SPRING FORECASTING EXPERIMENT 2014

Conducted by the

EXPERIMENTAL FORECAST PROGRAM

of the

NOAA/HAZARDOUS WEATHER TESTBED

http://hwt.nssl.noaa.gov/Spring_2014/

HWT Facility – National Weather Center

5 May - 6 June 2014

Program Overview and

Operations Plan

Mike Coniglio[2], Israel Jirak[1], Adam Clark[2,3], James Correia[1,3], Steven J. Weiss[1], Chris Melick[1,3], Patrick Marsh[1], Andy Dean[1], Kent Knopfmeier[2,3], David Imy[1]

(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

1. Introduction

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

The 2014 Spring Forecasting Experiment (SFE2014), a cornerstone of the EFP, will be conducted 5 May – 6 June. Building upon successful experiments of previous years, a main emphasis of SFE2014 will be the generation of probabilistic forecasts of severe weather valid over shorter time periods than current 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, beginning days in advance and continuing into the event occurrence. As in previous experiments, a suite of new and improved experimental mesoscale and CAM guidance will be central to the generation of these forecasts.

This operations plan summarizes the core interests of SFE2014 and provides information on the operations of the experiment. Detailed information on the organizational structure of the HWT and information on various forecast tools and diagnostics can also be found in this document. The remainder of the operations plan is organized as follows: Section 2 provides details on a number of new products and models being introduced during SFE2014 and Section 3 describes the core interests and new concepts being introduced for SFE2014. 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 SFE2014 forecasting activities will be to test two methods of generating probabilistic forecasts of severe weather that are valid over shorter time windows than current SPC operational products. As in SFE2013, severe weather forecasts over three-hour time windows will be produced, but this year we will be producing probabilities of individual hazards (wind, hail, tornadoes) for these three-hourly forecasts. Three separate forecasts will be valid over three-hour time windows between 1800 UTC and 0300 UTC (1800-2100, 2100-0000, and 0000-0300 UTC). As an alternative method to generate outlooks with high temporal specificity, a different group of SFE2014 participants will be producing combined forecasts of any type of severe wind, severe hail, and tornadoes (“total severe”) *valid over hourly time windows* over the 1800 – 0300 UTC period. The goal of testing these two different methods is to explore different ways of introducing probabilistic severe weather forecasts on scales that are currently addressed operationally with categorical forecast products (mesoscale discussions and convective watches). These experimental forecasts begin to explore ways of seamlessly merging probabilistic severe weather outlooks with probabilistic severe weather warnings as part of the NOAA Warn-on-Forecast initiative (Stensrud et al. 2009).

Another new addition to the suite of forecast products in SFE2014 will be the generation of separate probabilistic forecasts of wind, hail, and tornadoes that cover the Day 2 period (1200 UTC on Day 2 to 1200 UTC on Day 3). Only probabilistic forecasts of combined wind, hail and tornadoes (total severe) are currently produced in SPC operational forecasts that cover this period.

Generating the forecasts described above will be intensive and will thus rely on deterministic and ensemble CAM output for guidance and to generate first guesses for the severe weather probabilities. Most of the CAMs will be based on recent versions of the Advanced Research Weather Research and Forecasting (WRF-ARW) model. Included in the suite of CAM guidance will be a 20-member ensemble produced by the University of Oklahoma Center for Analysis and Prediction of Storms (CAPS), a ten-member ensemble produced by the Air Force Weather Agency (AFWA), and a newly developed nine-member WRF-ARW ensemble produced by NSSL. Additionally, the United Kingdom Meteorological Office (Met Office) will provide four CAM forecasts (two with 2.2 km grid spacing and two with 4.4 km grid spacing) over the CONUS that are based on a new version of their Unified Modeling System, two of which test a new "scale aware" parameterization scheme for turbulent mixing. Finally, the deterministic CAM output available to SPC operationally (or near operationally) generated by NSSL and the National Centers for Environmental Prediction (NCEP) Environmental Modeling Center (EMC), including the High-Resolution Rapid Refresh (HRRR) WRF-ARW based model, will be combined to produce a seven-member Storm Scale Ensemble of Opportunity (SSEO).

For the generation of first-guess forecasts from the CAM ensembles, it is important to extract variables in the forecasts that track proxies for 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 at the lowest model level can be effective for highlighting occurrences of severe weather in the model (Sobash et al. 2010, Kain et al. 2010, Clark et al. 2013). To support the SFE2014 goal of generating forecasts of individual hazards, there will be an effort to explore the ability of new model fields to delineate individual hazards, particularly for the size of hail and threat for tornadoes. While it is expected that individual hail categories in microphysical schemes will be a source for providing this information in the coming years, the SFE2014 will focus on other techniques to extract information regarding the likelihood of hail generation from model output, including the HAILCAST model (Adams-Selin 2012) implemented in WRF-ARW. Separate techniques for extracting the probability of hail based on the calibration of storm attributes combined with environment information and a machine-learning algorithm will also be explored.

