SPRING FORECASTING EXPERIMENT 2019

Conducted by the

EXPERIMENTAL FORECAST PROGRAM

of the

NOAA/HAZARDOUS WEATHER TESTBED

HWT Facility – National Weather Center

29 April - 31 May 2019

https://hwt.nssl.noaa.gov/sfe/2019/

Program Overview and Operations Plan

23 April, 2019

Adam Clark[2], Israel Jirak[1], Burkely T. Gallo[1,3], Andy Dean[1], Kent Knopfmeier[2,3], Brett Roberts[1,2,3], Louis Wicker[2], Makenzie Krocak[3,4], Patrick Skinner[2,3], Jessica Choate[2,3], Pam Heinselman[2], Katie Wilson[2,3], Jake Vancil[1,3], Kimberly Hoogewind[1,3], Nathan Dahl[1,3], Race Clark[1,3], Gerry Creager[2,3], Thomas Jones[2,3], Jidong Gao[1], Yunheng Wang[2,3], Scott Dembek[2,3], Steven J. Weiss[3]

(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 experimental convection-allowing (1 to 3 km grid-spacing) modeling systems (CAMs).

The 2019 Spring Forecasting Experiment (SFE 2019), a cornerstone of the EFP, will be conducted 29 April – 31 May with participation expected from about 100 forecasters, researchers, and model developers from around the world. Building upon successful experiments of previous years, a main emphasis of SFE 2019 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. Furthermore, for the fourth year, these contributions have been coordinated into a single ensemble framework called the Community Leveraged Unified Ensemble (CLUE; Clark et al. 2018). The 2019 CLUE is constructed by using a set of common model specifications (e.g., grid-spacing, vertical levels, domain size, etc.) so that the simulations contributed by each group can be used in carefully designed controlled experiments. This design will once again allow us to conduct several experiments geared toward identifying optimal configuration strategies for CAM-based ensembles. The 2019 CLUE includes 96 members using 3-km grid-spacing that will allow a set of several unique experiments. An additional feature of SFE 2019 will involve the continued testing of the Warn-on-Forecast prototype system (WoF Ensemble, hereafter), which will be used for the third year to issue very short lead-time outlooks, and learn how WoF Ensemble products are used and interpreted using surveys and real-time analytics. This activity will be expanded relative to previous experiments. Specifically, in addition to a 3-4 PM activity that is a regular part of the 8 AM-4 PM activities, an evening forecasting activity will be conducted from 4-8 PM that involves a small number of selected participants.

This operations plan summarizes the core interests of SFE 2019 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 models and products being introduced during SFE 2019 and Section 3 describes the core interests and new concepts being introduced for SFE 2019. 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 SFE 2019 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, named the Severe Hazards and

Innovation Desks, respectively, will issue different sets of convective outlooks for this testing. The Severe Hazards desk will issue Day 1 and 2 full-period outlooks (valid 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 conditional intensity forecasts within the Day 1 period for each hazard. The Innovation Desk will be issuing Day 1 and 2 full-period outlooks for total severe (i.e., outlook for combined hazards of severe hail, wind, or tornadoes), as well as Day 1 and Day 2 potential severe timing (PST) areas, which will indicate when the combined hazard probability will be $\geq$ 15% during the outlook day. These 4-h periods can occur at any time within the full-period outlook. These PSTs are designed to highlight areas and timing of severe weather occurrence to explore the feasibility of issuing a timing product to supplement current categorical forecast products (e.g., SPC Mesoscale Discussions and Severe Thunderstorm/Tornado Watches).

Finally, for the third year the Innovation desk will conduct a short-term forecasting activity using the WoF Ensemble, which will occur from 3-8 PM. During the 3-4 PM part of the activity, all participants at the Innovation Desk, as well as two forecasters selected specifically for the evening activity, will participate. From 4-8 PM, only the two evening forecasters (and possibly 1-2 other volunteers) will participate along with NSSL facilitators. Each hour, probabilistic total severe outlooks will be issued that are valid for short (1-h) and long (4-h) time windows. Additionally, each hour, outlooks will be issued for a “targeted” 1-h time window valid 0100-0200 UTC. During each hour, newly updated WoF Ensemble guidance will be available. The Severe Hazards desk will use these forecasts to update their hazard forecasts for the full period valid 2100–1200 UTC, but focused on refinements valid over the next few hours. These activities are the third year the WoF Ensemble has been tested in the EFP, and explores the potential utility of WoF products for issuing guidance between the watch and warning time scales (i.e. 0.5 to 6-h lead times). These activities represent efforts to explore 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. These efforts also support efforts to transition to higher temporal resolution forecasts at the SPC.

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 this CAM output is a part of the 96-member CLUE. The CAM output will include recent versions of the Advanced Research Weather Research and Forecasting (WRF-ARW) model, two ensembles based on the United Kingdom Met Office’s Unified Modeling System, and several configurations of the Finite-Volume Cubed-Sphere model (FV3). The FV3 was selected to replace the GFS as part of the Next Generation Global Prediction System (NGGPS) program. Furthermore, NOAA plans for the FV3 to be the foundation of a unified modeling suite encompassing all prediction time and space scales currently under the purview of NOAA’s Environmental Modeling Center (EMC). The FV3 runs examined during SFE 2019 will include global configurations with high resolution nests over the CONUS, and stand-alone-regional configurations (SARFV3) with domains over the CONUS and lateral boundary conditions provided by forecasts from another modeling system (e.g., NAM, GFS, etc).

In addition to the ensemble subsets contained within the 96-member CLUE system, several versions of the High-Resolution Ensemble Forecast system Version 2.1 (HREFv2.1) will be examined, which is a formalized implementation of the Storm Scale Ensemble of Opportunity (SSEO) with all members run at EMC using 3-km grid-spacing. The 10-member, HREFv2.1 includes the recent addition of two High Resolution Rapid Refresh (HRRR) members, one of which is 6-h time lagged. The HREF became operational in November of 2017. SFE 2019 will compare the HREFv2.1 to alternative configurations with different combinations of members including addition of a SARFV3 configuration provided by EMC and elimination of NMMB members. Results from this evaluation will be used to make recommendations for future versions of HREF.

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 have shown that fields like hourly-maximum updraft helicity (UH) and hourly-maximum wind speed near the surface can be effective for highlighting the likelihood of severe weather in CAMs (Kain et al. 2010, Sobash et al. 2011, Clark et al. 2013, 2018; Gallo et al. 2016, 2018, 2019; Sobash et al. 2016a and b, 2017). To support the goal of SFE 2019 to generate forecasts of individual hazards, there will be further efforts to explore the ability of new model fields and diagnostics to delineate individual hazards, particularly for the size of hail. An ensemble-subsetting method developed by researchers at Texas Tech University (TTU) will also be tested during SFE 2019, in which ensemble sensitivity to desired fields will be used to eliminate ensemble members that are performing poorly in sensitive areas. Probabilities from the full ensemble and the subset of ensemble members will then be compared to determine the effect of the subsetting method. In SFE 2018, the ensemble subsetting activity used a CAM ensemble run at TTU, however, for SFE 2019 most of the CLUE system will be used for subsetting.

Finally, new methods of real-time verification will continue to take place during SFE 2019. A particular focus will be an experimental, CAM scorecard being developed jointly with scientists at the National Center for Atmospheric Research (NCAR) Research Applications Laboratory (RAL), the Developmental Testbed Center (DTC), NSSL, and SPC. The purpose of this scorecard is to follow recommendations to unify verification systems between NOAA partner labs and the DTC where possible. The Scorecard is based on the enhanced Model Evaluation Tools (METplus) software package, and includes metrics specific to CAM ensemble, such as surrogate severe probabilities generated using UH. This verification will be applied to a subset of deterministic and ensemble forecasts, and will update in real-time throughout the experiment. The rest of this section provides further details on each modeling system utilized in SFE 2019.

a) The 2019 Community Leveraged Unified Ensemble (CLUE)

The CLUE is a carefully designed ensemble with subsets of members contributed by NSSL, the OU Center for Analysis and Prediction of Storms (CAPS) and Multi-scale data Assimilation and Predictability (MAP) groups, EMC, ESRL/GSD, NCAR, NOAA's Geophysical Fluid Dynamics Laboratory, and the United Kingdom Meteorology Office (UK Met). All members are initialized weekdays at 0000 UTC with 3-km grid-spacing covering a CONUS domain, except for the UK Met members which use 2.2 km grid-spacing and include 6-h time-lagged members. Depending on the CLUE subset, forecast lengths range from 36 to 120 h. Table 1 summarizes all contributions to the 2019 CLUE. Then, specifications for the members within each subset are detailed in the subsequent tables.

Table 1 Summary of the 19 unique subsets that comprise the 2019 CLUE.

Clue Subset	# of mems	IC/LBC perturbations	Mixed Physics	Data Assimilation	Model Core	Agency
fv3-ens	9	SREF	yes	cold start	FV3	CAPS (OU)
fv3-phys	7	none	yes	cold start	FV3	CAPS (OU)
wrf-exp	4	none	no	3DVAR	ARW	CAPS (OU)
caps-enkf	10	EnKF (CAPS)	yes	EnKF	ARW	CAPS (OU)
HRRRv3	1	none	no	GSI Ens-Var	ARW	ESRL/GSD
HRRRv4	1	none	no	GSI Ens-Var	ARW	ESRL/GSD
gsd-sarfv3	1	none	no	cold start	FV3	ESRL/GSD
hrrre	9	EnKF	no	EnKF	ARW	ESRL/GSD
ncar	10	EAKF (DART)	no	EAKF (DART)	ARW	NCAR
map-hybrid	10	EnKF-Var hybrid (GSI)	no	EnKF-Var hybrid (GSI)	ARW	MAP (OU)
map-ICpert	10	EnKF-Var hybrid (GSI) w/ GEFS	no	EnKF-Var hybrid (GSI) w/ GEFS	ARW	MAP (OU)
hrrre-nospp	9	EnKF	no	EnKF	ARW	NSSL
nssl-fv3	1	none	no	cold start	FV3	NSSL
nssl-sarfv3	1	none	no	cold start	FV3	NSSL
gfdl-fv3	1	none	no	cold start (GFS)	FV3	GFDL
ukmet-sphys	9	MOGREPS-G	no	cold start	UM	UK Met Office
ukmet-mphys	9	MOGREPS-G	yes	cold start	UM	UK Met Office
emc-fv3	1	none	no	cold start	FV3	EMC
emc-sarfv3	1	none	no	cold start	FV3	EMC

Table 2 Specifications of the fv3-ens CLUE members, which use mixed-physics and perturbed ICs/LBCs. These members use SARFV3 and are run with IC perturbations extracted from members of the 2100 UTC initializations of the Short-Range Ensemble Forecast System (SREF) run at EMC and added to the 0000 UTC, 12-km grid-spacing North American Mesoscale Model (NAM) analyses. The specific SREF members are indicated in the "ICs" column. Corresponding SREF member forecasts are used for LBCs. These runs are contributed by OU/CAPS using Stampede2 at the Texas Advanced Computing Center (TACC).

Members: fv3-ens	ICs	LBCs	Micro-physics	PBL	LSM	Radiation	Model
core-ctrl	NAMa	NAMf	Thompson	saMYNN	NOAH	RRTMG	FV3
pert-pbl1	NAMa+SREF arwn1	SREF arwn1	Thompson	saShinHong	NOAH	RRTMG	FV3
pert-pbl2	NAMa+SREF arwp2	SREF arwp2	Thompson	EDMF	NOAH	RRTMG	FV3
pert-mp1	NAMa+SREF arwp1	SREF arwp1	NSSL	saMYNN	NOAH	RRTMG	FV3
pert-mp2	NAMa+SREF arwn2	SREF arwn2	Morrison	saMYNN	NOAH	RRTMG	FV3
pert-lsm	NAMa+SREF arwp3	SREF arwp3	Thompson	saMYNN	RUC	RRTMG	FV3
pert-sfc1	NAMa+SREF arwn3	SREF arwn3	Thompson	saMYNN	RUC	RRTMG	FV3

Table 3 Specifications of the fv3-phys CLUE members. The first 7 members listed use SARFV3 with mixed-physics, and NAM analyses and forecasts for ICs and LBCs, respectively. The last 2 members listed use a global configuration of FV3 with a 3-km nest over the CONUS, and GFS analyses and forecasts for ICs and LBCs, respectively. These runs are contributed by OU/CAPS using Stampede2 at TACC. Note, core-globalgfs and core-sargfs will not be examined as part of SFE activities.

Members: fv3-phys	**ICs**	**LBCs**	**Micro-physics**	**PBL**	**LSM**	**SFC Layer**	**Radiation**	**Model**
core-ctrl	NAMa	NAMf	Thompson	saMYNN	NOAH	GFS	RRTMG	FV3
core-pbl1	NAMa	NAMf	Thompson	saShinHong	NOAH	GFS	RRTMG	FV3
core-pbl2	NAMa	NAMf	Thompson	EDMF	NOAH	GFS	RRTMG	FV3
core-mp1	NAMa	NAMf	NSSL	saMYNN	NOAH	GFS	RRTMG	FV3
core-mp2	NAMa	NAMf	Morrison	saMYNN	NOAH	GFS	RRTMG	FV3
core-lsm	NAMa	NAMf	Thompson	saMYNN	RUC	GFS	RRTMG	FV3
core-sfc1	NAMa	NAMf	Thompson	saMYNN	RUC	MYNN	RRTMG	FV3
core-globalgfs	GFS	n/a	Thompson	saMYNN	NOAH	GFS	RRTMG	FV3
core-sargfs	GFS	GFS	Thompson	saMYNN	NOAH	GFS	RRTMG	FV3

Table 4 Specifications of the wrf-exp CLUE members. The first 3 members listed use WRF-ARW Version 4.0.3. WSR-88D data, along with available surface and upper air observations, are analyzed using ARPS 3DVAR/Cloud-analysis system with 12-km grid-spacing NAM (members 1-3) and RAP (member 4) analyses as the background. The arw_m4 member uses WRF-ARW Version 3.9 as provided by ESRL/GSD. Additionally, the arw_m4 member using stochastic parameter perturbations in the Thompson microphysics scheme (namelist setting: spp_mp=7). These runs are contributed by OU/CAPS using Stampede2 at TACC. Note, these members will not be examined as part of SFE activities.

