hail-size forecasts derived from a subset of the CLUE members using a machine-learning algorithm will also be tested. For diagnosing the potential for anti-cyclonically rotating storms, negative UH will be examined for the first time and

other UH-based diagnostics may be introduced for diagnosing tornado potential. The rest of this section provides further details on each modeling system.

a) The Community Leveraged Unified Ensemble (CLUE)

The CLUE is a carefully designed ensemble with subsets of members contributed by NSSL, CAPS, the University of North Dakota, NOAA's Earth Systems Research Laboratory/Global Systems Division (ESRL/GSD), and NCAR. All members are initialized weekdays at 0000 UTC with 3-km grid-spacing covering a CONUS domain. All WRF-ARW members have 1680 grid-points in the east-west direction, 1152 north-south, and 51 vertical levels[1] with a model top of 50 hPa and use RRTMG short and long wave radiation. ALL NMMB members have 1568 grid-points in the east-west direction, 1120 north-south, and 50 vertical levels with a model top of 50 hPa and use the RRTM short and long wave radiation. Depending on the CLUE subset, forecast lengths range from 36 to 60 h. Specifications for the members within each subset are detailed in the tables below.

Table 1. Specifications for the "core" members of the CLUE that use the ARW dynamic core, mixed physics, perturbed ICs and LBCs, and radar data assimilation using ARPS 3DVAR. NAMa and NAMf refer to 12 km NAM analysis and forecast, respectively. 3DVAR refers to ARPS 3DVAR and cloud analysis. RAPa is the 13 km Rapid Refresh model analysis. Under the IC column, the model names appended with "pert" refer to perturbations extracted from a 16 km grid-spacing SREF member. For members core03-10 under the BC column, names refer to SREF member forecasts.

(1) CAPS: mixed phys + radar						
Members	**IC**	**BC**	**Microphysics**	**LSM**	**PBL**	**Model**
core01	NAMa+3DVAR	NAMf	Thompson	NOAH	MYJ	arw
core02	RAPa+3DVAR	GFSf	Thompson	RUC	MYNN	arw
core03	core01+arw-p1_pert	arw-p1	P3	NOAH	YSU	arw
core04	core01+arw-n1_pert	arw-n1	MY	NOAH	MYNN	arw
core05	core01+arw-p2_pert	arw-p2	Morrison	NOAH	MYJ	arw
core06	core01+arw-n2_pert	arw-n2	P3	NOAH	YSU	arw
core07	core01+nmmb-p1_pert	nmmb-p1	MY	NOAH	MYNN	arw
core08	core01+nmmb-n1_pert	nmmb_n1	Morrison	NOAH	YSU	arw
core09	core01+nmmb-p2_pert	nmmb-p2	P3	NOAH	MYJ	arw
core10	core01+nmmb-n2_pert	nmmb-n2	Thompson	NOAH	MYNN	arw

Table 2. Specifications for the "s-phys-rad" members of the CLUE that use the ARW dynamic core, single physics, perturbed ICs and LBCs, and radar data assimilation using ARPS 3DVAR. The blue shaded member is repeated from Table 1 because it is also the control member of the s-phys-rad ensemble subset.

(2) CAPS: single phys + radar						
Ens Members	**IC**	**BC**	**Microphysics**	**LSM**	**PBL**	**Model**
core01	NAMa+3DVAR	NAMf	Thompson	NOAH	MYJ	arw
s-phys-rad02	core01+arw-p1_pert	arw-p1	Thompson	NOAH	MYJ	arw
s-phys-rad03	core01+arw-n1_pert	arw-n1	Thompson	NOAH	MYJ	arw
s-phys-rad04	core01+arw-p2_pert	arw-p2	Thompson	NOAH	MYJ	arw
s-phys-rad05	core01+arw-n2_pert	arw-n2	Thompson	NOAH	MYJ	arw
s-phys-rad06	core01+arw-p3_pert	arw-p3	Thompson	NOAH	MYJ	arw
s-phys-rad07	core01+nmmb-p1_pert	nmmb-p1	Thompson	NOAH	MYJ	arw
s-phys-rad08	core01+nmmb-n1_pert	nmmb-n1	Thompson	NOAH	MYJ	arw
s-phys-rad09	core01+nmmb-p2_pert	nmmb-p2	Thompson	NOAH	MYJ	arw
s-phys-rad10	core01+nmmb-n2_pert	nmmb-n2	Thompson	NOAH	MYJ	arw

[1] WRF-ARW sigma levels are set to: 1.0, 0.998, 0.994, 0.987, 0.975, 0.959, 0.939, 0.916, 0.892, 0.865, 0.835, 0.802, 0.766, 0.727, 0.685, 0.64, 0.592, 0.542, 0.497, 0.4565, 0.4205, 0.3877, 0.3582, 0.3317, 0.3078, 0.2863, 0.267, 0.2496, 0.2329, 0.2188, 0.2047, 0.1906, 0.1765, 0.1624, 0.1483, 0.1342, 0.1201, 0.106, 0.0919, 0.0778, 0.0657, 0.0568, 0.0486, 0.0409, 0.0337, 0.0271, 0.0209, 0.0151, 0.0097, 0.0047, 0.0.

Table 3. Specifications for the "caps-enkf" members of the CLUE. This ensemble is run as follows: The 3-km GSI-EnKF system will be initialized at 1800 UTC each day, and will assimilate the RAP/HRRR GSI data stream hourly (except satellite data) from 1800-0000 UTC over the CONUS domain. Radar data will be assimilated every 15 minutes from 2300-0000 UTC using the CAPS EnKF system. The ensemble consists of 40 ARW members with initial perturbations and mixed physics options to provide input for the EnKF ensemble analyses. Each member uses Thompson microphysics, although with varied parameter settings. A 9-member ensemble forecast (run for 60 h) follows using the final EnKF analyses at 0000 UTC using the same multi-physics configurations as are used for the core members. In addition, one deterministic forecast from the ensemble mean analysis at 0000 is also produced. The specifications for the 10 forecast members are listed below.

(3) CAPS: EnKF						
Ens Members	**IC**	**BC**	**Microphysics**	**LSM**	**PBL**	**Model**
caps-enkf01	enkf_m01a	NAMf	Thompson	NOAH	MYJ	arw
caps-enkf02	enkf_mna	GFSf	Thompson	NOAH	MYNN	arw
caps-enkf03	enkf_m08a	arw-p1	P3	NOAH	YSU	arw
caps-enkf04	enkf_m10a	arw-n1	MY	NOAH	MYNN	arw
caps-enkf05	enkf_m17a	arw-p2	Morrison	NOAH	MYJ	arw
caps-enkf06	enkf_m23a	arw-n2	P3	NOAH	YSU	arw
caps-enkf07	enkf_m24a	nmmb-p1	MY	NOAH	MYNN	arw
caps-enkf08	enkf_m12a	nmmb-n1	Morrison	NOAH	YSU	arw
caps-enkf09	enkf_m09a	nmmb-p2	P3	NOAH	MYJ	arw
caps-enkf10	enkf_m6a	nmmb-n2	Thompson	NOAH	MYNN	arw

Table 4. Specifications for the "caps-nmmb-rad" member of CLUE that uses the NMMB dynamic core with radar data assimilation and cloud analysis using ARPS 3DVAR.

(4) CAPS: NMMB + radar						
Ens Members	**IC**	**BC**	**Microphysics**	**LSM**	**PBL**	**Model**
caps-nmmb-rad	NAMa+3DVAR	NAMf	Ferrier-Aligo	NOAH	MYJ	nmmb

Table 5. Specifications for the "caps-nmmb" members of the CLUE that use the NMMB dynamic core, single physics, IC and LBC perturbations extracted from SREF members, and no radar data assimilation.

(5) CAPS: NMMB (no radar)						
Ens Members	**IC**	**BC**	**Microphysics**	**LSM**	**PBL**	**Model**
caps-nmmb01	NAMa+arw-p3_pert	arw-p3	Ferrier-Aligo	NOAH	MYJ	nmmb
caps-nmmb02	NAMa+nmmb-p1_pert	nmmb-p1	Ferrier-Aligo	NOAH	MYJ	nmmb
caps-nmmb03	NAMa+nmmb-n1_pert	nmmb-n1	Ferrier-Aligo	NOAH	MYJ	nmmb
caps-nmmb04	NAMa+nmmb-p2_pert	nmmb-p2	Ferrier-Aligo	NOAH	MYJ	nmmb
caps-nmmb05	NAMa+nmmb-n2_pert	nmmb-n2	Ferrier-Aligo	NOAH	MYJ	nmmb

Table 6. Specifications for the "s-phys-norad" members of the CLUE that use the ARW dynamic core, single physics, perturbed ICs and LBCs from SREF members, and no radar data assimilation.

(6) NSSL: single phys						
Ens Members	**IC**	**BC**	**Microphysics**	**LSM**	**PBL**	**Model**
s-phys-norad01	NAMa	NAMf	Thompson	NOAH	MYJ	arw
s-phys-norad02	NAMa+arw-p1_pert	arw-p1	Thompson	NOAH	MYJ	arw
s-phys-norad03	NAMa+arw-n1_pert	arw-n1	Thompson	NOAH	MYJ	arw
s-phys-norad04	NAMa+arw-p2_pert	arw-p2	Thompson	NOAH	MYJ	arw
s-phys-norad05	NAMa+arw-n2_pert	arw-n2	Thompson	NOAH	MYJ	arw
s-phys-norad06	NAMa+arw-p3_pert	arw-p3	Thompson	NOAH	MYJ	arw
s-phys-norad07	NAMa+nmmb-p1_pert	nmmb-p1	Thompson	NOAH	MYJ	arw
s-phys-norad08	NAMa+nmmb-n1_pert	nmmb-n1	Thompson	NOAH	MYJ	arw
s-phys-norad09	NAMa+nmmb-p2_pert	nmmb-p2	Thompson	NOAH	MYJ	arw
s-phys-norad10	NAMa+nmmb-n2_pert	nmmb-n2	Thompson	NOAH	MYJ	arw

Table 7. Specifications for the "nssl_nmmb" members of the CLUE that use the NMMB dynamic core, single physics, IC and LBC perturbations extracted from SREF members, and no radar data assimilation.

