Final Report

National Weather Service Program Office (NWSPO)

NOAA Award NA16NWS4680002

A Partnership to Develop and Evaluate Optimized Realtime Convective-Scale Ensemble Data Assimilation and Prediction Systems for Hazardous Weather: Toward the Goals of a Weather-Ready Nation

PIs: Drs. **Ming Xue**, Fanyou Kong, Keith Brewster, Youngsun Jung, Nathan Snook

Center for Analysis and Prediction of Storms (CAPS)
The University of Oklahoma
120 David L. Boren Blvd, Norman, OK 73072

Reporting Period: 05/01/2016 - 04/30/2020

July 2020

Contents

1. ACCOMPLISHMENTS

The realtime convective-scale ensemble data assimilation and forecasting performed under this NOAA CSTAR project, together with retrospective analyses using the real time data, directly addresses the priority of the CSTAR program: "Improving the lead time and accuracy of forecasts and warnings for high impact weather events", and in particular "Utilizing of convection-allowing models and a storm scale ensemble system to advance Warn on Forecast capabilities". By producing and evaluating realtime storm-scale ensemble forecasts and participating NOAA testbeds HWT and HMT, this project aims to: 1) Improve application of NWP information in the forecast and warning process at Day 1 to Day 3 time scales; 2) Improve the use of ensemble predictions systems to: enable more effective forecaster assessment of uncertainty and historical context of potential high impact events; and develop probabilistic hazard information; 3) Provide direct guidance for the optimal design of an operational High Resolution Ensemble Forecast (HREF) system for NWS, as well as for determining the optimal DA strategy at the convective scale.

1.1 Summary of accomplishments

- Realtime SSEF contributed to NOAA 2016, 2017, 2018, 2019 HWT SFE
- Realtime SSEF contributed to NOAA 2016, 2017, 2018 HMT FFaIR
- GSI EnKF based radar assimilation experiment and examination
- Implemented and contributed FV3 and SAR-FV3 SSEF at CAM resolution
- QPF verifications from realtime testbed experiments
- Developed optimal ensemble consensus products such as localized probability matched mean precipitation (LPM)
- Authored/Co-authored over 25 journal papers, and dozens of conference presentations

1.2 CAPS realtime SSEF production in supporting HWT SFE

The Center for Analysis and Prediction of Storm (CAPS) produced multi-model multi-physics storm-scale ensemble forecasts (SSEF) at convection-allowing horizontal grid spacing of 3-4 km in realtime every year since 2007 from late April to early June to support the NOAA Hazardous Weather Testbed (HWT) Spring Forecasting Experiment (SFE). The primary funding came from the NOAA CSTAR grant as well as other NOAA grants. During current three-year reporting period from 2016 to 2019, CAPS remained a key contributor to the Community Leveraged Unified Ensemble (CLUE) for HWT SFE, implemented GSI EnKF data assimilation and ensemble forecast system. As WRF-ARW model remained as the primary modeling system in the SSEF, the operational model core NMMB was included in 2016 season, and the newly designated NGGPS

Finite-Volume Cubed-Sphere (FV3) model system, and its reginal stand-alone version SAR-FV3, were included since 2017 HWT SFE. Table 1 outlines the SSEF highlights from 2016 to 2019 for HWT SFEs under this CSTAR grant.

Table 1. CAPS HWT SSEF highlights for the reporting period

	2016	2017	2018	2019
member	24 (12)	24 (10)	40 (10)	15 (10)
Domain (grid)	CONUS (3 km)	CONUS (3 km)	CONUS (3 km)	CONUS (3 km)
Lead time	36-60 h	36-60 h	36-60 h	60 h
NWP models	ARW (v3.7.1) NMMB (6)	ARW (v3.8.1) FV3 (1)	ARW (v3.9.1.1) FV3 (12)	ARW (v4.0.3) SAR-FV3 (15)
EnKF	GSI+EnKF full domain	GSI+EnKF full domain	GSI+EnKF full domain	GSI+EnKF full domain

* Numbers in brackets in the member row refer to GSI+EnKF ensemble member; Non-ARW model forecast member counts are shown in brackets in the NWP row.

1.2.1 CAPS SSEF overview for NOAA/HWT SFE2016

The CAPS 2016 Storm-Scale Ensemble Forecast for the SFE2016 started on 18 April through 3 June 2016, encompassing the NOAA HWT 2016 Spring Forecasting Experiment that was officially between 2 May and 3 June. The regular 0000 UTC 3-km ensembles consist of 18 WRF-ARW members initialized with a onetime 3DVAR analysis, with the forecast lead time of 60 hours Figure 1 shows the model domains (both ARW and NMMB) used in the 2016 season.

Operational NMMB model core was added in 2016 season, with 6 members. Only one NMMB member (nmmb_cn) had radar data analysis. Other five non-radar NMMB members had IC and LBC perturbations provided from SREF perturbed members. All NMMB members used a fixed set of physics configuration matching the NCEP operational high-res NMMB runs.

The CAPS non-cycled 3DVAR-based SSEF and the cycled GSI+EnKF based SSEF during NOAA/HWT SFE2016 also contribute into a larger Community Leveraged Unified Ensemble (CLUE) coordinated among various groups including NSSL, SPC, CAPS, NCAR, UND, EMC, GSD, and DTC, in an effort to provide guidance to the design of near-future operational SSEF systems.

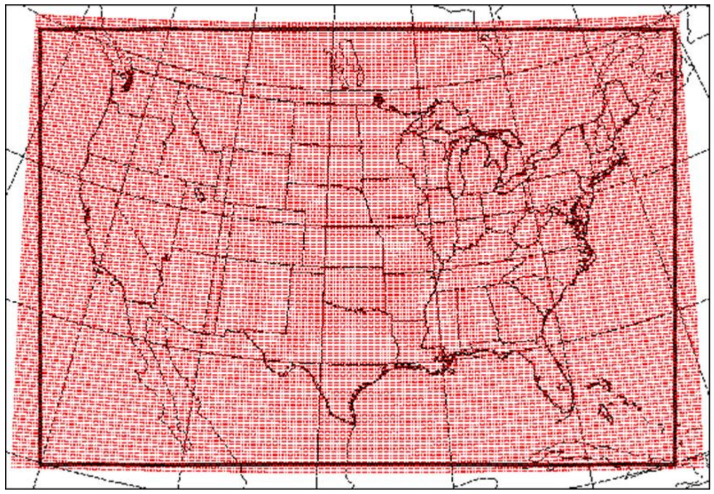

Figure 1. The 2016 CAPS Spring Forecast Experiment domains, with ARW domain marked by thick lines (1680x1152 at 3 km) and the NMMB domain marked as red dots.

Tables 1 lists member configurations for the 3DVAR initialized SSEF ensemble (non-cycled 3DVAR-based SSEF). As part of CLUE (Clark et al. 2016), all 3DVAR based ARW members are grouped into two sub-ensembles. The control member cn and members m3-m10 are members with both IC/LBC perturbation and physics variations. This group is called *mixed* sub-ensemble. The group containing members cn and m11 – m19 is called single physics (*s_phys*) sub-ensemble, among them m11-m19 do have IC and LBC perturbations. In Table 1, NAMa and NAMf refer to the NCEP operational 12 km NAM analysis and forecast, respectively. ARPSa refers to analysis after ARPS 3DVAR and Cloud Analysis using NAMa as the background.

Table 2 lists member configurations for the NMMB members. Only one NMMB member (nmmb_cn) has radar data analysis. Other five non-radar NMMB members have IC and LBC perturbations. All NMMB members use a fixed set of physics configuration matching the operational high-res NMMB runs.

The mixed ensemble members run for 60 h in forecast duration; whereas s_phys members and all NMMB members run for 36 h.

Table 2. Configurations of ARW members for SFE2016. 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_m3	arw_cn + arw-p1_pert	21Z SREF arw-p1	yes	P3	Noah	YSU

arw_m4	arw_cn + arw-n1_pert	21Z SREF arw-n1	yes	MY	Noah	MYNN
arw_m5	arw_cn + arw-p2_pert	21Z SREF arw-p2	yes	Morrison	Noah	MYJ
arw_m6	arw_cn + arw-n2_pert	21Z SREF arw-n2	yes	P3	Noah	YSU
arw_m7	arw_cn + nmmb-p1_pert	21Z SREF nmmb-p1	yes	MY	Noah	MYNN
arw_m8	arw_cn + nmmb-n1_pert	21Z SREF nmmb-n1	yes	Morrison	Noah	YSU
arw_m9	arw_cn + nmmb-p2_pert	21Z SREF nmmb-p2	yes	P3	Noah	MYJ
arw_m10	arw_cn + nmmb-n2_pert	21Z SREF nmmb-n2	yes	Thompson	Noah	MYNN
arw_m11	arw_cn + arw-p1_pert	21Z SREF arw-p1	yes	Thompson	Noah	MYJ
arw_m12	arw_cn + arw-n1_pert	21Z SREF arw-n1	yes	Thompson	Noah	MYJ
arw_m13	arw_cn + arw-p2_pert	21Z SREF arw-p2	yes	Thompson	Noah	MYJ
arw_m14	arw_cn + arw-n2_pert	21Z SREF arw-n2	yes	Thompson	Noah	MYJ
arw_m15	arw_cn + arw-p3_pert	21Z SREF arw-p3	yes	Thompson	Noah	MYJ
arw_m16	arw_cn + nmmb-p1_pert	21Z SREF nmmb-p1	yes	Thompson	Noah	MYJ
arw_m17	arw_cn + nmmb-n1_pert	21Z SREF nmmb-n1	yes	Thompson	Noah	MYJ
arw_m18	arw_cn + nmmb-p2_pert	21Z SREF nmmb-p2	yes	Thompson	Noah	MYJ
arw_m19	arw_cn + nmmb-n2_pert	21Z SREF nmmb-n2	yes	Thompson	Noah	MYJ

** For all members: ra_lw_physics= RRTMG; ra_sw_physics=RRTMG; cu_physics= NONE.*

Table 3. Configurations of NMMB members for SFE2016

member	IC	BC	Radar data	mp_phy	lw_phy	sw-phy	sf_phy
nmmb_cn	00Z ARPSa	00Z NAMf	yes	Ferrier-Aligo	RRTMG	RRTMG	Noah
nmmb_m1	00Z NAMa+ arw-p3_pert	21Z SREF arw-p3	no	Ferrier-Aligo	RRTMG	RRTMG	Noah

nmmb_m2	00Z NAMa+ nmmb-p1_pert	21Z SREF nmmb-p1	no	Ferrier-Aligo	RRTMG	RRTMG	Noah
nmmb_m3	00Z NAMa+ nmmb-n1_pert	21Z SREF nmmb-n1	no	Ferrier-Aligo	RRTMG	RRTMG	Noah
nmmb_m4	00Z NAMa+ nmmb-p2_pert	21Z SREF nmmb-p2	no	Ferrier-Aligo	RRTMG	RRTMG	Noah
nmmb_m5	00Z NAMa+ nmmb-n2_pert	21Z SREF nmmb-n2	no	Ferrier-Aligo	RRTMG	RRTMG	Noah

**For all members: pbl_physics=MYJ; cu_physics= NONE.*

1.2.2 CAPS SSEF overview for NOAA/HWT SFE2017

The CAPS 2017 Storm-Scale Ensemble Forecast for the SFE2017 started on 11 April through 2 June 2016, encompassing the NOAA HWT 2017 Spring Forecasting Experiment that was officially between 1 May and 2 June. The CAPS SSEF consisted of a 23-member (consisting of ARW) SSEF initiated at 0000 UTC with non-cycled 3DVAR analysis and running 60-h (36-h for single physics members and the MP-only members) forecasting, and a GSI+EnKF cycled ensemble procedure followed by a 10-member 48-h ARW SSEF starting at 0000 UTC.

The 23-member 3DVAR-based SSEF and 10-member GSI+EnKF based SSEF in the HWT SFE2017, as well as one single FV3 forecast, also contribute into the large coordinated Community Leveraged Unified Ensemble (CLUE), the “grand ensemble”, with other contributing institutions from NSSL, NCAR, and GSD, in an effort to provide guidance to the design of near-future operational SSEF systems. CAPS contributed 4 ensemble groups out of total 9 CLUE groups in 2017 season and one single FV3 3-km CONUS forecast nested in a 13-km global forecast, using CAPS implemented Thompson microphysics.

Figure 2 shows the model domains used in the 2017 season. The slight shrinking of 3-km domain for ARW (from 1680x1152 in 2016 to 1620x1120 in 2017) is to have the domain to fit within the operational RAP analysis data coverage in order to extract initial condition background for one HRRR member (m2 in Table 3). The HRRR member was run with ARPS3DVAR radar analysis, in comparison with the operational HRRR that uses MRMS radar mosaic with latent heat analysis.

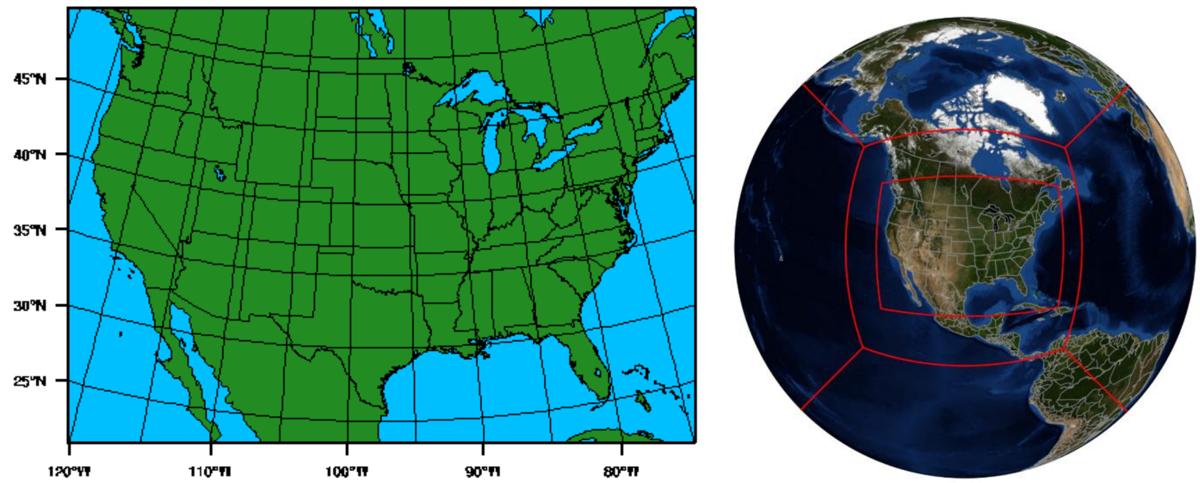

Figure 2. The 2017 CAPS Spring Forecast Experiment domains. Left: The ARW domain at 3-km grid spacing, consisting of 1620×1120 horizontal grid points; Right: FV3 cubed-sphere grid nesting domains, stretched to provide higher resolution on the face covering CONUS, and a nested 1728×1296 domain (inner red box) with ~3 km grid spacing.

Tables 3 lists member configurations for the 3DVAR initialized SSEF ensemble (non-cycled 3DVAR-based SSEF). As part of CLUE, all 3DVAR based ARW members are grouped into three sub-ensembles. The control member cn and members m2-m10 are members with both IC/LBC perturbation and physics variations. This group is called *mixed* sub-ensemble. The group containing members m02 and m11 – m19 is called single physics (*s_phys*) sub-ensemble, among them m11-m19 do have IC and LBC perturbations. The group m20 – m23 is called mixed microphysics that only has microphysics variation. In Table 3, NAMa and NAMf refer to the NCEP operational 12 km NAM analysis and forecast, respectively. RAPa and GFSf refer to 00Z RAP analysis and 18Z GFS forecast, respectively. ARPSa refers to analysis after ARPS 3DVAR and Cloud Analysis using NAMa as the background.