Finally, a cornerstone of HWT Spring Forecasting Experiments has been the systematic testing of schemes in WRF-ARW that parameterize sub-grid-scale processes, primarily in the way microphysical schemes depict convective processes and the performance of turbulent-mixing schemes in the planetary boundary layer in moist convective boundary layers. This theme will continue in the SFE2014 as new parameterization schemes will be tested including double-moment microphysics schemes developed by Hugh Morrison, and Jason Milbrandt. Furthermore, we will evaluate the Mellor-Yamada-Nakinishi-Niino (MYNN) turbulent mixing scheme [the scheme now implemented in the HRRR and parent Rapid Refresh (RAP) model], and the effect of smaller grid spacing on three-dimensional Smagorinsky-type mixing will be evaluated using the Met Office CAM forecasts.

a) CAPS Storm Scale Ensemble Forecast (SSEF) Systems

As in previous years, CAPS will provide a 0000 UTC and 1200 UTC-initialized 4-km grid-spacing Storm Scale Ensemble Forecast (SSEF) system. This year's 0000 UTC SSEF system has 20 members and is run at the University of Tennessee National Institute for Computational Sciences (NICS). New this year will be SSEF forecasts that extend to 60 hours to support the Day 2 forecasts. The members of the ensemble that will be evaluated in the SFE2014 use the WRF-ARW core and all forecasts use 51 vertical levels. WRF code was modified by CAPS to allow initial hydrometeor fields generated from a 3DVAR/ARPS Cloud analysis of WSR-88D radar reflectivity to pass into WRF initial conditions, and to write out a reflectivity field every 5 min. As in previous years, the 0000 UTC NAM analyses available on the 12 km grid (AWIPS 218) are used for initialization of the control and non-perturbed members, and as a first guess for initialization of perturbed members with the initial condition perturbations coming directly from the operational SREF system. WSR-88D data, along with available surface and upper air observations, are analyzed using ARPS 3DVAR/Cloud-analysis system. Forecast output at hourly intervals (with higher time frequency output for a limited selection of 2D fields) are archived at the NICS mass storage system (HPSS). Specifications for the members to be evaluated in the SFE2014 are provided in Table 1.

The basic strategy in constructing the SSEF system is to have a set of members accounting for as many error sources as possible that can be used to generate reliable forecast probabilities (non-shaded members in Table 1). These "core" members have IC/LBC perturbations as well as varied physics and model cores. Other sets of members were configured to allow for various sensitivity experiments (shaded members in Table 1). Three members are configured identically except for their microphysics parameterizations and two members are configured identically except for their turbulent-mixing (PBL) parameterizations.

Similar to last year, a SSEF system initialized at 1200 UTC, and run at the Oklahoma Super Computing and Education Center (OSCER), will be available for use in the forecasting activities. The basic strategy in constructing the 1200 UTC SSEF system is to have a subset of members that have the same configuration as in the 0000 UTC ensemble. Resources for running the 1200 UTC members in real time are more limited than those for the 0000 UTC ensemble, so only 8 members from the 0000 UTC ensemble are run at 1200 UTC (identified by members with an asterisk in Table 1). This will allow for a direct comparison of the change in skill between this set of eight members for forecasts initialized 12 hours apart. Furthermore, the selection of eight members was chosen to be similar to the number of members in the SSEO (see below) so that a comparison of the spread and skill characteristics of these sets of forecasts can be compared fairly.

Table 1. Configurations for the legacy SSEF WRF-ARW members. NAMa and NAMf refer to 12 km NAM analysis and forecast, respectively. ARPSa refers to ARPS 3DVAR and cloud analysis.

Member	IC	BC	Radar data	Microphy	LSM	PBL
*arw_cn	00Z ARPSa	00Z NAMf	yes	Thompson	Noah	MYJ

arw_c0	00Z ARPSa	00Z NAMf	no	Thompson	Noah	MYJ
arw_m3	arw_cn + em-p1_pert	21Z SREF em-p1	yes	Morrison	RUC	YSU
*arw_m4	arw_cn + em-n2_pert	21Z SREF em-n2	yes	Thompson	Noah	QNSE
*arw_m5	arw_cn + nmm-p1_pert	21Z SREF nmm-p1	yes	WDM6	Noah	MYNN
*arw_m6	arw_cn + nmmb-n1_pert	21Z SREF nmmb-n1	yes	WDM6	RUC	MYJ
arw_m7	arw_cn – nmmb-p1_pert	21Z SREF nmmb-p1	yes	MY2	Noah	YSU
*arw_m8	arw_cn + em-n1_pert	21Z SREF em-n1	yes	WDM6	Noah	QNSE
arw_m9	arw_cn – em-p2_pert	21Z SREF em-p2	yes	MY2	Noah	MYNN
*arw_m10	arw_cn – nmmb-n3_pert	21Z SREF nmmb-n3	yes	WDM6	Noah	YSU
*arw_m11	arw_cn – nmmb-p3_pert	21Z SREF nmmb-p3	yes	Thompson	Noah	YSU
*arw_m12	arw_cn – em-p3_pert	21Z SREF em-p3	yes	Thompson	Noah	MYNN
arw_m13	arw_cn – nmm-p2_pert	21Z SREF nmm-p2	yes	Morrison	Noah	QNSE
arw_m14	00Z ARPSa	00Z NAMf	yes	Thompson	Noah	MYNN
arw_m15	00Z ARPSa	00Z NAMf	yes	Thompson	Noah	YSU-T

arw_m16	00Z ARPSa	00Z NAMf	yes	Thompson	Noah	YSU
arw_m17	00Z ARPSa	00Z NAMf	yes	MY2	Noah	MYJ
arw_m18	00Z ARPSa	00Z NAMf	yes	MY	Noah	MYJ
arw_m19	00Z ARPSa	00Z NAMf	yes	Milbrant-Morrison	Noah	MYJ
arw_m20	00Z ARPSa	00Z NAMf	yes	Morrison	Noah	MYJ