(7) NSSL: NMMB						
Ens Members	**IC**	**BC**	**Microphysics**	**LSM**	**PBL**	**Model**
nssl-nmmb01	NAMa	NAMf	Ferrier-Aligo	NOAH	MYJ	nmmb
nssl-nmmb02	NAMa+arw-p1_pert	arw-p1	Ferrier-Aligo	NOAH	MYJ	nmmb
nssl-nmmb03	NAMa+arw-n1_pert	arw-n1	Ferrier-Aligo	NOAH	MYJ	nmmb
nssl-nmmb04	NAMa+arw-p2_pert	arw-p2	Ferrier-Aligo	NOAH	MYJ	nmmb
nssl-nmmb05	NAMa+arw-n2_pert	arw-n2	Ferrier-Aligo	NOAH	MYJ	nmmb

Table 8. Specification for the "HRRR36" member of CLUE that uses the ARW dynamic core with HRRR physics and data assimilation.

(8) GSD: HRRR-36h						
Ens Members	**IC**	**BC**	**Microphysics**	**LSM**	**PBL**	**Model**
HRRR36	RAP	GFSf	Thompson	RUC	MYNN	arw

Table 9. Specifications for the "ncar-enkf" members of the CLUE. This ensemble provides forecasts to 48 h and uses NCAR's DART (Data Assimilation Research Testbed) software with ARW version 3.6.1 and slightly different horizontal domain than other CLUE members. The analysis system is comprised of 50 members that are continuously cycled using the ensemble adjustment Kalman filter (EAKF). New analyses are produced every 6 h with 15-km grid-spacing. Other specifications include: 51 vertical levels with a 50 hPa top (same vertical levels as other ARW CLUE members), a horizontal localization of 1270 km and vertical localization of 2 scale heights, adaptive prior inflation, adaptive localization, sampling error correction, and freely-evolving soil states. The following observational sources are utilized: MADIS ACARS, METARs, radiosondes, NCEP MARINE, CIMMS cloud-track winds, Oklahoma Mesonet, and GPS radio occultation. All members have constant physics, which include Tiedtke cumulus parameterization, Thompson microphysics, MYJ PBL, NOAH land-surface model, and RRTMG shortwave and longwave radiation with aerosol and ozone climatologies. The 10-member forecasts are initialized daily at 0000 UTC with ICs provided by downscaled members of 0000 UTC WRF/DART EAKF analyses (described above). Perturbed LBCs from GFS forecasts are used. The physics are the same as from the data assimilation system, but without cumulus parameterization (detailed below).

(9) NCAR: EnKF						
Ens Members	**IC**	**BC**	**Microphysics**	**LSM**	**PBL**	**Model**
ncar-enkf01	anal01	GEFS01	Thompson	NOAH	MYJ	arw
ncar-enkf02	anal02	GEFS02	Thompson	NOAH	MYJ	arw
ncar-enkf03	anal03	GEFS03	Thompson	NOAH	MYJ	arw
ncar-enkf04	anal04	GEFS04	Thompson	NOAH	MYJ	arw
ncar-enkf05	anal05	GEFS05	Thompson	NOAH	MYJ	arw
ncar-enkf06	anal06	GEFS06	Thompson	NOAH	MYJ	arw
ncar-enkf07	anal07	GEFS07	Thompson	NOAH	MYJ	arw
ncar-enkf08	anal08	GEFS08	Thompson	NOAH	MYJ	arw
ncar-enkf09	anal09	GEFS09	Thompson	NOAH	MYJ	arw
ncar-enkf10	anal10	GEFS10	Thompson	NOAH	MYJ	arw

Table 10. Specifications for the "mp" members of CLUE that use ARPS 3DVAR radar data assimilation and cloud analysis and varied microphysics.

(10) UND: Microphysics						
Ens Members	**IC**	**BC**	**Microphysics**	**LSM**	**PBL**	**Model**
core01	NAMa+3DVAR	NAMf	Thompson	NOAH	MYJ	arw
mp2	NAMa+3DVAR	NAMf	Morrison	NOAH	MYJ	arw
mp3	NAMa+3DVAR	NAMf	MY	NOAH	MYJ	arw
mp4	NAMa+3DVAR	NAMf	P3	NOAH	MYJ	arw
mp5	NAMa+3DVAR	NAMf	WSM6	NOAH	MYJ	arw

The configuration of the CLUE will allow for eight unique experiments that have been designed to examine issues immediately relevant to the design of a NCEP/EMC operational CAM-based ensemble. These experiments are listed below:

1) ARW vs. NMMB: A direct comparison of subjective and objective skill of ARW and NMMB dynamic cores will be conducted. Computational cost will also be tracked, which is extremely important for operational applications. This experiment involves comparing the ***caps-nmmb (Table 5) and nssl-nmmb (Table 7) members to the s-phys-norad (Table 6) members***. Specifically, comparisons will be made between the 10-member NMMB and ARW ensembles with the same ICs/LBCs.

2) Multi-core vs. single core ensemble design: Three ensembles will be compared to test the effectiveness of a single core vs. multi-core configuration. The first ensemble uses ***5 ARW and 5 NMMB members [sphys-norad06-10 (Table 6) and nssl-nmmb (Table 7)], the second uses 10 ARW members [sphys-norad (Table 6)], and the third uses 10 NMMB members [caps-nmmb (Table 5) and nssl-nmmb (Table 7)].***

3) Single Physics vs. Multi-Physics: An ensemble with perturbed initial and lateral boundary conditions will test whether there is a noticeable advantage when using multiple physical PBL and microphysics parameterizations vs. common physics in all members. This involves comparing the mixed physics ***core members (Table 1) to the single physics s-phys-rad members (Table 2).***

4) Ensemble Radar vs. Ensemble No Radar: A single physics ensemble will test the influence of assimilating radar data. An important question is whether the radar data influence extends to longer forecast lengths relative to deterministic forecasts. This involves comparing the ***single physics, no-radar data assimilation s-phys-norad members (Table 6) to the single physics s-phys-rad members (Table 2) that use ARPS 3DVAR radar data assimilation and cloud analysis.***

5) 3DVAR vs. EnKF: These two methods for data assimilation will be compared, which involves comparing the ***core members (Table 1; 3DVAR), to the caps-enkf members (Table 3; EnKF), to the ncar-enkf members (Table 9; EnKF).***

6) GSD Radar vs. CAPS Radar Assimilation: This experiment will test two methods for assimilating radar data, and involves comparing ***member core02 (Table 1) to HRRR36 (Table 8).*** These members are identically configured except they use different methods to assimilate radar data.

7) Microphysics Sensitivities: The impact from different microphysical parameterizations on the resulting convective storm forecasts will be examined, which involves examining the forecasts from the ***mp members (Table 10).***

8) Ensemble Size Experiment: A comparison of the mixed-core ensembles with equal contributions of NMMB and ARW members using 6, 10, and 20 members will be conducted to examine the impact of ensemble size. ***The 6-member ensemble uses caps-nmmb03-05 (Table 5) and s-phys-norad01-03 (Table 6). The 10-member uses caps-nmmb (Table 5) and s-phys-norad06-10 (Table 6). The 20-members uses caps-nmmb (Table 5), s-phys-norad (Table 6), and nssl-nmmb (Table 7).***

To ensure consistent post-processing, visualization, and verification for subsets of CLUE ensemble members contributed by different collaborators, all groups will utilize the same post-processing software to output the same set of model output fields on the same grid. Thus, NSSL worked closely with scientists at the Developmental Testbed Center (DTC) and EMC to modify the most recent version of Unified Post-Processor (UPP) software to output a set of 107 output fields from each CLUE member (Table 11). These fields (output in grib2 format) are the same as the 2D fields output by the operational HRRR, and were chosen because of their relevance to a broad range of forecasting needs, including aviation, severe weather, and precipitation. Additional output fields, which were requested by NCEP's Weather Prediction Center, Storm Prediction Center, and Aviation Weather Center, were also included. All CLUE collaborating groups are providing this set of 107 fields, but they can also add additional diagnostics based on their own research interests. Furthermore, CAPS and NSSL will output six additional experimental diagnostics (Table 12).

Table 11. The set of 107 required output diagnostics for the CLUE members, which are output at hourly intervals.

Number	Level/Layer	Parameter	Description
001	entire atmosphere	REFC	Composite reflectivity [dB]
002	cloud top	RETOP	Echo Top [m]
003	entire atmosphere	VIL	Vertically Integrated Liquid [kg/m^2]
004	surface	VIS	Visibility [m]
005	1000 m above ground	REFD	Reflectivity [dB]
006	4000 m above ground	REFD	Reflectivity [dB]
007	surface	GUST	Wind Speed (Gust) [m/s]
008	500 mb	HGT	Geopotential Height [gpm]
009	500 mb	TMP	Temperature [K]
010	500 mb	DPT	Dew Point Temperature [K]
011	500 mb	UGRD	U-Component of Wind [m/s]
012	500 mb	VGRD	V-Component of Wind [m/s]
013	700 mb	HGT	Geopotential Height [gpm]
014	700 mb	TMP	Temperature [K]
015	700 mb	DPT	Dew Point Temperature [K]
016	700 mb	UGRD	U-Component of Wind [m/s]
017	700 mb	VGRD	V-Component of Wind [m/s]
018	850 mb	HGT	Geopotential Height [gpm]
019	850 mb	TMP	Temperature [K]
020	850 mb	DPT	Dew Point Temperature [K]
021	850 mb	UGRD	U-Component of Wind [m/s]
022	850 mb	VGRD	V-Component of Wind [m/s]
023	925 mb	TMP	Temperature [K]
024	925 mb	DPT	Dew Point Temperature [K]
025	925 mb	UGRD	U-Component of Wind [m/s]
026	925 mb	VGRD	V-Component of Wind [m/s]

