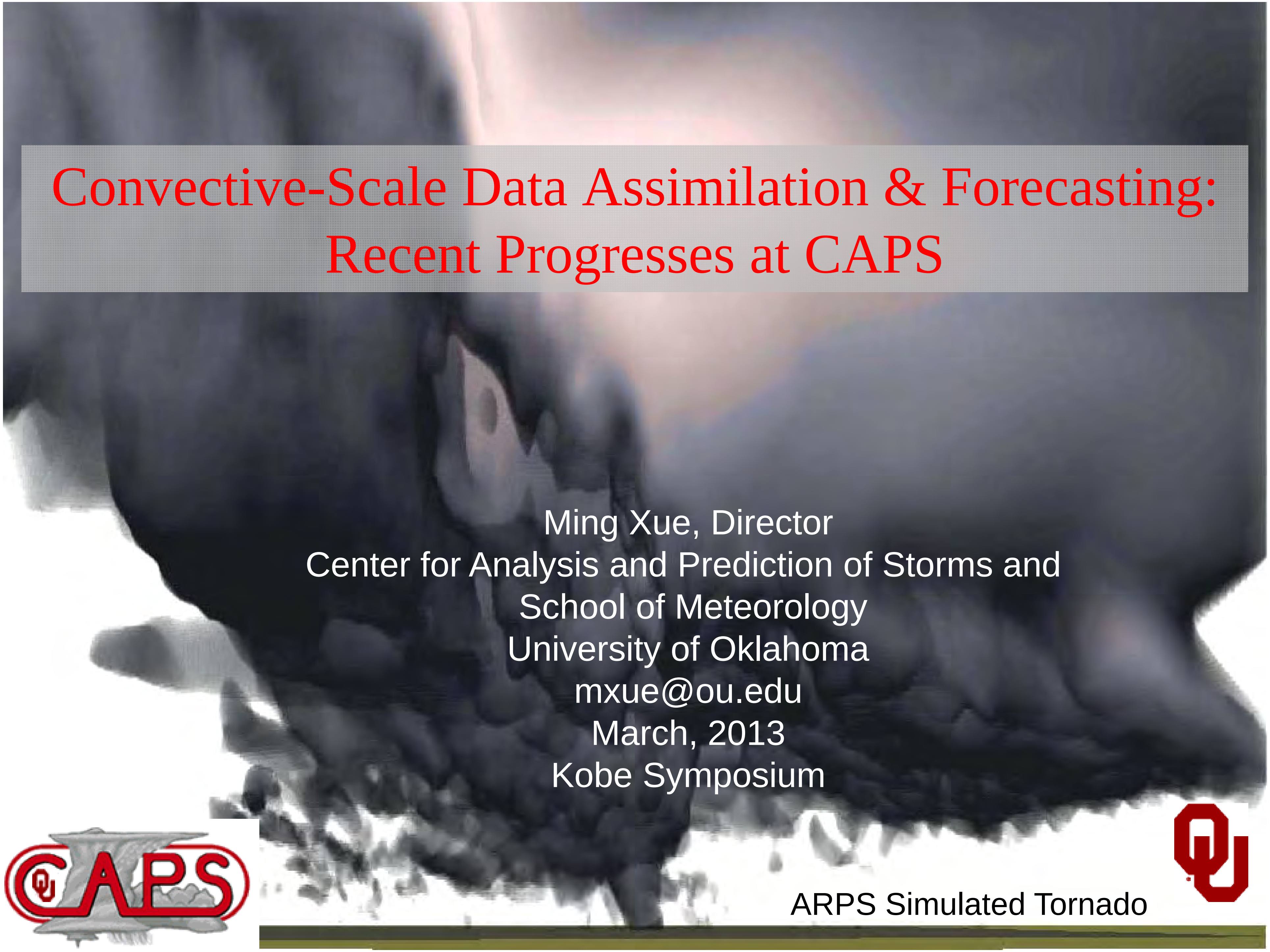
Convective-Scale Data Assimilation & Forecasting: Recent Progresses at CAPS

Ming Xue, Director
Center for Analysis and Prediction of Storms and
School of Meteorology
University of Oklahoma
mxue@ou.edu
March, 2013
Kobe Symposium



ARPS Simulated Tornado

Center for Analysis and Prediction of Storms

- Established in 1989 in the first group of National Science Foundation Science and Technology Centers (STC)

- CAPS's mission is to develop and demonstrate methods/techniques for the numerical analysis and prediction of high-impact local weather and environmental conditions.