Members: wrf-exp	**ICs**	**LBCs**	**Micro-physics**	**PBL**	**LSM**	**Radiation**	**Model**
arw_cn	ARPSa	NAMf	Thompson	MYJ	NOAH	RRTMG	ARW
arw_m2	ARPSa (no radial wind)	NAMf	Thompson	MYJ	NOAH	RRTMG	ARW
arw_m3	ARPSa (qv adjustment)	NAMf	Thompson	MYJ	NOAH	RRTMG	ARW
arw_m4	RAPa+3DVAR	18Z GFSf	Thompson	MYNN	RUC	RRTMG	ARW

Table 5 Specifications for the caps-enkf CLUE members. This 3-km GSI-EnKF system is initialized at 1800 UTC each day, and assimilates the RAP/HRRR GSI data stream hourly (except satellite data) from 1800-0000 UTC and radar data every 15 minutes from 2300-0000 UTC over the CONUS domain. The ensemble consists of 40 WRF-ARW members (Version 4.0.3) with initial perturbations and mixed physics options to provide input for the EnKF ensemble analyses. Each member uses Thompson microphysics, although with varied graupel density among members. A 10-member ensemble forecast (run for 48 h) follows using the final EnKF analyses at 0000 UTC using the same multi-physics configurations. These runs are contributed by OU/CAPS using Bridges at the Pittsburgh Supercomputing Center.

Members: caps-enkf	**ICs**	**LBCs**	**Micro-physics**	**PBL**	**LSM**	**Radiation**	**Model**
caps-enkf01	enkf_m01a	NAMf	Thompson	MYJ	NOAH	RRTMG	ARW
caps-enkf02	enkf_m02a	arw-p1	NSSL	YSU	NOAH	RRTMG	ARW
caps-enkf03	enkf_m15a	arw-n1	NSSL	MYNN	NOAH	RRTMG	ARW
caps-enkf04	enkf_m40a	nmmb-p1	Morrison	MYJ	NOAH	RRTMG	ARW
caps-enkf05	enkf_m08a	nmmb-n1	P3	YSU	NOAH	RRTMG	ARW
caps-enkf06	enkf_m26a	arw-p2	NSSL	MYJ	NOAH	RRTMG	ARW
caps-enkf07	enkf_m39a	arw-n2	Morrison	YSU	NOAH	RRTMG	ARW
caps-enkf08	enkf_m12a	nmmb-p2	Thompson	MYNN	NOAH	RRTMG	ARW
caps-enkf09	enkf_34a	arw-n6	Thompson	YSU	NOAH	RRTMG	ARW
caps-enkf10	enkf_38a	arw-p6	NSSL	MYNN	NOAH	RRTMG	ARW

Table 6 Specification for the hrrrv3 CLUE member. HRRRv3 became operational in July 2018 and uses GSI hybrid data assimilation with the 80-member GDAS (GFS) as the background and uses WRF-ARW Version 3.9. The 0000 UTC initializations out to 36-h are considered a part of the CLUE, but 36-h forecasts are also initialized at 0600, 1200, and 1800 UTC, and at all other hours 18-h forecasts are initialized. The HRRRv3 is initialized with an hour of 3-D radar reflectivity using a latent-heating specification technique including some refinements in this latent-heating from the parent RAPv4 model. The HRRRv3 uses GSI hybrid GFS ensemble-variational data assimilation of conventional observations. Building upon the advancements in the operational HRRRv2 at NCEP, HRRRv3 includes assimilation of TAMDAR aircraft observations, refines assimilation of surface observations for improved lower-tropospheric temperature, dewpoint (humidity) winds and cloud base heights and places more weight on the ensemble contribution to the data assimilation. HRRRv3 also adds assimilation of lightning flash rates as a complement to radar reflectivity observations through a similar conversion to specified latent heating rates during a one-hour spin-up period in the model. Numerous model changes within the HRRRv3 include transition to a hybrid sigma-pressure vertical coordinate for improved tropospheric temperature, dewpoint and wind forecasts, along with a higher resolution (15 second) land use dataset. Physics enhancements have also been made to the MYNN PBL scheme and RUC land surface model along with additional refinements to shallow cumulus/sub-grid-scale cloud parameterizations including enhanced interactions with the radiation and microphysics schemes for greater retention of cloud features.

Member: HRRRv3	**ICs**	**LBCs**	**Microphysics**	**PBL**	**LSM**	**Radiation**	**Model**
hrrrv3	RAP	GFSf	Thompson	MYNN	RUC	RRTMG	ARW

Table 7 Specifications for the hrrrv4 CLUE member. The next and final update to the deterministic Rapid Refresh, version 5 (RAPv5), and HRRRv4, is currently scheduled for an operational implementation in early-mid 2020. The physics suite for HRRRv4 continues to use actively-developed versions of Thompson aerosol-aware microphysics, MYNN PBL scheme, RUC land surface model and RRTMG SW/LW radiation schemes. Enhancements have been made to the MYNN PBL scheme to further improve both representation of sub-grid-scale clouds and their effects on the local environment (reducing model bias of incoming radiation and temperature/moisture fields). Gravity-wave drag enhancements have been made to improve representation of the effects of sub-grid terrain on the horizontal flow. Land surface model and state changes include installation of an inland lake model for improved lake-temperature prediction, higher-resolution MODIS albedo and inland lake datasets, use of fractional sea-ice data and FVCOM dynamic specification of temperature and ice concentrations for the Great Lakes. Finally, VIIRS-based fire-radiative power detections are used to specify wildfire-driven injection of particulate matter for 3-D advection and deposition of smoke plumes. Enhancements to numerics in HRRRv4 include a reduction in magnitude of the 6th order filter for momentum, thermodynamic and hydrometeor fields to improve depiction of weaker small-scale cloud and precipitation features. A new implicit-explicit vertical advection scheme is also being tested for inclusion in HRRRv4 that permits larger vertical motion in intense convection to facilitate improved diagnosis of rotational features such as mesocyclones. For data assimilation, The HRRRv4 uses an updated version of GSI and includes assimilation of additional datasets including lightning data from GOES (GLM), aircraft and RAOB moisture observations above 300 mb, tropical cyclone central pressure estimates from TCvitals for improved position and structure of tropical systems, and potentially some additional radiance data from GOES-16. A storm-scale ensemble data assimilation system (HRRRDAS) is also being tested in the CONUS HRRRv4 that uses 36 hourly-cycled CONUS HRRR members with assimilation of conventional, radar and satellite observations through GSI-EnKF. This system is designed to improve use of observations during data assimilation with better representation of meso-to-storm scale covariances when compared with the comparatively coarse global ensemble (GDAS) used in HRRRv3. More accurate retention and evolution of meso-to-storm scale features, particularly in the early forecast hours, are intended benefits of HRRRDAS use. The HRRRDAS system, while intended to improve deterministic HRRRv4 forecasts, also forms the basis for HRRR ensemble forecasts described in the HRRRE section.

Member: HRRRv4	**ICs**	**LBCs**	**Microphysics**	**PBL**	**LSM**	**Radiation**	**Model**
hrrrv4	RAP	GFSf	Thompson	MYNN	RUC	RRTMG	ARW

Table 8 Specifications for the gsd-sarfv3 CLUE member. This member uses the same ICs/LBCs and physics as hrrrv4, except with the SARFV3 dynamical core. The hrrr physics are implemented through the Common Community Physics Package (CCPP) interface. No data assimilation is being applied within this forecast.

Member: HRRRv3	**ICs**	**LBCs**	**Microphysics**	**PBL**	**LSM**	**Radiation**	**Model**
gsd-sarfv3	RAP	GFSf	Thompson	MYNN	RUC	RRTMG	FV3

Table 9 Specifications for the hrrre CLUE members. The experimental HRRR ensemble (HRRRE) is initialized daily from a combination of GFS for atmospheric ensemble mean, GDAS for atmospheric perturbations, and RAP/HRRR for land surface. 36 HRRRE members (15-km grid-spacing) are cycled hourly for 24 hours with GSI-EnKF to assimilate conventional and radar-reflectivity observations. The hourly cycling also includes cloud-clearing and cloud-building procedures. Posterior inflation during the hourly cycling, random boundary-condition perturbations, and stochastic parameter perturbations (SPP) applied to the land-surface, PBL, and microphysics schemes contribute to ensemble spread. Forecasts initialized from the first 9 members of the HRRRE are advanced as follows: 36-h forecast at 0000 UTC, 24-h forecast at 1200 UTC, and 18-h forecast at 1800 UTC. The 36 HRRRE analyses and 9 HRRRE forecasts provide initial conditions and boundary conditions for the experimental Warn-on-Forecast system. The most significant changes in the 2019 HRRRE relative to previous versions are the expansion of the 3-km grid to the HRRR full-CONUS grid, the use of WRF-ARW version 3.9, the cloud clearing and cloud building, and the use of stochastic physics (SPP).

Members: hrrre	**ICs**	**LBCs**	**Microphysics**	**PBL**	**LSM**	**Radiation**	**Model**
hrrre01	enkf_m01b	GFS	Thompson	MYNN	RUC	RRTMG	ARW
hrrre02	enkf_m02b	GFS	Thompson	MYNN	RUC	RRTMG	ARW
hrrre03	enkf_m03b	GFS	Thompson	MYNN	RUC	RRTMG	ARW
hrrre04	enkf_m04b	GFS	Thompson	MYNN	RUC	RRTMG	ARW
hrrre05	enkf_m05b	GFS	Thompson	MYNN	RUC	RRTMG	ARW
hrrre06	enkf_m06b	GFS	Thompson	MYNN	RUC	RRTMG	ARW
hrrre07	enkf_m07b	GFS	Thompson	MYNN	RUC	RRTMG	ARW
hrrre08	enkf_m08b	GFS	Thompson	MYNN	RUC	RRTMG	ARW
hrrre09	enkf_m09b	GFS	Thompson	MYNN	RUC	RRTMG	ARW

Table 10 Specifications for the ncar CLUE members. This ensemble provides forecasts to 60 h at 0000 UTC and 36 h at 1200 UTC and uses NCAR's DART (Data Assimilation Research Testbed) software with ARW version 3.8 (HRRR v3 code base) with the same horizontal domain as the hrrre CLUE members. The mesoscale analysis system is comprised of 80 members that are continuously cycled using the ensemble adjustment Kalman filter (EAKF). New analyses are produced every 1 h with 15-km grid-spacing. These analyses are downscaled to the 3-km forecast grid and used to initialize forecasts at 0000 UTC. Additional analyses are downscaled once daily at 0600 UTC on the 3-km forecast grid, also with hourly cycling, for a window of 6 hours, with the addition of assimilating radar reflectivity observations on the 3-km mesh. These 3-km analyses are used to initialize CLUE forecasts at 1200 UTC, nested within forecasts initialized from the 15-km analysis domain. Other specifications include: 51 vertical levels with a 15 hPa top, a horizontal localization of 635 km and vertical localization of 0.5 scale heights, relaxation to prior spread posterior inflation (1.1), sampling error correction, spread restoration, and freely-evolving soil states. The following observational sources are utilized: PREPBUFR ACARS, METARs, radiosondes, profilers and marine, CIMMS cloud-track winds, Oklahoma Mesonet, and GPS radio occultation; on the 3-km domain for the 1200 UTC initialized forecasts radar reflectivity drawn from MRMS composites are also assimilated. All analysis members have constant physics, which include the new Tiedtke cumulus parameterization (from WRF V3.9.1.1), Thompson microphysics, MYJ PBL, Noah-MP land-surface model, and RRTMG shortwave and longwave radiation with aerosol and ozone climatologies. The 10-member forecasts are initialized daily at 0000 and 1200 UTC with ICs provided by the first ten ensemble analysis members of the WRF/DART EAKF analyses (as described above). LBCs from GEFS forecasts are used for the ensemble forecasts. The physics for the ensemble forecasts is the same as from the data assimilation system, with no cumulus scheme on the 3-km domain.