Table 3. Configurations of ARW members for SFE2017. NAMa and NAMf refer to 12 km NAM analysis and forecast, respectively. RAPa and GFSf refer to 00Z RAP analysis and 18Z GFS 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_m2	RAPa+3DVAR	18Z GFSf	yes	Thompson	RUC	MYNN
arw_m3	arw_cn + arw-p1_pert	21Z SREF arw-p1	yes	P3	Noah	YSU
arw_m4	arw_cn + arw-n1_pert	21Z SREF arw-n1	yes	MY	Noah	MYJ

arw_m5	arw_cn + nmmb-p1_pert	21Z SREF nmmb-p1	yes	Morrison	Noah	MYJ
arw_m6	arw_cn + nmmb-n1_pert	21Z SREF nmmb-n1	yes	P3	Noah	YSU
arw_m7	arw_cn + arw-p2_pert	21Z SREF arw-p2	yes	MY	Noah	YSU
arw_m8	arw_cn + arw-n2_pert	21Z SREF arw-n2	yes	Morrison	Noah	YSU
arw_m9	arw_cn + nmmb-p2_pert	21Z SREF nmmb-p2	yes	P3	Noah	MYJ
arw_m10	arw_cn + nmmb-n2_pert	21Z SREF nmmb-n2	yes	Thompson	Noah	MYNN
arw_m11	00Z ARPSa	00Z NAMf	yes	Thompson	RUC	MYNN
arw_m12	arw_cn + arw-p1_pert	21Z SREF arw-p1	yes	Thompson	RUC	MYNN
arw_m13	arw_cn + arw-n1_pert	21Z SREF arw-n1	yes	Thompson	RUC	MYNN
arw_m14	arw_cn + nmmb-p1_pert	21Z SREF nmmb-p1	yes	Thompson	RUC	MYNN
arw_m15	arw_cn + nmmb-n1_pert	21Z SREF nmmb-n1	yes	Thompson	RUC	MYNN
arw_m16	arw_cn + arw-p2_pert	21Z SREF arw-p2	yes	Thompson	RUC	MYNN
arw_m17	arw_cn + arw-n2_pert	21Z SREF arw-n2	yes	Thompson	RUC	MYNN
arw_m18	arw_cn + nmmb-p2_pert	21Z SREF nmmb-p2	yes	Thompson	RUC	MYNN
arw_m19	arw_cn + nmmb-n2_pert	21Z SREF nmmb-n2	yes	Thompson	RUC	MYNN
arw_m20	RAPa+3DVAR	18Z GFSf	yes	Morrison	RUC	MYNN
arw_m21	RAPa+3DVAR	18Z GFSf	yes	MY	RUC	MYNN
arw_m22	RAPa+3DVAR	18Z GFSf	yes	P3	RUC	MYNN
arw_m23	RAPa+3DVAR	18Z GFSf	yes	WSM6	RUC	MYNN

** For all members: ra_lw_physics= RRTMG; ra_sw_physics=RRTMG; cu_physics= NONE.*

Table 4. Configurations FV3 forecast for SFE2017

member	IC	BC	Radar data	mp_phy	lw_phy	sw-phy	LSM	PBL
fv3_cn	00Z GFSa	-	no	Thompson	GFDL	GFDL	Noah	YSU

Table 4 lists member configuration for the single FV3 core member ran during SFE2017.

The mixed ensemble members run for 60 h in forecast duration; whereas s_phys members and all mixed microphysics members run for 36 h. FV3 member is planned to run 120 h during HWT and 84-h during HMT.

1.2.3 CAPS SSEF overview for NOAA/HWT SFE2018

The CAPS 2018 Storm-Scale Ensemble Forecast for the SFE2018 started on 30 April through 1 June 2018. Three ensemble suites were configured and run for NOAA/HWT SFE2018: a 28-member (consisting of ARW v3.9.1.1) SSEF initiated at 0000 UTC with non-cycled 3DVAR analysis and running 60-h (36-h for single physics members and the stochastic members) forecasting, a GSI+EnKF cycled ensemble procedure followed by a 10-member 48-h ARW SSEF starting at 0000 UTC, and a newly FV3 core based multi-physics ensemble of 12 members. The ARW domain was the same as in 2017 season (see Figure 2, left panel). Figure 3 shows the model domain for FV3 used in 2018 season.

In HWT SFE2018, CAPS contributed 5 ensembles out of total 8 CLUE ensemble groups. Newly added is a 12-member FV3 core 3.5-km CONUS ensemble forecasts nested in a 13-km global forecast, using CAPS implemented Thompson and NSSL microphysics and MYNN and YSU PBL schemes.

The 3DVAR-based ensembles and FV3 core ensemble were run on TACC Stampede2 for HWT SFE2018 and for FFaIR2018. The EnKF-based ensembles were run on Bridges at Pittsburg Supercomputing Center (PSC).

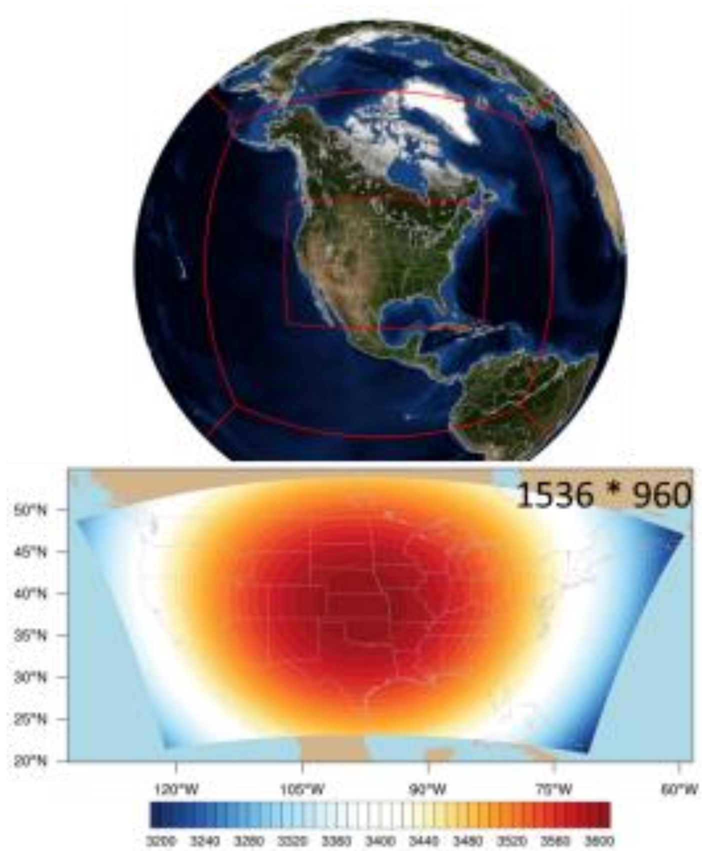

Figure 3. FV3 domain in 2018 season. Top: Uniform (13 km) global domain, with the innermost domain marking the fine nested domain area; Bottom: Coverage of the nested domain (~3.5 km), with color shades indicating grid spacing in meter.

Tables 5 lists member configurations for the 3DVAR initialized WRF SSEF ensemble (non-cycled 3DVAR-based SSEF). As part of CLUE, all 3DVAR based ARW members are grouped into three sub-ensembles. The control member cn and members m2-m10 are members with both IC/LBC perturbation and physics variations. This group is called *mixed physics* + *radar* sub-ensemble. The group containing members m11 – m18 is called *single physics* + *radar* sub-ensemble. The group containing members m19 – m26 is called *stochastic physics* + *radar* sub-ensemble. All three groups are with IC and LBC perturbations. In Table 5, NAMa and NAMf refer to the NCEP operational 12 km NAM analysis and forecast, respectively. RAPa and GFSf refer to 00Z RAP analysis and 18Z GFS forecast, respectively. ARPSa refers to analysis after ARPS 3DVAR and Cloud Analysis using NAMa as the background.

Table 6 lists member configuration for the newly implemented FV3 core ensemble members.

The mixed ensemble members run for 60 h in forecast duration; whereas single physics and stochastic physics members run for 36 h. FV3 member is planned to run 84 h during HWT and HMT.

Table 5. Configurations of ARW members for SFE2018. NAMa and NAMf refer to 12 km NAM analysis and forecast, respectively. RAPa and GFSf refer to 00Z RAP analysis and 18Z GFS forecast, respectively. ARPSa refers to ARPS 3DVAR and cloud analysis

Member	IC	BC	Radar data	Microphy	LSM	PBL
arw_cn0*	00Z ARPSa	00Z NAMf	yes	Thompson	Noah	MYJ
arw_cn	00Z ARPSa	00Z NAMf	yes	Thompson	Noah	MYJ
arw_m2	RAPa+3DVAR	18Z GFSf	yes	Thompson	RUC	MYNN
arw_m3	arw_cn + arw-p1_pert	21Z SREF arw-p1	yes	NSSL	Noah	YSU
arw_m4	arw_cn + arw-n1_pert	21Z SREF arw-n1	yes	NSSL	Noah	MYNN
arw_m5	arw_cn + nmmb-p1_pert	21Z SREF nmmb-p1	yes	Morrison	Noah	MYJ
arw_m6	arw_cn + nmmb-n1_pert	21Z SREF nmmb-n1	yes	P3	Noah	YSU
arw_m7	arw_cn + arw-p2_pert	21Z SREF arw-p2	yes	NSSL	Noah	MYJ
arw_m8	arw_cn + arw-n2_pert	21Z SREF arw-n2	yes	Morrison	Noah	YSU
arw_m9	arw_cn + nmmb-p2_pert	21Z SREF nmmb-p2	yes	P3	Noah	MYNN
arw_m10	arw_cn + nmmb-n2_pert	21Z SREF nmmb-n2	yes	Thompson	Noah	MYNN
arw_m11	00Z ARPSa	00Z NAMf	yes	Thompson	RUC	MYNN
arw_m13	arw_cn + arw-p1_pert	21Z SREF arw-p1	yes	Thompson	RUC	MYNN
arw_m14	arw_cn + arw-n1_pert	21Z SREF arw-n1	yes	Thompson	RUC	MYNN
arw_m15	arw_cn + nmmb-p1_pert	21Z SREF nmmb-p1	yes	Thompson	RUC	MYNN
arw_m16	arw_cn + nmmb-n1_pert	21Z SREF nmmb-n1	yes	Thompson	RUC	MYNN
arw_m17	arw_cn + arw-p2_pert	21Z SREF arw-p2	yes	Thompson	RUC	MYNN
arw_m18	arw_cn + arw-n2_pert	21Z SREF arw-n2	yes	Thompson	RUC	MYNN

arw_m19	00Z ARPSa	00Z NAMf	yes	Thompson	RUC	MYNN
arw_m20	RAPa+3DVAR	18Z GFSf	yes	Thompson	RUC	MYNN
arw_m21	arw_cn + arw-p1_pert	21Z SREF arw-p1	yes	Thompson	RUC	MYNN
arw_m22	arw_cn + arw-n1_pert	21Z SREF arw-n1	yes	Thompson	RUC	MYNN
arw_m23	arw_cn + nmmb-p1_pert	21Z SREF nmmb-p1	yes	Thompson	RUC	MYNN
arw_m24	arw_cn + nmmb-n1_pert	21Z SREF nmmb-n1	yes	Thompson	RUC	MYNN
arw_m25	arw_cn + arw-p2_pert	21Z SREF arw-p2	yes	Thompson	RUC	MYNN
arw_m26	arw_cn + arw-n2_pert	21Z SREF arw-n2	yes	Thompson	RUC	MYNN
arw_m27	RAPa+3DVAR	18Z GFSf	yes	Thompson	RUC	MYNN

Note 1: For all members: ra_lw_physics= RRTMG; ra_sw_physics=RRTMG; cu_physics=none
Note 2: arw_cn0 is the same as arw_cn, except with non-HRRR vertical levels and no smoothing
Note 3: arw_m19 ~ arw_m26 (dark shading) are with stochastic perturbation turned on (spp_mp=1, spp_pbl=1)
Note 4: arw_m27 is with Thompson stochastic setting on (spp_mp=7, spp_pbl=0)
Note 5: arw_12 is the same as arw_cn, so counted as one member

Table 6. Configurations of FV3 members for SFE2018

member	IC	mp_phy	PBL	Cumulus
fv3_m01	00Z GFSa	Thompson	MYNN-SA	Tiedtke
fv3_m02	00Z GFSa	Thompson	MYNN	Tiedtke
fv3_m03	00Z GFSa	Thompson	YSU-SA	Tiedtke
fv3_m04	00Z GFSa	Thompson	YSU	Tiedtke
fv3_m05	00Z GFSa	Thompson	EDMF	Tiedtke
fv3_m06	00Z GFSa	NSSL	MYNN-SA	Tiedtke

fv3_m07	00Z GFSa	NSSL	MYNN	Tiedtke
fv3_m08	00Z GFSa	NSSL	YSU-SA	Tiedtke
fv3_m09	00Z GFSa	NSSL	YSU	Tiedtke
fv3_m10	00Z GFSa	NSSL	EDMF	Tiedtke
fv3_m11	00Z GFSa	Thompson	MYNN-SA	SA-SAS
fv3_m12	00Z GFSa	GFDL	EDMF	SA-SAS

1.2.4 CAPS SSEF overview for NOAA/HWT SFE2019

NOAA HWT Spring Forecast Experiment (NOAA/HWT SFE2019) formally spanned from April 29 to May 31, 2019, essentially beyond this CSTAR project duration from May 1, 2016 to April 30, 2019. However, a one year NCE (No Cost Extension) was requested and granted to extend SFE2019 realtime support. The justification for the NCE is "to use the remaining funding to support HWT Spring Experiment forecasts in May and June 2019, and evaluate the forecasts".

In 2019, CAPS's effort to contribute to HWT and HMT was shifting from WRF-ARW model core to SAR-FV3 (the standalone reginal FV3 version) based ensembles. With the new development of the SAR-FV3 model core, there was a strong desire by NWS to switch the testing of FV3 CAM forecasting to SAR-FV3. During April 15 through May 31, 2019, CAPS contributed several physics schemes via CCPP into SAR- FV3 and tested them before 2019 HWT SFE. During NOAA/HWT SFE2019, CAPS ran 15 SAR FV3 3-km CONUS domain forecasts, with 9 member using multiple physics and the same IC, and 6 members including SREF-derived IC and LBC perturbations.

Figure 4 shows the model domains used in the 2019 season.

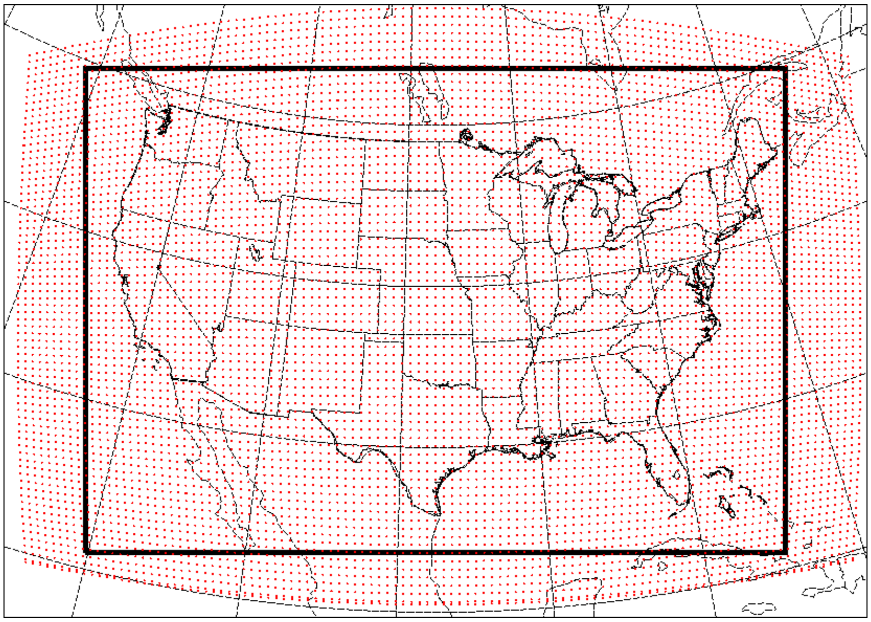

Figure 4. The computational domains for the 2019 Season. The inner rectangular is the WRF-ARW (CLUE) domain at 3-km grid spacing, consisting of 1620×1120 horizontal grid points (the same as in Figure 2 left panel). The SAR-FV3 3-km CONUS domain is marked with red dots, consisting of 1921×1297 horizontal grid points (at roughly 3-km grid spacing).