Note 1: For all members: *ra_lw_physics*= RRTMG; *ra_sw_physics*=RRTMG; *cu_physics*=none

b) NSSL WRF

SPC forecasters have used output from an experimental 4 km WRF-ARW produced by NSSL since the fall of 2006. Currently this WRF model is run twice daily at 00 UTC and 12 UTC throughout the year over a full CONUS domain with forecasts to 36 hours. Output is also available online at http://www.nssl.noaa.gov/wrf/ and http://www.nssl.noaa.gov/wrf/12Z/.

New to the experimental numerical guidance for this year's experiment is the addition of eight 4 km WRF-ARW runs that – along with the regular NSSL-WRF – will comprise a nine member NSSL-WRF based ensemble. The additional eight members are initialized at 0000 UTC and use 3 h SREF forecasts initialized at 2100 UTC for initial conditions and corresponding SREF member forecasts as lateral boundary conditions. The physics parameterizations for each member are identical to the regular NSSL-WRF. Although the unvaried physics will have lower spread than a varied physics ensemble, SPC forecaster are much more familiar with the behavior of the NSSL-WRF physics and this will allow us to isolate the contribution of spread from ICs/LBCs.

c) SPC Storm Scale Ensemble of Opportunity (SSEO)

The SSEO is a 7-member, convection-allowing ensemble consisting of deterministic models available operationally to SPC. This "poor man's ensemble" provides a practical alternative to a formal/operational storm-scale ensemble which will not be available in the next few years because of computational limitations. Similar to the other SSEF systems, hourly maximum storm-attribute fields, such as simulated reflectivity, updraft helicity, and 10-m wind speed are produced from the SSEO. Member specifications are provided in Table 2. Members marked with "-12h" in the Model column are 12h time-lagged members, initialized 12h earlier than the other members. All members are initialized with a "cold-start" from the operational NAM – i.e., no radar data assimilation or cloud model is used to produce ICs.

Table 2. SSEO member specifications.

Member #	Model	Grid-spacing	Agency	PBL	Microphysics	LSM
sseo01	WRF-ARW	4-km	NSSL	MYJ	WSM6	Noah
sseo02	Hi-Res Window WRF-ARW	5.15-km	NCEP/EMC	YSU	WSM3	Noah
sseo03	Hi-Res Window WRF-ARW -12h	5.15-km	NCEP/EMC	YSU	WSM3	Noah
sseo04	CONUS WRF-NMM	4-km	NCEP/EMC	MYJ	Ferrier	Noah
sseo05	Hi-Res Window WRF-NMM	4-km	NCEP/EMC	MYJ	Ferrier	Noah
sseo06	Hi-Res Window WRF-NMM -12h	4-km	NCEP/EMC	MYJ	Ferrier	Noah
sseo07	NMMB Nest	4-km	NCEP/EMC	MYJ	Ferrier+	Noah

d) Air Force Weather Agency (AFWA) 4-km ensemble

The U.S. Air Force Weather Agency (AFWA) runs a real-time 10-member 4-km WRF-ARW ensemble. Forecasts are initialized at 0000 UTC and 1200 UTC using 6 or 12 hour forecasts from three global models, an AFWA version of the Met Office Unified Model, the NCEP Global Forecast System (GFS), and the Canadian Meteorological Center Global Environmental Multiscale (GEM) Model. Diversity in the AFWA ensemble is achieved through IC/LBCs from the different global models and varied microphysics and boundary layer parameterizations. SPC is currently ingesting the AFWA grids in their real-time data feed and these forecasts will be available for examination during SFE2014.

Table 3. AFWA ensemble member specifications.

Member #	ICs/LBCs	LSM	Micro-physics	PBL
afwa01	18Z UM	Noah	WDM6	ACM2
afwa02	18Z GFS	Noah	Morrison	BouLac
afwa03	12Z GEM	Noah	WDM6	YSU
afwa04	12Z GEM	Noah	Morrison	BouLac
afwa05	18Z UM	Noah	Thompson	YSU
afwa06	18Z GFS	PX	WDM6	ACM2
afwa07	12Z GEM	PX	Thompson	BouLac
afwa08	18Z GFS	PX	Morrison	ACM2
afwa09	18Z UM	PX	WDM6	YSU
afwa10	18Z GFS	PX	Thompson	ACM2

e) UK-Met Office convection allowing models

The Met Office Unified Model (UM) is the name given to the suite of numerical modeling software used by the Met Office. Two nested limited-area high-resolution versions of the UM (4.4 km and 2.2 km horizontal resolution) running once per day will be supplied to the 2014 Spring Forecast Experiment. These operational models, the Global and nested hi-res versions, will incorporate the latest, soon to be released, UM dynamical core.