Members: ncar	**ICs**	**LBCs**	**Microphysics**	**PBL**	**LSM**	**Radiation**	**Model**
ncar01	anal01	GFS p01	Thompson	MYJ	NOAH-MP	RRTMG	ARW
ncar02	anal02	GFS p02	Thompson	MYJ	NOAH-MP	RRTMG	ARW
ncar03	anal03	GFS p03	Thompson	MYJ	NOAH-MP	RRTMG	ARW
ncar04	anal04	GFS p04	Thompson	MYJ	NOAH-MP	RRTMG	ARW
ncar05	anal05	GFS p05	Thompson	MYJ	NOAH-MP	RRTMG	ARW
ncar06	anal06	GFS p06	Thompson	MYJ	NOAH-MP	RRTMG	ARW
ncar07	anal07	GFS p07	Thompson	MYJ	NOAH-MP	RRTMG	ARW
ncar08	anal08	GFS p08	Thompson	MYJ	NOAH-MP	RRTMG	ARW
ncar09	anal09	GFS p09	Thompson	MYJ	NOAH-MP	RRTMG	ARW
ncar10	anal10	GFS p10	Thompson	MYJ	NOAH-MP	RRTMG	ARW

Table 11 Specifications for the map-hybrid CLUE members. These 3-km grid-spacing ensemble forecasts are run with WRF ARW and initialized by a GSI-based hybrid EnVar DA system directly assimilating both conventional and radar reflectivity observations (Johnson et al. 2015, Wang and Wang 2017). The ensemble for data assimilation has 41 members. The LBCs are provided by re-centering GEFS and SREF around the GFS control. The system assimilates the operational RAP/HRRR in-situ data stream hourly during 1800-0000 UTC and radar reflectivity every 20-min from 2300 to 0000 UTC over the CONUS CLUE domain. The control member is updated by GSI-based hybrid EnVar where both the ensemble covariance and the newly developed storm-scale static covariance are combined. A 10-member ensemble forecast is initialized at 0000 UTC and advanced for 36 hours, including one forecast (map-hybrid01) initialized from the GSI based hybrid EnKF-Var control analysis and 9-member re-centered GSI EnKF analyses. Same physics schemes as listed below are adopted for all members in both data assimilation and ensemble forecasts. The stochastic physics perturbations are additionally applied to ensemble free forecast.

Members: map-hybrid	**ICs**	**LBCs**	**Micro-physics**	**PBL**	**LSM**	**Radiation**	**Model**
map-hybrid01	hybrid EnVar	GFS-control	Thompson	MYNN	RUC	RRTMG	ARW
map-hybrid02	rEnKF_m1	GEFS	Thompson	MYNN	RUC	RRTMG	ARW
map-hybrid03	rEnKF_m2	GEFS	Thompson	MYNN	RUC	RRTMG	ARW
map-hybrid04	rEnKF_m3	GEFS	Thompson	MYNN	RUC	RRTMG	ARW
map-hybrid05	rEnKF_m4	GEFS	Thompson	MYNN	RUC	RRTMG	ARW
map-hybrid06	rEnKF_m5	GEFS	Thompson	MYNN	RUC	RRTMG	ARW
map-hybrid07	rEnKF_m6	GEFS	Thompson	MYNN	RUC	RRTMG	ARW
map-hybrid08	rEnKF_m7	GEFS	Thompson	MYNN	RUC	RRTMG	ARW
map-hybrid09	rEnKF_m8	GEFS	Thompson	MYNN	RUC	RRTMG	ARW
map-hybrid10	rEnKF_m9	GEFS	Thompson	MYNN	RUC	RRTMG	ARW

Table 12 Specifications for the map-ICpert CLUE members. A 10-member, 36-hr free forecast is initialized at 0000 UTC. These forecasts are initialized by re-centering the 0000 UTC GEFS analyses around the GSI hybrid EnVar control analysis (map-hybrid01). Consistent with "map-hybrid" members, same physics configuration including parameterization schemes and stochastic physics option is employed.

Members: map-hybrid	**ICs**	**LBCs**	**Micro-physics**	**PBL**	**LSM**	**Radiation**	**Model**
map-ICpert01	hybrid EnVar	GFS-control	Thompson	MYNN	RUC	RRTMG	ARW
map-ICpert02	rGEFS_m1	GEFS	Thompson	MYNN	RUC	RRTMG	ARW
map-ICpert03	rGEFS_m2	GEFS	Thompson	MYNN	RUC	RRTMG	ARW
map-ICpert04	rGEFS_m3	GEFS	Thompson	MYNN	RUC	RRTMG	ARW
map-ICpert05	rGEFS_m4	GEFS	Thompson	MYNN	RUC	RRTMG	ARW
map-ICpert06	rGEFS_m5	GEFS	Thompson	MYNN	RUC	RRTMG	ARW
map-ICpert07	rGEFS_m6	GEFS	Thompson	MYNN	RUC	RRTMG	ARW
map-ICpert08	rGEFS_m7	GEFS	Thompson	MYNN	RUC	RRTMG	ARW
map-ICpert09	rGEFS_m8	GEFS	Thompson	MYNN	RUC	RRTMG	ARW
map-ICpert10	rGEFS_m9	GEFS	Thompson	MYNN	RUC	RRTMG	ARW

Table 13 Specifications for the hrrre-nospp CLUE members. These members are the same as the hrrre member, except all stochastic physics are turned off.

Members: hrrre	**ICs**	**LBCs**	**Microphysics**	**PBL**	**LSM**	**Radiation**	**Model**
hrrre-nospp01	enkf_m01b	GFS	Thompson	MYNN	RUC	RRTMG	ARW
hrrre-nospp02	enkf_m02b	GFS	Thompson	MYNN	RUC	RRTMG	ARW
hrrre-nospp03	enkf_m03b	GFS	Thompson	MYNN	RUC	RRTMG	ARW
hrrre-nospp04	enkf_m04b	GFS	Thompson	MYNN	RUC	RRTMG	ARW
hrrre-nospp05	enkf_m05b	GFS	Thompson	MYNN	RUC	RRTMG	ARW
hrrre-nospp06	enkf_m06b	GFS	Thompson	MYNN	RUC	RRTMG	ARW
hrrre-nospp07	enkf_m07b	GFS	Thompson	MYNN	RUC	RRTMG	ARW
hrrre-nospp08	enkf_m08b	GFS	Thompson	MYNN	RUC	RRTMG	ARW
hrrre-nospp09	enkf_m09b	GFS	Thompson	MYNN	RUC	RRTMG	ARW

Table 14 Specifications for the nssl-fv3 CLUE member. This member uses the latest release of NEMSfv3gfs in the NOAA VLab and it run with a 25-km grid-spacing global mesh and high-resolution nest over the CONUS (3.3 km grid-spacing). The nssl-fv3 is initialized from 0000 UTC GFS analyses with forecasts to 60 h.

Member: HRRRv3	**ICs**	**LBCs**	**Microphysics**	**PBL**	**LSM**	**Radiation**	**Model**
nssl-fv3	GFS	n/a	Thompson	MYNN	NOAH	RRTMG	FV3

Table 15 Specifications for the nssl-sarfv3 CLUE member. This member uses the same physics as nssl-fv3, but with a stand-alone-regional configuration instead of global-with-nest.

Member: HRRRv3	**ICs**	**LBCs**	**Microphysics**	**PBL**	**LSM**	**Radiation**	**Model**
nssl-sarfv3	GFS	n/a	Thompson	MYNN	NOAH	RRTMG	FV3

Table 16 Specifications for the gfdl-fv3 CLUE member. The GFDL configuration uses a combination of grid nesting (Harris and Lin, 2013) and stretching (Harris et al 2016) to refine a 13-km global grid to a 3-km nested grid covering the CONUS region. This model consists of FV3 coupled to a modified form of the GFS Physics (Chen et al 2019 and references therein) and the Noah land model. We replace the existing GFS PBL and precipitation schemes with schemes better-suited for kilometer-scale prediction, including the Yonsei University PBL scheme (Hong et al 2006) and the six-category single-moment GFDL microphysics (Zhou et al 2019). The deep convective scheme is disabled on the nested grid. Initialization is a cold-start from regridded GFS real-time analyses. GFDL will provide simulations run daily at 00Z out to 126 hours to demonstrate the potential for medium-range prediction of convective-scale events.

Member: HRRRv3	**ICs**	**LBCs**	**Microphysics**	**PBL**	**LSM**	**Radiation**	**Model**
gfdl-fv3	GFS	n/a	GFDL	YSU	NOAH	RRTMG	FV3

Table 17 Specifications for the ukmet-sphys CLUE members. This ensemble uses single physics and is closely aligned with the UK ensemble, MOGREPS-UK. It uses the Met Office Unified Model (UM) and Joint UK Land Environment Simulator (JULES) with 2.2-km grid-spacing and 70 vertical levels across a slightly sub-CONUS domain. The members are downscaled from the Met Office global ensemble system, MOGREPS-G, and cold-start initialized at 1800 (6-h time lag) and 0000 UTC to produce forecasts to 48-h. Model uncertainty is depicted by the Random Parameter (RP) scheme, which stochastically perturbs a subset of physics parameters from the PBL and microphysics schemes throughout the forecast. The PBL scheme consists of a 3D turbulent mixing scheme using a locally scale-dependent blending of Smagorinsky and non-local K-profile boundary layer mixing schemes, and the Smith cloud scheme is used, where partial cloudiness is diagnosed assuming a trianglular moisture distribution with a width that is a universally specified function of height only. There is no convection-parameterization and single-moment microphysics is used. The "RA config" column refers to the "Regional Atmosphere" configuration. For these members "RA2M" is used, which designed for mid-latitudes.

Members: ukmet-sphys	ICs/LBCs	Init. time (UTC)	RA config	Micro-physics	PBL	Cloud Scheme	Model
ukmet-sphys01	MOGREPS-G01	0000	RA2M	single-mom.	Smag. blended	Smith	UM
ukmet-sphys02	MOGREPS-G02	0000	RA2M	single-mom.	Smag. blended	Smith	UM
ukmet-sphys03	MOGREPS-G03	0000	RA2M	single-mom.	Smag. blended	Smith	UM
ukmet-sphys04	MOGREPS-G04	0000	RA2M	single-mom.	Smag. blended	Smith	UM
ukmet-sphys05	MOGREPS-G05	0000	RA2M	single-mom.	Smag. blended	Smith	UM
ukmet-sphys06	MOGREPS-G06	0000	RA2M	single-mom.	Smag. blended	Smith	UM
ukmet-sphys07	MOGREPS-G18	1800	RA2M	single-mom.	Smag. blended	Smith	UM
ukmet-sphys08	MOGREPS-G19	1800	RA2M	single-mom.	Smag. blended	Smith	UM
ukmet-sphys09	MOGREPS-G20	1800	RA2M	single-mom.	Smag. blended	Smith	UM
ukmet-sphys10	MOGREPS-G21	1800	RA2M	single-mom.	Smag. blended	Smith	UM
ukmet-sphys11	MOGREPS-G22	1800	RA2M	single-mom.	Smag. blended	Smith	UM
ukmet-sphys12	MOGREPS-G23	1800	RA2M	single-mom.	Smag. blended	Smith	UM

Table 18 Specifications for the ukmet-mphys CLUE members. This ensemble uses mixed-physics, with 6 of the members shared from ukmet-sphys and 6 members that use an alternative configuration known as RA2T, which is designed for the tropics. The RA2T configurations use a Prognostic Cloud Scheme (PC2) where partial cloudiness is prognosed, with sources and sinks being calculated from all other parameterization schemes that modify temperature or moisture. The updated clouds are then advected by the wind. The other difference in the RA2T configurations is that the stochastic perturbations in the PBL are turned off.

Members: ukmet-mphys	ICs/LBCs	Init. time (UTC)	RA config	Micro-physics	PBL	Cloud Scheme	Model
ukmet01_ra2t	MOGREPS-G01	0000	RA2T	single-mom.	Smag. blended	PC2	UM
ukmet02_ra2t	MOGREPS-G02	0000	RA2T	single-mom.	Smag. blended	PC2	UM
ukmet03_ra2t	MOGREPS-G03	0000	RA2T	single-mom.	Smag. blended	PC2	UM
ukmet-sphys04	MOGREPS-G04	0000	RA2M	single-mom.	Smag. blended	Smith	UM
ukmet-sphys05	MOGREPS-G05	0000	RA2M	single-mom.	Smag. blended	Smith	UM
ukmet-sphys06	MOGREPS-G06	0000	RA2M	single-mom.	Smag. blended	Smith	UM
ukmet07_ra2t	MOGREPS-G18	1800	RA2T	single-mom.	Smag. blended	PC2	UM
ukmet08_ra2t	MOGREPS-G19	1800	RA2T	single-mom.	Smag. blended	PC2	UM
ukmet09_ra2t	MOGREPS-G20	1800	RA2T	single-mom.	Smag. blended	PC2	UM
ukmet-sphys10	MOGREPS-G21	1800	RA2M	single-mom.	Smag. blended	Smith	UM
ukmet-sphys11	MOGREPS-G22	1800	RA2M	single-mom.	Smag. blended	Smith	UM
ukmet-sphys12	MOGREPS-G23	1800	RA2M	single-mom.	Smag. blended	Smith	UM

Table 19 Specifications for the emc-fv3 CLUE member. Forecasts to 60-h are run daily using initial conditions from the 0000 UTC GFSv15 system currently under parallel testing by NCO.

Member: HRRRv3	ICs	LBCs	Microphysics	PBL	LSM	Radiation	Model
emc-fv3	GFSv15	n/a	GFDL	EDMF	NOAH	RRTMG	FV3

Table 20 Specifications for the emc-sarfv3 CLUE member. Forecast to 60-h are run daily using initial conditions from the 0000 UTC GFSv15 system currently under parallel testing by NCO. Lateral boundary conditions are specified every 3 hours from the GFSv15 forecasts.