027	1000 mb	TMP	Temperature [K]
028	1000 mb	DPT	Dew Point Temperature [K]
029	1000 mb	UGRD	U-Component of Wind [m/s]
030	1000 mb	VGRD	V-Component of Wind [m/s]
031	400-1000 mb above ground	MAXUVV	Hourly Max upward Vertical Velocity - lowest 400hPa [m/s]
032	400-1000 mb above ground	MAXDVV	Hrly Max downward Vertical Velocity - lowest 400hPa [m/s]
033	0.5-0.8 sigma layer	DZDT	Vertical Velocity (Geometric) [m/s]
034	mean sea level	PRMSL	Pressure Reduced to MSL [Pa]
035	1000 mb	HGT	Geopotential Height [gpm]
036	1000 m above ground	MAXREF	Hourly Max of Simulated Reflectivity at 1 km AGL [dB]
037	5000-2000 m above ground	MXUPHL	Hrly Max Updraft Helicity - 2km to 5 km AGL [m^2/s^2]
038	entire column	TCOLG	Total Column Integrate Graupel [kg/m^2]
039	surface	LTNG	Lightning [non-dim]
040	80 m above ground	UGRD	U-Component of Wind [m/s]
041	80 m above ground	VGRD	V-Component of Wind [m/s]
042	surface	PRES	Pressure [Pa]
043	surface	HGT	Geopotential Height [gpm]
044	surface	TMP	Temperature [K]
045	0 m underground	MSTAV	Moisture Availability [%]
046	surface	WEASD	Water Equivalent of Accumulated Snow Depth [kg/m^2]
047	surface	SNOWC	Snow Cover [%]
048	surface	SNOD	Snow Depth [m]
049	2 m above ground	TMP	Temperature [K]
050	2 m above ground	SPFH	Specific Humidity [kg/kg]
051	2 m above ground	DPT	Dew Point Temperature [K]
052	10 m above ground	UGRD	U-Component of Wind [m/s]
053	10 m above ground	VGRD	V-Component of Wind [m/s]
054	10 m above ground	WIND	Wind Speed [m/s]
055	surface	CPOFP	Percent frozen precipitation [%]
056	surface	PRATE	Precipitation Rate [kg/m^2/s]
057	surface	APCP	Total Precipitation [kg/m^2]
058	surface	WEASD	Water Equivalent of Accumulated Snow Depth [kg/m^2]
059	surface	APCP	Precipitation [kg/m^2] – hourly total
060	surface	WEASD	Water Equivalent of Accumulated Snow Depth [kg/m^2]
061	surface	CSNOW	Categorical Snow [-]
062	surface	CICEP	Categorical Ice Pellets [-]
063	surface	CFRZR	Categorical Freezing Rain [-]
064	surface	CRAIN	Categorical Rain [-]

065	surface	VGTYP	Vegetation Type [Integer(0- 13)]
066	500-1000 mb	LFTX	Surface Lifted Index [K]
067	surface	CAPE	Convective Available Potential Energy [J/kg]
068	surface	CIN	Convective Inhibition [J/kg]
069	entire column	PWAT	Precipitable Water [kg/m^2]
070	low cloud layer	LCDC	Low Cloud Cover [%]
071	middle cloud layer	MCDC	Medium Cloud Cover [%]
072	high cloud layer	HCDC	High Cloud Cover [%]
073	entire atmosphere	TCDC	Total Cloud Cover [%]
074	cloud base	PRES	Pressure [Pa]
075	cloud base	HGT	Geopotential Height [gpm]
076	cloud ceiling	HGT	Geopotential Height [gpm]
077	cloud top	PRES	Pressure [Pa]
078	cloud top	HGT	Geopotential Height [gpm]
079	top of atmosphere	ULWRF	Upward Long-Wave Rad. Flux [W/m^2]
080	surface	DSWRF	Downward Short-Wave Radiation Flux [W/m^2]
081	3000-0 m above ground	HLCY	Storm Relative Helicity [m^2/s^2]
082	1000-0 m above ground	HLCY	Storm Relative Helicity [m^2/s^2]
083	0-6000 m above ground	USTM	U-Component Storm Motion [m/s]
084	0-6000 m above ground	VSTM	V-Component Storm Motion [m/s]
085	0-1000 m above ground	VUCSH	Vertical U-Component Shear [1/s]
086	0-1000 m above ground	VVCSH	Vertical V-Component Shear [1/s]
087	0-6000 m above ground	VUCSH	Vertical U-Component Shear [1/s]
088	0-6000 m above ground	VVCSH	Vertical V-Component Shear [1/s]
089	180-0 mb above ground	4LFTX	Best (4 layer) Lifted Index [K]
090	180-0 mb above ground	CAPE	Convective Available Potential Energy [J/kg]
091	180-0 mb above ground	CIN	Convective Inhibition [J/kg]
092	surface	HPBL	Planetary Boundary Layer Height [m]
093	lifted condensation level	HGT	Geopotential Height [gpm]
094	90-0 mb above ground	CAPE	Convective Available Potential Energy [J/kg]
095	90-0 mb above ground	CIN	Convective Inhibition [J/kg]
096	255-0 mb above ground	CAPE	Convective Available Potential Energy [J/kg]
097	255-0 mb above ground	CIN	Convective Inhibition [J/kg]
098	equilibrium level	HGT	Geopotential Height [gpm]
099	255-0 mb above ground	PLPL	Pressure of level from which parcel was lifted [Pa]
100	surface	LAND	Land Cover (0=sea, 1=land) [Proportion]
101	surface	ICEC	Ice Cover [Proportion]
102	250 mb	UGRD	U-component of wind [m/s]

103	250 mb	VGRD	V-component of wind [m/s]
104	250 mb	HGT	Geopotential Height [gpm]
105	250 mb	TMP	Temperature [K]
106	700 mb	VVEL	Vertical Velocity [m/s]
107	-10 C	REFD	Reflectivity [dB]

Table 12 The set of six experimental output diagnostics for the CLUE members contributed by CAPS and NSSL, which are output at hourly intervals.

Number	Level/Layer	Parameter	Description
001	surface	HAIL1	Maximum hail size from HAILCAST [mm]
002	surface	HAIL2	Maximum hail size from Thompson method [mm]
003	3000-0 m above ground	MXUPHL	Hrly Max Updraft Helicity - 0km to 3 km AGL [m^2/s^2]
004	5000-2000 m above ground	MNUPHL	Hrly Min Updraft Helicity – 2km to 5 km AGL [m^2/s^2]
005	3000-0 m above ground	MNUPHL	Hrly Min Updraft Helicity – 0km to 3 km AGL [m^2/s^2]
006	1 km AGL	VVEL	Vertical Velocity [m/s]

b) The NSSL-WRF Ensemble

SPC forecasters have used output from an experimental 4-km WRF-ARW run produced by NSSL (NSSL-WRF) since the fall of 2006. Currently, the NSSL-WRF is run twice daily at 0000 UTC and 1200 UTC throughout the year over a full CONUS domain with forecasts to 36 hours. Output is also available online at the website: http://www.nssl.noaa.gov/wrf/newsite.

In addition to the deterministic NSSL-WRF, nine 4-km WRF-ARW runs have also been run since 2014 and together these runs comprise the 10-member NSSL-WRF ensemble. The additional nine members are initialized at 0000 UTC and use 3-h SREF forecasts initialized at 2100 UTC for initial conditions and corresponding SREF member forecasts as lateral boundary conditions. There is also a member that used the 0000 UTC GFS analysis for ICs and corresponding GFS forecasts for LBCs. The physics parameterizations for each member are identical to the regular NSSL-WRF. Although the unvaried physics will typically have lower spread than a varied physics ensemble, SPC forecasters are familiar with the behavior of the NSSL-WRF physics, and this type of configuration allows us to isolate the contribution of spread from different ICs/LBCs. The ensemble configuration is provided in Table 13. The NSSL-WRF ensemble has been tested during SFE2014 and SFE2015, with favorable results.

Table 13. Configuration of the NSSL-WRF ensemble

Members	ICs/LBCs	Microphysics	LSM	PBL	Model
1	00Z NAM	WSM6	NOAH	MYJ	arw
2	00Z GFS	WSM6	NOAH	MYJ	arw
3	21Z arw_ctl	WSM6	NOAH	MYJ	arw
4	21Z nmmb_ctl	WSM6	NOAH	MYJ	arw
5	21Z arw_p1	WSM6	NOAH	MYJ	arw
6	21Z arw_n1	WSM6	NOAH	MYJ	arw
7	21Z nmmb_p1	WSM6	NOAH	MYJ	arw
8	21Z nmmb_n1	WSM6	NOAH	MYJ	arw
9	21Z arw_p2	WSM6	NOAH	MYJ	arw
10	21Z nmmb_p2	WSM6	NOAH	MYJ	arw

c) SPC Storm Scale Ensemble of Opportunity (SSEO)

The SSEO is a 7-member, convection-allowing ensemble consisting of deterministic CAMs available routinely to SPC, with output available to forecasters since 2011 This "poor man's ensemble" provides a practical alternative to a formal/operational storm-scale ensemble Similar to the other convection-allowing ensemble systems, hourly maximum storm-attribute fields, such as simulated reflectivity, updraft helicity, and 10-m wind speed are produced from the SSEO. Member specifications are provided in Table 14. Members marked with "-12h" in the Model column are 12h time-lagged members, initialized 12h earlier than the other members. All members are initialized with a "cold-start" from the indicated modeling system – i.e., no radar data assimilation or cloud model is used to produce ICs. Forecasts to 36h are available at 0000 and 1200 UTC. Calibrated severe weather probability guidance created from a combination of SSEO-based explicit thunderstorm attributes and SREF-based environmental fields will again be available.

Table 14. SSEO Specifications.

Members	**ICs/LBCs**	**Microphysics**	**Grid-spacing**	**PBL**
NSSL-WRF	NAM/NAM	WSM6	4 km	MYJ
EMC HRW ARW	RAP/GFS	WSM6	4.2 km	YSU
EMC HRW ARW; -12h	RAP/GFS	WSM6	4.2 km	YSU
EMC HRW NMMB	RAP/GFS	Ferrier updated	3.6 km	MYJ
EMC HRW NMMB; -12h	RAP/GFS	Ferrier updated	3.6 km	MYJ
EMC CONUS WRF-NMM	NAM/NAM	Ferrier	4 km	MYJ
EMC CONUS NAM NEST	NAM/NAM	Ferrier-Aligo	4 km	MYJ

d) UK-Met Office convection allowing models (credit: Humphrey Lean)

The Met Office Unified Model (UM) is the name given to the suite of numerical modelling software used by the Met Office. Three fully (or quasi) operational, nested limited-area high-resolution versions of the UM (two at 2.2 km and one at 1.1 km horizontal resolution) running once per day will be supplied to SFE2016. These operational nested hi-res versions will incorporate the latest UM settings that are used over the UK.