Tables 7 and 8 list SAR-FV3 member configurations for the newly implemented Standalone Reginal FV3 core ensemble members.

Table 7. Configurations of physics-perturbation-only SAR-FV3 SSEF members for SFE2019. The first 7 members used SAR-FV3 with operational NAM analysis and forecasts as IC and LBCs. The last two members use a 3 km grid nested within a global grid and using GFS analysis as IC

Member	IC/LBC	Microphysics	PBL	SFC layer	LSM	Radiation
Core-ctrl	NAM	Thompson	saMYNN	GFS	NOAH	RRTMG
Core-pbl1	NAM	Thompson	saShinHong	GFS	NOAH	RRTMG
Core-pbl2	NAM	Thompson	EDMF	GFS	NOAH	RRTMG
Core-mp1	NAM	NSSL	saMYNN	GFS	NOAH	RRTMG
Core-mp2	NAM	Morrison-G.	saMYNN	GFS	NOAH	RRTMG
Core-lsm	NAM	Thompson	saMYNN	GFS	RUC	RRTMG

Core-sfcl	NAM	Thompson	saMYNN	MYNN	RUC	RRTMG
Core-globalgfs	GFS	Thompson	saMYNN	GFS	NOAH	RRTMG
Core-sargfs	GFS	Thompson	saMYNN	GFS	NOAH	RRTMG

Table 8. Configurations of mixed SAR-FV3 SSEF members for SFE2019, including both physics and IC/LBC perturbations. The IC and LBC perturbations were derived from the operational SREF forecasts

Member	IC/LBC	Microphysics	PBL	SFC layer	LSM	Radiation
Pert-pbl1	NAM+SREF arwn1	Thompson	saShinHong	GFS	NOAH	RRTMG
Pert-pbl2	NAM+SREF arwp2	Thompson	EDMF	GFS	NOAH	RRTMG
Pert-mp1	NAM+SREF arwp1	NSSL	saMYNN	GFS	NOAH	RRTMG
Pert-mp2	NAM+SREF arwn2	Morrison-G.	saMYNN	GFS	NOAH	RRTMG
Pert-lsm	NAM+SREF arwp3	Thompson	saMYNN	GFS	RUC	RRTMG
Pert-sfcl	NAM+SREF arwn3	Thompson	saMYNN	MYNN	RUC	RRTMG

1.3 CAPS realtime SSEF production in supporting HMT FFaIR

Since 2016, CAPS actively participated the NOAA Hydrometeorological Testbed (HMT) Flash Flood and Intensive Rainfall (FFaIR) Experiment by contributing a convection-allowing 3-km grid SSEF of 60-84 h long from June through July each summer during the 4-week-long FFaIR experiment period (two weeks before and two weeks after the July 4th week, with the July 4th week off). Ensemble products (mainly QPF and PQPF products) were provided to the HMT in GEMPAK format, and GRIB2 since 2018. They include neighborhood probability of QPF exceedances of Flash Flood Guidance (FFG) and Recurrence Intervals (RI), as well as QPF probability matched means (PM) of various accumulation lengths (3-, 6-, 12-, and 24-h). CAPS also developed a novel localized PM QPF algorithm for HMT FFaIR, called LPM, which possesses superior characteristics to conventional PM and remains computationally efficient (Snook et al. 2019, 2020).

This CSTAR award reporting period covered CAPS HMT FFaIR efforts for 2016, 2017, and 2018.

The NOAA/HMT FFaIR2016 CAPS SSEF consist of 15 members of non-cycled 3DVAR-based ensemble configured with 13 ARW and 2 NMMB members (Tables 9 and 10). All model domains and data flow are the same as in the HWT component.

Table 9. ARW members for the NOAA/HMT FFaIR2016. 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_m2	arw_cn + arw-p1_pert	21Z SREF arw-p1	yes	Morrison	Noah	MYNN
arw_m3	arw_cn + arw-n1_pert	21Z SREF arw-n1	yes	MY	Noah	MYNN
arw_m4	arw_cn + arw-p2_pert	21Z SREF arw-p2	yes	Morrison	Noah	MYJ
arw_m5	arw_cn + arw-n2_pert	21Z SREF arw-n2	yes	Thompson	Noah	MYNN
arw_m6	arw_cn + nmmb-p1_pert	21Z SREF nmmb-p1	yes	MY	Noah	MYNN
arw_m7	arw_cn + nmmb-n1_pert	21Z SREF nmmb-n1	yes	Morrison	Noah	MYNN
arw_m8	arw_cn + nmmb-p2_pert	21Z SREF nmmb-p2	yes	Morrison	Noah	MYJ
arw_m9	arw_cn + nmmb-n2_pert	21Z SREF nmmb-n2	yes	Thompson	Noah	MYNN
arw_m10	00Z ARPSa	00Z NAMf	yes	P3	Noah	MYJ
arw_m11	00Z ARPSa	00Z NAMf	yes	Morrison	Noah	MYJ
arw_m12	00Z ARPSa	00Z NAMf	yes	MY	Noah	MYJ
arw_m13	arw_cn + arw-n2_pert	21Z SREF arw-n2	yes	Thompson	Noah	MYJ

Table 10. NMMB members for NOAA/HMT FFaIR2016

member	IC	BC	Radar data	mp_phy	lw_phy	sw-phy	sf_phy
nmmb_cn	00Z ARPSa	00Z NAMf	yes	Ferrier-Aligo	RRTMG	RRTMG	Noah

nmmb_m1	00Z NAMa+ arw-p3_pert	21Z SREF arw-p3	no	Ferrier-Aligo	RRTMG	RRTMG	Noah

The 2016 FFaIR Report ranked CAPS SSEF on the top with respect to Day 1 Subjective QPF scores and a close second to the NAMRR for the Day 2 QPF scores.

The NOAA/HMT FFaIR2017 CAPS SSEF consist of 10-member (ARW) non-cycled 3DVAR-based ensemble consisting IC/LBC perturbations and mixed physics options, and one single FV3 core forecast that is the same as in HWT (Tables 11 and 12). The FV3 forecasts were 84 h long.

Table 11. ARW members for the NOAA/HMT FFaIR2017. 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_m2	arw_cn + arw-p1_pert	21Z SREF arw-p1	yes	Morrison	Noah	MYNN
arw_m3	arw_cn + arw-n1_pert	21Z SREF arw-n1	yes	MY	Noah	MYJ
arw_m4	arw_cn + arw-p2_pert	21Z SREF arw-p2	yes	Morrison	Noah	MYJ
arw_m5	arw_cn + arw-n2_pert	21Z SREF arw-n2	yes	Thompson	Noah	MYNN
arw_m6	arw_cn + nmmb-p1_pert	21Z SREF nmmb-p1	yes	MY	Noah	YSU
arw_m7	arw_cn + nmmb-n1_pert	21Z SREF nmmb-n1	yes	Morrison	Noah	MYNN
arw_m8	arw_cn + nmmb-p2_pert	21Z SREF nmmb-p2	yes	Morrison	Noah	MYJ
arw_m9	arw_cn + nmmb-n2_pert	21Z SREF nmmb-n2	yes	Thompson	Noah	MYNN
arw_m10	arw_cn + arw-n3_pert	21Z SREF arw-n3	yes	Thompson	Noah	MYJ

Table 12. FV3 members for NOAA/HMT FFaIR2017

member	IC	BC	Radar data	mp_phy	lw_phy	sw-phy	LSM	PBL
fv3_cn	00Z GFSa	-	no	Thompson	GFDL	GFDL	Noah	YSU

Table 13 lists SSEF members for the NOAA/HMT FFaIR2018, with a total 15-member (13 ARW and 2 FV3) non-cycled 3DVAR-based ensemble consisting IC/LBC perturbations and mixed physics options. All model domains and data flow are the same as in the HWT component.

Table 13. SSEF members for the NOAA/HMT FFaIR2018. 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_m2	arw_cn + arw-p1_pert	21Z SREF arw-p1	yes	NSSL	Noah	YSU
arw_m3	arw_cn + arw-n1_pert	21Z SREF arw-n1	yes	NSSL	Noah	MYNN
arw_m4	arw_cn + arw-p2_pert	21Z SREF arw-p2	yes	NSSL	Noah	MYJ
arw_m5	arw_cn + arw-n2_pert	21Z SREF arw-n2	yes	Morrison	Noah	YSU
arw_m6	arw_cn + nmmb-p1_pert	21Z SREF nmmb-p1	yes	Morrison	Noah	MYJ
arw_m7	arw_cn + nmmb-n1_pert	21Z SREF nmmb-n1	yes	P3	Noah	YSU
arw_m8	arw_cn + nmmb-p2_pert	21Z SREF nmmb-p2	yes	P3	Noah	MYNN
arw_m9	arw_cn + nmmb-n2_pert	21Z SREF nmmb-n2	yes	Thompson	Noah	MYNN
arw_m10	arw_cn + arw-n3_pert	21Z SREF arw-n3	yes	Thompson	Noah	MYJ
arw_m11	00Z ARPSa	00Z NAMf	yes	Morrison	Noah	MYJ
arw_m12	00Z ARPSa	00Z NAMf	yes	P3	Noah	MYJ
arw_m13	00Z ARPSa	00Z NAMf	yes	NSSL	Noah	MYJ
fv3_m14	GFS	-	no	Thompson	Noah	MYNN
fv3_m15	GFS	-	no	NSSL	Noah	MYNN

An example flash flood forecast case during FFaIR is the West Virginia Flash Flood occurred during 23-24 June 2016. It measured max gauge of 9.37 inches at Maxwelton, WV. The Elk River high reached 33.37 ft. The consequence is 23 fatalities with 15 in Greenbrier Co alone. 44 of 55 WV counties were placed in State Emergency. The CAPS HMT FFaIR2016 phase did capture the entire episode during the regular operation period.

Figure 5 shows the Multi-Radar Multi-Sensor (MRMS) precipitation estimation (QPE) of the 12 h, and 24 h accumulated precipitation valid at 00 UTC June 24, 2016. Figures 6 and 7 plot the 24 h and 12 h QPF in the form of probability matching mean from CAPS HMT SSEF. The forecasted heavy 24 h precipitation maxima over WV in Figure 6 are 378 and 337 mm, respectively, compared to MRMS's 623 mm in Figure 6b. The PM forecast values are more close to the max gauge recorded (238mm) in Maxwelton, WV. Figures 8 and 9 are example neighborhood probabilities. They demonstrate great values for the CAPS SSEF's ability to predict the intensive flash flood occurrence 24 h or even 48 h in advance,

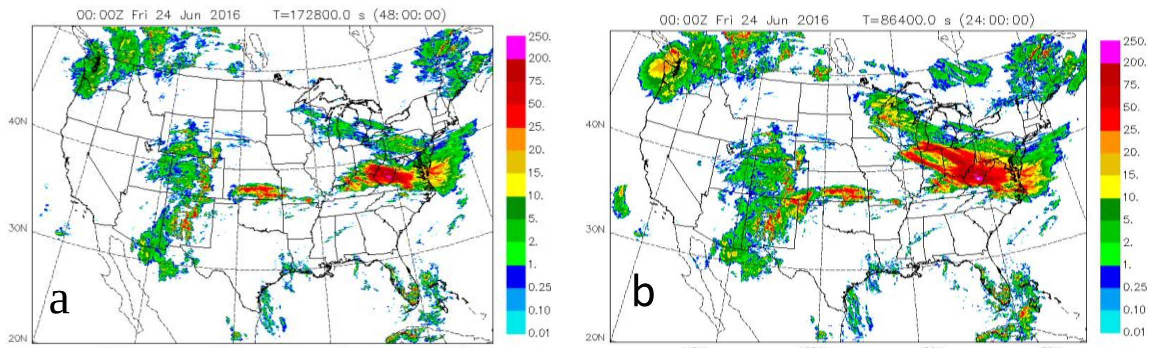

Figure 5. MRMS QPE: (a) 12 h accumulated precipitation, 12-00 UTC, (b) 24 h accumulated precipitation, 00-00 UTC, valid at 00 UTC June 24, 2016.

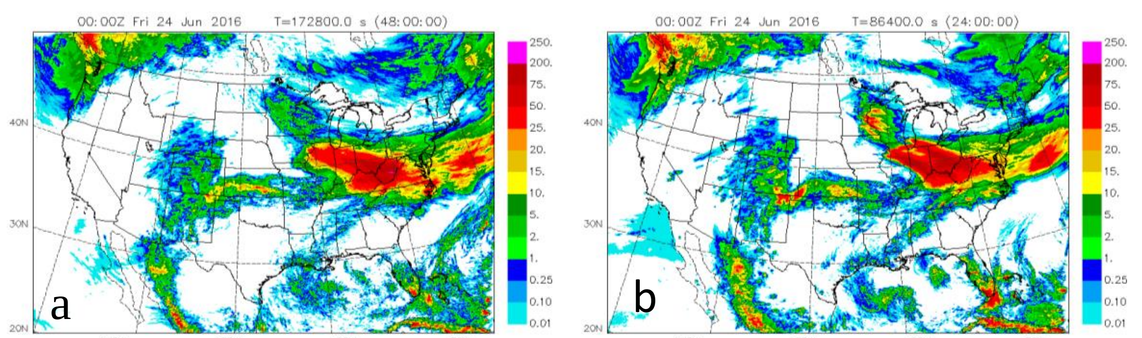

Figure 6.Probability matched mean forecast 24 h accumulated precipitation, valid at 00 UTC June 24, 2016. (a) 48 h forecast, (b) 24 h forecast

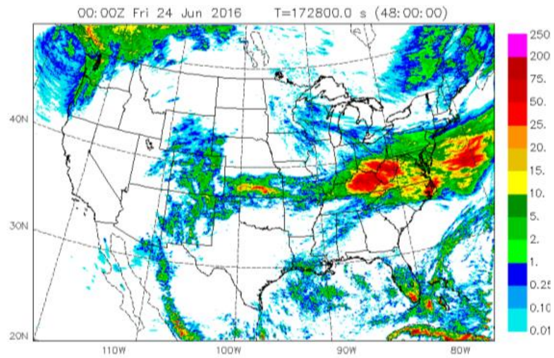

Figure 7. Same as Figure 6, except for 12 h accumulated precipitation.