Member: HRRRv3	ICs	LBCs	Microphysics	PBL	LSM	Radiation	Model
emc-sarfv3	GFSv15	GFSv15	GFDL	EDMF	NOAH	RRTMG	FV3

The configuration of the 2019 CLUE will allow for several 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) Stochastic physics perturbations: The HRRRE (Table 9) is a single physics ensemble, but uses stochastic parameter perturbations (SPP) applied to the land-surface, PBL, and microphysics schemes to account for model error. To evaluate the impact of SPP, an ensemble identical to the HRRRE, but run with SPP turned off (hrrre-nospp; Table 13) will be compared to the HRRRE. The ultimate goal is for a single physics ensemble with stochastic perturbations to be as good as or better than mixed-physics/mixed-model systems like HREFv2.1 in terms of forecast skill and reliability. The SPP approach is based on stochastic pattern generator that produces 2D perturbation fields with spatial and temporal correlations. The approach is analogous to the one used at European Center for Medium-Range Weather Forecasts (ECMWF) to perturb physics tendencies (Palmer 2009). The pattern is fully determined by four parameters specified by the user (with the namelist setting shown in parentheses): grid point standard deviation (gridpt_stddev and stddev_cutoff), length scales (lengthscale) and de-correlation time (timescale).

(2) UM single vs. mixed-physics: The Met Office is providing two 12-member ensembles to SFE 2019. The first ensemble, ukmet-sphys (Table 17), which is referred to as the "primary-ensemble" for the Met Office, uses the same physics configuration in each member and is closely aligned with the regional UK ensemble known as MOGREPS-UK. The second ensemble, ukmet-mphys (Table 18), is a mixed-physics ensemble that contains two different physics configurations: 6 members are shared from ukmet-sphys, and the other 6 members use an alternative configuration. The main goal of the UM ensembles is to compare the primary ensemble (ukmet-sphys) to other ensembles within the CLUE, and to explore the impact of using two physics configurations on the ensemble skill and spread by comparing the mixed-physics ensemble against the primary ensemble, which uses single-physics.

(3) FV3 physics sensitivities: CAPS is providing an ensemble of SARFV3 members that use the same set of ICs/LBCs, but different physics parameterizations (fv3-phys; Table 3). Three microphysics schemes

(Thompson, NSSL, and Morrison), three PBL schemes (saMYNN, saShinHong, and EDMF), and two land surface models (RUC and NOAH) are being tested in these members. This will be the first systematic testing of different physics configurations with SARFV3 using the Common Community Physics Package (CCPP) interface.

(4) Global-with-nest vs. SAR FV3: Two pairs of CLUE members with matching physics configurations will be run where one member uses a global configuration of FV3 with a high-resolution nest over the CONUS, and the other member uses SAR FV3 with ICs/LBCs provided by the GFS. The goal of these comparisons is to make sure that SAR FV3 performance is similar to that of the global-with-nest, and to see if there is any degradation at later forecast hours because of errors introduced by the LBCs in SAR FV3. These comparisons are between the nssl-fv3 and nssl-sarfv3 members provided by NSSL and listed in Tables 14 and 15, respectively, and the emc-fv3 and emc-sarfv3 members provided by EMC and listed in Tables 19 and 20, respectively. Note, the emc-fv3 and emc-sarfv3 comparisons are also being conducted outside the SFE, and comparison web graphics between these two configurations can be found at the following EMC web page: https://www.emc.ncep.noaa.gov/mmb/bblake/fv3/00z/nest/main.php.

(5) HRRRv3 vs. HRRRv4 vs. gsd-sarfv3: HRRRv4 (Table 7) will be the last operational implementation of the HRRR that uses the WRF-ARW dynamical core. After its implementation, subsequent versions will use SAR FV3 as soon as SAR FV3 forecast skill is determined to be similar or better than HRRRv4. Thus, to gauge the readiness of SAR FV3, ESRL/GSD is providing forecasts from a configuration of SAR FV3 (gsd-sarfv3; Table 8) that uses the same ICs/LBCs and physics configurations as HRRRv4. Additionally, comparisons will be made between HRRRv3 and HRRRv4.

(6) Data assimilation comparisons: Similar to the 2018 CLUE, there are several ensemble subsets that use various data assimilation strategies with an EnKF component. These ensembles include the map-hybrid (Table 11), map-ICpert (Table 12), ncar (Table 10), HRRRE (Table 9), and caps-enkf (Table 5). The most controlled comparison among these members is between the map-hybrid and map-ICpert ensembles. These ensembles using the same physics and share the same control member, which has ICs generated using GSI-based hybrid EnVar where the ensemble covariance and storm-scale static covariance are combined. In map-hybrid, the ensemble member ICs are generated from re-centered GSI EnKF analyses, and in map-ICpert the ensemble member ICs are generated by re-centering the 0000 UTC GEFS analyses around the control member. The goal of this comparison is to test the effectiveness of two different strategies for IC perturbations in an ensemble using GSI-based hybrid EnVar. The other ensembles listed above (ncar, HRRRE, and caps-enkf) allow for less controlled comparisons because there are differences in their configurations in addition to the different data assimilation strategies employed, but they will still be compared to assess overall differences in performance characteristics.

To ensure consistent post-processing, visualization, and verification for subsets of CLUE ensemble members contributed by different collaborators, all groups will generally utilize similar post-processing software to output the same set of model output fields on the same grid. For WRF-ARW members, the Unified Post-Processor software (UPP; available at http://www.dtcenter.org/upp/users/downloads/index.php) is used to output a minimum set of 123 output fields from each CLUE member (Table 21). These fields (output in grib2 format) are the same as the 2D fields output by HRRRv3 and were chosen because of their relevance to a broad range of forecasting needs, including aviation, severe weather, and precipitation. The UM ensembles will output a much smaller set of fields limited to low-level temperature, dewpoint, and winds; lowest model level and composite reflectivity, hourly maximum 2-5 km AGL updraft helicity, and total precipitation. Finally, the FV3 runs will use UPP software developed at EMC and output as many of the fields as output by the WRF-ARW runs as possible. The FV3 output fields will include storm attributes like updraft helicity and hail size.

Table 21 The minimum set of 123 output diagnostics for the WRF-ARW 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	Radar-Simulated 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	100-1000 mb above ground	MAXUVV	Hourly Max upward Vertical Velocity - lowest 100hPa [m/s]
032	100-1000 mb above ground	MAXDVV	Hrly Max downward Vertical Velocity - lowest 100hPa [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 Integrated 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]
108	-10 C	REFD	Hourly maximum of -10C reflectivity
109	5000-2000 m above ground	MNUPHL	Hrly Min Updraft Helicity - 2km to 5 km AGL [m^2/s^2]
110	2000-0 m above ground	MXUPHL	Hrly Max Updraft Helicity - 0km to 2 km AGL [m^2/s^2]
111	2000-0 m above ground	MNUPHL	Hrly Min Updraft Helicity - 0km to 2 km AGL [m^2/s^2]
112	3000-0 m above ground	MXUPHL	Hrly Max Updraft Helicity - 0km to 3 km AGL [m^2/s^2]
113	3000-0 m above ground	MNUPHL	Hrly Min Updraft Helicity - 0km to 3 km AGL [m^2/s^2]
114	2000-0 m above ground	RELV	Hrly Max Rel. Vort. – 0km to 2km AGL [1/s]
115	1000-0 m above ground	RELV	Hrly Max Rel. Vort. – 0km to 1km AGL [1/s]
116	entire column	HAIL	Hrly Max of Hail/Graupel Diameter [m]
117	0.1 sigma	HAIL	Hrly Max of Hail/Graupel Diameter [m]
118	5000-2000m AGL	UPHL	Updraft Helicity (instantaneous) [m^2/s^2]
119	6000-1000m AGL	UPHL	Updraft Helicity (instantaneous) [m^2/s^2]
120	top of atmos	SBT123	Simulated Brightness T for GOES 12 Ch. 3 [K]
121	top of atmos	SBT124	Simulated Brightness T for GOES 12 Ch. 4 [K]
122	top of atmos	SBT113	Simulated Brightness T for GOES 11 Ch. 3 [K]
123	top of atmos	SBT114	Simulated Brightness T for GOES 11 Ch. 4 [K]

b) High Resolution Ensemble Forecast (HREFv2.1) System

HREFv2.1 is a 10-member CAM ensemble currently running at EMC with forecasts that can be viewed at: http://www.spc.noaa.gov/exper/href/. HREFv2 was implemented operationally on 1 November 2017 and was recently updated to include two HRRR members (one 6-h time lagged). The design of HREFv2.1 originated from the SSEO, which demonstrated skill during the previous six years in the HWT and SPC prior to HREFv2 operational implementation. All members, except for the NAM CONUS Nest and HRRR, are initialized with a “cold-start”. Forecasts to 36 h, including storm-attribute hourly maximum fields (HMFs), are produced at 0000 and 1200 UTC. HREFv2.1 specifications are listed below in Table 21.

Table 22 Model specifications for HREFv2.1.

HREFv2.1	ICs	LBCs	Micro-physics	PBL	dx (km)	Vertical Levels	Included in HREF hours
HRRR	RAP -1h	RAP -1h	Thompson	MYNN	3.0	50	0 – 36
HRRR -6h	RAP -1h	RAP -1h	Thompson	MYNN	3.0	50	0 – 30
HRW ARW	RAP	GFS -6h	WSM6	YSU	3.2	50	0 – 48
HRW ARW -12h	RAP	GFS -6h	WSM6	YSU	3.2	50	0 – 36
HRW NMMB	RAP	GFS -6h	Ferrier-Aligo	MYJ	3.2	50	0 – 48
HRW NMMB -12h	RAP	GFS-6h	Ferrier-Aligo	MYJ	3.2	50	0 – 36
HRW NSSL	NAM	NAM -6h	WSM6	MYJ	3.2	40	0 – 48
HRW NSSL -12h	NAM	NAM -6h	WSM6	MYJ	3.2	40	0 – 36
NAM CONUS Nest	NAM	NAM	Ferrier-Aligo	MYJ	3.0	60	0 – 48
NAM CONUS Nest -12h	NAM	NAM	Ferrier-Aligo	MYJ	3.0	60	0 – 48

c) NSSL Experimental Warn-on-Forecast System (WoFS)

The NSSL Experimental Warn-on-Forecast System (WoFS) is a 36-member WRF-based ensemble data assimilation system used to produce very short-range (0-6 h) probabilistic 18-member forecasts of hazardous weather phenomena, such as supercell thunderstorm rotation, hail, high winds, and flash flooding. The starting point for each day's experiment will be the experimental HRRRE (Table 16) provided by ESRL/GSD. The full ensemble is updated by hourly GSI-EnKF data assimilation of conventional observations and Multi-Radar/MultiSensor (MRMS) radar reflectivity from 0300 UTC to 1800 UTC Day 1. A 36-h ensemble forecast launched from the 1200 UTC HRRRE analysis is used to provide boundary conditions for the WoFS system for the period 1800 UTC Day 1 – 0300 UTC Day 2. Similarly, a 1-h ensemble forecast launched from the 1700 UTC HRRRE analysis is used to provide initial conditions for the WoFS at 1800 UTC.

The daily WoFS domain location will target the primary region where severe weather is anticipated and cover a 900-km wide region with very frequent 15-min updates. All ensemble members utilize the NSSL 2-moment 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. MRMS radar reflectivity and NEXRAD Level II radial velocity data, cloud water path retrievals from the GOES-16 imager, and Oklahoma Mesonet observations (when available) will be assimilated every 15 min also using the GSI-EnKF method, beginning at 1800 UTC each day. ASOS data will also be assimilated at 15 minutes past each hour. 6-h (3-h) ensemble forecasts will be initialized from the WoFS analysis each hour (half-hour) from 1900 UTC Day 1 through 0300 UTC Day 2 for HWT product evaluation from 2000 – 2100 UTC. These forecasts will be viewable using the web-based WoFS Forecast Viewer (https://www.nssl.noaa.gov/projects/wof/WoFS/realtime/). Table 16 shows the differences in model specifications between the HRRRE and WoFS, and Figure 5 shows an example of a SPC Day 1 convective outlook and corresponding WoFS grid with WSR-88D radars used for data assimilation overlaid.

Table 23 HRRRE and WoFS configuration comparison.

	HRRRE	NEWS-e
Model Version	WRF-ARW v3.8+	WRF-ARW v3.8+
Grid Points	1800 x 1060 x 50	300 x 300 x 50
Grid Spacing	3 km	3 km
EnKF Cycling	36 mem. w/ GSI-EnKF **every 1 hr**	36 mem. w/ GSI-EnKF **every 15 min**
Observations	**-Conventional obs: T, q_v, u, v, and p from rawinsonde, aircraft, surface (land and marine), profiler; -MRMS radar reflectivity**	**-Surface: ASOS, Oklahoma Mesonet (when available); -Doppler velocity from ~20-25 WSR-88D sites; -MRMS reflectivity > 20 dBZ; radar 'zeroes'; -Cloud-water path (GOES-16)**
Radiation LW/SW	RRTMG/RRTMG	Dudhia/RRTM, RRTMG/RRTMG
Microphysics	**Thompson (aerosol aware)**	**NSSL 2-moment**
PBL	**MYNN**	**YSU, MYJ, or MYNN**
LSM	RUC (Smirnova)	RUC (Smirnova)

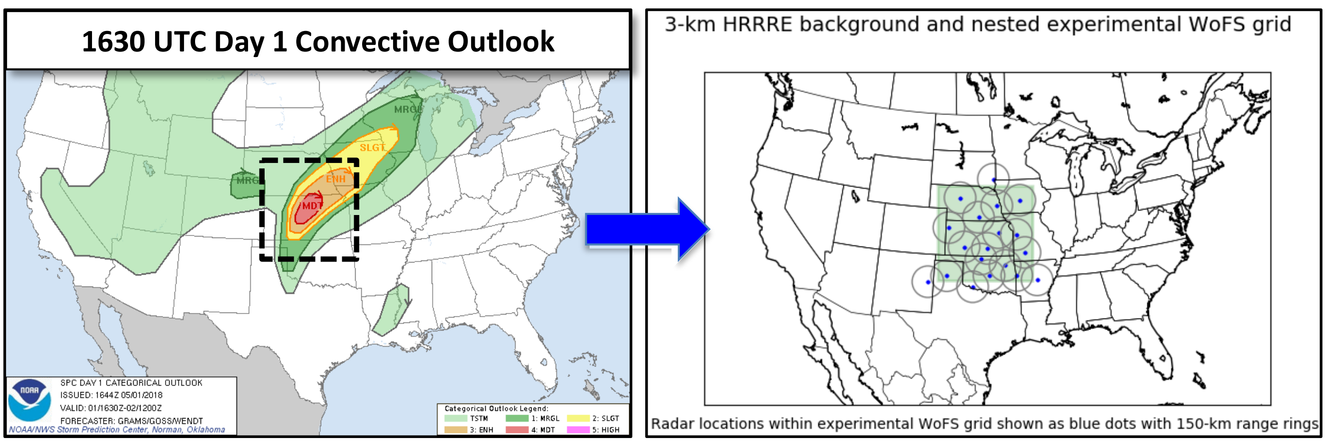

Figure 1 SPC 1630 UTC issued Day 1 convective outlook (left) and corresponding WoFS grid (right).