The 2.2-km version has 70 vertical levels (spaced between 5m and 40 km) across a slightly sub-CONUS domain. Taking its initial and lateral boundary conditions from the 00z 17-km horizontal grid-spacing global configuration of the UM, the 2.2-km model initializes without data assimilation and runs out to T+48. This model configuration uses a 3D turbulent mixing scheme using a locally scale-dependent blending of Smagorinsky and boundary layer mixing schemes, stochastic perturbations are made to the low-level resolved-scale temperature field in conditionally unstable regimes (to encourage the transition from subgrid to resolved scale flows) and the microphysics is single moment. Partial cloudiness is diagnosed assuming a triangular moisture distribution with a width that is a universally specified function of height only. There is no convection parametrization in this or any of the high resolution UM configurations.

The 1.1-km horizontal resolution version of the UM is nested within the 2.2-km model and runs over a 1600 km by 1800 km domain centered on eastern Oklahoma. The 1.1-km model takes its initial and lateral boundary conditions from the T+3 step of the 00z 2.2-km run, thus reducing spin-up time within the 1.1-km model, and runs out to T+33 (this may be extended to T+45 at request). As with the 2.2-km model, the 1.1-km model initializes without data assimilation and uses the same 70 vertical

level spacing as the 2.2-km. The 1.1-km model has identical planetary boundary layer and microphysics schemes as the 2.2-km model.

Finally, a parallel version of the 2.2-km model is being run with a new scheme, which addresses the moisture conservation issues in the model. This is because it appears that the excessive rain-rates sometimes observed in the cores of convective clouds are, at least in part, due to a problem with moisture conservation in the semi-lagrangian dynamics in the case of poorly resolved convective updrafts. An a posteriori correction scheme is used that enforces a moisture budget, i.e. Change in total domain moisture mass = net inward lateral boundary flux. The lateral boundary flux is estimated using the Pseudo-Lateral boundary Flux (PLF) scheme developed by Aranami et al (2015). This accounts for inflow and outflow of moisture through the lateral boundaries of the LAM by advecting lateral boundary and interior terms separately.

e) ESRL High Resolution Rapid Refresh (HRRR) model

The 3-km grid-spacing HRRR model developed by the NOAA/Earth Systems Research Laboratory (ESRL) will continue to be examined in SFE2016. Both the NCEP parallel HRRR (HRRRv2) and the ESRL developmental HRRR (HRRRv3) will be examined. The parallel HRRR is expected to replace the NCEP operational HRRR (HRRRv1) by late June. Thus, the current operational HRRRv1 will not be examined formally during SFE2016. The parallel HRRR (HRRRv2) has been tested during 2014-15 and is nested within the 13-km grid-spacing parallel RAPv3 (Benjamin et al. 2016), which is also expected to become operational this summer. HRRRv2 uses ARW version 3.6.1 and will provide 18 h forecasts over a full CONUS domain. HRRRv3 uses ARW version 3.7.1 and is nested within the ESRL developmental RAPv4, which has several improvements relative to RAPv3 (see Table 15). The HRRRv3 also has several improvements to HRRRv2 including improved depiction of upper-level clouds in the Thompson microphysics, improved convection via changes made to the Grell-Freitas shallow cumulus scheme, an update to the MYNN PBL scheme that includes better sub-grid clouds and mesoscale environment, among others (Table 15). Also, HRRRv3 includes forecasts that are extended to 24 h.

Table 15. Summary of changes to RAPv4 and HRRRv3 (provided by Curtis Alexander).

ESRL RAPv4/HRRRv3: 2016 Development and Testing

	Model	Data Assimilation
RAPv4 (13 km)	WRF-ARW v3.7.1+ incl. physics changes **Physics changes:** Thompson microphysics – improved upper-level clouds MYNN PBL update – better sub-grid clouds, meso env Thomp. aerosols + MYNN cloud-fraction – improved C&V RUC LSM update – more surface info, better lower bdy **Numerics changes:** Improved vertical coordinate – reduced noise, better turb	Merge with GSI trunk **New Observations:** Assimilation of new VAD winds – better winds Assimilation of Himawari satellite radiances Assimilation of GOES-R GLM **DA Methods:** More ensemble weight in hybrid DA – better winds Variational/hybrid cloud analysis – better ceiling/vis
HRRRv3 (3 km)	WRF-ARW v3.7.1+ incl. physics changes **Physics changes:** Thompson microphysics – improved upper-level clouds Grell-Freitas shallow cu – improved convection MYNN PBL update – better sub-grid clouds, meso env Thomp. aerosols + MYNN cloud-fraction – improved C&V RUC LSM update – more surface info, better lower bdy **Numerics changes:** Improved vertical coordinate – reduced noise, better turb	Merge with GSI trunk **New Observations:** Assimilate new VAD winds – better winds Assimilate GOES cloud-top cooling – convection Assimilate radar radial velocity – better convection **DA Methods:** More ensemble weight in hybrid DA – better winds Full atmospheric cycling – better 0-4 hr convection Variational/hybrid cloud analysis – better ceiling/vis

SPC-GSD • RAPv4/HRRRv3 06 Apr 2016 • 3

f) North American Mesoscale Rapid Refresh (NAMRR) System (Credit: Jacob Carley)

The NCEP experimental North American Mesoscale Rapid Refresh system (NAMRR) is an hourly-updated version of the North American Mesoscale (NAM) forecast system and its data assimilation system. The current version of the NAM, upgraded in August 2014, provides 84 h forecasts for its large 12 km domain and 60 h forecasts for its four high-resolution nested domains 4 times per day at 0000, 0600, 1200, and 1800 UTC. The NAM Data Assimilation system currently performs an atmospheric analysis every 3 hours and on its 12 km North American domain - there is no cycling of its nested domains. The NAMRR, in contrast, features an hourly forecast and assimilation cycle, including assimilation of radar data using the ESRL Diabatic Digital Filter Initialization (DDFI) technique. Furthermore, the NAMRR also features a data assimilation cycle for its 3 km CONUS and Alaska nested domains in addition to its North American 12 km domain. All forecasts from the NAMRR are a minimum of 18 hours in length, and to maintain compatibility with the current operational NAM, 60 h nested and 84 h large domain forecasts are issued at 0000, 0600, 1200, and 1800 UTC. The NAMRR will be examined during the SFE2016 forecast process and evaluations will also be conducted comparing the NAMRR to the HRRR.

g) NCAR's Model for Prediction Across Scales (MPAS; credit: Bill Skamarock)

For the second year, forecasts from NCAR's Model for Prediction Across Scales (Skamarock et al. 2012) will be examined. MPAS is one of two finalists for the Next Generation Global Prediction System (NGGPS), which will be the foundation for US global weather prediction for the next several decades (http://www.nws.noaa.gov/ost/nggps/). MPAS will produce daily 0000 UTC initialized forecasts at 3-km grid-spacing over the CONUS with forecasts to 120 h (5 days). The MPAS horizontal mesh is based on Spherical Centriodal Voronoi Tesselations (SCVTs). These meshes allow for both quasi-uniform discretization of the sphere and local refinement with smoothly varying mesh spacing between regions with differing resolutions. Importantly, the smoothly-varying mesh eliminates the major problems encountered with mesh transitions in forecast systems using traditional grid-nesting. The C-grid discretization, where the normal component of velocity on cell edges is prognosed, is especially well-suited for higher-resolution mesoscale and convective-scale atmosphere simulations where horizontally divergent motions (e.g. convection) are the least-well resolved. Idealized convective tests, in addition to real-data hindcasts tests on 3-km global meshes and performance during SFE2015, have shown that the MPAS produces convective storm realizations similar to that of the ARW model.

During SFE2015 MPAS transitioned from a 50km mesh over the globe to a 3-km mesh over the US. During SFE2016, there will be a transition from 15 km to3km, instead. Other changes that have been made to MPAS since the SFE2015 include: (1) the PBL scheme has been changed from YSU to the MYNN scheme that is running in HRRRv3. (2) the microphysics has been changed from WSM6 to Thompson. (3) the Grell-Freitas scale aware convection parameterization has been updated, which is in response to a high bias for tropical precipitation. Figure 1 shows the MPAS mesh that will be used for SFE2016. MPAS forecasts will be viewed using a webpage maintained by NCAR (http://wrf-model.org/plots/realtime_mpas.php or http://www2.mmm.ucar.edu/imagearchive/mpas/images.php).