The 4.4 km version has 70 vertical levels (spaced between 5m and 40 km) across a domain covering the CONUS and adjacent areas. Taking its initial and lateral boundary conditions from the 00z 17 km horizontal resolution global configuration of the UM, the 4.4 km model initializes without data assimilation and runs out to T+48. This model configuration uses a convective parameterization scheme that limits the convection-scheme activity, and the mixing scheme used is 2D Smagorinsky in the horizontal and the boundary layer mixing scheme in the vertical with single moment microphysics.

The 2.2 km horizontal resolution version of the UM is nested with the 4.4 km model and runs over a slightly sub-CONUS domain. The 2.2 km model takes its initial and lateral boundary conditions from the T+3 step of the 00z 4.4 km run, thus reducing spin-up time within the 2.2 km model, and runs out to T+45. As with the 4.4 km model, the 2.2 km model initializes without data assimilation and uses the same 70 vertical level spacing as the 4.4 km. The 2.2 km model has identical planetary boundary layer and microphysics schemes as the 4.4 km model but does not have the constrained convection scheme of the 4.4 km and instead has no convection parameterization.

Finally, parallel versions of both the 4.4 km and 2.2 km models are also being run with a new blended 'scale aware' planetary boundary layer mixing scheme. The aim here is for the scale aware parameterizations to make a smooth transition from unresolved to resolved scales.

f) ESRL High Resolution Rapid Refresh (HRRR) model

We will continue to include output from the HRRR model developed by the NOAA/Earth Systems Research Laboratory (ESRL). The experimental 3-km grid-spacing HRRR model is nested within the hourly development version of the 13 km RAP model, which provides ICs/LBCs for the HRRR. The HRRR uses a version of the WRF-ARW. A unique aspect of the RAP is the hourly Gridpoint Statistical Interpolation (GSI) data assimilation system that incorporates a wide array of observational datasets including radar reflectivity via the radar-Diabatic Digital Filter Initialization. The HRRR integration is run hourly over a full CONUS domain with forecasts to 15 hours.

New this year will be an experimental version of the HRRR run out to 48 h for select analysis cycles (0100 UTC and 1300 UTC) to support the production of Day 2 forecasts by the SFE participants. In addition, ESRL will provide time-lagged ensemble severe weather probabilities for the Day 1 period to support the high-temporal resolution outlooks.

3. SFE2014 Core Interests/Daily Activities

a. Forecast Products and Activities

Similar to previous years, the experimental forecasts this year will continue to explore our ability to add temporal specificity to longer-term convective outlooks. We will continue to split the participants into two teams which will be called the "SPC Desk" and the "Development Desk", as the SPC desk, led by SPC personnel, will examine products and techniques that are closer to operational implementation while the Development Desk, led by NSSL personnel, will explore additional cutting-edge products and techniques. For the SPC Desk, the first forecast will mimic the SPC operational Day 1 convective outlooks by producing probability forecasts of large hail, damaging wind, and/or tornadoes within 25 miles (40 km) of a point valid 1600 UTC to 1200 UTC the next day. This is new to the SFE2014; past experiments only produced combined probabilities of hail, wind, and tornadoes ("total severe") over this time period. On the Development Desk, a separate Day 1 forecast will be made for total severe probabilities valid over the same period. A text product will accompany each Day 1 forecast that describes the meteorology of the day and the models used and why they were used to create the severe storm forecasts.

Each desk will then manually stratify their respective Day 1 forecasts into periods with higher temporal resolution. The SPC Desk will generate separate probability forecasts of large hail, damaging wind, and/or tornadoes valid for three periods: 1800-2100 UTC, 2100-0000 UTC, and 0000-0300 UTC. As an alternative way of stratifying the Day 1 forecast, the Development Desk will generate probability forecasts of total severe valid *every hour* from 1800-0300 UTC. The goal of testing these two methods is to explore different ways of introducing probabilistic severe weather forecasts on time scales that are currently addressed with mostly categorical forecast products (mesoscale discussions and convective watches) and to begin to explore ways of seamlessly merging probabilistic severe weather outlooks with probabilistic severe weather warnings as part of the NOAA Warn-on-Forecast initiative.

Similar to previous experiments in which calibrated severe storm guidance from the SSEO was used to temporally disaggregate a 1600-1200 UTC period human forecast, the same technique will be applied to the SPC Desk forecasts during SFE2014. This year, however, calibrated guidance for the individual hazards will be derived from the SREF (environment information) and SSEO (explicit storm attributes) forecasts and will be used in the temporal disaggregation technique. The 1600-1200 UTC human forecasts for the SPC Desk will be temporally disaggregated into the 3-h periods (1800-2100 UTC, 2100-0000 UTC and 0000-0300 UTC) to provide a first guess for the three forecast periods. In addition, hourly calibrated probabilities of total severe will be generated from the SSEO, NSSL-WRF, and CAPS SSEF ensembles to serve as first guess fields for the human-generated forecasts at the Development Desk.