3. SFE 2019 Core Interests/Daily Activities

a. Forecast products and activities

The experimental forecasts this year will focus on our ability to add temporal specificity to convective outlooks within the Day 1 and Day 2 time period. Additionally, we will explore the feasibility of providing more precise information on the intensity of specific hazards. We will continue to split the participants into two desks, with those at the Innovation Desk forecasting total severe threat (combining hail, wind, and tornado hazards) and those at the Severe Hazards Desk forecasting individual severe hazards. The experimental forecasts will 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.

For the Severe Hazards Desk, the first forecast will be done as a group and will mimic the SPC operational Day 1 Convective Outlooks by producing individual probabilistic 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. Then, the Severe Hazards Desk will issue conditional intensity forecasts of tornado, wind, and hail, in which areas are delineated with reports that are expected to follow a "normal", "hatched", or "double-hatched" distribution. These conditional intensity forecasts are new to SFE 2019; thus, the following provides some background for their generation: When generating Day 1 Convective Outlooks, SPC forecasters draw probabilities that represent the chance of each hazard occurring within 25 miles of a point. Forecasters can also delineate "hatched" areas, which represent regions with a 10% chance or greater of significant severe weather (EF-2 or greater tornadoes, winds $\geq$ 65 kts, or hail $\geq$ 2-in.) within 25 miles of a point. Research by the SPC has shown that, as the forecast coverage of a hazard increases, the expected intensity of the verifying reports also increases. For instance, on days where a "hatched" area is drawn and the maximum tornado coverage is 10 or 15%, 17% of the observed tornadoes are significant. When a "hatched" area is drawn and the maximum tornado coverage is 30% or higher, 32% of observed tornadoes are significant. In other words, as the forecast tornado coverage increases, the observed tornadoes grow progressively more intense, regardless of how many tornadoes occur; preliminary results show a similar pattern for wind and hail. Therefore, current coverage forecasts include intensity information that is not explicitly communicated to users, so coverage forecasts and intensity forecasts could be better labeled/communicated. These results have been used to identify three conditional intensity probability distributions that can be forecast via examination of the atmospheric environment: "normal", "hatched", and "double-hatched". In plain language, "normal" refers to a typical severe weather day, where significant severe weather is unlikely, "hatched" areas indicate where significant severe weather is possible, and "double-hatched" areas indicate where high-impact significant severe weather is expected. After these two sets of forecasts have been completed as a group, the 1600–1200 UTC outlook for individual hazards will be temporally disaggregated into 4-h periods (hourly through the end of the convective day) using HREF/SREF calibrated hazard guidance to provide automated timing information on the severe weather threat, as has been provided in previous SFEs. Then, the Severe Hazards Desk participants will split into five groups and use a web interface to generate their own set of coverage and intensity forecasts using Google Chromebooks. Each group will be assigned a specific CLUE subset to use for this task. The subsets will include caps-enkf, HRRRE, fv3-ens, map-hybrid, and HREFv2.1.

For the Innovation Desk, the first forecast will also cover the 1600 to 1200 UTC time period and be conducted as a group. Rather than individual hazards, the Innovation Desk will issue probabilistic outlooks for total severe (combined tornado, hail, and wind). Similar to SFE 2018, the Innovation Desk will then create a product aimed toward the emergency management community, designating the 4-h periods when severe weather is expected throughout the day. These potential severe timing areas (PSTs) will occur within areas of 15% probability as indicated by the Day 1 full-period outlook previously generated by the Innovation Desk. Despite the different end products, the goals of the activities are the same as in prior years – 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), and to begin to explore ways of seamlessly bridging probabilistic severe weather outlooks and probabilistic severe weather warnings as part of the NOAA WoF and FACETS initiatives.

The PSTs will be issued by the lead forecaster at the Innovation Desk on the N-AWIPS machines. Meanwhile, the Innovation Desk participants will split into five groups and use a web interface to generate their own set of PSTs using Google Chromebooks. Similar to the Severe Hazards Desk activity, each group will be assigned a specific CLUE subset to use for this task that will include caps-enkf, HRRRE, fv3-ends, map-hybrid, and HREFv2.1. After issuing the PSTs, the Innovation Desk will regroup and discuss the forecasts and behavior of the CLUE subsets. This approach is planned to engage the participants more directly with the CLUE subsets, since in prior year participants only interacted with CLUE subsets through facilitator-led discussions.

After both desks have issued all their morning outlooks, there will be a map discussion open to all tenants of the National Weather Center summarizing forecast challenges and highlighting interesting findings from the previous day. Each day of the week will also feature a brief discussion of a special topic, which can be found on the daily schedule of activities (Table 24).

After lunch, the Severe Hazards Desk will issue Day 2 full-period (i.e., 1200 to 1200 UTC the next day) probabilistic forecasts of tornado, wind, and hail over a regional area of interest, which will be done as a group activity. Similarly, the Innovation Desk will issue full period probabilistic forecasts of total severe, as well as 4-h PSTs, which will both be conducted as a group.

Later in the afternoon, scientific evaluations will take place (summarized in the next section). For the final activity of the day on Tuesday through Friday, forecasting activities using the WoF Ensemble will be conducted on both desks from 3-4pm. On Mondays, a training activity for the WoF activity will occur from 3-4pm for SFE participants at both desks. At the Severe Hazards Desk, on Tuesday-Friday, participants will update their full period (2100 – 1200 UTC) hazard probability and conditional intensity forecasts in the same small groups as the morning activity and using the Chromebooks. At the Innovation Desk, forecasts will be drawn by facilitators (Adam Clark and Burkely Gallo) and informed by small groups of participants interrogating the WoF system data on their Chromebooks or personal laptops. Probabilistic total severe outlooks will be issued that are valid for short (1-h) and long (4-h) time windows. Additionally, an outlook will be issued that covers a "targeted" 1-h time window valid 8-9pm (0100 – 0200 UTC). The short (1-h) outlook will be issued during the 3-3:30pm time period using the 1930 UTC initialization of the WoF system and will be valid 2100 – 2200 UTC (4-5pm). The long (4-h) outlook will be issued during the 3:30-4pm time period using the 2000 UTC initialization of the WoF system and will be valid 2200 – 0200 UTC. Finally, the targeted (1-h) outlook will also be issued during the 3:30-4pm time period using the 2000 UTC WoF initialization and will be valid 0100 – 0200 UTC. After this set of outlooks at both desks has been issued, SFE activities will conclude for the majority of participants. However, for two forecasters that have been selected for a WoF-based evening activity, additional sets of outlooks will be issued each hour from 4-8pm. The evening forecasters will issue these outlooks individually on the Chromebooks. Additionally, evaluation activities will be conducted at 6, 7, and 8pm. NSSL facilitators will be on hand every evening to assist in the forecast generation and evaluation process. Table 25 summarizes both the 3-4pm WoF activities and the 4-8pm WoF activities at the Innovation Desk.

These WoF activities are the third year the WoF Ensemble has been tested in the EFP, and explores the potential utility of WoF products for issuing guidance between the watch and warning time scales (i.e. 0.5 to 6-h lead times). These activities represent efforts to explore 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. These efforts also support the transition to higher temporal resolution forecasts at the SPC.

Table 24 SFE 2019 Schedule of Daily Activities.

Severe Hazards Desk	*Innovation Desk*
0800 – 0845: **Evaluation of Experimental Forecasts & Guidance** Subjective rating relative to radar evolution/characteristics, warnings, preliminary reports, and MRMS MESH and rotation tracks	
• Days 1 & 2 full-period probabilistic forecasts of tornado, wind, and hail • Day 1 4-h period temporal disaggregation and guidance for tornado, wind, and hail	• Days 1 & 2 full-period probabilistic forecast of total severe • Days 1 & 2 4-h potential severe timing areas • Day 1 1-h and 4-h total severe outlooks
0845 – 0915: **Map Analysis** Hand analysis of 12Z upper-air & surface maps, discussion, and domain selection (from two areas)	
0915 – 1130: **Convective Outlook Generation**	
• Day 1 full-period probabilistic forecasts of tornado, wind, and hail valid 16-12Z over mesoscale area of interest* • Day 1 full-period (16-12Z) conditional intensity forecasts of tornado, wind, and hail using CLUE subsets*	• Day 1 full-period probabilistic forecast of total severe valid 16-12Z over mesoscale area of interest • Day 1 4-h potential severe timing (PST) areas (16-12Z) for full-period total severe ≥15% using CLUE subsets*
1130 – 1200: **Map Discussion** Brief discussion of today's forecast challenges and products Topic of the day: Ens. Subsetting (M), FV3 (T), WoF system (W), Met Office (R), CAM scorecard (F)	
1200 – 1300: **Lunch**	
1300 – 1345: **Convective Outlook Generation**	
• Day 2 full-period probabilistic forecasts of tornado, wind, and hail valid 12-12Z over mesoscale area of interest	• Day 2 full-period prob. forecast of total severe valid 12-12Z over mesoscale area of interest & 4-h PST areas (≥15% prob.)
1345 – 1500: **Scientific Evaluations**	
• Mesoscale Analyses • CLUE: CAM Ensembles • HREF Configurations w/FV3 • CLUE: Physics & IC Perturbations • Hail Guidance • Sensitivity-Based Ensemble Subsetting	• Ensemble Object-Based Probabilities • Deterministic CAMs (FV3 Nest, SAR) • Deterministic CAMs (HRRR, UM) • CLUE: FV3 Physics • WoF System Evaluation • WoF-based Outlook Evaluation
1500 – 1600 (2000 for WoF participants): **Short-term Outlook Update**	
• Update full-period prob/intensity forecasts of tornado, wind, and hail valid 21-12Z using observations and WoF system*	• Utilize obs. and WoF system to generate short (1-h), long (4-h), and targeted (1-h) probabilistic forecasts of total severe*
* Denotes forecasts also made by participants using the web drawing tool on Chromebooks.	

Table 25 Details on when experimental outlooks and evaluations based on the WoF ensemble will be issued, their valid time, and the latest WoF Ensemble initialization that will be available. The yellow, green, and blue shaded cells indicate the short, long, and targeted outlooks, respectively, while the unshaded cells indicate times at which subjective evaluations of earlier forecasts will be conducted.

	Experiment Time	Outlook Valid Time	WoF Ensemble Initialization
All Participants (Tues – Fri)	3:00 – 3:30 PM	4:00 – 5:00 PM	2:30 PM (1930 UTC)
	3:30 – 4:00 PM	5:00 – 9:00 PM	3:00 PM (2000 UTC)
	3:30 – 4:00 PM	8:00 – 9:00 PM	3:00 PM (2000 UTC)
Evening Participants (Mon – Thurs)	4:00 – 4:30 PM	5:00 – 6:00 PM	3:30 PM (2030 UTC)
	4:30 – 5:00 PM	6:00 – 10:00 PM	4:00 PM (2100 UTC)
	4:30 – 5:00 PM	8:00 – 9:00 PM	4:00 PM (2100 UTC)
	5:00 – 5:30 PM	6:00 – 7:00 PM	4:30 PM (2130 UTC)
	5:30 – 6:00 PM	7:00 – 11:00 PM	5:00 PM (2200 UTC)
	5:30 – 6:00 PM	8:00 – 9:00 PM	5:00 PM (2200 UTC)
Evaluation	6:00 PM	4:00 – 5:00 PM	2:30 PM (1930 UTC)
	6:00 – 6:30 PM	7:00 – 8:00 PM	5:30 PM (2230 UTC)
	6:30 – 7:00 PM	8:00 PM – 12:00 AM	6:00 PM (2300 UTC)
	6:30 – 7:00 PM	8:00 – 9:00 PM	6:00 PM (2300 UTC)
Evaluation	7:00 PM	5:00 – 6:00 PM	3:30 PM (2030 UTC)
	7:00 – 7:30 PM	8:00 – 9:00 PM	6:30 PM (2330 UTC)
	7:30 – 8:00 PM	9:00 PM – 1:00 AM	7:00 PM (0000 UTC)
Evaluation	8:00 PM	6:00 – 7:00 PM	4:30 PM (2130 UTC)

b. Formal Evaluation Activities

There will be two periods of formal evaluations during SFE 2019 (not including evening activities). The first will occur during the morning on Tuesday through Friday, 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 (LSRs), NWS warnings, and Multi-Radar Multi-Sensor (MRMS) radar-estimated hail sizes over the same time periods. Both raw LSRs and "practically perfect" fields (Hitchens et al. 2013) will be used. 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 updated 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.