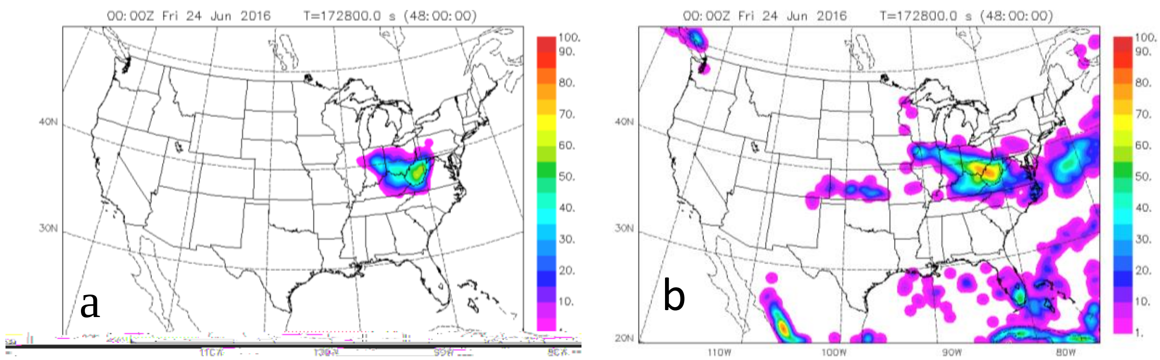

Figure 8. 48 h forecast of neighborhood probability of 24 h QPF exceeding (a) 24 h Flash Flood Guidance (FFG), and (b) exceeding 3 inches, valid at 00 UTC June 24, 2016.

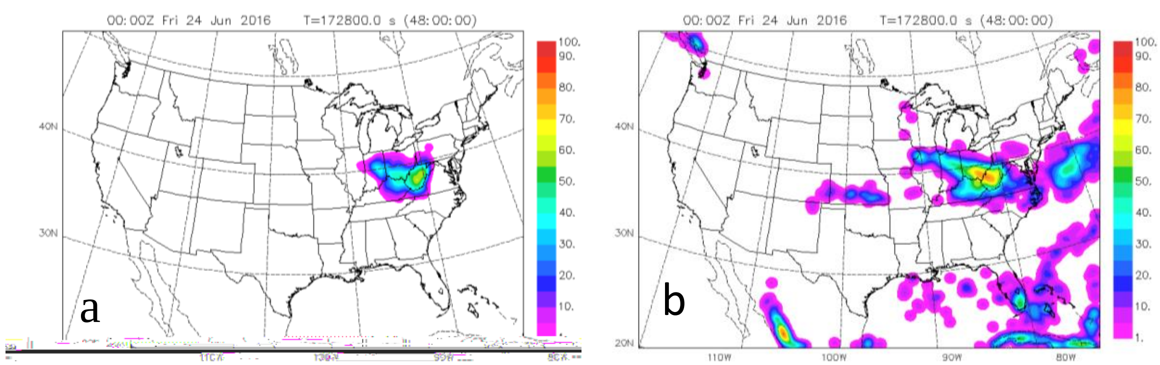

Figure 9. 48 h forecast of neighborhood probability of 24 h QPF exceeding (a) 24 h Flash Flood Guidance (FFG), and (b) exceeding 3 inches, valid at 00 UTC June 24, 2016.

1.4 GSI+EnKF based data assimilation and ensemble forecast

From 2016 to 2019 SFEs, CAPS continued realtime EnKF based data assimilation and storm-scale ensemble forecasting experiment. Different from the years in 2014 and 2015 when only CAPS EnKF was used in assimilating radar data as well as surface and profiler data in a one hour window from 2300 UTC to 0000 UTC, the operational GSI EnKF was combined with CAPS EnKF since 2016 SFE.

In 2016 SFE, the GSI EnKF was used hourly in a 6-hour long cycling period (1800 – 0000 UTC) with RAP/HRRR GSI data stream (except satellite data and Mesonet1 data), followed a one hour frequent radar data assimilation in 15 min interval using CAPS EnKF, to further improve the ensemble initial conditions. First, a 40-member WRF-ARW ensemble was initiated at 1800 UTC over the same 3-km grid CONUS domain, using the 18Z NAM analysis with perturbations retrieved from SREF members. This ensemble was configured with initial perturbations and mixed physics options to provide input for EnKF analysis. Unlike in previous years, each member used Thompson microphysics (while in 2014 & 2015 seasons when WSM6 was used) with different parameter settings in graupel density. No radar data was analyzed for this set of runs until 2300 UTC. RAP/HRRR GSI data stream (except satellite data and Mesonet1 data) were assimilated hourly from 1900 to 0000 UTC using the GSI EnKF system. Radar reflectivity and radial velocity data were assimilated using CAPS EnKF system from 2300 to 0000 UTC every 15 min. A 12-member ensemble forecast of 60 h long followed using the 0000 UTC final GSI+EnKF analyses in 2016 SFE. Among them, nine were initiated using selected ensemble member analyses with mixed IC/LBC perturbations and physics options, and three were deterministic forecasts from the 0000 UTC ensemble mean analysis with three different microphysics schemes. Figure 10 is a workflow diagram showing the CAPS GSI+EnKF cycling process in 2016 SFE.

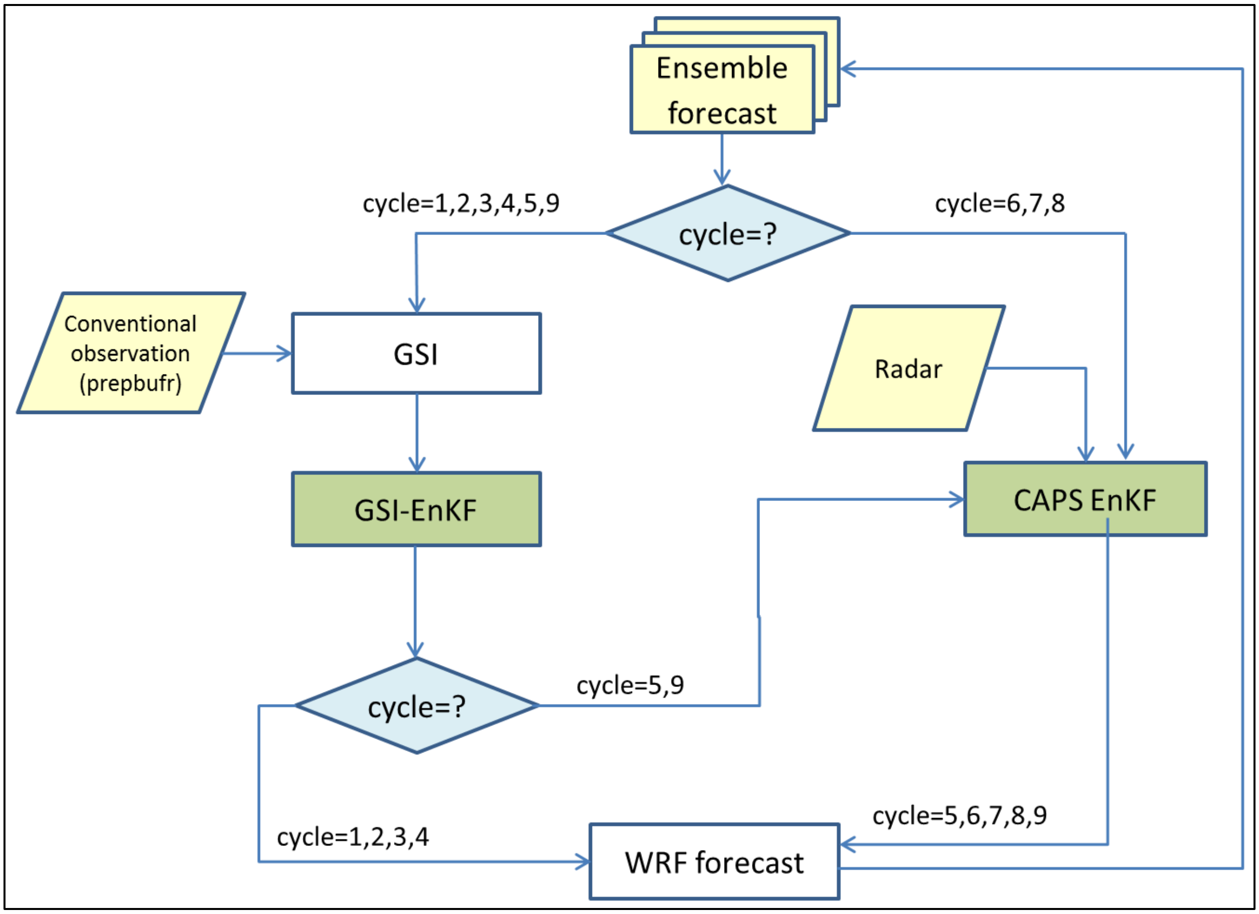

Figure 10. Diagram showing GSI+EnKF cycles in 2016 SFE.

In 2017 and 2018 SFEs, the same GSI+EnKF workflow as in 2016 was utilized, except that a 10-member ensemble forecast (run for 48 hours) was following using the final EnKF analyses at 0000 UTC using multi-physics multi-microphysics WRF-ARW configurations (see Tables 14 and 15).

Table 14. Configuration of the GSI+EnKF-based 0000 UTC ensemble forecasts for 2017 SFE

Member	IC	BC	Microphysics	LSM	PBL
enkf_m1	enk_m1a	00Z NAMf	Thompson	Noah	MYJ
enkf_m2	enk_m2a	21Z SREF arw-p1	Morrison	Noah	YSU
enkf_m3	enk_m15a	21Z SREF arw-n1	MY	Noah	MYNN
enkf_m4	enk_m40a	21Z SREF nmmb-p1	Morrison	Noah	MYJ

enkf_m5	enk_m8a	21Z SREF nmmb-n1	Thompson	Noah	YSU
enkf_m6	enk_m36a	21Z SREF arw-p2	MY	Noah	MYNN
enkf_m7	enk_m39a	21Z SREF arw-n2	MY	Noah	YSU
enkf_m8	enk_m17a	21Z SREF nmmb-p2	NSSL	Noah	MYJ
enkf_m9	enk_mn	00Z NAMf	Thompson	Noah	MYJ
enkf_m10	enk_mn	00Z NAMf	NSSL	Noah	MYJ

Table 15. Configuration of the GSI+EnKF-based 0000 UTC ensemble forecasts for 2018 SFE

Member	IC	BC	Microphysics	LSM	PBL
enkf_m01	enk_m01a	00Z NAMf	Thompson	Noah	MYJ
enkf_m02	enk_m02a	21Z SREF arw-p1	NSSL	Noah	YSU
enkf_m03	enk_m15a	21Z SREF arw-n1	NSSL	Noah	MYNN
enkf_m04	enk_m40a	21Z SREF nmmb-p1	Morrison	Noah	MYJ
enkf_m05	enk_m8a	21Z SREF nmmb-n1	P3	Noah	YSU
enkf_m06	enk_m26a	21Z SREF arw-p2	NSSL	Noah	MYJ
enkf_m07	enk_m39a	21Z SREF arw-n2	Morrison	Noah	YSU
enkf_m08	enk_m12a	21Z SREF nmmb-p2	P3	Noah	MYNN
enkf_m09	enk_mn	00Z NAMf	Thompson	Noah	MYJ
enkf_m10	enk_mn	00Z NAMf	NSSL	Noah	MYJ

In 2019 SFE, enhanced GSI-based EnKF system was used both to assimilate the RAP/HRRR GSI data stream (except for restricted data) at hourly intervals from at 1900 through 0000 UTC over the CONUS domain and in the final hour from 2300–0000 UTC to assimilate radar data every 15 minutes. This is unlike previous years when CAPS EnKF was used for radar data assimilation. The first 10 members of the EnKF analyses at 0000 UTC were used to initialize 10 ensemble forecasts that were run to 48 hours, and the WRF forecast model used different physics combinations to account for model error. LBCs were from the SREF forecasts from the 2100 UTC cycle. Table 16 shows the configurations of the 10 forecast members.

Table 16. Configuration of the EnKF ensemble forecasts for 2019 SFE

Member	IC	BC	Microphysics	LSM	PBL
enkf_m01	enk_m01a	00Z NAMf	Thompson	Noah	MYJ
enkf_m02	enk_m02a	21Z SREF arw-p1	NSSL	Noah	YSU
enkf_m03	enk_m15a	21Z SREF arw-n1	NSSL	Noah	MYNN
enkf_m04	enk_m40a	21Z SREF nmmb-p1	Morrison	Noah	MYJ
enkf_m05	enk_m8a	21Z SREF nmmb-n1	P3	Noah	YSU
enkf_m06	enk_m26a	21Z SREF arw-p2	NSSL	Noah	MYJ
enkf_m07	enk_m39a	21Z SREF arw-n2	Morrison	Noah	YSU
enkf_m08	enk_m12a	21Z SREF nmmb-p2	Thompson	Noah	MYNN
enkf_m09	enk_m34a	21Z SREF arw_n6	Thompson	Noah	YSU
enkf_m10	enk_m37a	21Z SREF arw-p6	NSSL	Noah	MYNN

Development and testing of GSI-based hybrid capabilities and the ability to directly assimilate radar data within the hybrid system has been performed in the reduced domain because of the computational efficiency of the GSI system for the full-resolution radar data. The computation domain used in the test is approximately 1/36 of the size of the CONUS domain, which contains a mesoscale convective system (MCS). 30-minute forecasts from the final analysis of GSI 3DVar, EnKF, En3DVar, and hybrid En3DVar with 80% flow-dependent ensemble covariance are compared in Fig. 11, which are valid at 0030 UTC. While all four experiments maintains the main convective line structure in the MCS, two convective systems on the west of the domain exhibit rapid deterioration in terms of intensity and structure in all experiments except for hybrid En3DVar. This is mainly due to spread deficiency in EnKF and En3DVar and the lack of cross-covariance in 3DVar. The results shows the benefit of static background error covariance in the hybrid system.

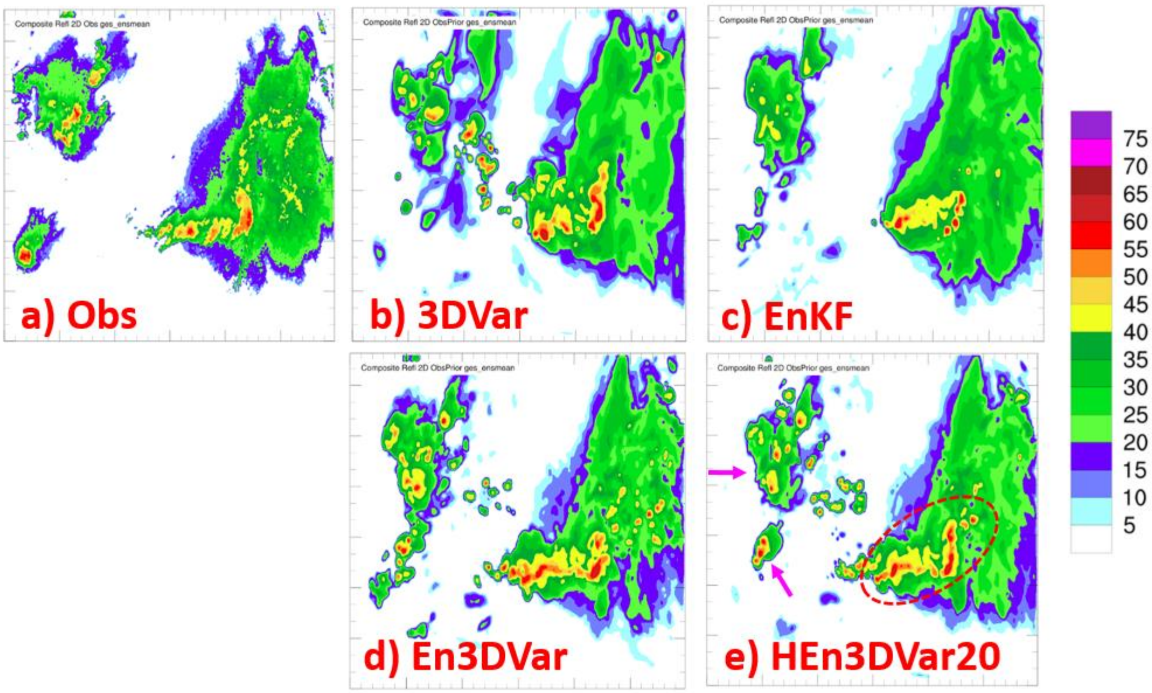

Figure 11. Comparisons of (a) the observed and (b-e) simulated composite reflectivity for (b-d) backgrounds and (e-g) analyses at 0030 UTC May 27 2016/30-min forecast that directly assimilate reflectivity in (b) 3DVar, (c) EnKF, (d) En3DVar, and (e) hybrid En3DVar with 20% weight given to the static background error covariance.