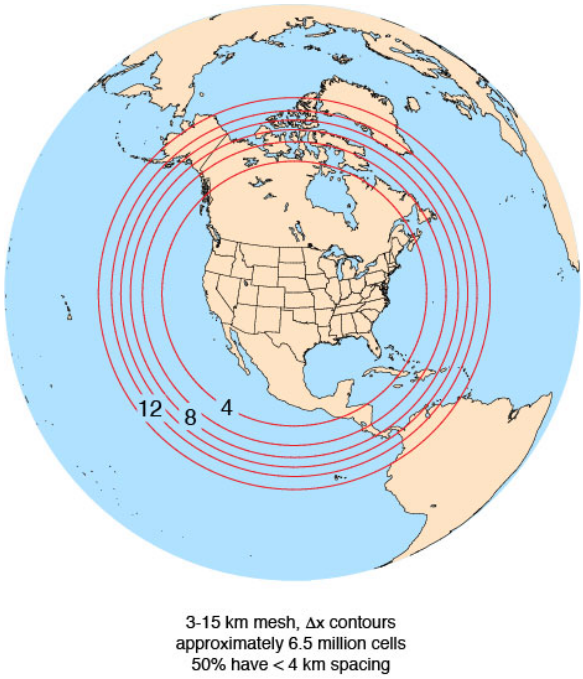

Fig. 1. MPAS grid being used for SFE2016.

h) NSSL Experimental Warn-on-Forecast System for ensembles (NEWS-e; Credit: Dustan Wheatley)

Although not a formal part of SFE2016, an experimental ensemble that is a prototype of NSSL's Warn-on-Forecast System will be run during the afternoons. SFE2016 participants may have an opportunity to examine and evaluate these forecasts near the end of (or after) formal SFE daily activities. This ensemble is described as follows: The NSSL Experimental Warn-on-Forecast System for ensembles (NEWS-e) is a 36-member WRF-based ensemble data assimilation system used to produce very short-range (0-3 h) probabilistic forecasts of supercell thunderstorm rotation, high winds, and flash flooding. The starting point for each day's experiment will be an hourly cycled experimental High-Resolution Rapid Refresh Ensemble (HRRRE) under development at the Global Systems Division (GSD). The 20-member HRRRE is comprised of 15-km (fixed) and 3-km WRF ensembles nested within the GFS (or RAP) ensemble, which are initialized at 2100 UTC the previous day with HRRR physics. The 3-km domain is chosen nominally on a weekly basis (by weather pattern), and encompasses the central plains, south-central, southeast, or northeast region of the United States. The full ensemble is updated by hourly EnKF assimilation of conventional observations from 2200 UTC Day 0 to 1800 UTC Day 1. An 18-member, 12-hr forecast of the HRRRE 3-km domain at 1500 UTC will provide initial and boundary conditions for the NEWS-e.

The daily NEWS-e domain location will target the primary region where severe weather is anticipated and cover a 1000-km wide region with very frequent 30-min updates. All ensemble members utilize the Thompson microphysics parameterization and the RAP land-surface model, but the PBL and radiation physics options are varied amongst the ensemble members to address uncertainties in model physics. Multi-Radar/Multi-Sensor (MRMS) radar reflectivity and Level II radial velocity data, cloud water path retrievals from the GOES imager, and Oklahoma Mesonet observations (when available) will be assimilated every 15 min using an EnKF approach. A 3 h ensemble forecast will be initialized from the resultant storm-scale analyses every 30 minutes during each hour of the real-

time experiment. These forecasts will be viewable using the web-based Probabilistic Hazard Information (PHI)-tool developed by NSSL.

3. SFE2016 Core Interests/Daily Activities

a. Forecast products and activities

Similar to previous years, the experimental forecasts this year will continue to explore our ability to add temporal specificity to longer-term convective outlooks. We will continue to split the participants into two desks: one forecasting total severe and the other forecasting individual hazards. For the hazards desk, the first forecast will mimic the SPC operational Day 1 Convective Outlooks by producing separate probability forecasts of large hail, damaging wind, and tornadoes within 25 miles (40 km) of a point valid 1600 UTC to 1200 UTC the next day. This is the third year this desk has issued outlooks for individual hazards. The first forecast for the total severe desk will also cover the 1600 to 1200 UTC time period, but cover only combined severe (tornado, hail and wind). A text product will accompany each experimental Day 1 outlook that describes the meteorology of the day and the usefulness of the suite of model guidance during the creation of the severe storm forecasts. A separate Day 1 forecast will be made at each desk. The experimental forecasts cover a limited-area domain typically covering the primary severe threat area with a center-point selected based on existing SPC outlooks and/or where interesting convective forecast challenges are expected.

Each desk will then manually stratify the experimental Day 1 outlooks into periods with higher temporal resolution. The individual hazards desk will generate separate probability forecasts of large hail, damaging wind, and/or tornadoes in two 4-h periods: 1800-2200 UTC and 2200-0200 UTC. As an alternative way of stratifying the Day 1 outlook, the other desk will generate five 4-h period outlooks of total severe at two-hour intervals covering the time periods: 1800-2200 UTC, 2000-0000 UTC, 2200-0200 UTC, 0000-0400 UTC, and 0200-0600 UTC. Additionally, the total severe desk will draw isochrones of severe weather at two-hour intervals on top of the full period Day 1 total severe probabilities to delineate the start-time of the 4-h time window with the highest total severe probabilities. For example, the area between the 2000 UTC and 0000 UTC isochrones would expect total severe probabilities to be highest during the 2000-0000 UTC period. We are exploring the feasibility of drawing isochrones to add more detailed timing information to outlooks as an alternative (or supplement) to issuing more frequent outlooks valid for shorter time periods. The goals of testing these different methods is to explore multiple ways of introducing probabilistic severe weather forecasts on time/space scales that are currently addressed with mostly categorical forecast products (SPC Mesoscale Discussions and Tornado/Severe Thunderstorm Watches), and to begin to explore ways of seamlessly merging probabilistic severe weather outlooks with probabilistic severe weather warnings as part of the NOAA WoF and FACETS initiatives.

During previous experiments, calibrated probabilistic severe guidance from the SREF/SSEO (Jirak et al. 2014) was used to temporally disaggregate a 1600-1200 UTC period human forecast. A scaling factor was formulated by matching the full-period calibrated severe SSEO guidance to the human forecast, then this scaling factor (unique at every grid point) was applied to the SSEO calibrated severe guidance for each individual period, and finally consistency checks were conducted to arrive at the final temporally disaggregated forecasts (Jirak et al. 2012). These automated forecasts from SFE2012 - SFE2015 fared favorably both in terms of objective metrics (e.g., CSI, FSS) and subjective impressions when compared to manually drawn forecasts. Similarly for SFE2016, the 1600-1200 UTC

human forecasts for the individual hazards will be temporally disaggregated into the 4-h periods (1800-2200 UTC and 2200-0200 UTC) to provide a first guess for the two forecast periods.

The first set of short-time-window forecasts and timing forecasts will be issued in the morning by both desks. At the individual hazards desk, the lead forecaster will generate the short-time-window forecasts on the N-AWIPS machines. However, the participants will split into five groups and use a web-based tool to generate their own short-time window probability forecasts using Google Chromebooks and a web-based tool known as the Probabilistic Hazard Information (PHI) tool. The PHI tool will have first guess probability fields generated from CAM forecasts, as well as other important observational and model fields for participants to utilize in the forecast generation process. At the total severe desk, both the lead forecaster and participants will issue the full period and short-time-window forecasts as a group. However, the isochrones will be drawn by five small groups using Google Chromebooks. The lead forecaster will also draw independent isochrones of severe weather. After the teams issue the short-time-window forecasts, there will be a map discussion summarizing forecast challenges and highlighting interesting findings from the previous day. .

After lunch, both desks will transition to examining the Day 2 period. Teams will examine operational guidance as well as experimental CAM guidance that will extend into Day 2 and generate probability forecasts similar to Day 1, but for the next convective day (1200-1200 UTC). This is the third year of testing Day 2 outlooks for individual hazards. The total severe desk will also produce a Day 3 forecast if time allows. A text product will accompany each Day 2 or Day 3 outlook that describes the meteorology of the day and the usefulness of the suite of model guidance during the creation of the severe storm forecasts.

Later in the afternoon, scientific evaluations will take place (summarized in the next section). Finally at the end of the day, each team will examine observational trends and morning/afternoon model guidance to update the short-time-window forecasts made earlier in the day. Because the forecast process for these updates will begin at 3:15 pm, only the forecasts valid from 2200-0200 UTC will be updated for the individual hazards. The total severe desk will update their 4 h time window forecasts, starting with the 2200-0200 UTC period. Finally, the total severe desk will update their total severe isochrones, contouring the 2200, 0000, and 0200 UTC times. All these updated forecasts will be issued by 4 pm. Table 16, which is included below, presents a summary of the SFE2016 daily activities schedule.

Table 16. Summary of Daily Schedule

Severe Hazards Desk	*Total Severe Desk*
0800 – 0845: **Evaluation of Experimental Forecasts & Guidance** Subjective rating relative to radar evolution/characteristics, warnings, and preliminary reports and objective verification using preliminary reports and MESH	
• Day 1 & 2 full-period probabilistic forecasts of tornado, wind, and hail • Day 1 4-h period forecasts and guidance for tornado, wind, and hail	• Days 1, 2, & 3 full-period probabilistic forecast of total severe • Day 1 4-h period probability forecasts and isochrones
0845 – 1115: **Day 1 Convective Outlook Generation** Hand analysis of 12Z upper-air maps and surface charts	
• Day 1 full-period probabilistic forecasts of tornado, wind, and hail valid 16-12Z over mesoscale area of interest • Day 1 4-h probabilistic forecasts of tornado, wind, and hail valid 18-22 and 22-02Z*	• Day 1 full-period probabilistic forecast of total severe valid 16-12Z over mesoscale area • Day 1 4-h probabilistic total severe forecasts valid 18-22, 20-00, 22-02, 00-04, and 02-06Z. • Day 1 isochrones for 4-h periods (every 2 h) with highest probability of total severe*
1115 – 1130: **Break** Prepare for map discussion	
1130 – 1200: **Map Discussion** Brief discussion of today's forecast challenges and products Highlight findings from previous days	
1200 – 1300: **Lunch** Brief EWP participants at 1245 if needed	
1300 – 1345: **Day 2 Convective Outlook Generation**	
• Day 2 full-period probabilistic forecasts of tornado, wind, and hail valid 12-12Z over mesoscale area of interest	• Day 2 or Day 3 full-period probabilistic forecasts of total severe valid 12-12Z over mesoscale area of interest
1345 – 1515: **Scientific Evaluations**	
• CLUE (4): Radar data assimilation • CLUE (8): Ensemble size comparisons • CLUE: SSEO as baseline • NAMRR Nest, HRRRv2, HRRRv3 • Ensemble sensitivity (TTU)	• CLUE (1,2): Model core (det. & ens.) • CLUE (6): Radar data assimilation approaches • CLUE (7): Microphysics sensitivity • Explicit hail size forecast comparison
1515 – 1600: **Short-term Outlook Update**	
• Update 4-h probabilistic forecasts of tornado, wind, and hail valid 22-02Z*	• Update 4-h probabilistic forecasts of total severe for 22-02, 00-04, and 02-06Z. • Update total severe isochrones (22, 00, & 02Z)*

* Denotes forecasts also made by participants using the PHI tool on Chromebooks.

b. Formal Evaluation Activities

There will be two periods of formal evaluations during SFE2016. The first is during the morning when experimental outlooks from the previous day generated by both forecast teams will be examined. In these next-day evaluations, the team forecasts and the first-guess guidance will be compared to observed radar reflectivity, reports of severe weather, NWS warnings, and Multi-Radar Multi-Sensor (MRMS) radar-estimated hail sizes over the same time periods. The SFE participants will provide their subjective evaluations of the strengths and weaknesses of each of the forecasts. This evaluation will include examining and comparing calibrated guidance, temporal disaggregation first guess, and human initial and update forecasts. The goal is to determine the relative skill of the first-guess guidance and the human-generated forecasts over all periods, in part to assess the feasibility of issuing operational high-temporal resolution severe weather forecasts.