The first set of short-time-window forecasts will be issued by 11am local time (i.e., 1600 UTC) for both desks. Participants at the SPC Desk will jointly discuss and develop the forecast using the NMAP software on the N-AWIPS workstations. Participants at the Development Desk will generate their short-time-window forecasts differently. The lead forecaster at the Development Desk (Dave Imy) will generate a set of hourly probability forecasts over the 1800-0300 UTC period using N-AWIPS. The other team members will split into five groups and use a web-based tool to generate their own hourly probability forecasts of total severe over the same period. Google Chromebooks will be provided to the participants so that they can make their own forecasts (to serve as human ensemble members, with Dave Imy acting as the "human control member"). Because no more than ten participants will likely be on the development desk at any time, there should usually be no more than two people in each of the groups that generate hourly forecasts using the web tool. First guess hourly probability fields generated from CAM forecasts will be available in all of the product generation systems.

After the short-time-window forecasts are issued, both desks will transition to examining the Day 2 period. Teams will examine operational guidance as well as experimental CAM guidance that will extend into Day 2 and generate probability forecasts of severe weather within 25 miles (40 km) of a point valid from 1200 UTC the next day to 1200 UTC the following day (Day 3). The SPC Desk will create separate forecasts for large hail, damaging wind, and tornadoes while the Development Desk will produce forecasts for total severe. Producing any type of severe weather forecast into Day 2 is relatively new to the SFE2014. The goal is to explore the feasibility of issuing forecasts of individual severe storm hazards into Day 2; current SPC operational forecasts valid for Day 2 and beyond only consider probabilities of total severe. A text product will accompany each Day 2 outlook that describes the meteorology of the day and the usefulness of the suite of model guidance during the creation of the severe storm forecasts. These Day 2 forecasts will be made at each desk and will be issued by 12pm local time (1700 UTC).

Finally, each team will examine observational trends and morning/afternoon model guidance to update their respective short-time-window forecasts made earlier in the day. Because the forecast process for these updates will begin at 2:30pm, only the forecasts valid from 2100-0300 UTC will be updated. These updated forecasts will be issued by 4pm local time (2100 UTC).

To explore the process of enhancing collaboration between the HWT forecast and warning experimental activities, digitized SFE forecasts of severe convection over the short time windows will be provided to the Experimental Warning Program (EWP) for use in preparation for their operations. The SFE and EWP Experiments will overlap for four out of the five weeks (only 27-30 May will not have overlapping operations). During these weeks, collaboration activities will occur during the day depending on the timing of convection on that day. This activity is an early manifestation of the goal of providing probabilistic hazard forecasts on multiple scales from the synoptic scale to the storm scale.

b. Forecast and Model Evaluations

(1) Subjective Evaluation of Experimental Forecasts

In the next day evaluations, the individual period team forecasts and the first-guess guidance will be compared to observed radar reflectivity, reports of severe weather, NWS warnings, and radar-estimated hail sizes and rotation tracks over the same time periods. The SFE participants will provide their subjective evaluations of the strengths and weaknesses of each of the forecasts. This evaluation will include examining and comparing calibrated guidance, temporal disaggregation first guess, and human initial and update forecasts. The goal is to assess the skill of the first-guess guidance and the human-generated forecasts over all periods.

(2) Objective Evaluation of Experimental Forecasts

Similar to last year, experimental forecasts of total severe thunderstorm probabilities will be evaluated using Critical Success Index (CSI) and Fractions Skill Score (FSS) based on the local storm reports (LSRs) as the verification. For the first time, individual hazards of tornado, wind, and hail will also be considered separately. In addition, comparisons of results from the experimental forecasts to the first-guess automated fields will be possible. CSI will be calculated at a couple of fixed probability thresholds used in SPC operational outlooks. The utility of verification metrics in assessing forecast skill for long and short periods will continue to be explored by comparing the scores to the subjective impressions of the participants.

(3) Evaluation of Deterministic Convection-Allowing Models and Physics Sensitivities

Standard fields from convection-allowing model output (simulated reflectivity, 10-m wind speed and UH) and experimental simulated storm attributes will be generated from select CAMs. Day-2 and day-1 guidance for the SSEF and AFWA will be compared. The focus of these evaluations will be the general accuracy of the forecasts in predicting severe convection explicitly, as well as the impact of various physics options on the forecasts. There will be evaluations of the new microphysics schemes available in WRF-ARW and newly updated schemes provided to us by the developers. There will also be comparisons of the Met Office CAMs and the NSSL WRF ensemble control run using model soundings in the pre-convective environment as the Met Office office uses more vertical levels (70) and a different approach to parameterizing turbulent mixing than the WRF-ARW model. These types of comparisons have shown to be valuable in identifying systematic model deficiencies that can impact convective forecasts significantly (Clark et al. 2012a, Coniglio et al. 2013, Duda et al. 2014).