1.5 SSEF QPF Products and verifications

1.5.1 Localized probability matched mean (LPM)

An advanced and computationally efficient ensemble consensus QPF product, a Localized Probability-Matched mean or LPM mean, was developed by CAPS for the HMT FFaIRs (Snook et al. 2019, 2020). The standard Probability-Matched (PM) mean is obtained by replacing ensemble mean field with values from entire ensemble members ranked over the entire domain grids (globally). The “global” PM generally can improve results compared to the simple ensemble mean of the output value at each grid point averaged over all ensemble members. Some issues can arise from using the values over the entire domain, especially for large domains such us the CONUS domain, where rain amounts from one part of the domain may be mapped to grid points on the opposite side of the country. Even though the resulting PM mean field has been found to be more skillful than a simple ensemble mean, it also exhibits a loss of small-scale structure compared to individual ensemble members. When applied over a large forecast domain, the PM mean also has the drawback of combining precipitation information from very different mesoscale and geographic environments, such as coastal sea-breeze convection and stratiform precipitation over the northern plains. To ameliorate this issue, in the LPM mean, the domain is broken into overlapping subdomains and the process repeated for each subdomain thus preventing the distant reassignment.

The LPM algorithm developed by CAPS applies a PM mean algorithm over a series of local patches. To calculate the LPM, the domain is divided into a set of rectangular local patches. Each patch is centered within a larger, rectangular calculation area. The patches do not overlap, but the calculation areas of adjacent or nearby patches may overlap. A conceptual illustration of this setup is shown in Figure 12.

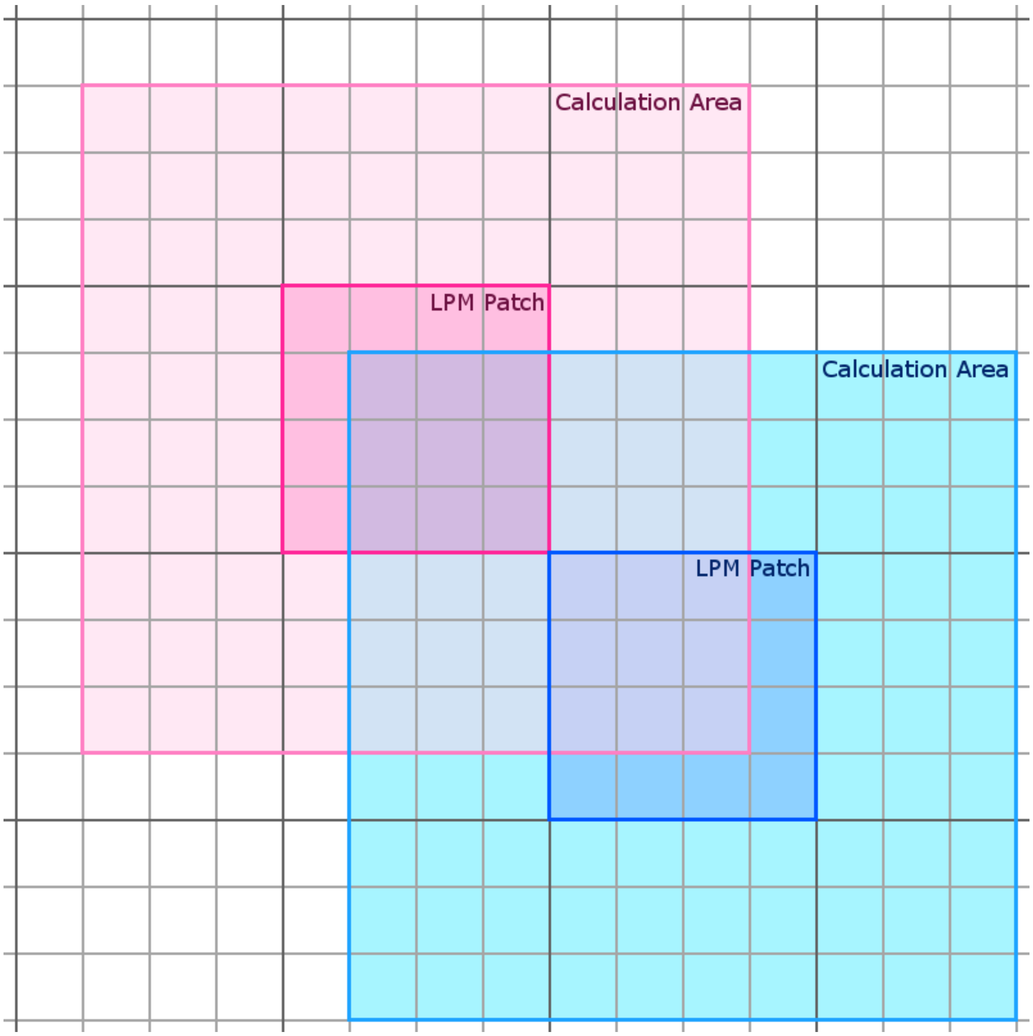

Figure 11. Conceptual illustration of the patches used to generate localized probability matched mean products. Highlighted are two patches (the darker magenta and blue shaded regions) along with their associated calculation areas (the lighter magenta and blue shaded regions surrounding the patches). The gray lines indicate the edges of model grid cells.

To generate the LPM, the PM mean is calculated for each patch over the subdomain contained within the calculation region associated with that patch. The portion of the PM mean field from the calculation area that falls within the patch itself is returned for each patch, and these patches are stitched together to form a single field for the full CONUS domain. Finally, a Gaussian smoother is applied to the stitched patches to minimize discontinuities along patch boundaries. Discontinuities are also minimized by the overlap of adjacent calculation areas and the use of relatively small patches; in preliminary testing, discontinuities were not an issue for patches smaller than approximately 10 by 10 grid points.

Figure 13 (cited from Snook et al. 2020) shows FSS scores (13a) and the spatial structure of 6-h accumulated precipitation from 18-24 h forecast in LPM, PM, and simple mean forecasts compared to that of observed 6-h precipitation (from Stage IV data) in terms of variance spectra (13b), analyzed for the full 2018 HWT SFE period. The simple mean performs quite poorly in terms of FSS, failing to attain a skillful FSS at any scale. The PM mean exhibits a minimum skillful scale of approximately 95 km, compared to 65-75 km (unsmoothed) and 80-100 km (smoothed) for the LPM means. The unsmoothed LPM outperforms the PM in terms of FSS at all scales; smoothing reduces FSS performance of the LPM slightly, though even the smoothed LPM outperforms PM at scales larger than the minimum skillful scale. The point-wise and patch-wise LPM configurations perform similarly, although when weak smoothing is applied, LPM calculated using larger local patches exhibits slightly higher FSS (Figure 13a). Figure 13b shows that the overly-smooth simple mean exhibits much lower variance than the observations at all scales smaller than 1000 km, while the PM and LPM means remain much closer to the observed spectrum. However, the PM mean over-predicts variance at scales larger than about 300 km, while the spectra of the LPM means remain very close to the observed spectrum even at the largest scales. It is also showing that LPM patch size has little impact on variance spectra for patches of up to 40×40 grid points. Difference in variance spectra among LPM configurations is largely attributable to smoothing; when smoothing is applied, variance is drastically reduced for scales smaller than about 40 km (this is expected, as smoothing using a Gaussian filter with a standard deviation of one grid point is designed to minimize near-grid-scale noise and patch-boundary discontinuities).

The LPM mean performs similarly to the PM mean in terms of ETS, and exhibits superior bias behavior, particularly when weak smoothing is applied, though for ensembles with different bias properties tuning of the smoother might be required. When spatial structure is considered, the LPM mean outperforms the PM and simple means, exhibiting a substantially smaller minimum skillful scale in terms of FSS, and higher FSS overall for scales greater than around 100 km. The LPM mean also exhibits superior precipitation variance spectra compared to the simple mean (which underpredicts variance at small scales) and the PM mean (which overpredicts variance at large scales). Varying the size of the LPM patch from a single point to 40 x 40 points resulted in only small changes to precipitation forecast performance for the objective skill metrics considered.

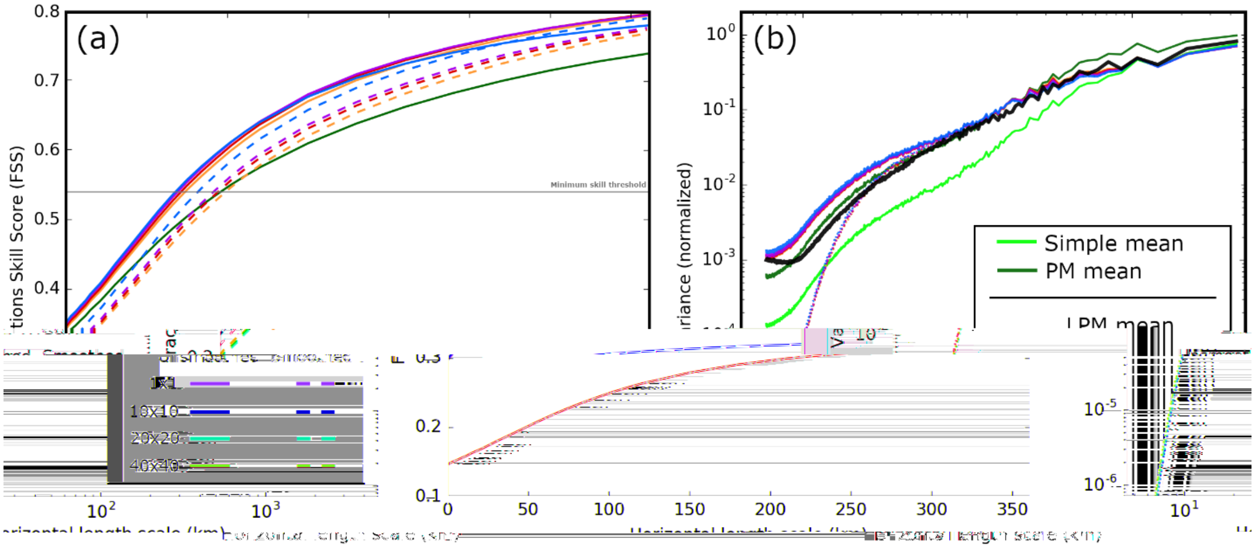

Figure 13. Mean (a) fractions skill score (FSS) and (b) normalized variance spectra, plotted as a function of horizontal length scale, calculated over the 2018 HWT SFE period, for 6-h accumulated precipitation in 18-24-h CAPS SSEF forecasts valid 0000 UTC the day after initialization. Shown are the simple mean (light green), PM mean (dark green), and various LPM mean configurations. LPM means include smoothed (dotted lines) and unsmoothed (solid lines) variants using both point-wise (orange) and patch-wise algorithms, including patches of three sizes (red: 10×10, purple: 20×20, blue: 40×40). The minimum skillful FSS value is indicated by the horizontal gray line in panel (a). The mean normalized variance spectrum of stage IV rainfall accumulation is shown (thick black line) for comparison in panel (b). [from Snook et al. 2020]

While the LPM mean, even in its patch-wise form, is more computationally expensive than the PM and simple means, its ability to retain local structures and magnitudes is valuable. Precipitation forecasts using the patch-wise LPM mean were produced by CAPS during the 2017 and 2018 Hydrometeorology Testbed Flash Flood and Intense Rainfall (FFaIR) experiments, and were subjectively rated as more useful than simple or PM mean forecasts by FFaIR participants. Based upon these tests, the patch-wise LPM mean was recommended for operational implementation and has been tested in a prototype version 3 of the High Resolution Ensemble Forecast (HREFv3), where it was found to exhibit good performance, particularly in terms of bias, though it underperformed PM for ETS at low precipitation amounts.

Figure 14 compares PM and LPM with point-wise and neighborhood verification using FFaIR2017 dataset. For point-wise verification the PM mean slightly outperformed the LPM at both the 0.01 and 0.5 inch thresholds (Figures not shown). For the 12 and 24 km neighborhoods the LPM slightly outperformed the PM mean at the lower threshold, but not at the higher rainfall threshold where the PM slightly outperformed the LPM at both neighborhoods examined.

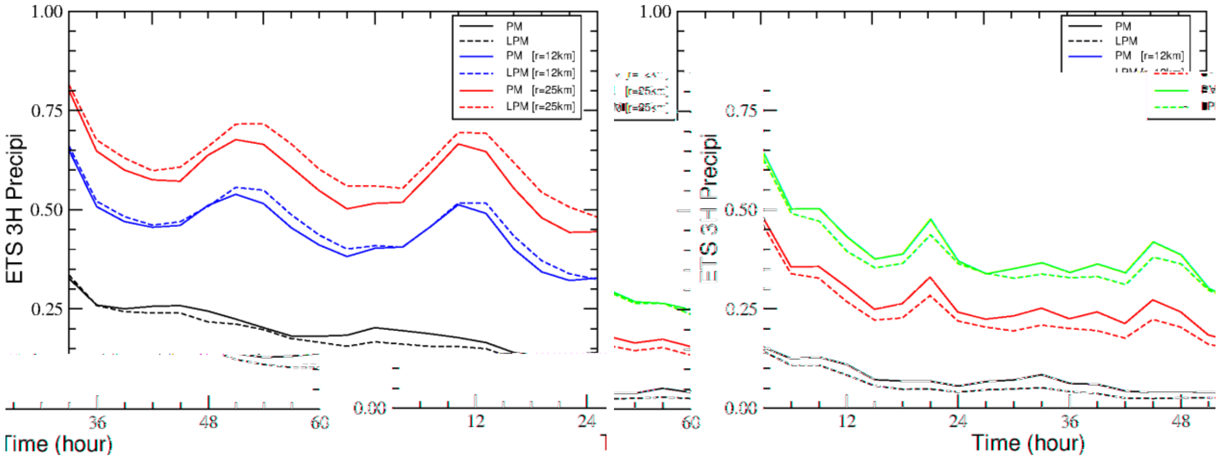

Figure 14. ETS scores comparing the LPM mean to the PM mean for point verification (black) and neighborhood verifications of 12 km (blue) and 24 km (red) neighborhoods for 3-h rainfall at 0.01 inch threshold (left) and 0.5 inch (right).

1.5.2 Verifications of different ensemble groups

Post season performance evaluation was performed by examining ensemble spread, RMSE, ETS, and Area under ROC. Verification of rainfall forecast products has been done using the Multi-Radar Multi-System (MRMS) rainfall estimates. There are two focuses in the evaluation phase: One is to examine ensemble characteristics and QPF performance among three CLUE WRF sub-ensembles CAPS contributed: the *mixed physics + radar* (called *mixed* hereafter), *single physics + radar* (*single*), and *stochastic physics + radar* (*stochastic*). The second is to examine performance of FV3 ensemble at convection-allowing grid and how it compares to WRF ones.

Figure 15 shows the domain and season averaged ensemble spread of some variables for the three CLUE WRF sub-ensembles. It shows that the mixed-physics, which is configured with IC/LBC perturbation and multi-physics members, ensemble has the largest spread over the other two ensembles. Interestingly, the ensemble with additional stochastic perturbations show little spread improvement or not at all (as in HGT500). Figure 16 shows RMSE and the ensemble spread of ensemble mean 1-h precipitation for the three ensembles. None of the sub-ensembles has its spread matching the mean error except morning hours when convection is inactive.