The afternoon evaluation period will involve comparisons of different ensemble diagnostics and CLUE ensemble subsets. The Innovation and Severe Hazards Desks will conduct two different sets of evaluations, which are summarized below.

Innovation Desk Evaluations

(1) Ensemble Object-Based Probabilities

The object-based probabilistic (OBPROB) approach to CAM ensemble verification and visualization is intended to objectively quantify ensemble probabilistic forecasts of convective scale details, such as storm mode and morphology, that can be lost when generating grid point or neighborhood-based ensemble mean or probability plots. Manually evaluating such details in each available ensemble member can be time consuming, and even then, might be difficult to synthesize into actionable quantitative guidance. The object identification and matching procedures (summarized in Appendix D) are similar to what was applied during the 2018 SFE. Two new developments (also in Appendix D), include the ability to look at the details of individual storms from the ensemble that are matched to each probabilistic object and calibration of the object-based probabilities. The OBPROB verification and visualization framework will be evaluated by SFE participants in order to determine whether/how the OBPROB products affect the forecast process (e.g., raising new questions about the forecast, answering questions of the day about the forecast, concisely summarizing what was already inferred, etc.).

(2) Deterministic CAMs: FV3 Nest and SARFV3

This activity will focus on assigning ratings to gauge the skill and utility of several pairs of deterministic CAMs that use FV3 global-with-nest and SARFV3. Particular attention will be given to simulated storm structure, convective evolution, and location/coverage of storms. Storm surrogate fields, like hourly maximum updraft helicity, will also be examined to gauge their utility for forecasting severe storms. The main goal of these comparisons is to make sure that SAR FV3 performance is similar to that of the global-with-nest, and to see if there is any degradation at later forecast hours because of errors introduced by the LBCs in SAR FV3 (additional details provided in section 2a, CLUE Experiment 4).

(3) Deterministic CAMs: HRRR, ukmet-sphys01, core-cntr, and gsd-sarfv3

This activity will focus on assigning ratings to gauge the skill and utility of deterministic CAMs including HRRRv3, HRRRv4, ukmet-sphys01 (Table 17), core-cntr (version of SARFV3 provided by CAPS and initialized with the NAM; Table 3), and gsd-sarfv3 (version of SARFV3 provided by ESRL/GSD and initialized with RAP; Table 8). Particular attention will be given to simulated storm structure, convective evolution, and location/coverage of storms. Storm surrogate fields, like hourly maximum updraft helicity, will also be examined to gauge their utility for forecasting severe storms. The main goals for these comparisons are to gauge differences/improvements in HRRRv4 relative to HRRRv3, and to gauge performance of deterministic UM and SARFV3 systems.

(4) FV3 Physics Evaluation

The contribution by CAPS of seven different SARFV3 runs initialized from the NAM allows for testing of the forecast sensitivity to different parameterization schemes (Table 3). Three microphysics, three PBL, and two land surface model schemes are being tested. This is the first systematic testing of different physics configurations with SARFV3 using the Common Community Physics Package (CCPP) interface. Additional details provided in section 2a, CLUE Experiment 3).

(5) Warn-on-Forecast (WoF) Ensemble Evaluation

To gauge the performance of the WoF ensemble and measure performance relative to operationally available systems, WoF ensemble forecasts will be compared to HRRR time-lagged (HRRR-TL) forecasts and the HRRRE. This evaluation will focus on 1900 and 2100 UTC WoF Ensemble initializations and corresponding HRRR-TL and HRRRE (i.e., 1800 UTC) initializations.

(6) WoF-based Outlook Evaluations

The outlooks issued the previous day by the evening forecasters based on the WoF ensemble will be evaluated. A summary of these outlooks was provided in Table 25.

Severe Hazards Desk Evaluations

(1) Mesoscale Analyses

For the first time in the SFEs, an evaluation will be conducted focused on different mesoscale analysis systems. The analyses examined will include (1) SPC's surface Objective Analysis (sfcOA), (2) the 3D real-time mesoscale analysis (3D-RTMA) upscaled to the sfcOA 40-km grid, (3) the 3D-RTMA on its native 3-km grid, (4) the HRRRE 15-km mean analyses, and (5) the WoF Ensemble mean analyses. Displays of temperature and dewpoint with dots indicating the locations of surface stations that are sized and color coded to indicate differences with respect to the analysis system grids will be used for the evaluations. Various CAPE fields will also be examined.

(2) CLUE: CAM Ensembles

This comparison encompasses multiple elements. One aspect focuses on four ensembles within the CLUE that are configured with EnKF or hybrid data assimilation: the HRRRE (Table 8), the NCAR ensemble (Table 9), CAPS EnKF (Table 4), and the MAP Hybrid system (Table 10). A second aspect compares the 1200 UTC cycle of the HRRRE and NCAR ensembles to the HRRR-TL. All of these ensembles will be compared to HREFv2.1, which is used as the baseline.

(3) HREF Configurations with FV3

Various configurations of HREF will be evaluated, which will include the current version, HREFv2.1, and configurations with combinations of members without the two HRW NMMB members and with SARFV3 provided by EMC (emc-sarfv3; Table 20). Other combinations will include ones without the HRRR and HRW NMMB members and with the SARFV3 member.

(4) CLUE: Physics & IC Perturbations

This comparison will test three pairs of CLUE subsets that will all be compared against HREFv2.1. The first comparison will evaluate the impact of stochastic physics by comparing the HRRRE, which uses stochastic parameter perturbations (SPP) applied to several physics schemes, to a version of HRRRE with SPP turned off (hrrre-nospp; Table 13). The second comparison will compare the two UM ensembles with and without mixed-physics. The third comparison will compare two different methods for generating IC perturbations in a GSI-based hybrid-EnVar system. Additional details are provided in section 2a, CLUE Experiments 1, 2, and 6, respectively.

(5) Hail Guidance

Maximum hail size fields will be formally evaluated within the HRRRE. These hail size forecasts will include those derived by (1) the HAILCAST algorithm (Adams-Selin and Ziegler 2016), which predicts maximum hail size using a hail growth model coupled to WRF, (2) the Thompson method, which estimates hail size directly from the microphysics size distribution by finding the largest graupel or hail hydrometeor diameter that exceeds a specified number concentration, (3) neighborhood-based, probabilistic forecasts of UH exceeding a fixed threshold loosely calibrated to maximize the fractions skill score (FSS) for 1-inch hail, and (4) a machine-learning-based method that provides probabilistic hail size forecasts (Gagne et al. 2017). Comparisons will be made to hail LSRs and MRMS MESH.

(6) Texas Tech University Sensitivity-Based Ensemble Subsetting Evaluation

An ensemble-subsetting method developed by researchers at Texas Tech University (TTU) will be tested during SFE 2019, in which ensemble sensitivity to desired fields will be used to eliminate ensemble members that are performing poorly in sensitive areas. Probabilities from the full ensemble and the subset of ensemble members will then be compared to determine the effect of the subsetting method. In SFE 2018, the ensemble subsetting activity used a CAM ensemble run at TTU, however, for SFE 2019 most of the CLUE system will be used for subsetting. Further details on the ensemble subsetting are provided in Appendix E.

c. Other specialized activities

Blog posts will be made 2-3 times weekly during the experiment, as in the past two years. These blog posts will highlight interesting case studies, preliminary results from evaluations, and feature activities and discussions taking place during the SFEs. The blog is hoped to supplement the formal results produced during the 2019 SFE by providing a more informal look into the questions and discussions that take place within the framework of the formal forecasting and evaluation activities. The blog can be found at: https://blog.nssl.noaa.gov/efp/.

Appendix A: List of scheduled SFE 2019 participants (names in bold at the bottom denote evening forecasters for the 4-8pm WoF activity)

Week 1	Week 2	Week 3	Week 4	Week 5
April 29-May 3	**May 6-10**	**May 13-17**	**May 20-24**	**May 28-31**
Brian Ancell (TTU)	Brian Ancell (TTU)	Austin Coleman (TTU)	Michael Brennan (NHC)	Ben Blake (EMC)
Austin Coleman (TTU)	Austin Coleman (TTU)	David Gagne (NCAR) M-Th	Clark Evans (UWM)	Jamie Wolff (DTC)
Willy Sedlacek (USAF)	Geoff Manikin (EMC)	Lance Bosart (SUNYA)	Jason Otkin (UW/CIMSS)	Curtis Alexander (GSD)
Shawn Corvec (ECCC)	Lara Pagano (WPC)	Tyler Leicht (SUNYA)	Greg Thompson (NCAR)	David Walters (UKmet)
Tracey Dorian (EMC)	Lindsay Blank (DTC)	Alex Mitchell (SUNYA)	Bill Gallus (ISU)	Steve Willington (UKmet)
Eric Aligo (EMC)	Glen Romine (NCAR)	Logan Dawson (EMC)	Zach Hiris (ISU)	Gordon Brooks (USAF)
Terra Ladwig (GSD)	Trevor Alcott (GSD)	Alicia Bentley (EMC)	Jacob Carley (EMC)	Amanda Burke (OU)
Christina Kalb (DTC)	Shin-Ping Kuan (CWB)	Ryan Sobash (NCAR)	Craig Schwartz (NCAR; M-W)	David Imy (SPC Ret.)
Eric Loken (OU)	Ping-Hsiang Wang (CWB)	John Brown (GSD)	Kai-Yuan Cheng (GFDL)	John Boris (WFO APX)
David Jahn (SPC)	Jeff Duda (GSD)	Shin-Ping Kuan (CWB)	Ed Szoke (GSD)	Jeff Milne (SPC)
David Harrison (SPC)	Victor Gensini (NIU)	Ping-Hsiang Wang (CWB)	Jon Petch (UKmet)	Arianna Jordan (NERTO/Howard Univ.)
Patrick Gilchrist (WFO GGW)	John Allen (CMU)	Aurore Porson (UKmet)	Paul Davies (UKmet)	Andy Bollenbacher (WFO HNX)
Joseph Clark (WFO DTX)	Rachel North (UKmet)	Andy Hartley (UKmet)	Steve Willington (UKmet)	Brandt Maxwell (WFO SGX)
Jimmy Correia (NWS AFS)	James Varndell (UKmet)	David Hayter (UKmet)	Neil Armstrong (UKmet)	Michael Hill (WFO LIX)
Seongmook Kim (CAPS - KMA)	David Hayter (UKmet)	Katie Deroche (AWC)	Arianna Jordan (NERTO/Howard Univ.)	Dan Hofmann (WFO LWX)
	Becky Adams-Selin (AER)	David Stark (WFO OKX)	Don Van Dyke (WFO TAE)	David Dowell (GSD)
	Tom Hultquist (WFO MPX)	Brett Albright (WFO OAX)	Sarah Trojniak (WPC)	
	Andy Wilkins (OU)	Daniel Zumpfe (WFO MSO)	Harald Richter (BoM)	
	Tom Galarneau (CIMMS/OU)		Anders Jensen (NCAR)	
	John Henderson (AER; W-F)		Rob Hepper (AWC)	
	TBD (Edmonton MSC) T		*Rich Fulton (OWAQ) M-T*	*Chris Stammers (Winnipeg MSC) T*
	Chad Entremont (WFO JAN) T-Th		*Brian Oswiak (Toronto MSC) T*	
David Cox - WoF (WFO JAN)	**Brittany Newman - WoF (WFO GLD)**	**Andrew Moore - WoF (WFO FGF)**	**Larry Hopper - WoF (WFO PSR)**	**Brad McGavock - WoF (WFO TSA)**
Joseph Cebulko - WoF (WFO ALY)	**Suzanna Lindeman - WoF (WFO BOI)**	**Michael Hollan - WoF (WFO BIS)**	**Christina Leach - WoF (WFO JKL)**	**Aaron Mangels - WoF (WFO GID)**

SFE Facilitators: Adam Clark (NSSL), Israel Jirak (SPC), Steve Weiss (retired SPC), Burkely Gallo (CIMMS/SPC/NSSL), Kenzie Krojac (CIMMS/NSSL/OU), Brett Roberts (CIMMS/SPC/NSSL), Kimberly Hoogewind (CIMMS/SPC/NSSL), Kent Knopfmeier (CIMMS/NSSL), and Andy Dean (SPC).

WoF evening activity facilitators: Pam Heinselman (NSSL), Kimberly Hoogewind (CIMMS/SPC/NSSL), Patrick Skinner (CIMMS/NSSL), Katie Wilson (CIMMS/NSSL), Jessie Choate (CIMMS/NSSL), and Corey Potvin (CIMMS/NSSL).