(4) Comparison and Evaluation of Convection-Allowing Ensembles

A variety of ensemble fields (spaghetti, exceedance probability, maximum from any member) based on hourly maximum fields (HMFs) of simulated storm attributes, such as 1-km AGL simulated reflectivity, updraft speed, updraft helicity, and 10-m winds, will be available from four 0000 UTC convection-allowing ensembles (SSEO, CAPS SSEF, AFWA, and NSSL-WRF). These 0000 UTC ensembles will be compared and evaluated on their ability to provide useful severe weather guidance. Forecasts from 1200 UTC-initialized forecasts will also be available for three convection-allowing ensembles (SSEO, SSEF, and AFWA). Subjective and objective evaluations will be performed comparing each ensemble's 0000 UTC forecasts to the 1200 UTC forecasts and comparing the ensembles to one another. In addition, two of the 0000 UTC ensembles (SSEF and AFWA) will have forecasts extending out to 60 hours, which will allow for a first-time comparison of guidance on Day 2 versus Day 1 for these ensembles.

(5) Evaluation of model guidance for hail

There is interest in evaluating the ability of CAMs to predict hail size because of the increased focus on forecasting individual thunderstorm hazards in the SFE2014. A new product for this year is the HAILCAST algorithm implemented in WRF-ARW to predict hail size (Adams-Selin 2013), which is based on the algorithm described by Brimelow (2002) and Jewell and Brimelow (2009). Although microphysics schemes in WRF-ARW could be used to extract hail size, only limited success was found using the graupel size distributions (Adams-Selin 2012), which is consistent with our experience with exploring integrated graupel mixing ratios and lowest-model level maximum graupel size in various microphysics schemes in the SFE2013. Rather than predict hail size explicitly, the HAILCAST model uses convective cloud and updraft attributes to determine the growth of hail from initial embryos. The cloud attributes for the model are those predicted explicitly in the WRF-ARW forecasts and the snow, ice and graupel mixing ratios at the first level above the freezing level at which they exist are used to determine the initial embryo size. For more information on the implementation of this algorithm in WRF-ARW, see Adams-Selin (2013).

Explicit prediction of hail size from the HAILCAST model within the NSSL-WRF will be evaluated against storm reports and the WSR-88D-derived Maximum Expected Size of Hail (MESH) product developed by NSSL as part of the Warning Decision Support System – Integrated Information (WDSS-II) suite of algorithms. It will also be examined against forecasts of hourly maximum UH to determine if the HAILCAST model output gives more information on the potential for severe hail beyond what has been construed from UH tracks in the model.

c. SFE2014 Daily Activities Schedule

Scheduled activities are in local (CDT) time.

<u>SPC/Severe Desk</u>	<u>NSSL/Development Desk</u>

0800 – 0845: **Evaluation of Previous Day's Experimental Forecasts**

- Subjective rating relative to radar evolution/characteristics, warnings, and preliminary reports and objective verification using preliminary reports and MESH:
    - Day 1 & 2 full-period probabilistic forecasts of tornado, wind, and hail
    - Day 1 3-h period forecasts and guidance for tornado, wind, and hail
    - Day 1 & 2 full-period probabilistic forecasts of total severe
    - Day 1 1-h period forecasts and guidance of total severe

0845 – 1100: **Day 1 Convective Outlook Generation**

- After hand analyses of 1200 UTC upper-air maps and surface charts and discussion:
    - Prepare probability forecasts for tornado, wind, and hail valid 16-12Z over mesoscale area of interest
    - Adjust temporally disaggregated first guess for tornado, wind, and hail forecasts valid for 3-h periods: 18-21, 21-00, and 00-03Z; make these available to EWP
    - Prepare probability forecasts for total severe valid 16-12Z over mesoscale area of interest
    - Adjust first guess for total severe valid for 1-h periods: 18-03Z; make these available to EWP

1100 – 1200: **Day 2 Convective Outlook Generation**

- After analyses of 12Z model guidance:
    - Prepare probability forecasts for tornado, wind, and hail valid 12-12Z over mesoscale area of interest
    - Prepare probability forecasts for total severe valid 12-12Z over mesoscale area of interest

1200 – 1300: **Lunch**

1300 – 1330: **Briefing**

- Overview and discussion of today's forecast challenges and products
- Highlight interesting features/findings from yesterday including 3-D visualization

1330 – 1430: **Scientific Evaluations**

- Examine convection-allowing ensemble guidance: Day 2 vs Day 1
- Compare convection-allowing guidance (SSEO, SSEF, AFWA, and NSSL; 00Z and 12Z)

- Met Office convection-allowing runs
- Model guidance for hail & Microphysics Comparison