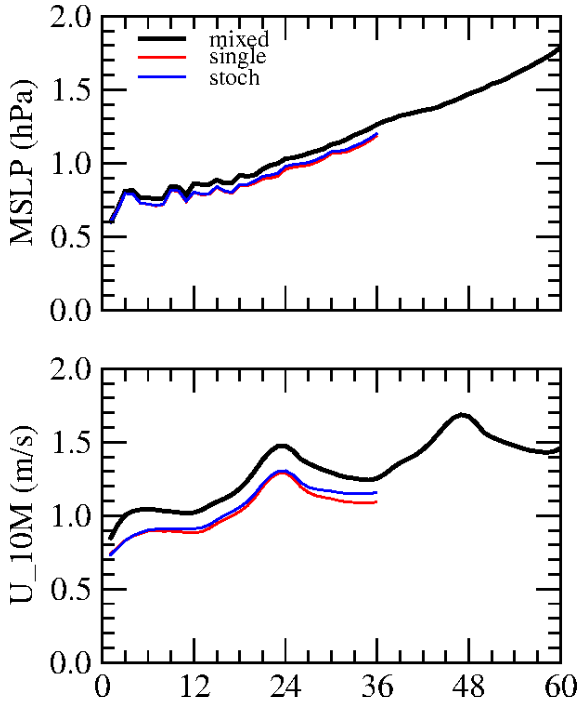

Figure 15. Ensemble spread averaged over entire domain and throughout the SFE2018 season for the three sub-ensembles: mixed, single, and stochastic.

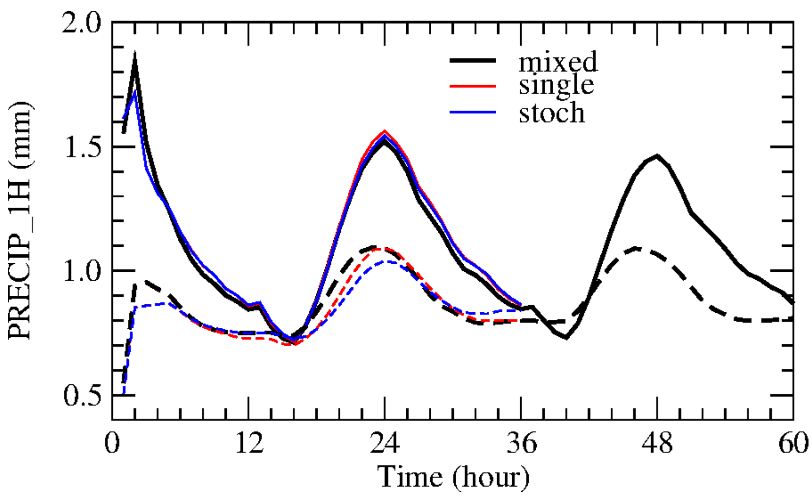

Figure 16. RMSE from ensemble mean (solid) and ensemble spread (dash) of the 1-h accumulated precipitation forecast from the three CLUE sub-ensembles: mixed, single, and stochastic.

Equitable Threat Scores (ETS) for 3-hour accumulated precipitation forecasts for the 0.01 and 0.5 inch thresholds are shown in Figure 17 from all individual members and in Figure 18 from probability matched mean. All three sub-ensembles have comparable ETS values individually (see Figure 17). While the probability matched mean ETS from the stochastic ensemble are the lowest except during the first 18 h for the 0.5 inch threshold (Figure 18).

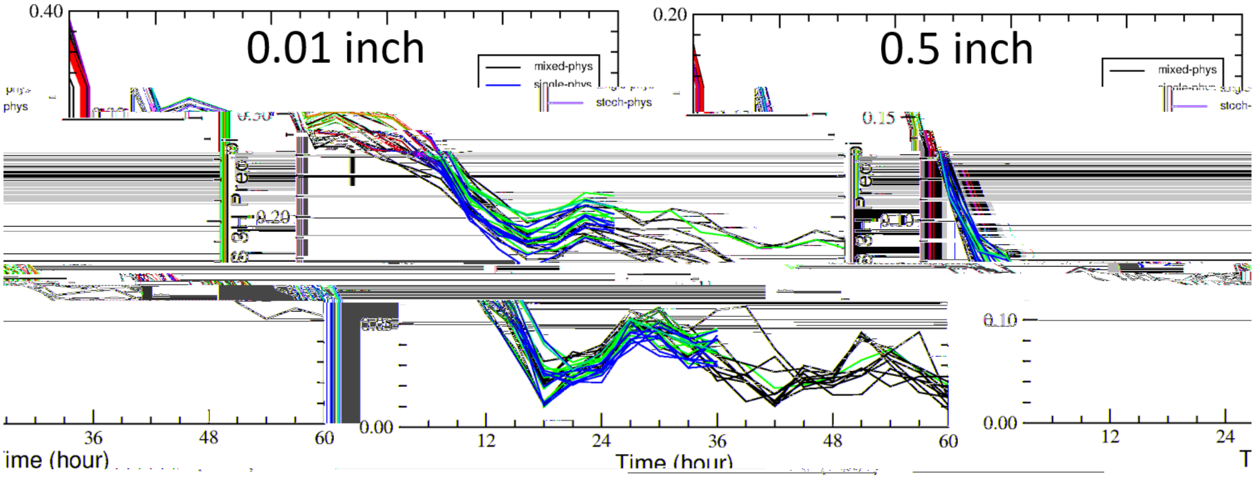

Figure 17. ETS of 3-h accumulated precipitation forecast from all members of the three CLUE sub-ensembles: mixed, single, and stochastic.

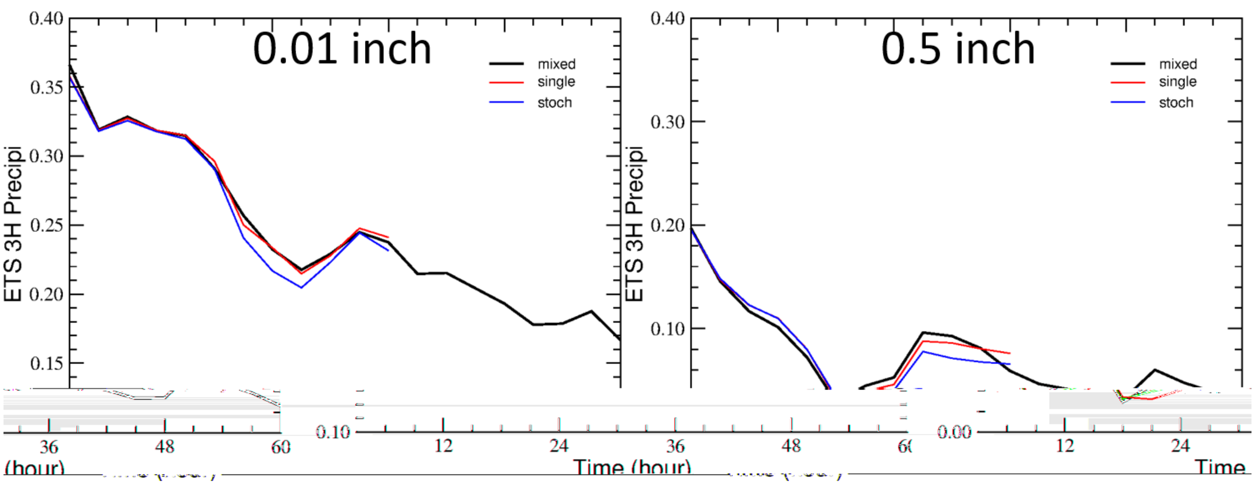

Figure 18. ETS of probability matched mean of 3-h accumulated precipitation forecast for the three CLUE sub-ensembles: mixed, single, and stochastic.

Figure 19 presents the area under ROC (also called AUROC) of 3-h accumulated precipitation probabilistic forecast for the three CLUE WRF sub-ensembles, with one set (dashed lines) considering a neighborhood of 25-km radius that removes some of the common slight malposition errors. Again, the mixed-physics ensemble has the highest AUROC among the three. The ensemble with additional stochastic perturbations doesn't show any performance benefit.

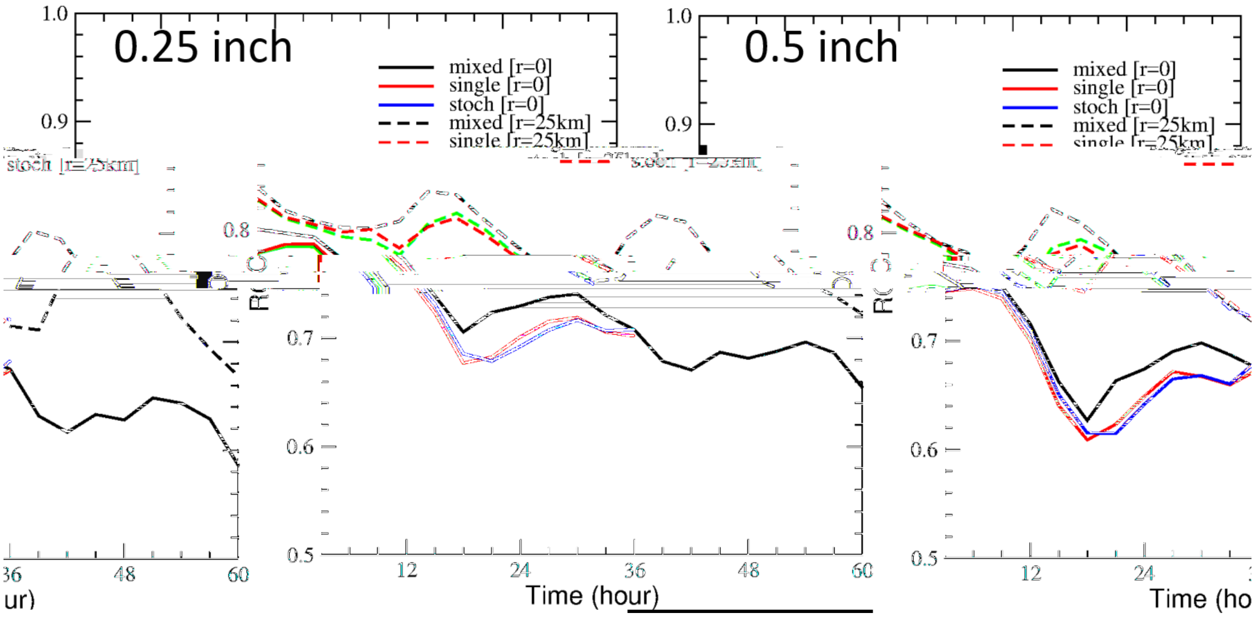

Figure 19. Area under ROC of 3-h accumulated precipitation forecast for the three CLUE sub-ensemble, with dashed lines representing a 25km neighborhood applied.

The precipitation forecasting skill of FV3 ensemble members with advanced physics schemes (see Table 6) is also evaluated to demonstrate how well the Finite-Volume Cubed-Sphere (FV3) model predicts precipitation over Contiguous United States (CONUS) when run at a high enough resolution to explicitly model convective storms. The physical parameterization schemes available in FV3 were mostly from GFS and not necessarily suitable for convective-scale predictions. CAPS scientists implemented into FV3 several advanced physics schemes taken from a more established and most widely used convective-scale model, the WRF model, including schemes for treating turbulence exchanges in atmospheric boundary layer (PBL) and those for representing cloud and precipitation processes (microphysics).

FV3 is shown to be capable of predicting precipitation with skill comparable to those of WRF. The precipitation forecast is somewhat sensitive to the microphysics schemes used, but not particularly sensitive to the PBL schemes tested. It is concluded that the newly selected FV3 model, when equipped with appropriate physics parameterization schemes, can serve as the foundation for the next-generation regional forecasting models of the NWS.

Figure 20 show the ETS of hourly precipitation with a 45km neighborhood range (labelled NETS in the figure caption text) from one FV3 control member, one WRF control member, and the 00UTC initiated operational HRRR version 2 forecast. The HRRR forecasts only extend to 18 hours. The 3-km operational HRRR forecasts has slightly higher NETSs for 99th percentile and higher NETSs between 12 and 18 hour forecasts for 99.9th percentile (Figure20a and b), while the CAPS WRF forecast and FV3 are comparable after hour 6. The differences of bias scores among the operational HRRR, CAPS FV3 and WRF are very small (Figure 20c and d).

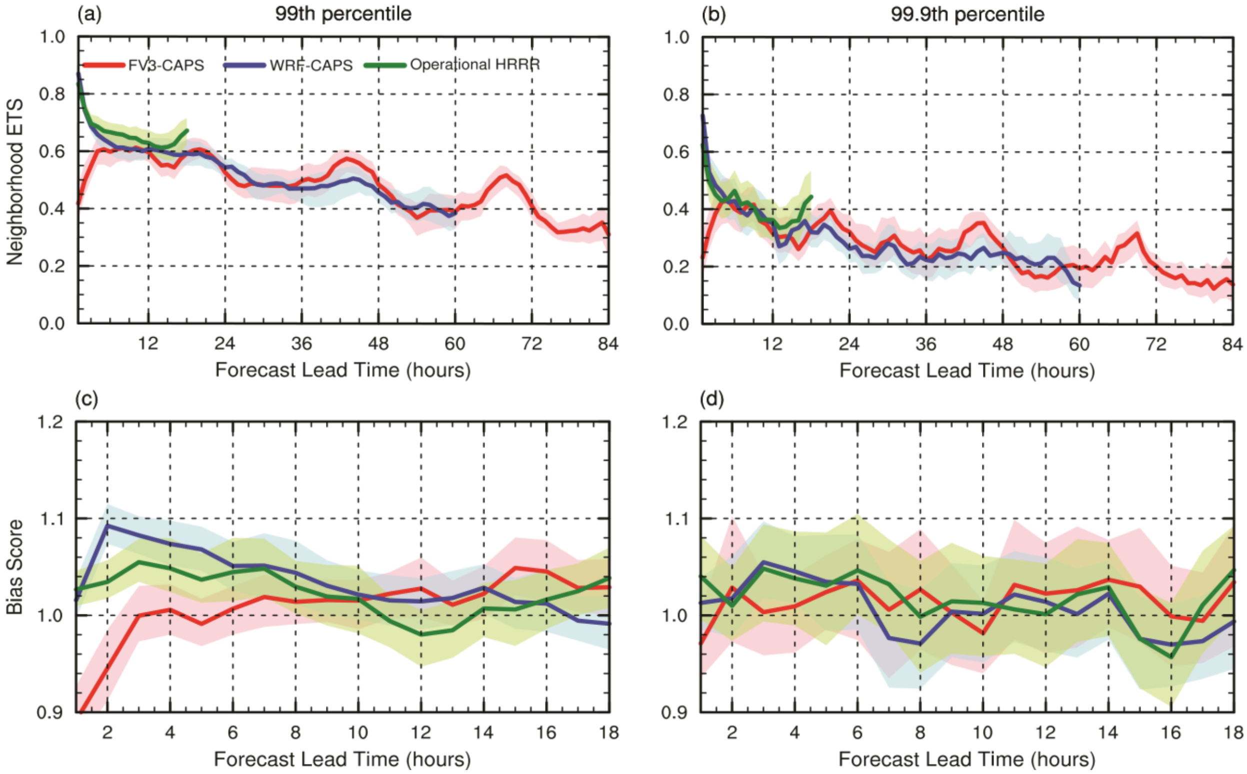

Figure 20. NETS (a, b) and bias (c, d) (using a 45 km r) for FV3 (FV3-CAPS), CAPS WRF (WRF-CAPS), and operational HRRR version 2. FV3-CAPS is the FV3 forecast using the Thompson and SA-MYNN schemes, while WRF-CAPS is the WRF forecast using the same physics package as the operational HRRR. The lines are mean values for the 25 cases during the 2018 HWT SFE. The shading indicates the 2.5th to 97.5th percentile range of possible mean values based on 10,000 bootstrap re-samplings from the 25 cases for each forecast. Both CAPS WRF and operational HRRR use advanced data assimilation for their initial conditions. CAPS WRF forecasts were run for 60 hours, while the operational HRRR was run for 18 hours. Only forecasts up to 18 hours are shown for bias (c, d). NETS and bias are calculated on the native grid for each forecast.

