INFORMATION EXTRACTION AND VERIFICATION OF CONVECTION-ALLOWING MODELS FOR TORNADO FORECASTING

Israel Jirak, *NOAA/Storm Prediction Center*

Harold Brooks, *NOAA/National Severe Storms Laboratory*

Matt Pyle, *NOAA/Environmental Modeling Center*

Robert Hepper, *CIMMS/SPC Research Associate*

Jeff Milne, *OU Grad Student and CIMMS/SPC Research Assistant*

Project Overview

- Collaborative effort among SPC, NSSL, and EMC to improve severe weather forecasting
- Direct involvement with the Hazardous Weather Testbed (HWT) for testing and evaluation of products/techniques
- Addresses NGGPS program priority of advancing forecasts for high-impact weather in days 0-3 by focusing on the operational utility of CAMs and CAM ensembles
- Two primary components :
  - Verification
    - Evaluate various convection-allowing ensemble configurations and assess current skill of CAMs using appropriate metrics
  - Information Extraction
    - Develop techniques/diagnostics to mine useful information embedded within convection-allowing models for severe weather forecasting

Verification:

Convection-Allowing Ensembles

- **Community-Leveraged Unified Ensemble (CLUE)**
  - *Inspired by the UMAC to provide evidence-based decision making with regard to the design of a future operational convection-allowing ensemble*
  - *Unprecedented effort to leverage several academic and government research institutions to help guide NOAA's 0-36 h operational prediction of convective storms*
  - GOAL: **Design experiments to provide more controlled datasets that can be better utilized to inform configuration of near-future operational systems**
  - Contributors agreed on a set of model specifications (e.g., model version, grid-spacing, domain, vertical levels, physics, input data).
  - Post-processing was also formalized. All groups output the same set of fields in grib2 format using a modified version of UPP.
  - **2016 Contributors:** CAPS (35), NCAR (10), University of North Dakota (4), NSSL (15), and GSD (1) – 65 Total CLUE Members

Verification:

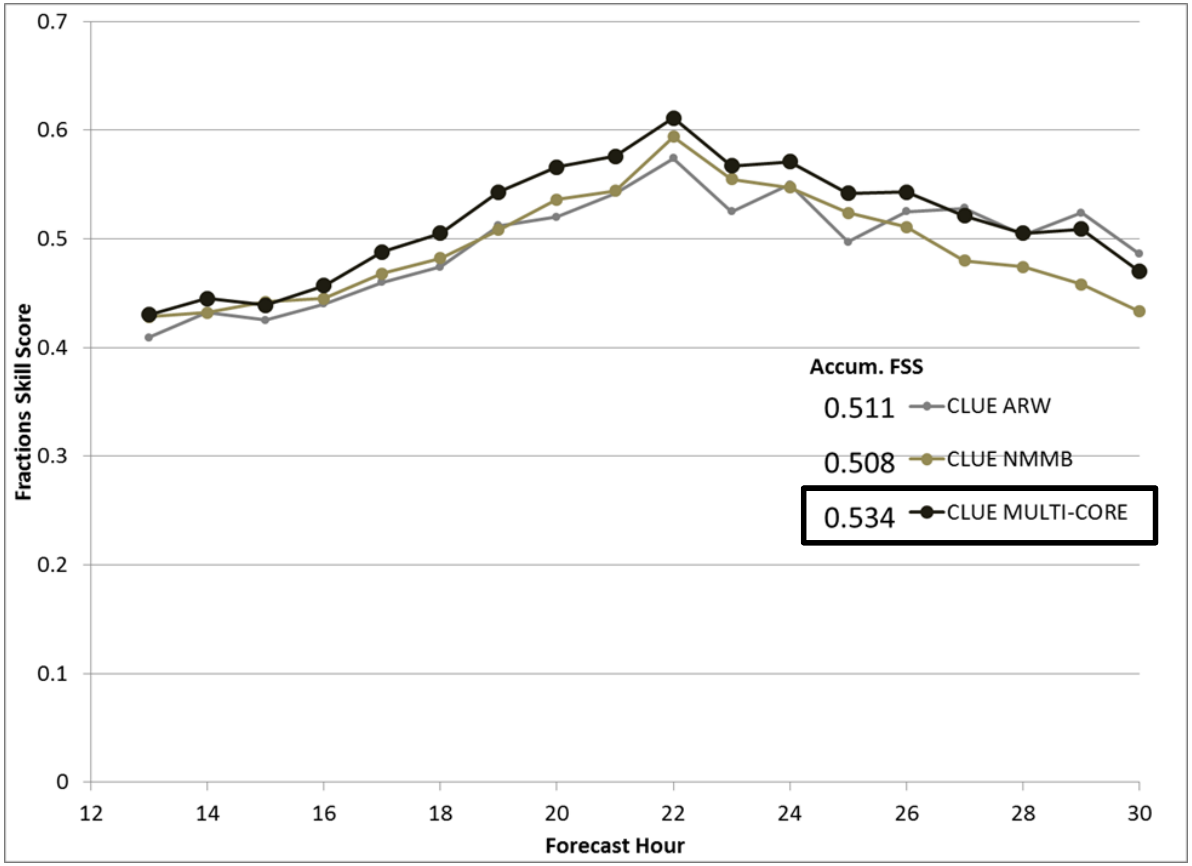
2016 HWT CLUE Experiments

- The design of the CLUE allowed for several controlled experiments:
  - Deterministic comparisons of WRF-ARW and NMMB
  - **Multi-core vs. single-core ensembles** – A comparison of three 10-member ensembles with the same IC/LBCs: 1) one with ARW core, 2) one with NMMB core, and 3) one with 5 ARW members and 5 NMMB members
  - **Single-physics vs. multi-physics ensembles** – A comparison of two 10-member ensembles with the same IC/LBCs: 1) one with constant physics and 2) one with varied physics
  - Ensembles with and without radar data assimilation
  - 3DVAR vs. EnKF data assimilation approaches
  - Microphysics sensitivities
  - **Ensemble size comparisons** - A comparison of the mixed-core ensembles with equal contributions of NMMB and ARW members using 2, 4, 6, 10, and 20 members.

Verification:

2016 HWT CLUE Experiments: ***Core***

- The ***multi-core ensemble*** generally provided improved probabilistic forecasts of simulated reflectivity over both of the single-core ensembles.