Appendix B: 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. C1). 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),

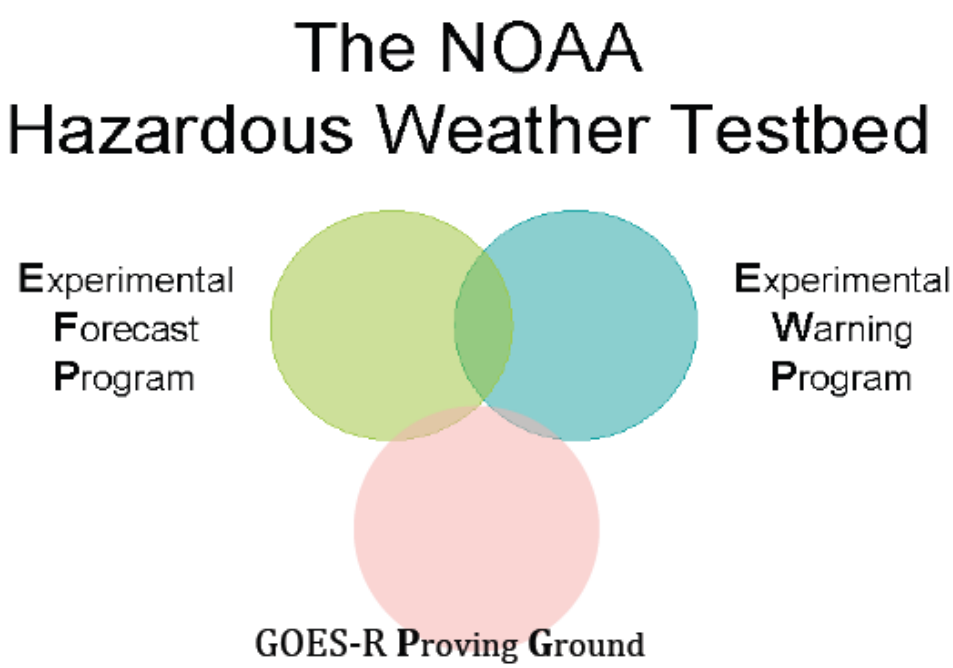

Figure C1: The umbrella of the NOAA Hazardous Weather Testbed (HWT) encompasses two program areas: The Experimental Forecast Program (EFP), the Experimental Warning Program (EWP), and the GOES-R Proving Ground (GOES-R).

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.

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), Clark et al. (2012; 2018), and Gallo et al. (2017).

Appendix C: Details on Object-based Probabilistic Forecasts (OBPROB)

To identify objects, a 2*dx (6 km) Gaussian convolution is first applied to composite reflectivity forecasts to remove grid scale noise, and a threshold of 35 dBZ is applied to the convolved field to identify a discrete set of storm objects. For each object, the attributes defined in Table D1 are calculated if the area of the object is at least 7*7=49 grid points (i.e. above the model's effective resolution of ~7*dx).

Table C1 Description of object attributes.

Attribute Name	Description
Centroid Location	Center of "mass" of the binary field of the object.
Area	Number of grid points contained in the object.
Aspect Ratio	W/L, where L=length of longest axis, W=length of axis perpendicular to L.
UH	90^{th} percentile of updraft helicity values within the object
10mspd	As in UH, except for 10m wind speed values
HAIL	As in UH, except for hail size values
PRECIP	As in UH, except for hourly accumulated precipitation

For object matching, A Total Interest (I) value is calculated to determine the overall similarity of any pair of objects (e.g., objects from different ensemble members). For simplicity, and ease of physical interpretation, we here use,

$$I = f_{a_1} * f_{a_2} * f_{a_3},$$

where is the similarity function in terms of each of Centroid Location difference, Area ratio, and Aspect Ratio difference. These attributes quantify the similarity of the storm objects' location, size and shape according to the interest functions in Figure D1.

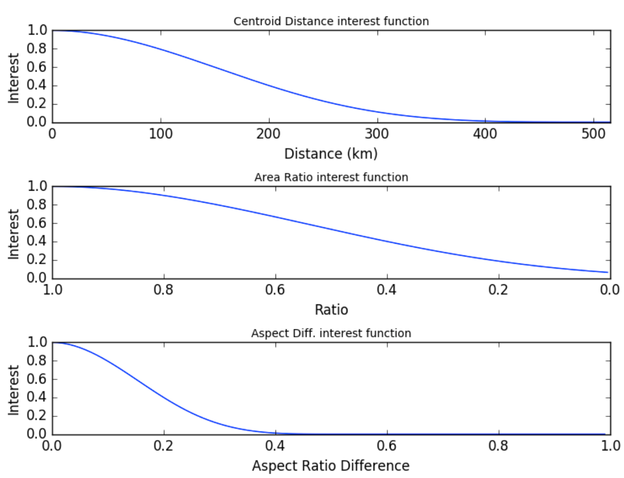

Figure C1 Interest functions to define similarity (i.e., interest) of the centroid location, area and aspect ratio attributes between two objects.

For computation of object probabilities, the probability that each forecast object will be matched by a similar observed object is taken to be the percentage of ensemble members with a matching object. In order to include the possibilities from the entire ensemble "envelope", the OBPROB product is constructed as follows:

<u>Step 1</u>: Calculate probability for every object from every ensemble member.
<u>Step 2</u>: Sort the objects based on probability.
<u>Step 3</u>: Add the highest probability object to the plot, shaded by probability.
<u>Step 4</u>: Remove this object, and all matching objects that contributed to its probability, from the list.
<u>Step 5</u>: Repeat from step 1 until the list is empty.

The resulting OBPROB product is conceptually similar to a paintball plot since every storm object is included. The difference is that many similar storms from different ensemble members can be represented by a single high-probability object. This makes the plot less busy, and easier to interpret. As an example, the ensemble forecast in Figure C2 could be interpreted to indicate a much higher probability that the storm(s) in Oklahoma will still have a cellular mode at this time, although there is a slight chance of upscale growth having occurred. Similarly, while both linear and cellular modes are likely in central Kansas into southern Nebraska, the linear systems have greater ensemble agreement. This information could be gathered qualitatively from the ensemble by evaluating each member individually or quantitatively from the OBPROB product (Fig. C2k).

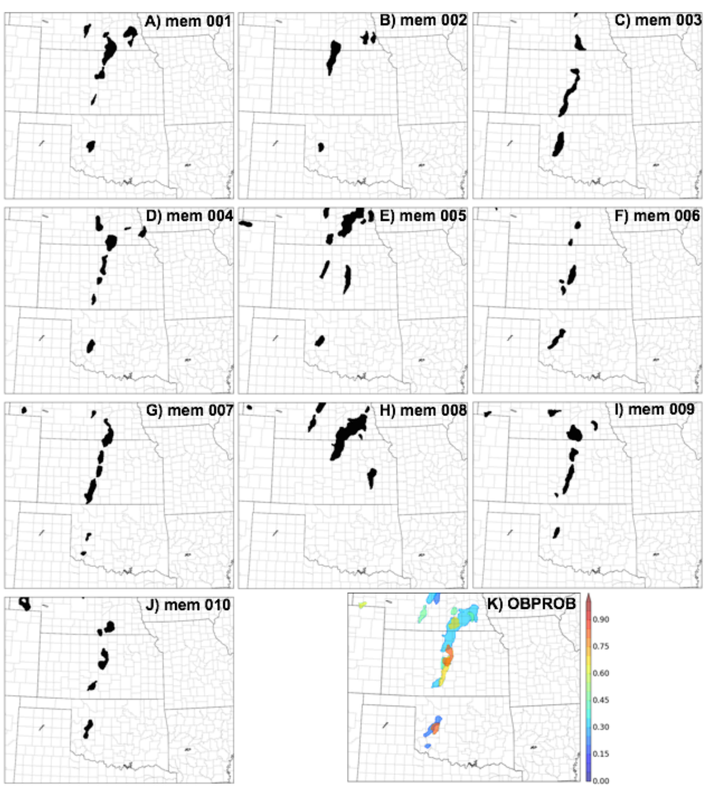

Figure C2 Objects forecast by the (a)-(j) each ensemble member, and (k) the OBPROB product. Forecast was initialized at 0000 UTC 16 May 2017, and valid at 0100 UTC 17 May 2017.

The OBPROB products for the OUMAP ensemble can be viewed at http://weather.ou.edu/~map/dev/interface_2019.php. The OBPROB forecasts can be viewed in terms of storm mode (i.e., including all objects) or in terms of objects producing strong rotation, strong surface winds, large hail or heavy precipitation, respectively, by only including objects meeting a minimum attribute (Table C1) threshold of 65 m^2s^{-2}, 20 ms^{-1}, 2.54 mm, or 12.7 mm h^{-1}, respectively.

An example of the OBPROB web interface to the OU MAP ensemble forecasts is shown in Figure C3. The two figures shown are the (left) uncalibrated and (right) calibrated OBPROB forecast. Each object is labelled with black text with an arbitrary number. Clicking on the number at the bottom of the page corresponding to an object brings up a figure showing only that object, together the individual ensemble member objects that were to matched to this object (e.g., Fig. C4). This functionality was suggested by participants during the OBPROB evaluations in the 2018 SFE, and allows the user to dig deeper into the interpretation of the OBPROB forecast.

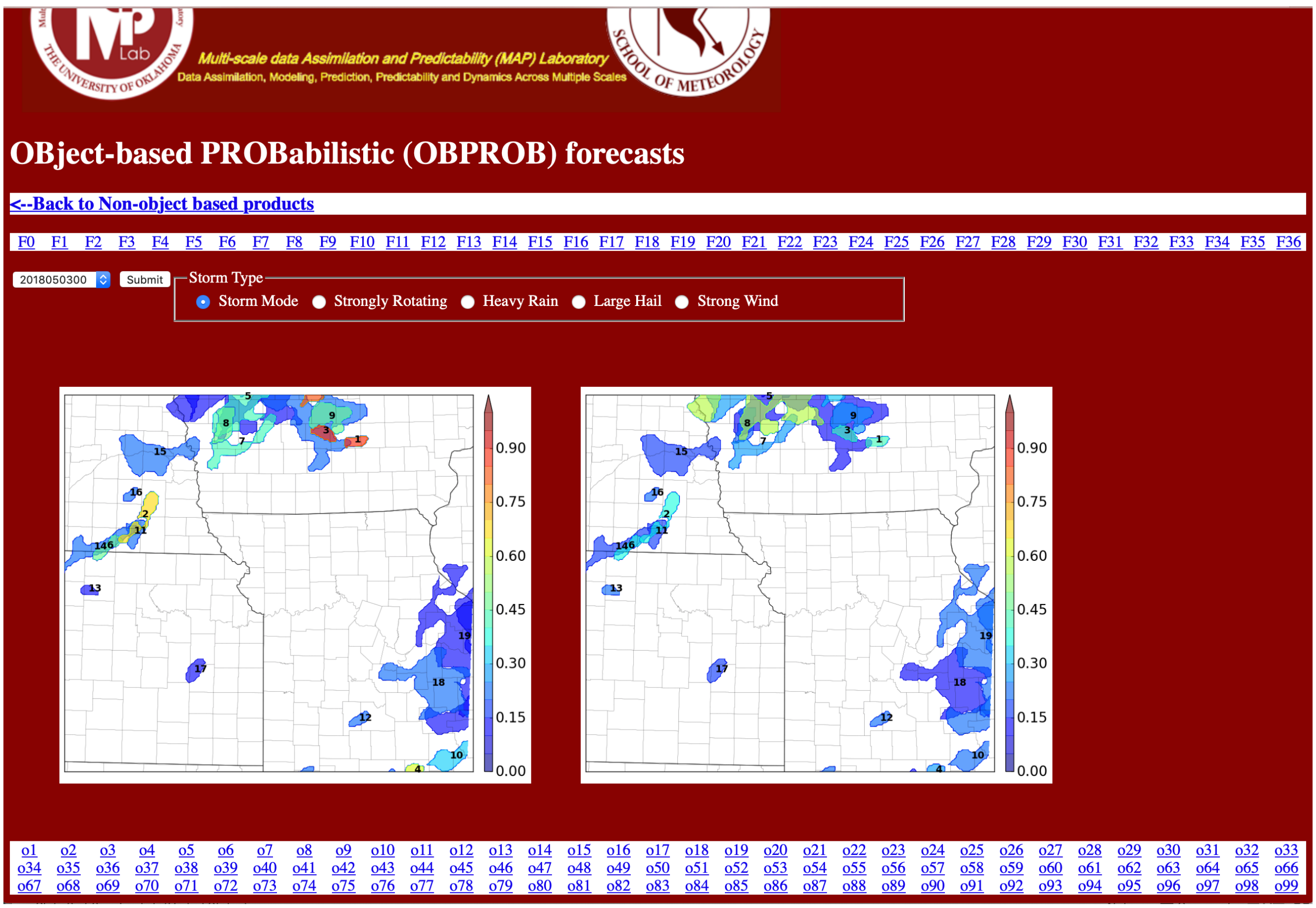

Figure C3 Example screenshot of the OBPROB web-interface for the SFE 2019.