- One of 8 University Strategic Organizations (USO) at OU

- Pioneered radar data assimilation and explicit storm-scale NWP

(http://www.caps.ou.edu)

Outline

- Convective scale/Radar DA problem/challenges;
- 3DVAR and EnKF DA for MCS and supercell tornado cases and Warn-on-Forecast concept;
- CONUS-scale storm-scale ensemble forecasting;
- 4DEnSRF, LETKF, parallel EnSRF, hybrid algorithm development and testing;
- Prototype EnKF and EnKF/hybrid for operational Rapid-Refresh system;
- Typhoon/hurricane radar DA.

Specific Challenges with Convective-Scale Data Assimilation

- Radar observes very few parameters (Vr and Z), no coverage in clear air region;
- Model constraint is essential to ‘retrieve’ unobserved variables;
- Analysis very sensitive to model errors, especially for the analysis of un-observed/retrieved parameters;
- Need accurate storm environment to support storm prediction – a multiscale/multi-data source problem;
- Much higher degree of freedom – more difficult for typical small-ensemble DA system to adequately sample essential directions to fit observations and retrieve unobserved quantities;

Specific Challenges with Convective-Scale Data Assimilation - continued

- Strongly nonlinear physics a serious obstacle for adjoint-based 4DVAR methods; Single-moment complex ice microphysics may not be enough;
- Ensemble system design/ensemble perturbation maintenance less well understood than larger-scale DA system;
- Much larger computational challenges;
- Radar data quality control remains a problem, QC is also flow-dependent.

Main Methods of Radar Data Assimilation

- 3DVAR
- 4DVAR
- EnKF
- En3DVAR/En4DVAR/Hybrid/4DEnKF/En4DVAR with and without adjoint, etc…
- Semi-empirical methods, e.g., complex cloud analysis
- Multi-step retrieval/analysis methods, e.g., Single-Doppler Velocity Retrieval (SDVR), thermodynamic retrieval techniques

Radial Velocity Assimilation

- Most times, only single-Doppler Vr data is available;
- Need to determine 3 wind components from one!
- Typically clear air data is un-useable;
- Radial velocity unfolding problem still a challenge;
- Analysis within 3DVAR straightforward but not necessarily accurate.

Reflectivity Data Assimilation

$$Z_e = Z_e(q_r, q_s, q_h, N_{0r}, N_{0s}, N_{0h}, etc.)$$

$$Z = 10\log(Z_e)$$

From Z, need to determine q_r, q_s, q_h, N_{0r}, N_{0s}, N_{0h}, and q_c, q_i, q_v, T, etc.

The problem is under-determined!

Need help! From physical constraints, e.g., numerical model.