Mesoscale Area of Interest
1-km AGL Reflect. >40 dBZ
20160502-20160603
00Z cycle; fh013-030

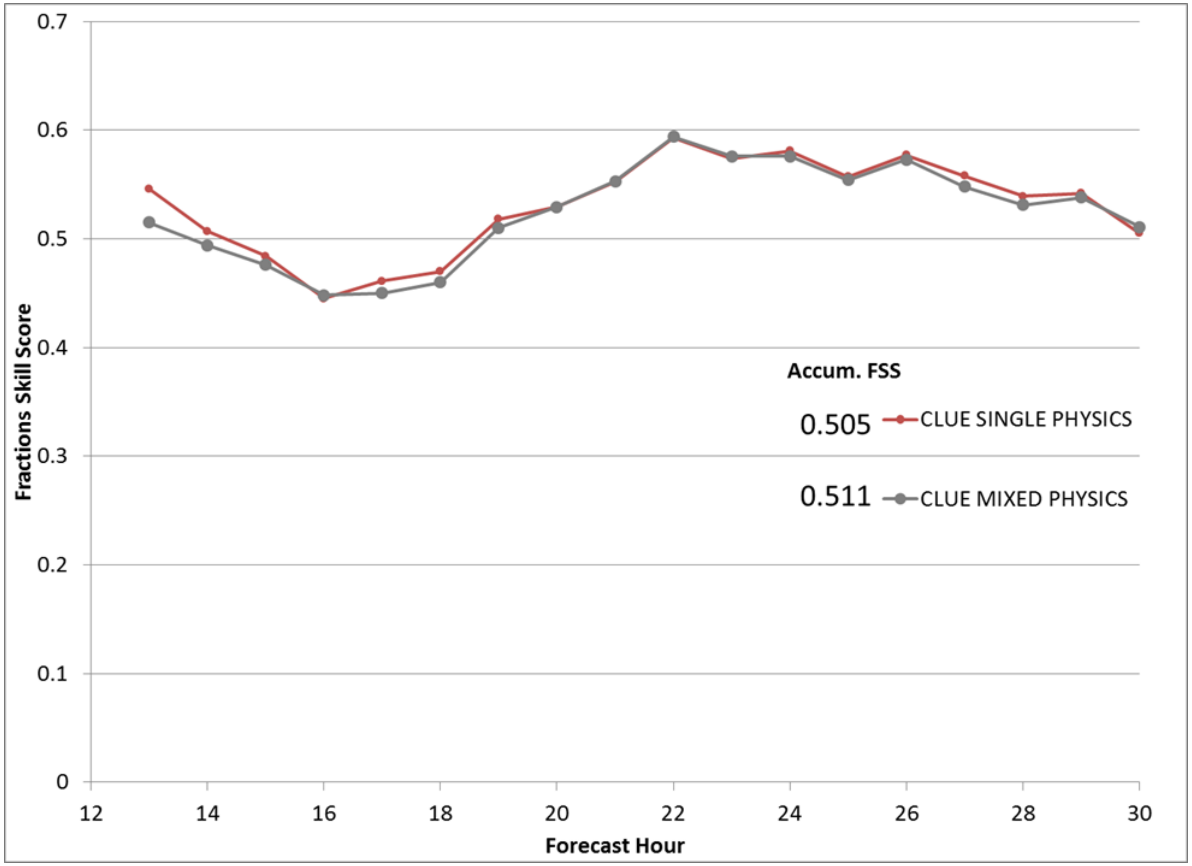
40-km Radius of Influence
10 grid-point Gaussian smoothing parameter

CLUE_ARW: 10 ARW
CLUE_NMMB: 10 NMMB
CLUE_MULT: 5 ARW,5 NMMB

Verification:

2016 HWT CLUE Experiments: ***Physics***

- The mixed-physics and ***single-physics*** ensembles produced statistically similar probabilistic forecasts of simulated reflectivity during the spring.



Mesoscale Area of Interest
1-km AGL Reflect. >40 dBZ
20160502-20160603
00Z cycle; fh013-030

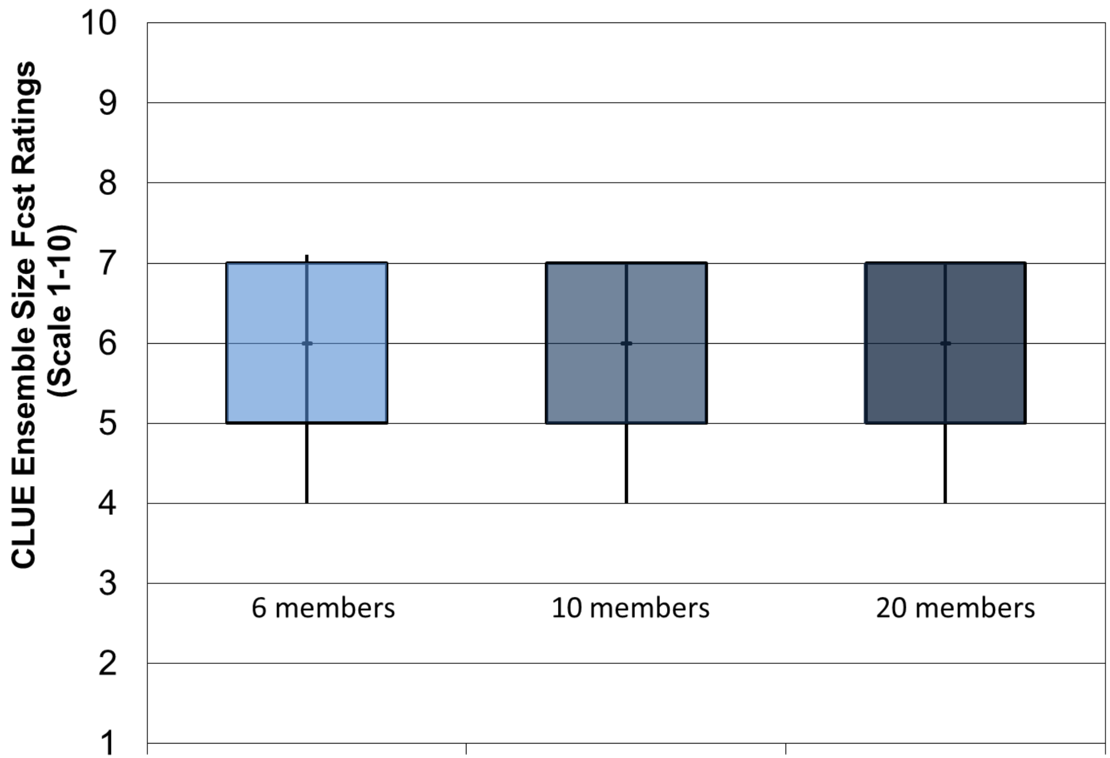
40-km Radius of Influence
10 grid-point Gaussian smoothing parameter

CLUE_SINGLE: 10 ARW
CLUE_MIXED: 10 ARW

Verification:

2016 HWT CLUE Experiments: ***Size***

- Subjective ratings from SFE participants indicated that the ensembles of different sizes produced ***very similar forecasts*** on most days (i.e., little practical difference).