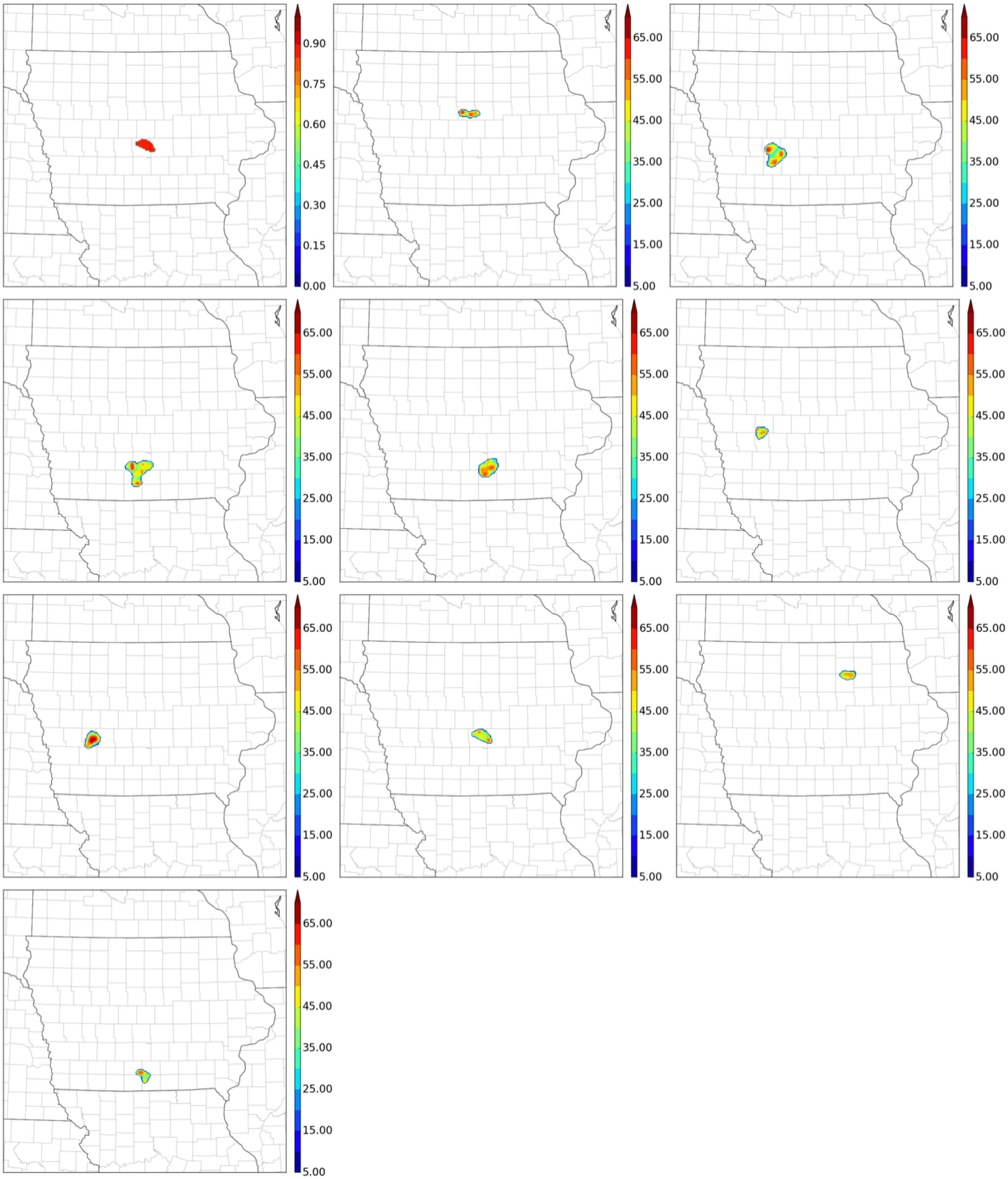

Figure C4 Example pop-up window showing further information about the ensemble of objects matched to the high-probability object number "3" from Figure D3.

An object-based evaluation of the 2018 OU MAP lab ensemble forecasts revealed large biases in some of the OBPROB probabilities. Therefore, a calibrated product is also provided this year. The calibration uses logistic regression, trained on the 2018 ensemble forecasts. Logistic Regression consists of fitting the following equation to a training set of predictors x_j and observed events y (taking values of 0 or 1 in the training data, predicted by the fitted equation to be in the range 0-1), where there are J+1 co-efficients fitted to the model using J predictors:

$$y = \left[1 + exp\left(-(\beta_0 + \sum_{j=1}^{J} \beta_j x_j)\right)\right]^{-1}$$

In the case of OBPROB calibration the predicted event, y, is whether the forecast object will be matched by an observation object. The predictors include the number of ensemble members with a matching object, the area of the object, the 10m wind speed attribute, the accumulated precipitation attribute, the hail size attribute, the updraft helicity attribute, object aspect ratio, centroid latitude and centroid longitude.

Appendix D: TTU Sensitivity-Based Ensemble Subsetting within the CLUE

Ensemble sensitivity is a statistical technique applied within an ensemble that identifies features in the flow at early forecast times that are related to the predictability of chosen severe storm characteristics later in the forecast. In other words, ensemble sensitivity reveals the flow features for which associated errors will grow rapidly to adversely affect the predictive skill of chosen severe storm aspects. It can thus be expected that ensemble members that have the least error in the most sensitive regions early in a forecast window will provide better forecasts than other members, allowing the generation of adjusted and improved probabilities well before the next extended forecast cycle. *The goal of this HWT 2019 activity is to evaluate ensemble sensitivity-based subsets within the CLUE to understand whether this technique can both improve probabilities over that from the full CLUE ensemble in a real-time environment*. The planned activity follows an evaluation at SFE 2018 of ensemble subsets only within the TTU ensemble, which suggested the majority of forecasts could be improved with the technique.

A daily evaluation of probabilities from the full CLUE against those based on 20-member CLUE subsets chosen objectively through the use of ensemble sensitivity information will be conducted. Each day, a response function location and time will be chosen through a web-based graphical user interface that identifies areas of Day 1 severe convection. The Day 1 response function will be chosen over a 6-hr period between 1800 UTC and 1200 UTC (next day) in areas where better predictions of severe convection are desired (e.g. areas of high uncertainty). Once the response time and location are chosen, the sensitivity of two independent response functions will be calculated: 1) number of grid points exceeding 50 m^2/s^2 2–5km updraft helicity, and 2) number of grid points exceeding 40 dBZ near-surface simulated reflectivity. These sensitivities will be generated completely within the TTU 42-member ensemble. The sensitivities of the two response functions will be calculated with respect to 300- and 500-hPa temperature, winds, and geopotential height, and 700-hPa temperature all with respect to the 6-hr forecast (valid 0600 UTC). CLUE members (interpolated to the TTU modeling grid) from the 0000 UTC run will then be chosen objectively based on their errors in the most sensitive regions (greatest 50% of sensitivity magnitudes) using the 0600 UTC analysis from either the 0600 UTC TTU ensemble initialization or the Rapid Refresh system.

Probability fields (specifically exceedance probabilities of updraft helicity) of Day 1 convection will be generated for the CLUE subset and will be compared against the full CLUE the following day after the severe event has occurred. SPC storm reports and the associated practically perfect probability field will serve as the observations against which both the full and subset CLUE probabilities are evaluated. The TTU 42-member ensemble system within which the sensitivities are generated is a DART WRF ensemble Kalman filter that assimilates numerous surface and upper-air observations on a 6-hr assimilation cycle. Assimilation is performed over a 12-km CONUS domain with downscaled 48-hr WRF forecasts run twice daily on a 4-km domain across the U.S. Midwest and South Plains. Real-time output from the Texas Tech ensemble can be viewed at http://www.atmo.ttu.edu/bancell/real_time_ENS/ttuenshome.php.

Appendix E: References

Adams-Selin, R.D. and C.L. Ziegler, 2016: Forecasting Hail Using a One-Dimensional Hail Growth Model within WRF. *Mon. Wea. Rev.,* **144**, 4919–4939,https://doi.org/10.1175/MWR-D-16-0027.1

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

Chen, J.-H., and S.-J. Lin, 2013: Seasonal predictions of tropical cyclones using a 25-km-resolution general circulation model. *J. of Climate*, **26:2**, 380-398.

Chen, J-H., and Coauthors, 2019: Advancements in Hurricane Prediction with NOAA's Next Generation Forecast System. Geophysical Research Letters. DOI:10.1029/2019GL082410.

Clark, A. J., and Coauthors, 2012: 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.

Clark, A.J. and Coauthors, 2018: The Community Leveraged Unified Ensemble (CLUE) in the 2016 NOAA/Hazardous Weather Testbed Spring Forecasting Experiment. *Bull. Amer. Meteor. Soc.,* **99**, https://doi.org/10.1175/BAMS-D-16-0309.1

Gagne, D. J., A. McGovern, S. E. Haupt, R. A. Sobash, J. K. Williams, and M. Xue, 2017: Storm-Based Probabilistic Hail Forecasting with Machine Learning Applied to Convection-Allowing Ensembles. *Wea. Forecasting*, **32**, 1819-1840.

Gallo, B. T., A. J. Clark, and S. R. Dembek, 2016: Forecasting Tornadoes Using Convection-Permitting Ensemble. *Wea. Forecasting*, **31**, 273-295.

Gallo, B.T., and Coauthors, 2017: Breaking New Ground in Severe Weather Prediction: The 2015 NOAA/Hazardous Weather Testbed Spring Forecasting Experiment. *Wea. Forecasting,* **32**, 1541–1568, https://doi.org/10.1175/WAF-D-16-0178.1

Gallo, B. T., A. J. Clark, B.T. Smith, R.L. Thompson, I. Jirak, and S.R. Dembek, 2018: Blended Probabilistic Tornado Forecasts: Combining Climatological Frequencies with NSSL–WRF Ensemble Forecasts. *Wea. Forecasting,* **33**, 443–460, https://doi.org/10.1175/WAF-D-17-0132.1.

Gallo, B. T., A. J. Clark, B. T. Smith, R. L. Thompson, I. L. Jirak, and S. R. Dembek, 2019: Incorporating UH occurrence time to ensemble-derived tornado probabilities. *Wea. Forecasting*, **34**, 151-164.

Han, J., Wang, W., Kwon, Y.C., Hong, S.Y., Tallapragada, V. and Yang, F., 2017: Updates in the NCEP GFS cumulus convection schemes with scale and aerosol awareness. *Wea. Forecasting*, **32**, 2005-2017.

Harris, L. M., and S.-J. Lin, 2013: A two-way nested global-regional dynamical core on the cubed-sphere grid. *Mon. Wea. Rev*., **141**, 283-306.

Harris, L.M., S.-J. Lin, and C.-Y. Tu, 2016: High-Resolution Climate Simulations Using GFDL HiRAM with a Stretched Global Grid. *Journal of Climate*, **29**, 4293-4314.

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

Hong, S.Y., Noh, Y. and Dudhia, J., 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.,* **34**, 2318-2341.

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.

Johnson, A., X. Wang, J.R. Carley, L.J. Wicker, and C. Karstens, 2015: A Comparison of Multiscale GSI-Based EnKF and 3DVar Data Assimilation Using Radar and Conventional Observations for Midlatitude Convective-Scale Precipitation Forecasts. *Mon. Wea. Rev.,* **143**, 3087–3108, https://doi.org/10.1175/MWR-D-14-00345.1

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.

Lin, S.-J., and R.B. Rood, 1996: Multidimensional Flux-Form Semi-Lagrangian Transport Schemes. *Mon. Wea. Rev.*, **124**, 2046-2070.

Lin, S.-J., and R.B. Rood, 1997: An explicit flux-form semi-lagrangian shallow-water model on the sphere. *QJRMS*, **123**, 2477-2498.

Lin, S.-J., 1997: A finite-volume integration method for computing pressure gradient force in general vertical coordinates. *QJRMS*, **123**, 1749-1762.

Lin, S.-J., 2004: A "vertically Lagrangian" finite-volume dynamical core for global models. *Mon. Wea. Rev.*, **132**, 2293-2307.

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.

Palmer, T.N., R. Buizza, F. DoblasReyes, T. Jung, M. Leutbecher, G. Shutts, M. Steinheimer, and A. Weisheimer, 2009: Stochastic parametrization and model uncertainty. ECMWF Tech. Memo. 598, 42 pp. [Available online at http://www.ecmwf.int/publications/.]

Putman, W M., and S.-J. Lin, 2007: Finite-volume transport on various cubed-sphere grids. *Journal of Computational Physics*, **227**, 55-78.

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.

Sobash, R. A., C. S. Schwartz, G. S. Romine, K. R. Fossess, and M. L. Weisman, 2016a: Severe Weather Prediction Using Storm Surrogates from an Ensemble Forecasting System. *Wea. Forecasting*, **31**, 255-271.

Sobash, R. A., G. S. Romine, C. S. Schwartz, D. J. Gagne, and M. L. Weisman, 2016b: Explicit Forecasts of Low-Level Rotation from Convection-Allowing Models for Next-Day Tornado Prediction. *Wea. Forecasting*, **31**, 1591-1614.

Sobash, R. A. and J. S. Kain, 2017: Seasonal variations in severe weather forecast skill in an experimental convection-allowing model. *Wea. Forecasting*, **32**, 1885-1902.

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

Thompson, G. and T. Eidhammer, 2014: A Study of Aerosol Impacts on Clouds and Precipitation Development in a Large Winter Cyclone. *J. Atmos. Sci.*, **71**, 3636–3658, https://doi.org/10.1175/JAS-D-13-0305.1

Wang, Y. and X. Wang, 2017: Direct Assimilation of Radar Reflectivity without Tangent Linear and Adjoint of the Nonlinear Observation Operator in the GSI-Based EnVar System: Methodology and Experiment with the 8 May 2003 Oklahoma City Tornadic Supercell. *Mon. Wea. Rev.,* **145**, 1447–1471, https://doi.org/10.1175/MWR-D-16-0231.1

Wang Y., X. Wang, and J. Carley, 2018a: GSI-based EnKF-Variational Hybrid Data Assimilation for NCEP NAMRR: System Development and Initial Testing. *Mon. Wea. Rev.* to be submitted.

Wang Y., X. Wang, and D. Dowell, 2018b: GSI-based EnKF-Variational Hybrid Data Assimilation for HRRR. *Mon. Wea. Rev. in preparation*.

Zhou, L., S.-J. Lin, J.-H. Chen, L. Harris, X. Chen, and S. L Rees, 2019: Toward Convective-Scale Prediction within the Next Generation Global Prediction System. *Bull. Amer. Meteor. Soc.* DOI:10.1175/BAMS-D-17-0246.1.