Objective verification metrics will also be computed for some of the experimental outlooks and first-guess guidance. As with the 2014 and 2015 SFE, experimental probabilistic forecasts of tornado, wind, and hail will be evaluated using Critical Success Index (CSI) and Fractions Skill Score (FSS) based on the local storm reports (LSRs) as the verification event. Similar to last year, supplemental observations for hail from multi-hourly MRMS-based Maximum Estimated Size of Hail (MESH) will be used in near real-time to calculate skill scores and gauge the usefulness of alternative sources for verification. A quality control measure will be applied to the hourly MESH grids, which ensures the existence of nearby CG lightning flashes. Further, only spatially filtered grids are considered to ensure the presence of contiguous swaths in the high-resolution MESH tracks (Melick et al. 2014). Similar to those created with LSRs, practically perfect hindcasts (Brooks et al. 1998) will be created from the MESH to provide valuable baselines to measure the skill of the probabilistic severe hail forecasts.

The afternoon evaluation will involve comparisons of different ensemble diagnostics and CLUE ensemble subsets. The total severe and individual hazards desks will conduct two different sets of evaluations.

i. Evaluations at the total severe desk

(1) Comparison of CAPS 3DVAR- and EnKF-based ensembles and deterministic forecasts

An evaluation activity will focus on the first 12 h of core02 (Table 1), caps-enkf01 (Table 3), and HRRR36 (Table 8). These are members that use different methods for assimilating radar data. The activity will focus on a regional area of interest and evaluate how well these members depict storms in the initial conditions and their subsequent evolution during the first 12 h of the forecast.

(2) Evaluation of model guidance for hail

Similar to SFE2014 and SFE2015, there is interest in evaluating the ability of CAMs to predict hail size because of the need to forecast individual thunderstorm hazards, which are included in the SPC operational Day 1 Convective Outlooks. Thus, for the third year, the HAILCAST algorithm implemented in ARW will be used to predict hail size (Adams-Selin 2013; Adams-Selin and Ziegler 2016), which is based on the algorithm in Brimelow (2002) and Jewell and Brimelow (2009). The HAILCAST model uses convective cloud and updraft attributes to determine the growth of hail from initial embryos. The cloud attributes for the model are those predicted explicitly in the ARW forecasts

and the snow, ice and graupel mixing ratios at the first level above the freezing level at which they exist are used to determine the initial embryo size.

During SFE2014, it was very apparent that HAILCAST over-predicted hail sizes. As a result, after SFE2014 changes were made to HAILCAST that resulted in more realistic hail size forecasts. Specifically, rime soaking and variable density options were added, and the dependency on microphysics scheme was removed by using 5 constant initial embryo sizes as opposed to those predicted in the schemes themselves. The changes were implemented in the NSSL-WRF and NSSL-WRF ensemble on 9 July 2014. Further changes were made in early 2016, which include variable hail density for both wet and dry growth regimes, an updraft multiplier simulating advection of the hail embryo across and updraft, temperature dependent ice collection efficiency, mass growth by vapor deposition or condensation, an improved liquid water shedding threshold, and enhanced melting during collisions with above-0 degree C liquid water. These changes are detailed in Adams-Selin and Ziegler (2016). The most up-to-date version of the HAILCAST algorithm will be available in all the runs conducted by NSSL and CAPS.

Additionally, a hail size diagnostic derived directly from the microphysics parameterizations will be examined, which was implemented by Greg Thompson of NCAR. This diagnostic will also be available from the simulations conducted by CAPS and NSSL.

Finally, probabilistic hail size forecasts derived from a machine learning algorithm developed by David Gagne of OU will be examined – these forecasts will be derived from the core members of the CLUE, and are described as follows: a hail size forecasting model generates hail size forecasts from a combination of image processing and machine learning models. Potential hailstorms are identified and tracked in the modeled hourly-max column-integrated graupel field and the MRMS MESH product using the enhanced watershed object identification method. Model and observed tracks are matched using a multidimensional distance function. The distribution of MESH values within an object is approximated by a gamma distribution. The gamma distribution parameters are simultaneously predicted by a single random forest machine learning model given information about each model storm and its environment. Another random forest model predicts whether or not an observed hailstorm will occur in the vicinity of the forecast hailstorm, which determines whether or not a particular hail object is kept. Finally, hail sizes are sampled from each predicted gamma distribution and are applied in rank order to each grid point within each hailstorm object. This process is repeated for each ensemble member, and ensemble neighborhood probabilities and other derived products can be produced with this information.

As part of the evaluation activity, the utility of probabilistic hail size forecasts using all three methods will be compared. The predictions of hail size will be evaluated against storm reports and MRMS MESH.

(3) Comparing the skill of NMMB versus ARW

This evaluation activity will involve comparing deterministic forecasts from core01 (Table 1) and caps-nmmb-rad (Table 4), which use the ARW and NMMB dynamic cores, respectively. We will have the ability to compare multiple fields over a regional domain of interest, but the focus will be on the reflectivity forecasts and whether aspects of forecast storms like mode, timing, and evolution are simulated correctly.

(4) Comparing the skill of an ARW-based ensemble vs. an NMMB-based ensemble vs. a multi-core ensemble

This evaluation activity will involve comparing probabilistic forecasts of simulated reflectivity derived from a 10-member, ARW-based ensemble, a 10 member NMMB-based ensemble, and 10-member ensemble with equal numbers of ARW and NMMB members.

(5) Microphysics comparisons

Since 2010, one component of model evaluation activities during annual SFEs has involved subjectively examining sensitivity to microphysics parameterizations used in the WRF model by comparing various forecast fields including simulated reflectivity, simulated brightness temperature, low-level temperature and moisture, and instability for a set of ensemble members with identical configurations except for their microphysical parameterization. In previous years, CAPS have conducted these simulations, but during SFSE2016 the University of North Dakota is providing them. During SFE2016, the following microphysics parameterizations will be examined: WSM6, Thompson, Morrison, Milbrandt and Yau (MY), and the Predicted Particle Properties (P3) scheme.

ii. Evaluations at the hazards desk

(1) Radar Data Assimilation

Two 10-member WRF-ARW ensembles with single physics will be compared where the only difference is the assimilation of radar data. The members of one ensemble will have radar data assimilated in the 0000 UTC NAM analysis using the CAPS 3DVar method, followed by applying SREF perturbations for IC/LBC diversity. The other ensemble will not include radar data assimilation and apply the same SREF perturbations to the 0000 UTC NAM analysis. This experiment should highlight the length of time that radar data assimilation has a noticeable impact on convection-allowing ensemble forecasts. The objective component of this evaluation will focus on ensemble neighborhood probability forecasts of simulated reflectivity compared to observed radar reflectivity while the subjective component will examine ensemble forecasts (ensemble maximum and neighborhood probabilities) of hourly maximum fields (HMFs) of UH, updraft speed, and 10-m wind speed relative to LSRs of hail, wind, and tornadoes.

(2) Ensemble size comparisons

Three multi-core ensembles (i.e., equal membership between WRF-ARW and NMMB) will be compared where the only difference is the number of ensemble members. A six-member ensemble subset will be compared to a ten-member ensemble subset and the full twenty-member multi-core ensemble with single physics (per core) and no radar data assimilation. This experiment should highlight the impact of increasing membership in improving the skill of forecasts from a convection-allowing ensemble, which is important information to consider given limits to NOAA operational high performance computing resources. The objective component of this evaluation will focus on ensemble neighborhood probability forecasts of simulated reflectivity compared to observed radar reflectivity while the subjective component will examine ensemble forecasts (ensemble maximum and neighborhood probabilities) of hourly maximum fields (HMFs) of UH, updraft speed, and 10-m wind speed relative to LSRs of hail, wind, and tornadoes.

(3) SSEO as Baseline

Various ensemble subsets from the CLUE, including a multi-core ensemble and two EnKF-based WRF-ARW ensembles, will be compared to the SSEO. Given the utility and success of the SSEO in forecasting hazardous weather since 2011, it can be used as a baseline to assess operational CAM ensemble performance, which should be able to generate forecasts at least as skillful as this "poor man's" ensemble. This experiment should highlight the current status of ensemble development and inform the design of the initial configuration of an operational convection-allowing ensemble. The objective component of this evaluation will focus on ensemble neighborhood probability forecasts of simulated reflectivity compared to observed radar reflectivity while the subjective component will examine ensemble forecasts (ensemble maximum and neighborhood probabilities) of hourly maximum fields (HMFs) of UH, updraft speed, and 10-m wind speed relative to LSRs of hail, wind, and tornadoes. Objective metrics will include FSS, which is computed by comparing simulated reflectivity forecasts against mosaic, hybrid-scan reflectivity from the MRMS system. As in prior years, the evaluation will be facilitated via webpages using spatial plots for distinct time frames as well as a table, which summarizes statistical results. For some of the summary type analyses, time-series plots will also be constructed for a more graphical representation of trends in the scores. This approach has proved effective since introduced during SFE2012 (Melick et al. 2012) for the participants to quickly evaluate the forecast verification metrics and provide feedback on the comparison between the objective results and their own subjective impressions.