Mesoscale Area of Interest
1-km AGL Reflect. >40 dBZ
20160502-20160603
00Z cycle; fh013-030

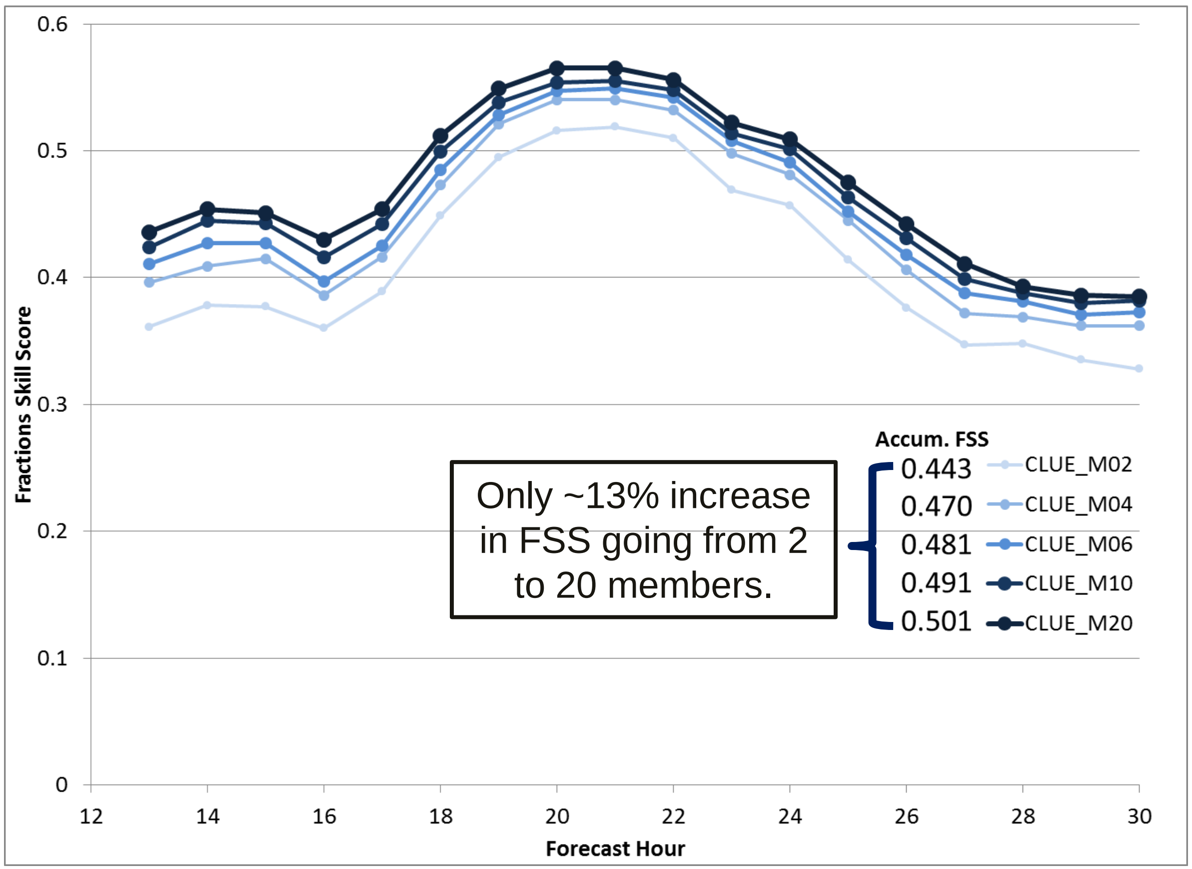
40-km Radius of Influence
10 grid-point Gaussian smoothing parameter

CLUE_M06: 3 ARW, 3 NMMB
CLUE_M10: 5 ARW, 5 NMMB
CLUE_M20:10 ARW,10 NMMB

Verification:

*2016 HWT CLUE Experiments: **Size***

- Increasing ensemble size only has a ***small*** positive impact on the skill of the probabilistic forecasts of simulated reflectivity over the CONUS during the spring.



CONUS
1-km AGL Reflect. >40 dBZ
20160502-20160603
00Z cycle; fh013-030

40-km Radius of Influence
10 grid-point Gaussian smoothing parameter

CLUE_M02: 1 ARW, 1 NMMB
CLUE_M04: 2 ARW, 2 NMMB
CLUE_M06: 3 ARW, 3 NMMB
CLUE_M10: 5 ARW, 5 NMMB
CLUE_M20:10 ARW,10 NMMB

Verification:

2017 HWT CLUE Experiments

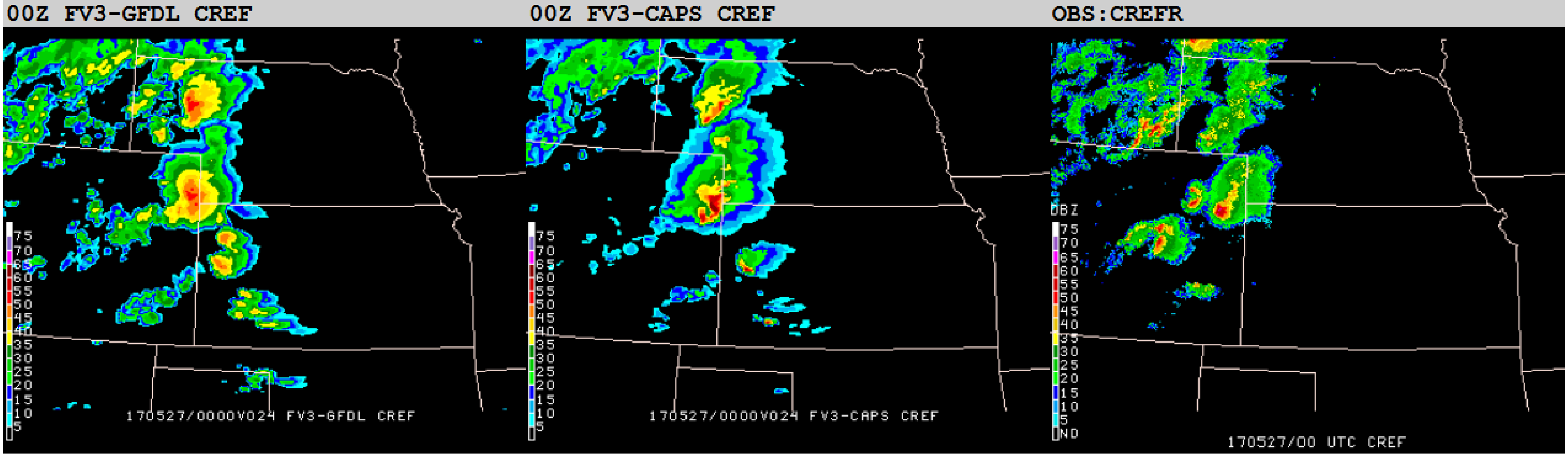
- The design of the 2017 CLUE allowed for several controlled experiments:
  - Multi-core and single-core ensemble comparisons
  - Physics perturbations, including a stochastic approach
  - 3DVAR vs. EnKF data assimilation approaches
  - Microphysics sensitivities
  - **FV3 -** Two different experimental versions of FV3 at 3-km grid spacing were examined and compared to existing real-time CAMs to gauge performance at convective scales*
    - FV3-GFDL: GFS physics with GFDL microphysics
    - FV3-CAPS: GFS physics with Thompson microphysics

* This required substantial effort from GFDL and CAPS to implement severe weather diagnostic variables into the FV3 code and to generate grib2 output

Verification:

2017 HWT CLUE Experiments: ***FV3***

- HWT SFE participants examined the deterministic FV3 reflectivity forecasts and subjectively rated the quality of the forecast from a severe weather perspective.