Solutions

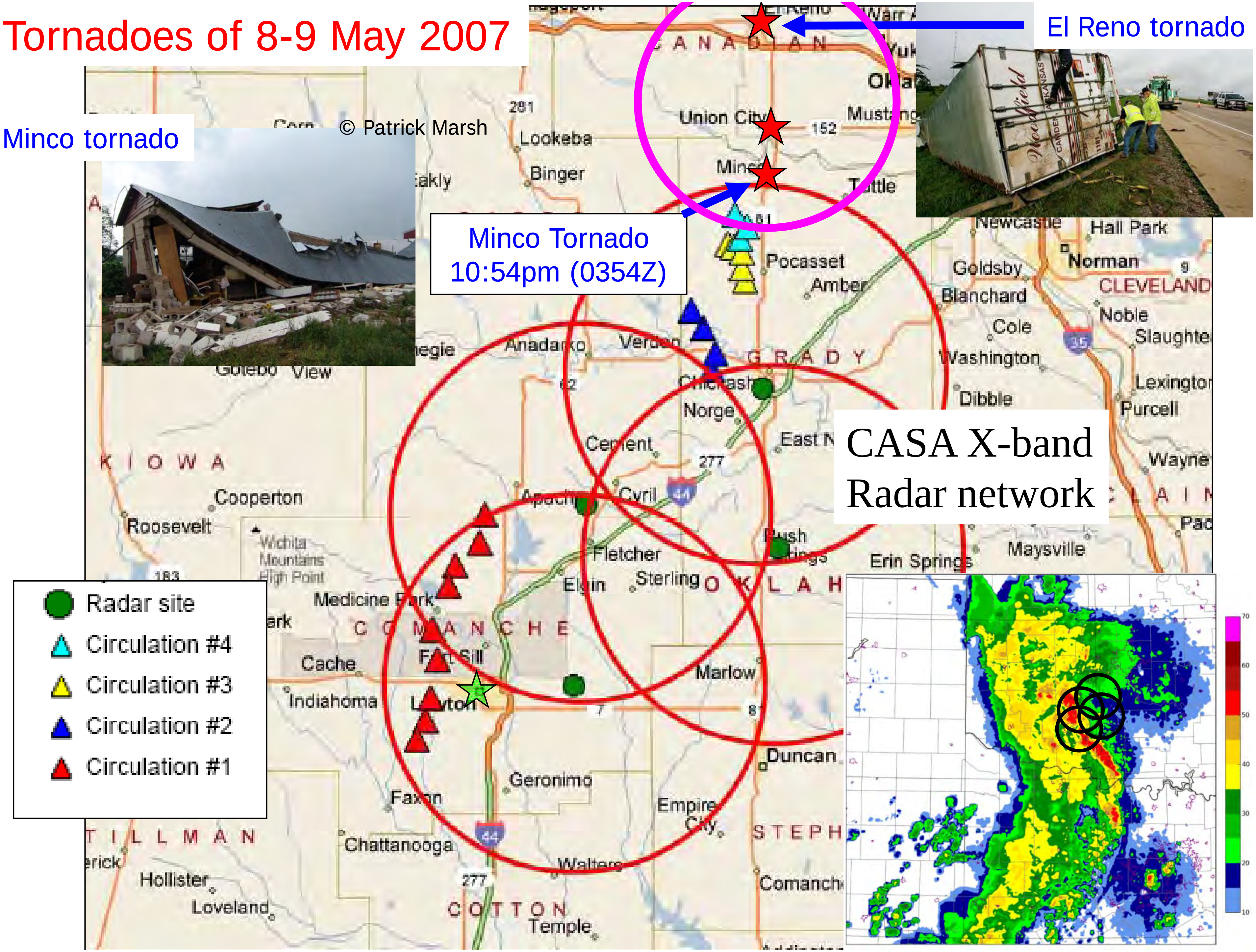
- 4DVAR uses model equations as a strong constraint (but has difficulties with nonlinear ice physics), and multi-time level data
- EnKF uses model constraints in a statistical representation, and accumulates info in time.
- Complex cloud analysis schemes employ physical models (e.g., parcel theory) and semi-empirical rules, and observations, and update those state variables they can.

May 7, 2007 Mesovortex Tornadoes within an MCS

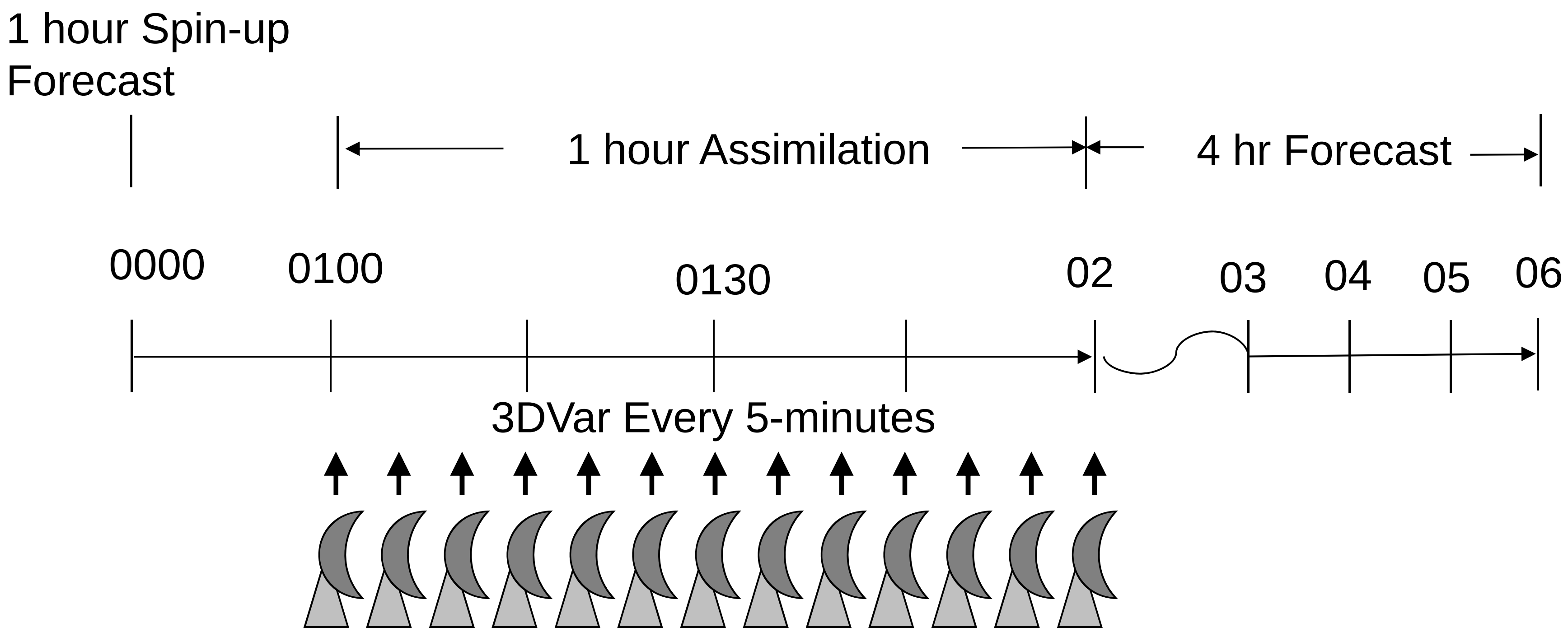
ARPS (Advanced Regional Prediction System)
3DVAR/Cloud analysis DA experiments
on 2 km, 400m grids
(Schenkmen et al. 2011a,b)