(4) NAMRR Nest and HRRR Comparison

The 1500 UTC 3-km hourly NCEP experimental NAMRR Nest will be compared to the next operational version of the HRRR (HRRRv2) and the ESRL/GSD developmental version of the HRRR (HRRRv3). This evaluation activity will focus on a regional area of interest and evaluate how well these deterministic runs depict storms in the initial conditions and their subsequent evolution during the 15-h forecast.

(5) TTU Ensemble Sensitivity Analysis Products

Texas Tech University (TTU) is running a 42-member Data Assimilation Research Testbed (DART) nested WRF-model ensemble Kalman filter (EnKF) over a fixed CONUS 12-km domain and a moveable 4-km nest that encompasses much of the Midwest and Eastern U.S. Data assimilation is performed on the 12-km grid on a 6-hr cycle, and 48-hr forecasts are produced twice daily from the 0000 and 1200 UTC initialization. Ensemble sensitivity of 13 response functions involving simulated reflectivity, updraft helicity, and surface wind speed will be produced within the real-time TTU ensemble forecasting system. Sensitivity of the response functions valid at forecast hours 12-48 will be calculated in real time at 3-hourly increments prior to the response function time with respect to 8 model state variables, such at 500-hPa geopotential height and sea level pressure. The location of the response function will be chosen daily prior to the following 0000 and 1200 UTC initializations of the ensemble to produce sensitivity fields that can be interrogated the following day. The response function location for the 0000 UTC initialization will be chosen to diagnose Day 1 events, while response locations for the 1200 UTC run will be chosen to reflect Day 2 events.

Ensemble sensitivity analysis is a process that can highlight the atmospheric flow features that are relevant to the predictability of severe weather events. Specifically, ensemble sensitivity reveals how a chosen forecast metric (e.g. maximum simulated reflectivity, or number of grid points with updraft helicity exceeding a threshold value over a specified area) responds to differences among

ensemble members at or before the time of the chosen forecast metric. By being aware of the features critical to the predictability of severe events, it may be possible to choose ensemble members, or even different ensemble systems, which are more likely to be skillful. This evaluation activity will be an initial assessment into the potential of ensemble sensitivity analysis to improve forecasts of severe weather.

Ensemble system website - http://www.atmo.ttu.edu/bancell/real_time_ENS/ttuenshome.php
Sensitivity product website - http://www.atmo.ttu.edu/bancell/real_time_ENS/hwt/ttu-hwt.php

c. Other specialized activities

For the third year in the HWT Spring Forecasting Experiments, CAM output in three-dimensional (3D) displays will be presented in real-time as part of the daily activities. CAPS will provide selected 3D model fields over the daily mesoscale region of interest at 10-minute output frequency for 18 – 30 h forecasts to interrogate using the VIS5D analysis and display system. The goal is to explore CAM storm characteristics like vertical vorticity, graupel mixing ratio, simulated reflectivity, and cold pools in 3D to learn more about how simulated storms are structured on ARW convection-allowing grids (see Figure 3 for an example of what this display looked like in SFE2014). The model storm structures will be compared to the structures of storms observed by the WSR-88D network and displayed within VIS5D. We will also examine characteristics of the storm environments in CAM forecasts like depth of water vapor mixing ratio in the PBL and depictions of low-level convergence boundaries and how they may play a role in the initiation of convection in the model. Since this is still an initial exploration into the detailed structures of storms and model forecast environments in real time as part of a forecast process, this activity will be somewhat informal and less structured than most SFE2016 activities.

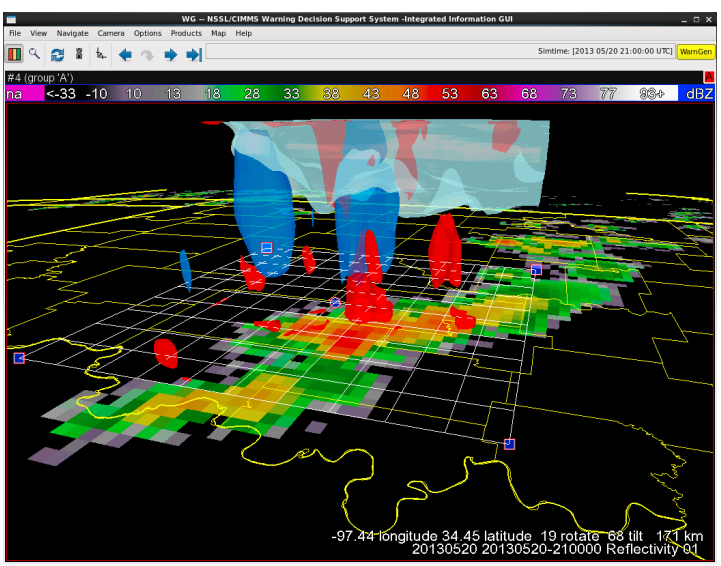

Figure 3. Example of how CAM forecasts will be interrogated for a select few runs from the CAPS SSEF system. The 2D field is the simulated reflectivity on the lowest model level (color scale near the top of the image) with 3D isosurfaces of vertical velocity x vertical vorticity ($w*\zeta$) > 2 m s^{-2} (red areas), graupel mixing ratio > 5 g kg^{-1} (dark blue areas), and snow mixing ratio > 2 g kg^{-1} shown within the box outlined in white.

Appendix A: List of scheduled SFE2016 participants. Facilitators/leaders for SFE2016 include: Adam Clark (NSSL), Kent Knopfmeier (CIMMS/NSSL), Israel Jirak (SPC), Dave Imy (retired SPC), James Correia Jr. (CIMMS/SPC), Chris Melick (CIMMS/SPC), Andy Dean (SPC), and Steve Willington (UKMO).

Week 1	Week 2	Week 3	Week 4	Week 5
May 2-6	**May 9-13**	**May 16-20**	**May 23-27**	**May 31-June 3**
Brian Ancell (TTU)	Brian Ancell (TTU)	Brock Burghardt (TTU)	Brock Burghardt (TTU)	Brian Ancell (TTU)
Brock Burghardt (TTU)	Brock Burghardt (TTU)	Jack Kain (NSSL)	Andy Taylor (WFO FGZ)	Aaron Kennedy (UND)
Mike Evans (WFO BGM)	Bill Gallus (Iowa State)	Tom Workhoff (FirstEnergy)	Pete Wolf (WFO JAX)	Brooke Hagenhoff (UND)
Madalina Surcel (McGill Univ.)	Brian Squitieri (Iowa State)	Nathan Hitchens (Ball State)	Nathan Wendt (SPC)	Joshua Markel (UND)
Greg Gallina (WPC)	Sean Stelten (Iowa State)	Lance Bosart (SUNYA)	Mark Rodwell (ECMWF)	Jingyu Wang (UND)
Ben Albright (WPC/HMT)	Jim Nelson (WPC)	Kyle Pallozzi (SUNYA)	Vince Agard (MIT)	John Stoppkotte (WFO LBF)
Mike McClure (WFO DVN)	Marc Chenard (WPC)	Bruno Ribeiro (SUNYA/Brazil)	Aaron Johnson (WFO DDC)	Binbin Zhou (EMC)
Pat Spoden (WFO PAH)	Stan Czyzyk (WFO VEF)	Bill Martin (WFO GSP)	Ivan Tsonevsky (ECMWF)	Jeff Beck (GSD)
Tom Holtquist (WFO MPX)	Robert Hepper (SPC)	Glen Romine (NCAR)	Matt Pyle (EMC)	Isidora Jankov (GSD)
Stephen Bieda (WFO AMA)	Jacob Carley (EMC)	Corey Guastini (EMC)	John Brown (GSD)	Jeff Milne (SPC)
Ryan Sobash (NCAR)	Trevor Alcott (GSD)	Curtis Alexander (GSD)	Ed Szoke (GSD)	Greg Thompson (NCAR) W-F
Geoff Manikin (EMC)	David Dowell (GSD) Th,F	David Dowell (GSD) M,T	Steve Willington (UK Met)	Hugh Morrison (NCAR) W-F
Terra Ladwig (GSD)	Becky Adams-Selin (USAF)	Anke Finnenkoetter (UK Met)	Mark Bevan (UK Met)	Jason Milbrandt (EC) W-F
Eric James (GSD)	Humphrey Lean (UK met)	Humphrey Lean (UK Met) M	Glenn White (EMC)	Steve Willington (UK Met)
Makenzie Krocak (OU)	Anke Finnenkoetter (UK Met) F	Tracey Dorian (EMC)	David John Gagne (OU)	Mark Conder (WFO LUB)
	Bill Skamarock (NCAR) W-F	Pamela Eck (SUNYA)	Nick Nauslar (SPC)	Ron Miller (WFO OTX)
	Corey Potvin (NSSL)	Paula Davidson (NWS) T,W	Bruce Entwistle (AWC)	Pam Heinselman (NSSL)

Appendix B: Experimental Severe Thunderstorm Forecasts

Severe weather graphics for the full-period Day 1 (1600-1200 UTC) and Day 2 (1200-1200 UTC) individual hazard probabilities will be in the same format as that used for the operational SPC day 1 outlooks (categorical and general thunderstorm outlooks will not be made). For reference, the Probability-to-Categorical conversion for individual hazards used for the SPC Day 1 Outlook is shown below. These same probabilities will be used for generating the individual hazard forecasts in the four-hour periods.

Day 1 Probability to Categorical Outlook conversions

Day 1 Outlook Probability	TORN	WIND	HAIL
2%	MRGL	Not Used	Not Used
5%	SLGT	MRGL	MRGL
10%	ENH	Not Used	Not Used
10% with Significant Severe	ENH	Not Used	Not Used
15%	ENH	SLGT	SLGT
15% with Significant Severe	MDT	SLGT	SLGT
30%	MDT	ENH	ENH
30% with Significant Severe	HIGH	ENH	ENH
45%	HIGH	ENH	ENH
45% with Significant Severe	HIGH	MDT	MDT
60%	HIGH	MDT	MDT
60% with Significant Severe	HIGH	HIGH	MDT

Total severe weather probabilities for the full period Day 1 (1600-1200 UTC) and Day 2 (1200-1200 UTC) total severe storm hazards will be in the same format as that used for the operational SPC Day 2 outlooks (5, 15, 30, 45, and 60 %). An area delineating potential for significant severe storms will be included when the probability for significant severe is 10% or greater. For reference, the Probability-to-Categorical conversion for total severe used for the SPC Day 2 Outlook, and is shown below. For the hourly probabilities of total severe, the severe weather probability within 25 miles of a point in any given hour is expected to be low, so the contours of 2% and 10% can be added to the probability contours that can be drawn.