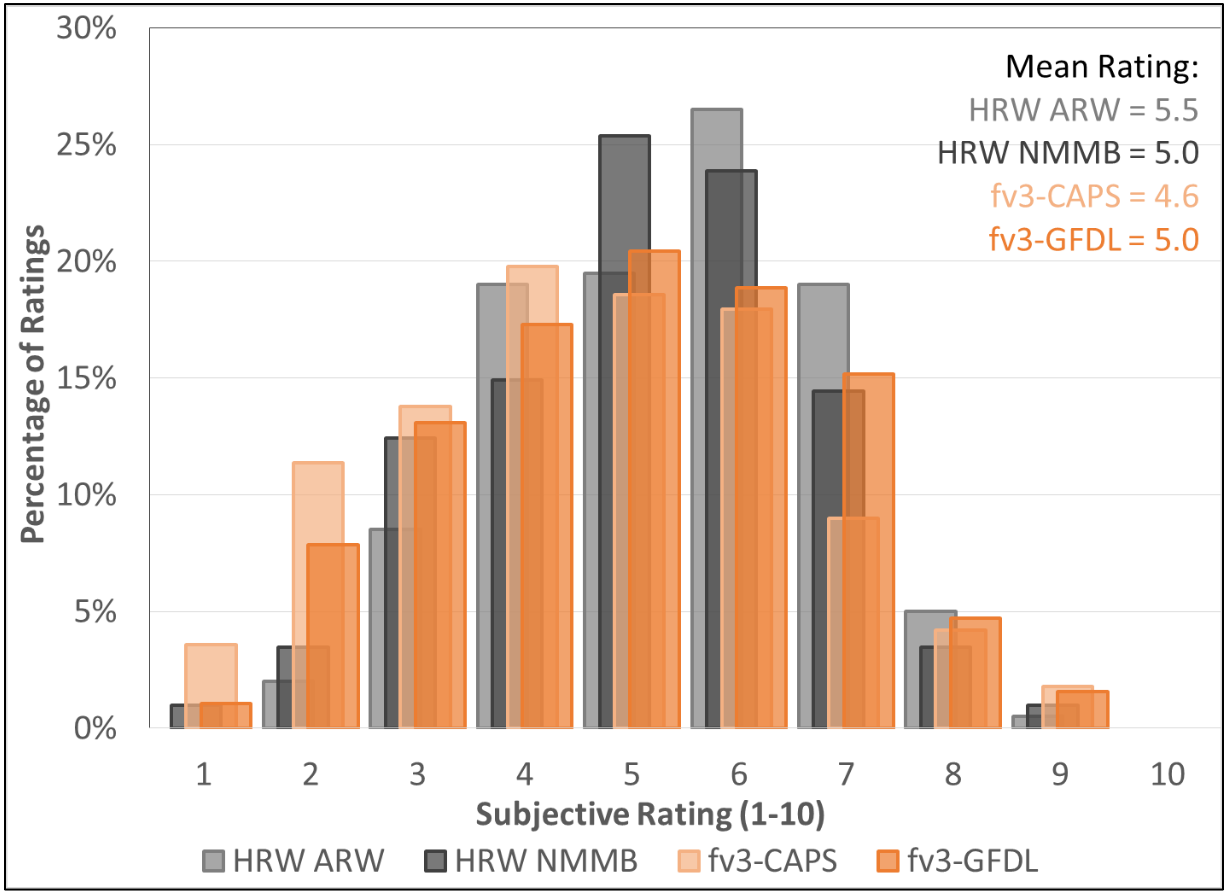
FV3-GFDL FV3-CAPS OBS

24-h Forecast of Composite Reflectivity
Valid 00Z on 27 May 2017

Verification:

2017 HWT CLUE Experiments: ***FV3***

- Subjective ratings from SFE participants indicated that the FV3 reflectivity forecasts **compared favorably** to operational CAMs, especially for initial 3-km FV3 versions.



Mesoscale Area of Interest
1-km AGL Reflectivity
20170501-20170602
00Z cycle; fh018-030

Operational HRW ARW
Operational HRW NMMB
FV3-CAPS
FV3-GFDL

Verification:

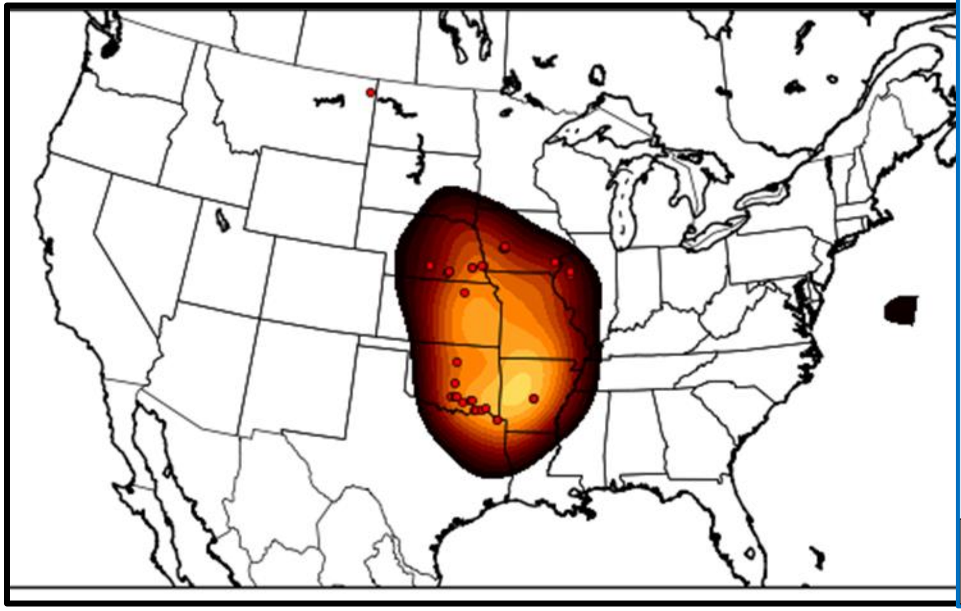
Current Skill of UH Forecasts for Tornado Forecasting

- Several recent studies (e.g., Jirak et al. 2014; Gallo et al. 2016; Sobash et al. 2016) have examined the utility of updraft helicity (UH) forecasts from CAM ensembles as an indirect proxy for tornado forecasting
- As a preliminary exercise, neighborhood probability forecasts of UH from the SSEO were examined for notable tornado events, as subjectively rated by SPC forecasters according to the quality of operational SPC outlooks:
  - "**good forecasts**" where SPC tornado outlooks verified well
  - "**overforecasts**" where SPC tornado probabilities were too high
  - "**underforecasts**" where SPC tornado probabilities were too low
- Provides a baseline for the current skill of CAM ensembles to provide proxy guidance for tornado prediction and how they relate to human-generated tornado outlooks