1430 – 1600: **Short-term Outlook Update**

- Update probability forecasts for tornado, wind, and hail valid 21-00 and 00-03Z; make these available to EWP
- Update hourly probability forecasts for total severe valid 21-03Z; make these available to EWP

d. Other specialized activities

For the first time in the HWT Spring Forecasting Experiments, an effort will be made to view CAM output in three-dimensional (3D) displays in real time as part of the Development Desk activities. Selected 3D model fields over our mesoscale region of interest at 10-minute output frequency for 18 – 30 h forecasts will be interrogated using the WDSS-II display system. The goal is to explore CAM storm characteristics like vertical vorticity, graupel mixing ratio, simulated reflectivity, and cold pools in 3D to learn more about how simulated storms are structured on WRF-ARW convection-allowing grids (see Figure 1 for an example of what this display will look like in real time). The model storm structures will be compared to the structures of storms observed by the WSR-88D network and displayed within WDSS-II. We will also examine characteristics of the storm environments in CAM forecasts like depth of water vapor mixing ratio in the PBL and depictions of low-level convergence boundaries and how they may play a role in the initiation of convection in the model. Since this is an initial exploration into the detailed structures of storms and model forecast environments in real time as part of a forecast process, this activity will be somewhat informal and less structured than the other SFE2014 activities.

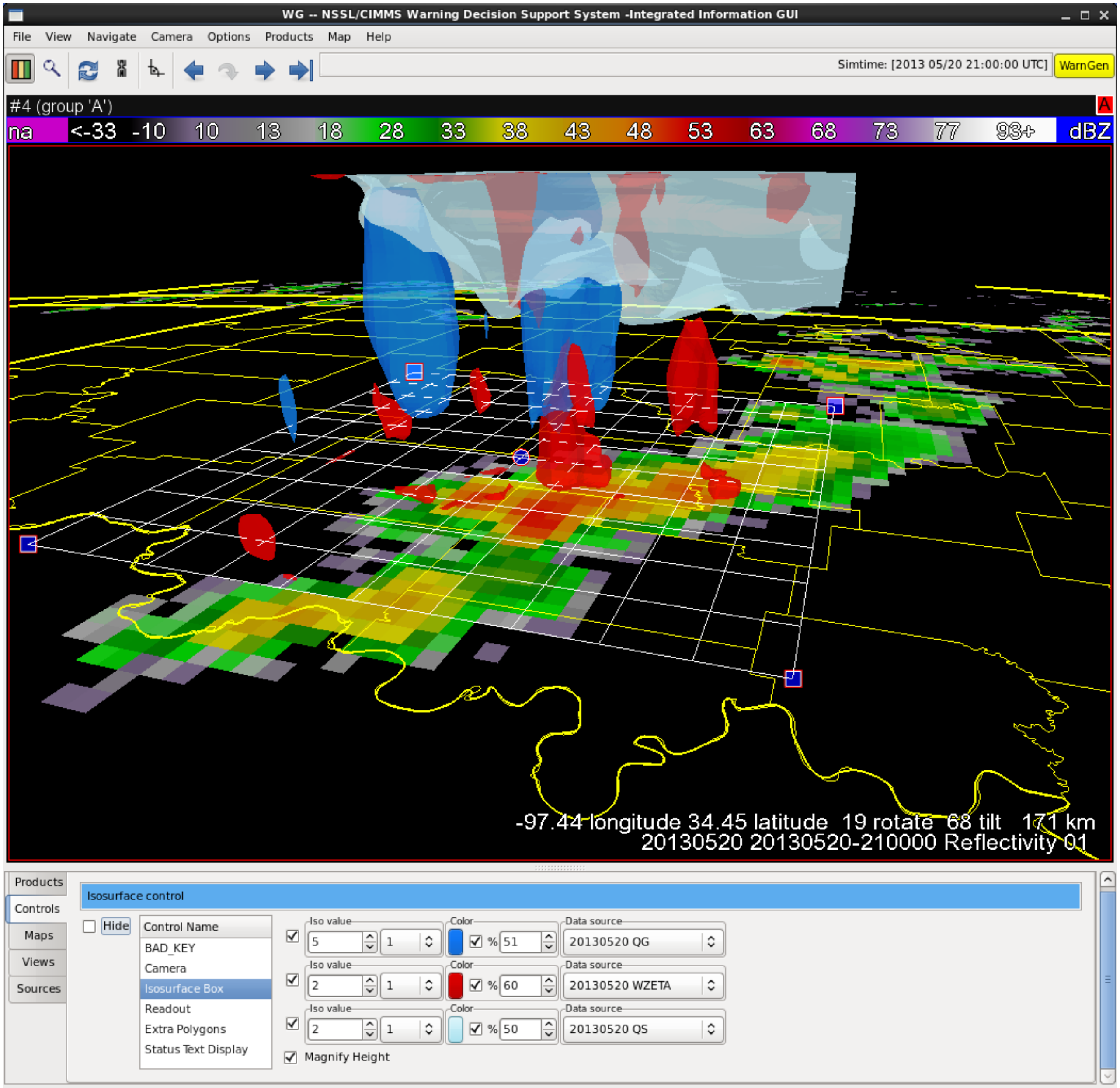

Figure 1. Example of how CAM forecasts will be interrogated for a select few runs from the CAPS SSEF system. The 2D field is the simulated reflectivity on the lowest model level (color scale near the top of the image) with 3D isosurfaces of vertical velocity x vertical vorticity ($w*\zeta$) > 2 m s^{-2} (red areas), graupel mixing ratio > 5 g kg^{-1} (dark blue areas), and snow mixing ratio > 2 g kg^{-1} shown within the box outlined in white. All fields are 21-h forecasts valid 2100 UTC 20 May 2013.