Day 2 Probability to Categorical Outlook conversions

Day 2 Outlook Probability	Combined TOR, WIND, HAIL
5%	MRGL
15%	SLGT
15% with Significant Severe	SLGT
30%	ENH
30% with Significant Severe	ENH
45%	ENH
45% with Significant Severe	MDT
60%	MDT
60% with Significant Severe	HIGH

Appendix C. Organizational structure of the NOAA/Hazardous Weather Testbed

NOAA's Hazardous Weather Testbed (HWT) is a facility jointly managed by the National Severe Storms Laboratory (NSSL), the Storm Prediction Center (SPC), and the NWS Oklahoma City/Norman Weather Forecast Office (OUN) within the National Weather Center building on the University of Oklahoma South Research Campus. The HWT is designed to accelerate the transition of promising new meteorological insights and technologies into advances in forecasting and warning for hazardous mesoscale weather events throughout the United States. The HWT facilities are situated between the operations rooms of the SPC and OUN. The proximity to operational facilities, and access to data and workstations replicating those used operationally within the SPC, creates a unique environment supporting collaboration between researchers and operational forecasters on topics of mutual interest.

The HWT organizational structure is composed of three overlapping programs (Fig. 4). The Experimental Forecast Program (EFP) is focused on predicting hazardous mesoscale weather events on time scales ranging from hours to a week in advance, and on spatial domains ranging from several counties to the CONUS. The EFP embodies the collaborative experiments and activities previously undertaken by the annual SPC/NSSL Spring Experiments. For more information see http://www.nssl.noaa.gov/projects/hwt/efp/.

The Experimental Warning Program (EWP) is concerned with detecting and predicting mesoscale and smaller weather hazards on time scales of minutes to a few hours, and on spatial domains from several counties to fractions of counties. The EWP embodies the collaborative warning-scale experiments and technology activities previously undertaken by the OUN and NSSL. For more information about the EWP see http://www.nssl.noaa.gov/projects/hwt/ewp/. A key NWS strategic goal is to extend warning lead times through the "Warn-on-Forecast" concept (Stensrud et al. 2009), which involves using frequently updated short-range forecasts (≤ 1h lead time) from convection-resolving ensembles. This provides a natural overlap between the EFP and EWP activities.

The GOES-R Proving Ground (established in 2009) exists to provide pre-operational demonstration of new and innovative products as well as the capabilities available on the next generation GOES-R satellite. The overall goal of the Proving Ground is to provide day-1 readiness once GOES-R launches in late 2015. The PG interacts closely with both product developers and NWS forecasters. More information about GOES-R Proving Ground is found at http://cimss.ssec.wisc.edu/goes_r/proving-ground.html.

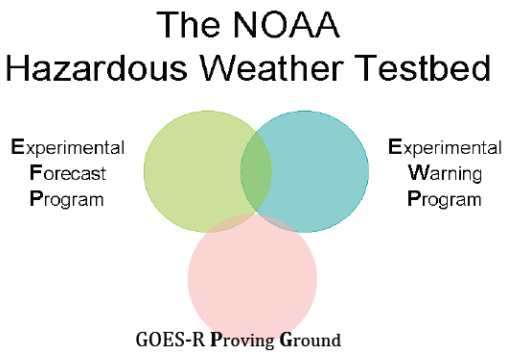

Figure 4: The umbrella of the NOAA Hazardous Weather Testbed (HWT) encompasses two program areas: The Experimental Forecast Program (EFP), the Experimental Warning Program (EWP), and the GOES-R Proving Ground (GOES-R).

Rapid science and technology infusion for the advancement of operational forecasting requires direct, focused interactions between research scientists, numerical model developers, information technology specialists, and operational forecasters. The HWT provides a unique setting to facilitate such interactions and allows participants to better understand the scientific, technical, and operational challenges associated with the prediction and detection of hazardous weather events. The HWT allows participating organizations to:

- Refine and optimize emerging operational forecast and warning tools for rapid integration into operations
- Educate forecasters on the scientifically correct use of newly emerging tools and to familiarize them with the latest research related to forecasting and warning operations
- Educate research scientists on the operational needs and constraints that must be met by any new tools (e.g., robustness, timeliness, accuracy, and universality)
- Motivate other collaborative and individual research projects that are directly relevant to forecast and warning improvement

For more information about the HWT, see http://www.nssl.noaa.gov/hwt/. Detailed historical background about the EFP Spring Experiments, including scientific and operational motivation for the intensive examination of high resolution NWP model applications for convective weather forecasting, and the unique collaborative interactions that occur within the HWT between the research and operational communities, are found in Weiss et al. (2010 – see http://www.spc.noaa.gov/publications/weiss/hwt-2010.pdf) and Clark et al (2012b).

Appendix D: References

Adams-Selin, R., 2012: Using the Advanced Research Weather Research and Forecasting (WRF-ARW) model to explicitly forecast hail. Preprints, *25th Conf. on Weather Analysis and Forecasting/21st. Conf. on Numerical Weather Prediction*, Montreal, QC, Canada, Amer. Met. Soc., 1D202.

Adams-Selin, R. 2013: In-line 1D WRF hail diagnostic. AFWA *Internal Tech. Memo*, SEMSD.21495.

Adams-Selin, R., C. L. Ziegler, 2016: Forecasting hail using a one-dimensional hail growth model within WRF. *Mon. Wea. Rev*., (in review).

Aranami, K., T. Davies, and N. Wood, 2015: A mass restoration scheme for limited-area models with semi-lagrangian advection. *QJRMS*, **141**, 1795-1803.

Benjamin, S. G., and Coauthors, 2016: A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. *Mon. Wea. Rev*., **144**, 1669-1694.

Brimelow, J.C., 1999: Modeling maximum hail size in Alberta thunderstorms. *Wea. Forecasting*, **17**, 1048-1062.

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

Clark, A. J., J. S. Kain, P. T. Marsh, J. Correia, Jr., M. Xue, and F. Kong, 2012a: Forecasting tornado path lengths using a 3-dimensional object identification algorithm applied to convection-allowing forecasts. *Wea. Forecasting*, **27**, 1090-1113.

Clark, A. J., and Coauthors, 2012b: An Overview of the 2010 Hazardous Weather Testbed Experimental Forecast Program Spring Experiment. *Bull. Amer. Meteor. Soc.*, **93**, 55–74.

Clark, A.J., J. Gao, P. T. Marsh, T. Smith, J. S. Kain, J. Correia, Jr., M. Xue, and F. Kong, 2013: Tornado path length forecasts from 2010-2011 using ensemble updraft helicity. *Wea. Forecasting*, **28**, 387-407.

Coniglio, M.C., J. Correia Jr., P. T. Marsh, F. Kong, 2013: Verification of convection-allowing WRF model forecasts of the planetary boundary layer using sounding observations. *Wea. Forecasting*, **28**, 842-862.

Duda, J.D., X. Wang, F. Kong, and M. Xue, 2014: Using varied microphysics to account for uncertainty in warm-season QPF in a convection-allowing ensemble. *Mon. Wea. Rev*., in press.

Hitchens, N. M., and H. E. Brooks, 2012: Evaluation of the Storm Prediction Center's Day 1 Convective Outlooks. *Wea. Forecasting*, **27**, 1580–1585.

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

Jewell, R., and J. Brimelow, 2009: Evaluation of Alberta Hail Growth Model using severe hail proximity soundings from the United States. *Wea. Forecasting*, **24**, 1592-1609.

Jirak, I. L., C. J. Melick, A. R. Dean, S. J. Weiss, and J. Correia, Jr., 2012: Investigation of an automated temporal disaggregation technique for convective outlooks during the 2012 Hazardous Weather Testbed Spring Forecasting Experiment. Preprints, *26th Conf. on Severe Local Storms*, Nashville, TN, Amer. Meteor. Soc., 10.2.

Jirak, I. L., C. J. Melick, and S. J. Weiss, 2014: Combining probabilistic ensemble information from the environment with simulated storm attributes to generate calibrated probabilities of severe weather hazards. Preprints, *27th Conf. on Severe Local Storms*, Madison, WI, Amer. Meteor. Soc., 2.5.

Kain, J.S., S.R. Dembek, S.J. Weiss, J.L. Case, J.J. Leit, 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.

Melick,C. J, I. L. Jirak, A. R. Dean, J. Correia Jr, and S. J. Weiss, 2012: Real Time Objective Verification of Convective Forecasts: 2012 HWT Spring Forecast Experiment. Preprints, *37th Natl. Wea. Assoc. Annual Meeting*, Madison, WI, Natl. Wea. Assoc*.,* P1.52.

Melick, C. J., I. L. Jirak, J. Correia Jr., A.R. Dean, and S.J. Weiss, 2014: Exploration of the NSSL Maximum Expected Size of Hail (MESH) Product for Verifying Experimental Hail Forecasts in the 2014 Spring Forecasting Experiment. *Preprints*, *27th Conf. Severe Local Storms*, Madison, WI.

Skamarock, W. C., Klemp, J. B., Duda, M., Fowler, L. D., Park, S.-H., & Ringler, T. (2012). A Multiscale Nonhydrostatic Atmospheric Model Using Centroidal Voronoi Tesselations and C-Grid Staggering. *Monthly Weather Review*, *140*(9), 3090–3105. doi:10.1175/MWR-D-11-00215.1

Sobash, R.A., J.S. Kain, D.R. Bright, A.R. Dean, M.C. Coniglio, and S.J. Weiss, 2010: Probabilistic forecast guidance for severe thunderstorms based on the identification of extreme phenomena in convection-allowing model forecasts. *Wea. Forecasting*, **26**, 714-728.

Stensrud, D. J., and Coauthors, 2009: Convective-Scale Warn-on-Forecast System. *Bull. Amer. Meteor. Soc.*, **90**, 1487–1499.