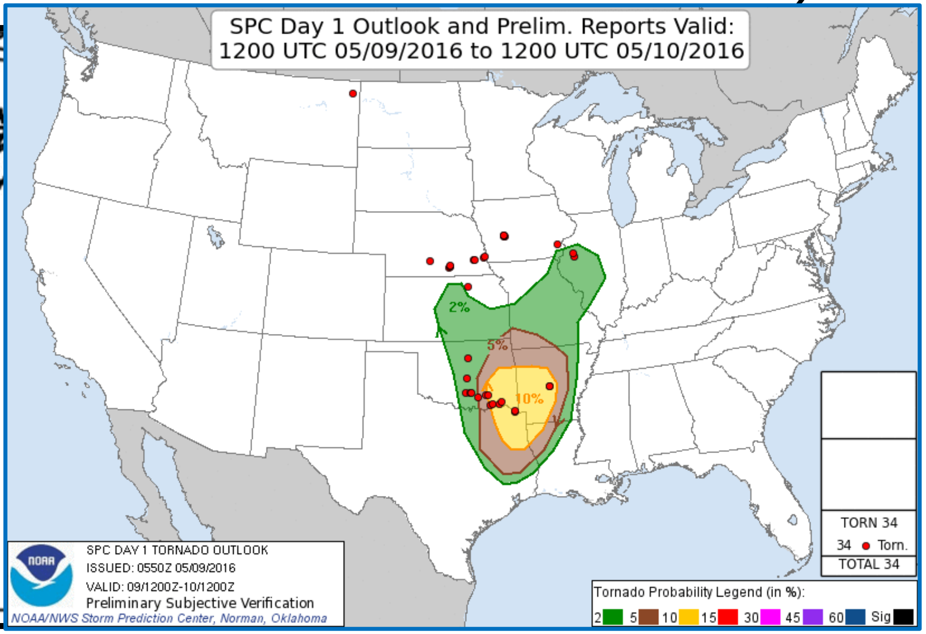
Verification:

Current Skill of UH Forecasts for Tornado Forecasting

- For most of the events examined, the SSEO UH forecasts fell into the same subjective category as the SPC tornado outlook (i.e., skill of forecasts and outlooks are correlated)



00Z SSEO 24-h Neighborhood Probability of Updraft Helicity ≥25 m^2s^{-2}
Valid 09 May 2016



SPC 06Z Day 1 Tornado Outlook
Valid 09 May 2016
"**Underforecast**" – significant tornadoes in south-central Oklahoma

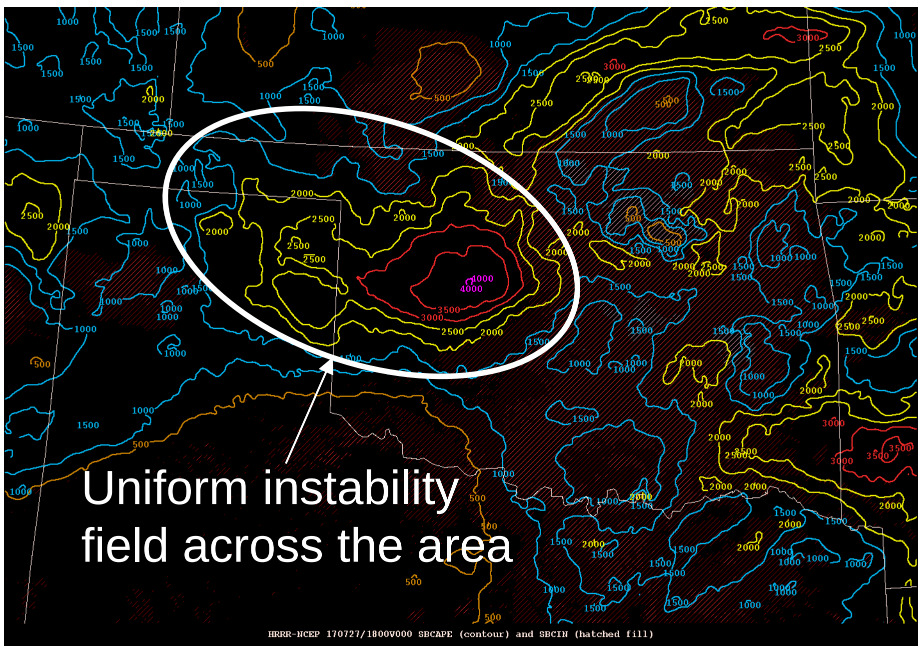
Information Extraction:
Refine UH diagnostic for tornado prediction

- While traditional UH output from 2-5 km AGL provides indication of mid-level rotation in simulated storms, more information needs to be considered for tornado potential
- Different approaches have been applied to highlight tornadic potential with UH:
  - Utilize probabilistic information on favorable tornadic *environment*
  - Filter out UH that is not within a favorable tornadic environment
  - Examine low-level UH (e.g., 0-3 km AGL)
- An in-depth exploration of UH is underway to best identify tornado potential in CAMs, including calculation over various layers, accounting for updraft tilt, and the sensitivity to horizontal and vertical resolution

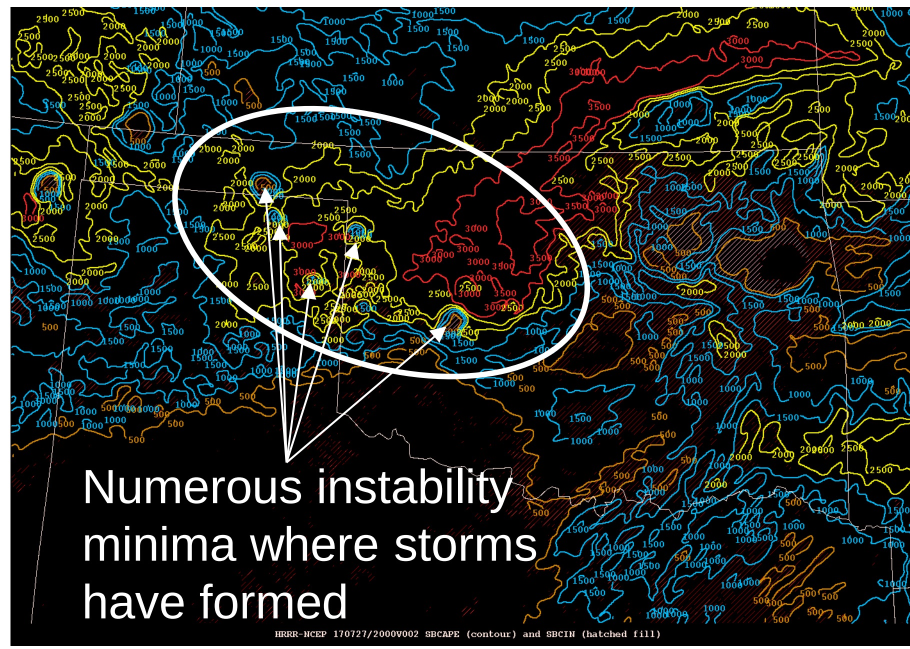
Information Extraction:

Explore the near-storm environment in CAMs

- Near-storm environment fields from CAMs are very detailed and strongly modulated after models initiate storms, so forecasters often prefer examining these fields from coarser resolution models
- Identify an optimal approach for extracting pre-convective and near-storm environment information from CAMs



18Z HRRR 0-h SBCAPE/SBCIN
Valid 18Z on 27 July 2017



18Z HRRR 2-h fcst SBCAPE/SBCIN
Valid 20Z on 27 July 2017

Summary

- The **Community-Leveraged Unified Ensemble (CLUE)** examined during the HWT Spring Forecasting Experiment (SFE) has been an effective way to bring the community together to work on convection-allowing ensemble design:
  - A **multi-core** ensemble generally provided improved probabilistic forecasts at the convection-allowing scale for severe weather applications than a single-core ensemble
  - A **single-physics** ensemble performed similarly to a mixed-physics ensemble with the benefit of lower development & maintenance costs
  - Adding more members had **limited benefit** for this ensemble configuration, as the ensemble forecasts were very under dispersive
- **FV3** forecasts were executed and examined at **convection-allowing** scale (3-km grid spacing) for the first time in real-time during the 2017 HWT SFE, providing comparable forecasts to other operational and experimental CAMs

BACKUP SLIDES

HWT Spring Forecasting Experiment:

2016 CLUE Size Experiment Configuration

- Single physics (per model core) ensemble without radar data assimilation. IC/LBCs from operational SREF members.