Appendix A: List of scheduled SFE2014 participants. Facilitators/leaders for SFE2014 include: Michael Coniglio (NSSL), Adam Clark (NSSL), Kent Knopfmeier (NSSL), Israel Jirak (SPC), James Correia Jr. (SPC), Chris Melick (SPC), Dave Imy (SPC-retired), and Steve Willington (Met Office).

5-9 May	12-16 May	20-24 May	28 May – 31 May	3-7 June
J. Blaes (NWS)	D. Dawson (CAPS)	R. Adams-Selin (AFWA)	J. Barnwell (NWS)	T. Alcott (NWS)
M. Chenard (NWS)	G. Dial (SPC)	C. Alexander (GSD)	J. Carley (EMC)	E. Aligo (EMC)
G. Creighton (NWS)	D. Gagne III (OU)	L. Bosart (SUNYA)	B. Entwistle (AWC)	B. Ancell (TTU)
M. Seltzer (Met Office)	B. Gallus (ISU)	A. Cohen[Tu-F] (SPC)	B. Etherton (GSD)	B. Burghardt (TTU)
J. Guyer (SPC)	A. Hill (TTU)	M. Evans (NWS)	L. Gilchrist (UKMO)	J. Brown (GSD)
S. Hitchcock (OU)	T. Hultquist (NWS)	L. Gilchrist (UKMO)	J. Grams (SPC)	C. Franks (NWS)
A. Kahraman (PSU)	J. Lawson (ISU)	C. Guastini (SUNYA)	D. Harris (UKMO)	M. Hirsch (NWS)
G. Romine (NCAR)	P. Manousos (FE)	D. Harris (UKMO)	S. Lack (NWS)	M. Pyle (EMC)
C. Schultz (NWS)	J. Milbrandt (EC)	H. Lean (UKMO)	A. Lese (NWS)	B. Rubin-Oster (WPC)
Y.-G. Skabar (SMN)	H. Morrison[M-W] (NCAR)	H. Richter (CAWAR)	M. Petty (GSD)	B. Smith (SPC)
E. Szoke (GSD)	C. Nowotarski (TAMU)	D. Van Dyke (NWS)	H. Richter (CAWAR)	C. Tubbs (UKMO)
	N. Roberts (UKMO)	M. Vaughn (SUNYA)	S. Roberts (NWS)	B. Twiest (PSU)
	S. Rogowski (NWS)	C. Ziegler (NSSL)	M. Wandishin (GSD)	C. Weiss (TTU)
	G. Thompson[M-W] (NCAR)			
	M. Weeks (UKMO)			

Appendix B: Experimental Severe Thunderstorm Forecasts

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

Day 1 Probability to Categorical Outlook Conversion

(SIGNIFICANT SEVERE area needed where denoted by hatching - otherwise default to next lower category)

Outlook Probability	TORN	WIND	HAIL
2%	SEE TEXT	NOT USED	NOT USED
5%	SLGT	SEE TEXT	SEE TEXT
10%	SLGT	NOT USED	NOT USED
15%	MDT	SLGT	SLGT
30%	HIGH	SLGT	SLGT
45%	HIGH	MDT	MDT
60%	HIGH	HIGH	MDT

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

Day 2 Probability to Categorical Outlook Conversion

(SIGNIFICANT SEVERE area needed where denoted by hatching - otherwise default to next lower category)

Outlook Probability	Combined TORN, WIND, and HAIL
5%	SEE TEXT
15%	SLGT
30%	SLGT
45%	MDT
60%	HIGH

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

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

The HWT organizational structure is composed of three overlapping programs (Fig. 1). The Experimental Forecast Program (EFP) is focused on predicting hazardous mesoscale weather events on time scales ranging from hours to a week in advance, and on spatial domains ranging from several counties to the CONUS. The EFP embodies the collaborative experiments and activities previously undertaken by the annual SPC/NSSL Spring Experiments. For more information see http://www.nssl.noaa.gov/projects/hwt/efp/.

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

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

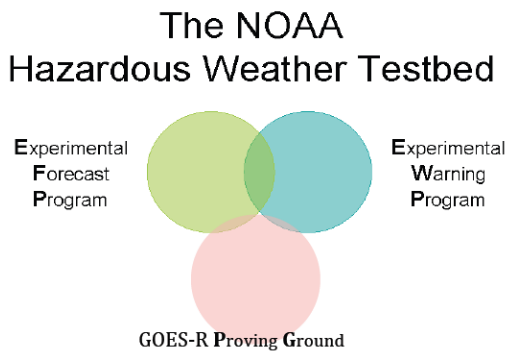

Figure 1: The umbrella of the NOAA Hazardous Weather Testbed (HWT) encompasses two program areas: The Experimental Forecast Program (EFP), the Experimental Warning Program (EWP), and the GOES-R Proving Ground (GOES-R).

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

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

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

Appendix D: References

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

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

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

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

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

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

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

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

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.

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.

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

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