1.5.3 Verification of GSI EnKF ensemble

Analysis was also performed to evaluate performance between the forecasts initiated using GSI+EnKF final analysis ensemble and using a single 3DVAR analysis with complex cloud analysis at 00 UTC. Further tuning on GSI and CAPS EnKF was performed during SFE2017. Figure 21 show ETS scores of the ensemble probability matched mean for both ensembles from 2016 and 2017 HWTs. The results show that the relative performance of the EnKF-based forecasts to the 3DVAR-based forecast is significantly improved in 2017 compared to those in 2016 in terms of the Equitable Threat Scores (ETSs) for the 1-hour rainfall accumulation exceeding 0.1 in and 0.25 in thresholds.

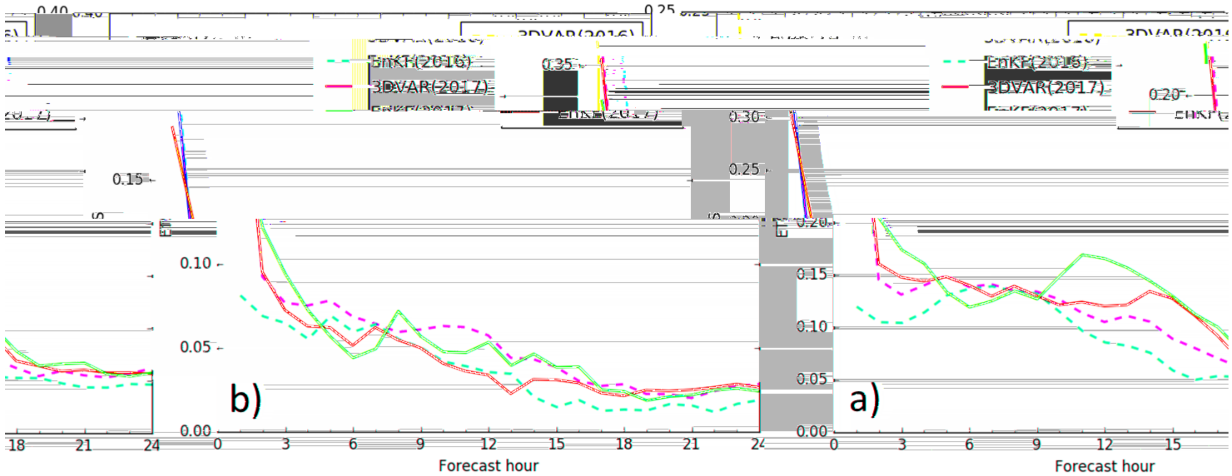

Figure 21. Equitable Threat Scores (ETSs) of 1-hour rainfall accumulation against Multi-Radar Multi-Sensor (MRMS) rainfall estimates ≥ (a) 0.1 and (b) 0.25 inches for the non-cycled 3DVAR-based and cycled GSI+EnKF-based members averaged over the 2016 and 2017 HWT SFE periods.

1.6 Journal publications and conference presentations

Journal Publications

Clark, Adam, Israel Jirak, Scott Dembek, Gerry Creager, Fanyou Kong, Kevin Thomas, Kent Knopfmeier, Burkely Gallo, Christopher Melick, Ming Xue, Keith Brewster, Youngsun Jung, Aaron Kennedy, Xiquan Dong, Joshua Markel, Glen Romine, Kathryn Fossell, Ryan Sobash, Jacob Carley, Brad Ferrier, Matthew Pyle, Curtis Alexander, Steven Weiss, John Kain, Louis Wicker, Gregory Thompson, Rebecca D. Adams-Selin, David Imy, 2018: The Community Leveraged Unified Ensemble (CLUE) in the 2016 NOAA/Hazardous Weather Testbed Spring Forecasting Experiment. Bulletin of the American Meteorological Society, 99, 1433-1448.

Corey K. Potvin, Jacob R. Carley, Adam Clark, Louis J. Wicker, Patrick S. Skinner, Anthony E. Reinhart, Burkely T. Gallo, John S. Kain, Glen Romine5, Eric Aligo, Keith A. Brewster, David C. Dowell, Lucas M. Harris, Israel L. Jirak, Fanyou Kong, Timothy A. Supinie, Kevin W. Thomas, Xuguang Wang, Yongming Wang, and Ming Xue, 2019: Systematic comparison of convection-allowing models during the 2017 NOAA HWT Spring Forecasting Experiment. Weather and Forecasting, 34, 1395-1416.

Duda, J. D., X. Wang, F. Kong, M. Xue and Judith Berner, 2016: Impact of a stochastic kinetic energy backscatter scheme on warm season convection-allowing ensemble forecasts. *Mon. Wea Rev.*, 144, 1887-1908. *doi: http://dx.doi.org/10.1175/MWR-D-15-0092.1*.

Dahl, N. and M. Xue, 2016: Prediction of the 14 June 2010 Oklahoma City extreme precipitation and flooding event in a multi-physics multi-Initial condition storm scale ensemble

forecasting system *Wea. Forecasting*, 31, 1215–1246. *doi: http://dx.doi.org/10.1175/WAF-D-15-0116.1*
Gagne, D. J., II, A. McGovern, S. E. Haupt, R. Sobash, J. K. Williams, and M. Xue, 2017: Storm-Based Probabilistic Hail Forecasting with Machine Learning Applied to Convection-Allowing Ensembles. Weather Forecasting, 32, 1819-1840.
Gallo, B. T., A. J. Clark, I. Jirak, J. S. Kain, S. J. Weiss, M. Coniglio, K. Knopfmeier, J. C. Jr., C. J. Melick, E. Iyer, A. R. Dean, M. Xue, F. Kong, Y. Jung, F. Shen, K. W. Thomas, K. Brewster, D. Stratman, G. Carbin, W. Line, R. Adams-Selin, and S. Willington, 2017: Breaking New Ground in Severe Weather Prediction: The 2015 NOAA/Hazardous Weather Testbed Spring Forecasting Experiment. Wea. Forecasting, 32, 1541-1568.
Grasso, Lewis, Daniel T. Lindsey, Yoo-Jeong Noh,Christopher O'Dell, Fanyou Kong, 2018: Improvements to cloud top brightness temperatures computed from the CRTM at 3.9 µm. submitted to Mon. Wea. Rev., 146, 3927-3944.
Griffin, Sarah M., Jason A. Otkin, Greg Thompson, Maria Frediani, Judith Berner, Fanyou Kong, 2020, Assessing the Impact of Stochastic Perturbations in Cloud Microphysics using GOES-16 Infrared Brightness Temperatures. Mon. Wea. Rev., 148 (8): 3111–3137.
Hardy, J., J. J Gourley, P-E.l Kirstetter, Y. Hong, F. Kong, and Z. L. Flamig, 2016: A Method for Probabilistic Flash Flood Forecasting, *J. of Hydrlogy*, 541, Part A, 480–494. *doi: http://dx.doi.org/10.1016/j.jhydrol.2016.04.007*
Iyer, E. R., A. Clark, M. Xue, and F. Kong, 2016: A comparison of 36-60 hour precipitation forecasts from convection-allowing and convection-parameterizing ensembles. *Wea. Forecasting*, 31, 647-661. *doi: http://dx.doi.org/10.1175/WAF-D-15-0143.1*.
Labriola, J., N. Snook, Y. Jung, and M. Xue, 2019: Explicit Ensemble Prediction of Hail in 19 May 2013 Oklahoma City Thunderstorms and Analysis of Hail Growth Pocesses with Several Multi-Moment Microphysics Schemes, *Mon. Wea. Rev.*, **147**, 1193-1213, https://doi.org/10.1175/MWR-D-18-0266.1.
Labriola, J., N. Snook, M. Xue, and K. Thomas, 2020: Evaluating Ensemble Kalman Filter Analyses of Severe Hail Storms on 8 May 2017 in Colorado: Effects of State Variable Updating and Multi-Moment Microphysics Schemes on State Variable Cross-Covariances. Mon. Wea. Rev., **148**, 2365-2389.
Liu, C., M. Xue, and R. Kong, 2019: Direct assimilation of radar reflectivity data using 3DVAR: Treatment of hydrometeor background errors and OSSE tests. Mon. Wea. Rev., 137, 17-29.
Loken, E., A. Clark, M. Xue; F. Kong, 2017: Impact of Horizontal Resolution on CAM-Derived Next-Day Probabilistic Severe Weather Forecasts, Wea. Forecasting, 32, 1403-1421.
Loken, E., A. Clark, M. Xue; F. Kong, 2019: Spread and Skill in Mixed- and Single-Physics Convection-Allowing Ensembles. Submitted to Wea. Forecasting, 34, 305-330.
McCaul, E. W. Jr., G. Priftis, J. L. Case, T. Chronis, P. Gatlin, S. J. Goodan, F. Kong, 2020, Sensitivities of the WRF Lightning Forecasting Algorithm to Parameterized Microphysics and Boundary Layer Schemes. Wea. Forecasting. 35 (4): 1545–1560.

Putnam, B. J., M. Xue, Y. Jung, G. Zhang, and F. Kong, 2017: Simulation of polarimetric radar variables from 2013 CAPS spring experiment storm scale ensemble forecasts and evaluation of microphysics schemes. *Mon. Wea. Rev.*, 145, 40-73.

Snook, N., Y. Jung, J. Brotzge, B. Putnam, and M. Xue, 2016: Prediction and ensemble forecast verification of hail in the supercell storms of 20 May 2013. *Wea. Forecasting*, **31**, 811-825. *doi: http://dx.doi.org/10.1175/WAF-D-15-0152.1*

Snook, N., M. Xue, and Y. Jung, 2019: Tornado-Resolving Ensemble and Probabilistic Predictions of the 20 May 2013 Newcastle-Moore EF5 Tornado. Mon. Wea. Rev., 147, 1215-1235.

Snook, N., F. Kong, K. Brewster, M. Xue, B. Albright, K. W. Thomas, and T. Supinie, 2019: Advances in Real-Time Ensemble Precipitation Prediction during the 2016 and 2017 Hydrometeorology Testbed Flash Flood and Intense Rainfall Experiments. Wea. and Forecasting, 37, 781-804.

Snook, N., F. Kong, A. Clark, B. Roberts, K. A. Brewster, and M. Xue, 2020: Comparison and Verification of Point-wise and Patch-wise Localized Probability Matched Mean Algorithms for Ensemble Consensus Precipitation Forecasts. Geophysical Research Letters, 47, e2020GL087839 DOI:10.1029/2020GL087839.

Skinner, Patrick S., Adam J. Clark, Jamie K. Wolff, Tara L. Jensen, John Halley Gotway, Randy Bullock, Ming Xue, Fanyou Kong, Keith A. Brewster, 2020: The impact of radar assimilation on short-term and next-day thunderstorm forecasts in the 2016 Community Leveraged Unified Ensemble (CLUE). Mon. Wea. Rev., conditional accepted.

Surcel, Madalina, Isztar Zawadzki, M.K. Yau, Ming Xue, Fanyou Kong, 2017: More on the scale-dependence of the predictability of precipitation patterns: Extension to the 2009-2013 CAPS Spring Experiment ensemble forecasts. *Mon. Wea. Rev.*, 145, 3625-3646.

Thompson, G., M. Tewari, K. Ikeda, S. Tessendorf, C. Weeks, J. Otkin, and F. Kong, 2016: Explicitly-Coupled Cloud Physics and Radiation Parameterizations and Subsequent Evaluation in WRF High-Resolution Convective Forecasts. *Atmos. Res.*, 168, 92-104. *doi: http://dx.doi.org/10.1016/j.atmosres.2015.09.005*

Zhang, Cunxi, Ming Xue, Timothy A. Supinie, Fanyou Kong, Nathan Snook, Kevin W. Thomas, Keith Brewster, Youngsun Jung, Lucas M. Harris, and Shian-Jiann Lin, 2019: How Well Does an FV3-based Model Predict Precipitation at a Convection-Allowing Resolution? Results from CAPS Forecasts for the 2018 NOAA Hazardous Weather Testbed with Different Physics Combinations. Geophysical Research Letters, 46, 3523-3531 DOI:10.1029/2018GL081702.

Conference Presentations

Brewster, K. A., F. Kong, N. Snook, T. Supinie, K. Thomas, and M. Xue, 2018: 2017 CAPS Storm-Scale Ensemble Forecasts (SSEFX) for the HMT FFaIR Experiment. Eighth Conference on Transition of Research to Operations, AMS, Austin, Texas.

Brewster, K. A., N. Snook, F. Kong, M. Xue T. Supinie, CZhang, K. Thomas, 2020: SAR-FV3 Storm-Scale Ensemble Forecasts (CAPS SSEF) and Ensemble Consensus Products for the 2019 HMT FFaIR Experiment. Tenth Conference on Transition of Research to Operations, AMS, Boston, 3A.4.

Clark, A., and Co-authors, 2020: The 2019 NOAA Hazardous Weather Testbed Spring Forecasting Experiment. 10th Conference on Transition of Research to Operations, AMS, Boston, 2.2

Jung, Y., G. Zhao, F. Kong, K. Thomas, T. Supinie, and M. Xue, 2016: CAPS Real Time Storm-Scale EnKF Data Assimilation and Forecasts System for the 2016 Hazardous Weather Testbed Spring Experiment. *2016 NJU-OU Symposium for Weather and Climate Research*, Nanjing, China

Jung, Y., F. Kong, G. Zhao, K. Thomas, T. Supine, N. Snook, and M. Xue, 2016: CAPS Real-time Storm-scale Ensemble System for the NOAA HWT Spring Experiment. 2nd Ensemble Design Workshop, College Park, MD, https://drive.google.com/drive/folders/0B6ikBQ6A--swSEFtUGJ3Mm1OcGs

Jung, Y., and Coauthors, 2017: CAPS real-time continental US domain 3-km EnKF data assimilation and forecasts for the NOAA hazardous weather testbed spring forecasting experiment. 7th WMO symposium on DA, Floriaonopolis, Brazil. (http://pocanga.cptec.inpe.br/inscricaoDas/program.xhtml#)

Jung, Y., M. Xue, G. Zhao, J. Luo, T. Supinie, C. Liu, R. Kong, F. Kong, and K. Thomas, 2018: Development of GSI-based EnKF and hybrid EnVar data assimilation capabilities for continental-scale 3-km convection-permitting ensemble forecasting and testing via NOAA hazardous weather testbed spring forecasting experiments, *29th Conf. Wea. Anal. Forecasting/25th Conf. Num. Wea. Pred.*, Denver, CO

Jung, Y., M. Xue, G. Zhao, F. Kong, K. Thomas, T. Supinie, K. Brester, and N. Snook, 2018: CAPS real-time storm-scale EnKF data assimilation and forecasts for the NOAA hazardous weather testbed spring forecasting experiments: Towards the goal of operational ensemble-variational data assimilation for convective-permitting models, *The 8th EnKF Workshop*, Montreal, Canada

Kong, Fanyou, M. Xue, Y. Jung, K. A. Brewster, N. Snook, K. W. Thomas, and G. Zhao, 2016: CAPS Storm-Scale Ensemble Forecast Experiment in Supporting 2016 NOAA HWT and HMT. 28th Conference on Severe Local Storms, Portland, OR., AMS

Kong, F., M. Xue, Y. Jung, K. Brewster, N. Snook, K. Thomas, T. Supinie, and G. Zhao, 2016: Storm-Scale Ensemble Forecasts for the 2016 NOAA HWT Spring Forecasting Experiment and HMT FFaIR Experiment: An Overview. *2016 NJU-OU Symposium for Weather and Climate Research*, Nanjing, China

Kong, F., M. Xue, Y. Jung, K. Brewster, N. Snook, K. Thomas, T. Supinie, and G. Zhao, 2016: CAPS Storm-scale Ensemble Forecasts for the NOAA/HWT SPE2016, *The 2nd Ensemble Design Workshop*, College Park, MD, https://drive.google.com/drive/folders/0B6ikBQ6A--swSEFtUGJ3Mm1OcGs

Kong, F., M. Xue, Y. Jung, K. Brewster, and N. Snook, 2017: Storm-Scale Ensemble Forecasts for the NOAA HWT Spring Forecasting Experiment and HMT FFaIR Experiment. 9th European Conference on Severe Storms, Pula, Crotia.