CLUE_M20
CLUE_M10
CLUE_M06
CLUE_M04
CLUE_M02

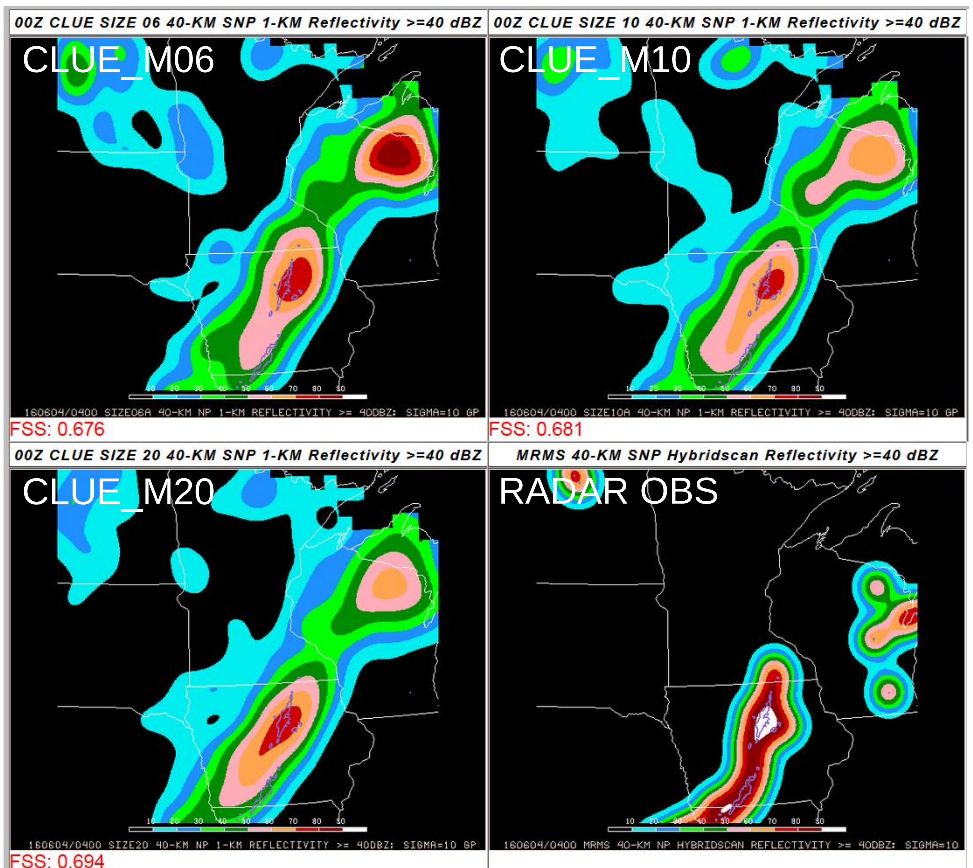
Model	IC	BC	Microphysics	LSM	PBL
ARW	NAMa	NAMf	Thompson	NOAH	MYJ
NMMB	NAMa	NAMf	Ferrier-Aligo	NOAH	MYJ
ARW	NAMa + arw-p1_pert	arw-p1	Thompson	NOAH	MYJ
NMMB	NAMa + nmmb-n1_pert	nmmb-n1	Ferrier-Alig[illegible]	NOAH	MYJ
ARW	NAMa + nmmb-p2_pert	nmmb-p2	The[illegible]	NOAH	MYJ
NMMB	NAMa + arw-p2_pert	arw-p2	[illegible]igo	NOAH	MYJ
ARW	NAMa + arw-n2_pert	arw-n2	[illegible]ompson	NOAH	MYJ
NMMB	NAMa + arw-n1_pert	[illegible]	Ferrier-Aligo	NOAH	MYJ
ARW	NAMa + nmmb-n1_pert	[illegible]	Thompson	NOAH	MYJ
NMMB	NAMa + nmmb-p2_[illegible]	[illegible]mmb-p2	Ferrier-Aligo	NOAH	MYJ
ARW	NAMa + arw[illegible]	arw-n1	Thompson	NOAH	MYJ
NMMB	NAM[illegible]_pert	arw-p3	Ferrier-Aligo	NOAH	MYJ
ARW	[illegible]w-p2_pert	arw-p2	Thompson	NOAH	MYJ
NMMB	[illegible]a + nmmb-p1_pert	nmmb-p1	Ferrier-Aligo	NOAH	MYJ
ARW	NAMa + arw-p3_pert	arw-p3	Thompson	NOAH	MYJ
NMMB	NAMa + nmmb-n2_pert	nmmb-n2	Ferrier-Aligo	NOAH	MYJ
ARW	NAMa + nmmb-p1_pert	nmmb-p1	Thompson	NOAH	MYJ
NMMB	NAMa + arw-p1_pert	arw-p1	Ferrier-Aligo	NOAH	MYJ
ARW	NAMa + nmmb-n2_pert	nmmb-n2	Thompson	NOAH	MYJ
NMMB	NAMa + arw-n2_pert	arw-n2	Ferrier-Aligo	NOAH	MYJ

Limited diversity using single-physics approach

Neighborhood Reflectivity Verification:

CLUE Size Comparison: Subjective Ratings

- During the SFE, participants subjectively compared the probability forecasts of simulated reflectivity >40 dBZ from the different ensembles.



Mesoscale Area of Interest
1-km AGL Reflect. >40 dBZ
20160603
00Z cycle; fh028

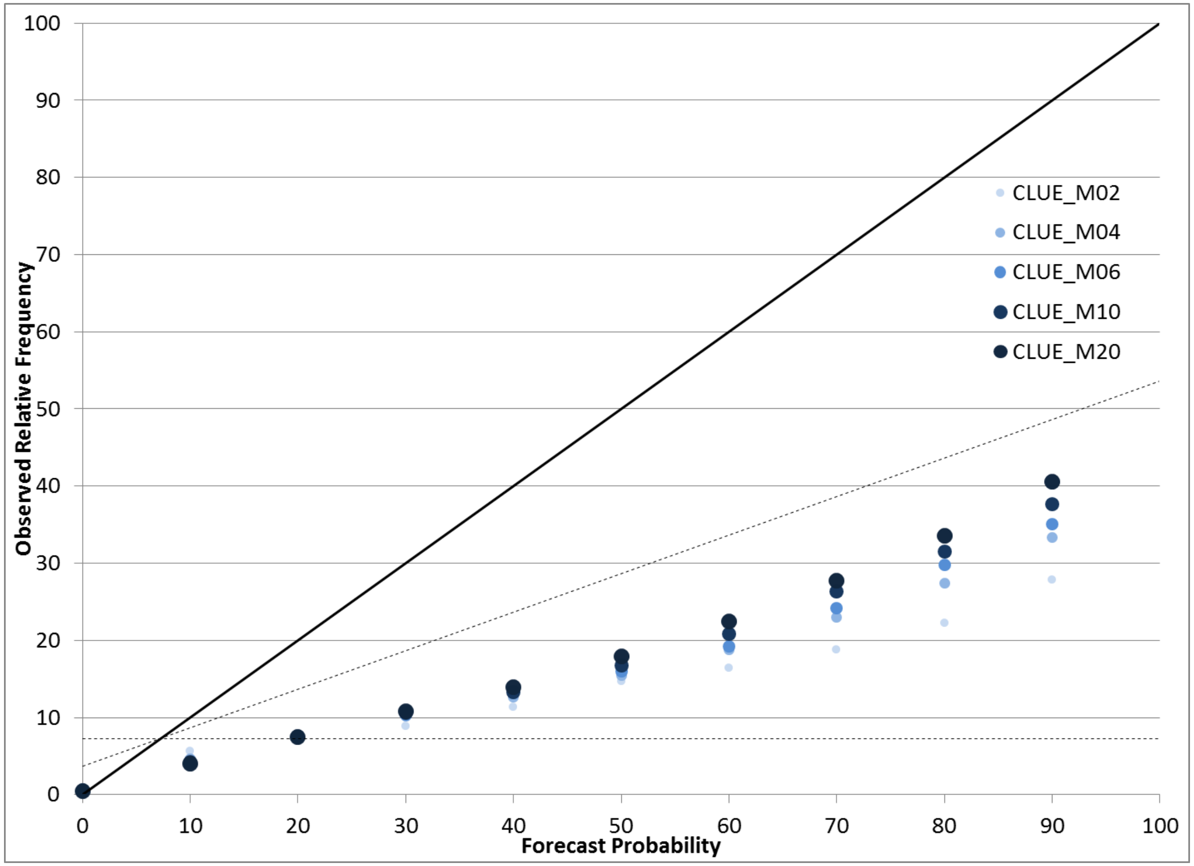
40-km Radius of Influence
10 grid-point Gaussian
smoothing parameter

CLUE_M06: 3 ARW, 3 NMMB
CLUE_M10: 5 ARW, 5 NMMB
CLUE_M20:10 ARW,10 NMMB

Neighborhood Reflectivity Verification:

CLUE Size Comparison: Reliability

- Reliability is similarly poor below 50% regardless of ensemble size. Some improvement in reliability for larger ensembles occurs in the higher probability bins.



CONUS
1-km AGL Reflect. >40 dBZ
20160502-20160603
00Z cycle; fh013-030

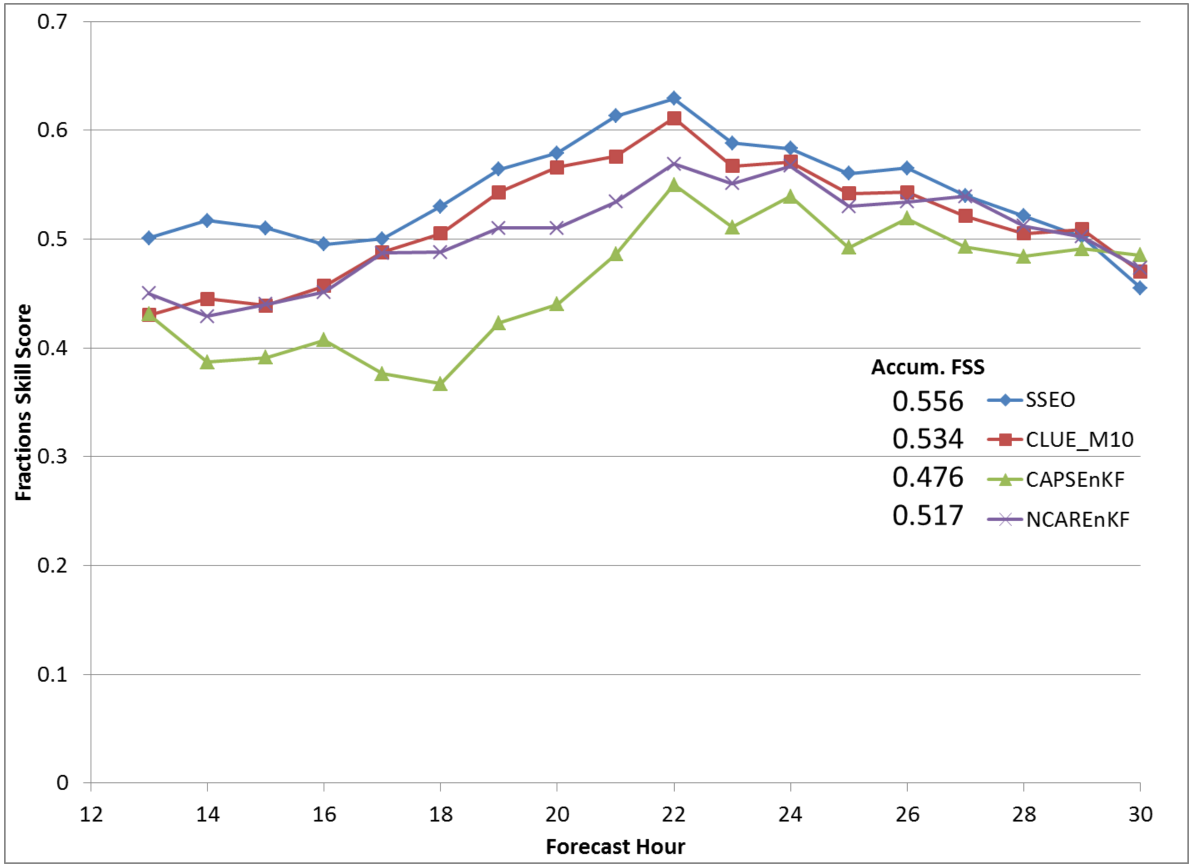
40-km Radius of Influence
10 grid-point Gaussian smoothing parameter

CLUE_M02: 1 ARW, 1 NMMB
CLUE_M04: 2 ARW, 2 NMMB
CLUE_M06: 3 ARW, 3 NMMB
CLUE_M10: 5 ARW, 5 NMMB
CLUE_M20:10 ARW,10 NMMB

Neighborhood Reflectivity Verification:

CLUE Comparison to SSEO: FSS

- For fractions skill score (FSS) of simulated reflectivity, the SSEO performed as well as any CLUE subset during SFE2016. The CLUE multi-core ensemble verified better than the single-core EnKF systems.