Kong, F., M. Xue, Y. Jung, K. A. Brewster, N. Snook, C. Zhang, T. A. Supinie, K. W. Thomas, A. J. Clark, and I. L. Jirak, 2018: CAPS Storm-Scale Ensemble Forecast Experiment in Supporting 2018 NOAA HWT SFE and HMT FFaIR. *The 29th Conf. on Severe Local Storms*, Stowe, VT, AMS.

Kong, F., M. Xue, T. Supinie, X. Hu., K. Brewster, Y. Jung, and N. Snook, 2019: Storm-Scale Ensemble Forecasting during NOAA 2019 HWT and HMT using SAR-FV3 with Multiple Physics. 10th European Conference on Severe Storms, Krkow, Poland.

Snook, N., F. Kong, M. Xue, K. Brewster, K. W. Thomas, and T. Supinie, 2018: Verification of CAPS Storm-scale Ensemble Forecasts from the 2017 HMT FFaIR Experiment and a Patchwise Localized Probability Matched Mean for Ensemble Precipitation Forecasts. 29th Conference on Weather Analysis and Forecasting (WAF)/25th Conference on Numerical Weather Prediction (NWP), Denver, CO, Amer. Meteor. Soc., 5 Jun. 2018.

Snook, N., K. Brewsterrrrr F. Kong, M. Xue, 2020: Verification and Optimization of a Patchwise Localized Probability Matched Mean for Ensemble QPF. 30th Conference on Weather Analysis and Forecasting (WAF)/26th Conference on Numerical Weather Prediction (NWP), AMS, Boston, 2B.1

Supinie, T., C. Zhang, M. Xue, Brewster, K. A., K. Thomas, F. Kong, 2020: SAR FV3 Based Storm-Scale Ensemble Implementation and Testing for the 2019 HWT and FFaIR Experiments. 30th Conference on Weather Analysis and Forecasting (WAF)/26th Conference on Numerical Weather Prediction (NWP), AMS, Boston, J59.2.

2. PROGRESS AGAINST MILESTONES/SCHEDULES

Itemized tasks for Year 2016-2017:

Task 1. Develop and test software to interface the NMMB model with the CAPS SSEF system software, including ingest of CAPS 3DVAR/cloud analysis fields for IC, and interpolation of NMMB forecasts to a common grid for post-processing.

Task 2. Run 3 km CONUS baseline SSEF realtime forecasts using CAPS 3DVAR/cloud analysis DA with radar data plus SREF IC and LBC perturbations.

- Daily 10-member SSEF forecasts to 60 hours from 0000 UTC for 7-10 weeks starting mid-April 2016, and contribute to a coordinated pool of 3-km CONUS domain ensembles
- Forecasts will use WRF-ARW and NMMB dynamic cores with multi-physics options
- Forecast products will be generated and sent to the HWT

Task 3. Develop and test new EnKF DA capabilities based on GSI and CAPS's parallel EnKF DA system for the 3 km CONUS domain

- Develop and test GSI-based EnKF capabilities for the 3 km CONUS domain with operational RAP/HRRR prepbufr data stream
- Enhance CAPS's EnKF to support better support Thompson microphysics scheme used by ARW-based HRRR, including the consistent reflectivity operators and update of total number concentrations
- Interface the CAPS parallel EnKF DA system with the GSI-based EnKF system to allow for sufficiently fast realtime assimilation of high-density high-frequency radar data (the parallelization strategy in GSI-based EnKF is not fast enough for assimilating millions of radar data)

Task 4. Perform cycled EnKF DA and ensemble forecasts based on ARW core during the 2016 HWT Spring Experiment in near realtime, with 15-min DA cycles for one hour.

- Hourly GSI-based EnKF assimilating operational RAP/HRRR data stream 1800-2400 UTC, with 15-minute interval radar DA using CAPS's EnKF system as a follow-on step to the GSI-based EnKF at 2300 and 2400 UTC and by itself at 2315, 2330 and 2345 UTC.
- Ensemble forecasts of ~20 members plus a deterministic forecast from final EnKF analyses at 0000 UTC up to 60 hours, using the same multi-physics WRF-ARW configurations as the baseline SSEF system.
- The EnKF-based ensemble forecast products made available to HWT for evaluation in quasi-realtime. The ensemble forecasting performance will be compared with that of baseline SSEF in post HWT evaluation.

The project is on schedule and all tasks proposed for the year 2016-2017 have been completed (Tables 16).

Table 46. Proposed and accomplished tasks.

Proposed tasks		Tasks completed
Task 1. NMMB model addition	Code development and testing, post-processing	Completed
Task 2. Run 3 km CONUS baseline SSEF realtime forecasts using CAPS 3DVAR/cloud analysis DA with radar data plus SREF IC and LBC perturbations, for 7 weeks	Daily 10-member SSEF forecasts to 60 hours for 7-10 weeks	Completed (24 members, 7 weeks)
	WRF-ARW and NMMB dynamic cores with multi-physics options	Completed
	products generated and sent to the HWT	Completed

Task 3. Development and testing of new ensemble DA capabilities	GSI-based EnKF capabilities for a 3 km CONUS domain with operational RAP/HRRR prepbufr data stream	Completed
	Enhance CAPS's EnKF to support better support Thompson microphysics	Completed
	Interface the CAPS-EnKF-based DA system with the above GSI-based EnKF system	Completed
Task 4. Quasi-realtime cycled EnKF DA and ensemble forecasts during the 2016 HWT Spring Experiment	Hourly GSI-based EnKF with 15-minute interval radar DA for an hour using CAPS's parallel EnKF system	Completed **in realtime**
	Ensemble forecasts of 20 members plus a deterministic forecast from EnKF analysis in non-realtime	Completed **in realtime** (12 members)
	The EnKF-based ensemble forecast products made available to HWT	Completed **in realtime**
	Performance comparison of the baseline and EnKF SSEFs in post season	Completed

Itemized tasks for Year 2017-2018:

Task 1. Perform evaluations on the baseline SSEFs contributed by CAPS, NSSL, NCAR, and compare the performance of sub-ensembles consisting of different members from the pool, and suggest optimal configurations based on limited ensemble size (~10) that may potentially be implemented operationally. Make recommendations for EMC and GSD/ESRL for such configurations.

Task 2. Coordinating with NSSL and other contributing organizations, design improved 3DVAR-initialized SSEFs of 10-20 members and run them in realtime, providing forecast products to HWT Spring Experiment of 2017 for evaluation.

Task 3. Perform cycled EnKF DA and ensemble forecasts during the 2017 HWT Spring Experiment in realtime, with 10-min radar data cycles for one hour.

- Hourly GSI-based EnKF assimilating operational RAP/HRRR data stream 1800-2400 UTC, with 10-minute interval radar DA using CAPS's EnKF system as a follow-on step to the GSI-based EnKF from 2300 through 0000 UTC.

- Ensemble forecasts of ~20 members plus a deterministic forecast from final EnKF analyses at 0000 UTC up to 60 hours.

- Subjectively and objectively evaluate and compare EnKF-based ensemble forecasts with the baseline ensemble forecasts.

Task 4. Develop GSI-based hybrid DA capabilities for the 3 km CONUS grid.

- Develop, test and tune GSI-based coupled EnKF-En3DVar hybrid capabilities for the 3 km CONUS domain with operational RAP/HRRR GSI data stream, and the ability to directly assimilate radar data within the hybrid system.

- Run non-realtime hybrid DA experiments and tests for cases selected from realtime experiments, and compare the hybrid performance against GSI 3DVAR and pure EnKF, and determine the advantage of the hybrid scheme (if any) over pure EnKF.

Tasks for the year 2017-2018 have been completed.

Table 17. Proposed and accomplished tasks.

Proposed tasks		Tasks completed
Task 1. Evaluations on the baseline SSEFs	Post-season evaluations	Completed
Task 2. Design, Run 3 km CONUS SSEF realtime forecasts using CAPS 3DVAR/cloud analysis DA with radar data plus SREF IC and LBC perturbations	Design and test 3DVAR-based SSEF (10-20 members) for 2017	Completed
	Realtime forecast runs and products generated and sent to the HWT	Completed
Task 3. Perform cycled EnKF DA and ensemble forecasts for 2017 HWT	GSI-based EnKF capabilities for a 3 km CONUS domain with operational RAP/HRRR prepbufr data stream	Completed
	Realtime runs of up to 20 members and provide product to HWT	Completed
	Subjectively and objectively evaluate	Completed
Task 4. Develop GSI-based hybrid DA capabilities	Develop, test and tune GSI-based coupled EnKF-En3DVar hybrid capabilities	Completed

	Run non-realtime hybrid DA experiments and tests for cases	Completed

Itemized tasks for Year 2018-2019:

Task 1. In collaboration with GSD/NSSL/EMC, implement computationally efficient parallelization strategies within the GSI-based EnKF system. For optimal efficiency, conventional data will likely continue to use the strategy currently employed by GSI-based EnKF while for radar and dense satellite data, the domain-decomposition strategy used in CAPS's EnKF system may be used.

Task 2. Fully execute realtime GSI-based EnKF and EnKF-En3DVar hybrid DA with RAP/HRRR GSI data stream and full volume radar data during 2018 HWT Spring Experiment.

- Hourly GSI-based 40-50 member (depending on available computational resources) EnKF (for updating perturbation ICs) and En3DVar hybrid (for updating control analysis that also replaces the ensemble mean of EnKF DA) DA assimilating RAP/HRRR GSI data stream from 1800-2400 UTC, with 10-minute interval radar DA as part of the GSI-based ensemble DA between 2300 and 2400 UTC

- Ensemble forecasts of up to 50 members from the ensemble DA ICs plus a deterministic forecast from the En3DVar control analysis up to 60 hours, using ARW and NMMB dynamic cores with multi-physics configurations.

- Further refined ensemble forecast products available to HWT for evaluation in realtime.

Task 3. Compare forecasts using EnKF and hybrid En3DVar DA methods.

Tasks for the year 2018-2019 were completed.

Table 18. Proposed and accomplished tasks.

Proposed tasks		Tasks completed
Task 1. Optimal GSI-based EnKF system	Computationally efficient parallelization strategies	Completed
Task 2. Perform realtime hybrid DA with RAP/HRRR GSI data stream and full volume radar data and ensemble forecasts for 2018 HWT	GSI-based EnKF capabilities for a 3 km CONUS domain with operational RAP/HRRR prepbufr data stream	Completed
	Realtime runs of up to 50 members and provide product to HWT	Modified for selected cases in non-realtime

	Refined ensemble forecast products	Completed
Task 3. Compare forecasts using EnKF and hybrid En3DVar DA methods	Post season evaluations	Completed

3. ISSUES DELAYING CURRENT OR FUTURE PROGRESS

Transition from supercomputer computer systems used in 2016 HWT to new systems in TACC Stampede-KNL and PSC Bridges caused some headache due to platform configuration issues related to hardware and software. Right before the start of 2017 HWT (on May 1st), CAPS was forced to move 3DVAR-based ensemble forecast (23 total members) to a different system, Lonestar5, due to the delay in TACC transition from Stampede to Stampede2. The Bridges at PSC is also a new machine for the project (for the GSI+EnKF based ensemble). Despite these difficulties, CAPS managed to complete pre-season testing work. Due to a number of unexpected hardware- and network-related issues on both Lonestar5 and Bridges, many member forecasts were failed to run during the HWT SFE period. Missing members were reproduced after the SFE period that resulted in delayed post-season evaluation. These mechanical problems also affected development and testing of the GSI-based hybrid capabilities. The autopilot of the CAPS real-time Task Manager was interrupted frequently, requiring manual restoration and resubmission of jobs. The CAPS Scientist in charge of development of hybrid system had to step in to help.

For 2018 HWT SFE, the CAPS EnKF system was still used to assimilate radar data instead of the GSI-based EnKF system because of the performance issue with the latter. The GSI alone, which is used to produce observation priors, takes about 15 minutes to process both reflectivity and radial velocity data from the entire WSR-88D network on a large memory node, whose capacity is much bigger than those typically available on many supercomputers (e.g. It can run only on the large memory node (LSM) with 3 TB RAM on the PSC bridges system). Adding to time required to pre-process observations and to run the GSI-based EnKF system, the current GSI-based EnKF system is not practical with dense radar observations. For real-time applications, severe observation thinning would be necessary. On the other hand, the CAPS EnKF system is very efficient and does not require pre-calculated observation priors nor observation thinning. Radar DA takes about 4.5 minute using the CAPS EnKF system on 1014 cores, which is more realistic for CONUS 3-km grid real-time forecasts. Comparison of EnKF and hybrid En3DVar DA methods will be performed using several selected cases over a reduced domain.

4. INTERRACTION WITH NOAA

Regular meeting every three weeks staring from January with NOAA HWT team prior to HWT Spring Forecasting Experiment to discuss/coordinate ensemble configuration and product list, and logistics (Adam Clark, Israel Jirak, Steve Weiss, Kent Knopfmeier, Chris Melick, etc); weekly meeting within CAPS HWT team;

PIs Participated daily weather briefing during the HWT SFE weeks

Participating NOAA organized Ensemble Design Workshop in NCEP in August 2016 (Youngsun Jung, Fanyou Kong)

Teleconference and email exchanges between WPC HMT team (Sarah Perfater, Ben Albright, James Nelson, Sarah Trojniak) and PIs to discuss science and logistic for the HMT FFaIR2019 experiment.

Nate Snook (June, 2017) and Keith Brewster (July 2017) participated in the FFaIR2017, and Keith Brewster participated in FFaIR2018 at WPC as external forecasters one week each and each presented a seminar covering the CAPS SSEF for the HWT and HMT and relevant ensemble products research.

5. PREVIOUSLY UNREPORTED CHANGES

With the newly chosen FV3 model core in NOAA NGGPS program, CAPS is closely collaborated with GFDL to test FV3 system with nested 3-km CONUS CAM forecast and to contribute the Thompson and NSSL microphysics and MYNN and YSU PBL schemes into FV3 via CCPP. One FV3 forecast was run during both HWT and HMT periods using the University of Oklahoma's local supercomputing facility Schooner at the courtesy support by the university. This work was not in the original proposal task list, but the effort by CAPS to run a new version of FV3 3-km forecast and provide realtime product to HWT is highly appreciated by HWT and HMT teams.

For the 2018 HWT SFE season, the number of ensemble members for the EnKF-based ensemble is reduced from 12 to 10 because the NSSL microphysics scheme needs more nodes (25 nodes/member) than other microphysics schemes (17 nodes/member). The Bridges in PSC has limited computing resources available to CAPS 2018 HWT runs.

6. OUTCOMES FOR TRANSITION TO OPERATIONS

SSEF forecast dataset in GRIB2 form were transferred to NOAA HWT in realtime

SSEF dataset in GEMPAK form and GRIB2 form were transferred to HMT FFaIR in realtime.

LPM implemented by NCEP in operational product suite.

7. BUDGET ISSUES

n/a