Mesoscale Area of Interest
1-km AGL Reflect. >40 dBZ
20160502-20160603
00Z cycle; fh013-030

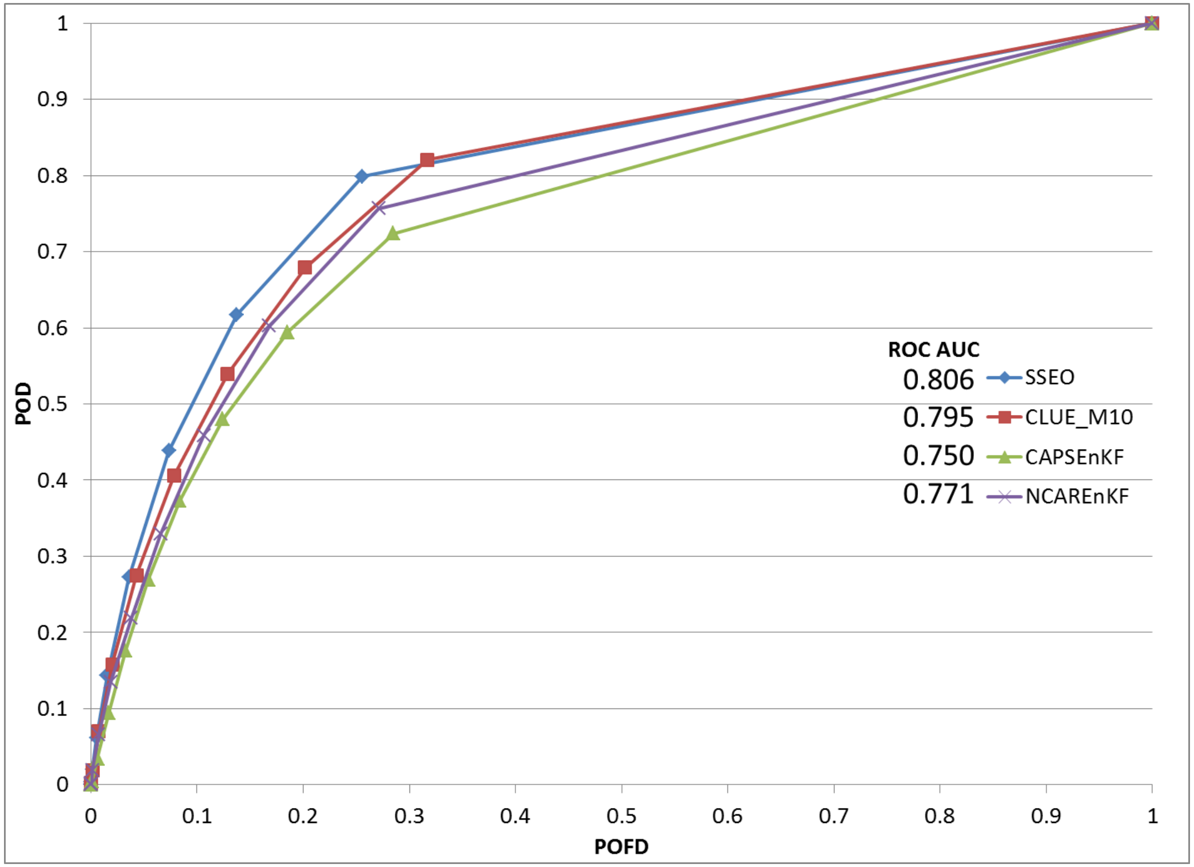
40-km Radius of Influence
10 grid-point Gaussian smoothing parameter

SSEO: 7-member multi-model
CLUE_M10: 5 ARW, 5 NMMB
CAPSEnKF: 9-member ARW
NCAREnKF:10-member ARW

Neighborhood Reflectivity Verification:

CLUE Comparison to SSEO: ROC

- The SSEO also had the largest area under the relative operating characteristic (ROC) curve for ensemble reflectivity forecasts during SFE2016, followed by the CLUE mixed-core ensemble.



Mesoscale Area of Interest
1-km AGL Reflect. >40 dBZ
20160502-20160603
00Z cycle; fh013-030

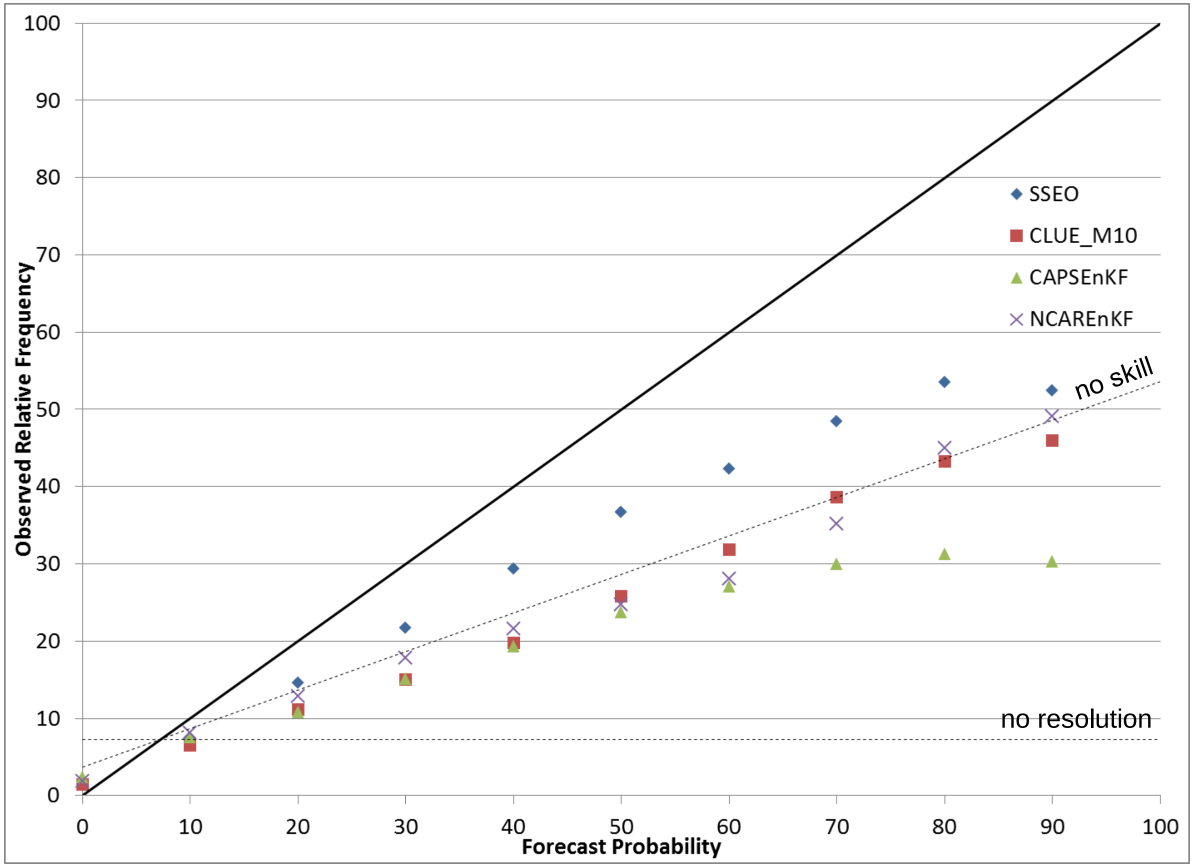
40-km Radius of Influence
10 grid-point Gaussian smoothing parameter

SSEO: 7-member multi-model
CLUE_M10: 5 ARW, 5 NMMB
CAPSEnKF: 9-member ARW
NCAREnKF:10-member ARW

Neighborhood Reflectivity Verification:

CLUE Comparison to SSEO: Reliability

- All ensembles produced overforecasts of reflectivity probabilities during SFE2016, though the SSEO was closer to being reliable than the other ensembles.



Mesoscale Area of Interest
1-km AGL Reflect. >40 dBZ
20160502-20160603
00Z cycle; fh013-030

40-km Radius of Influence
10 grid-point Gaussian smoothing parameter

SSEO: 7-member multi-model
CLUE_M10: 5 ARW, 5 NMMB
CAPSEnKF: 9-member ARW
NCAREnKF:10-member ARW