Diagnostic analysis on 100m nest
(Schenkman et al. 2012)

Tornadoes of 8-9 May 2007



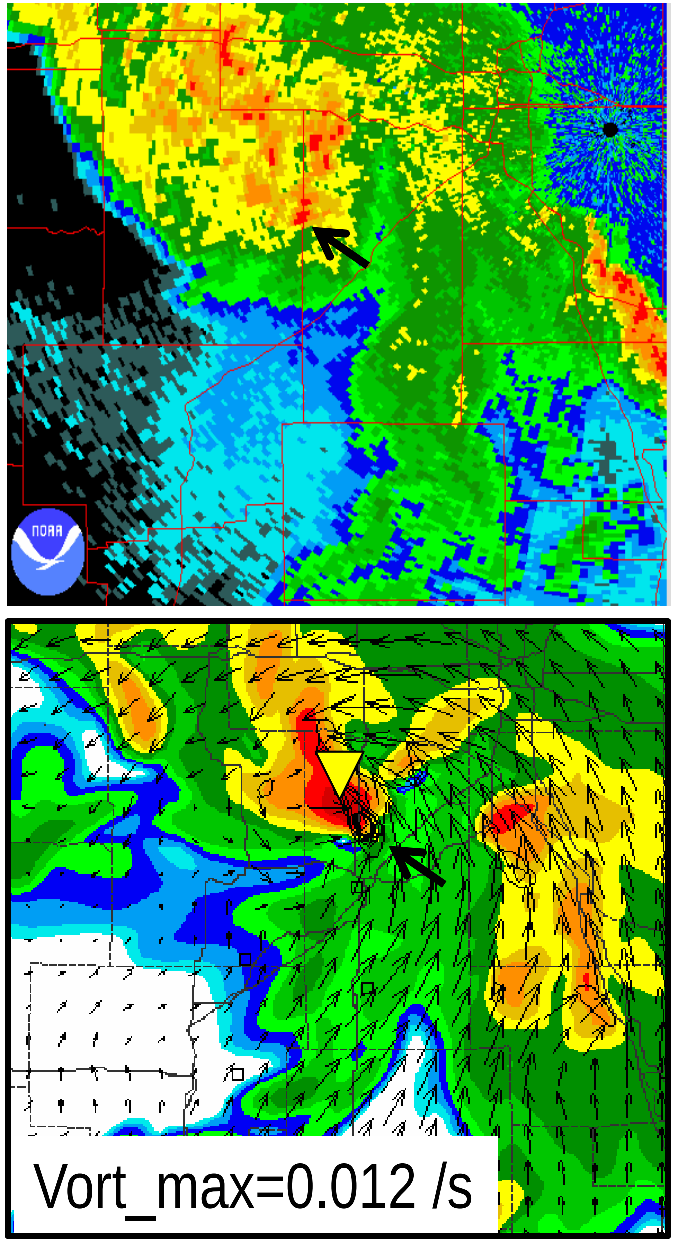
3DVar/Cloud Analysis Assimilation Experiments



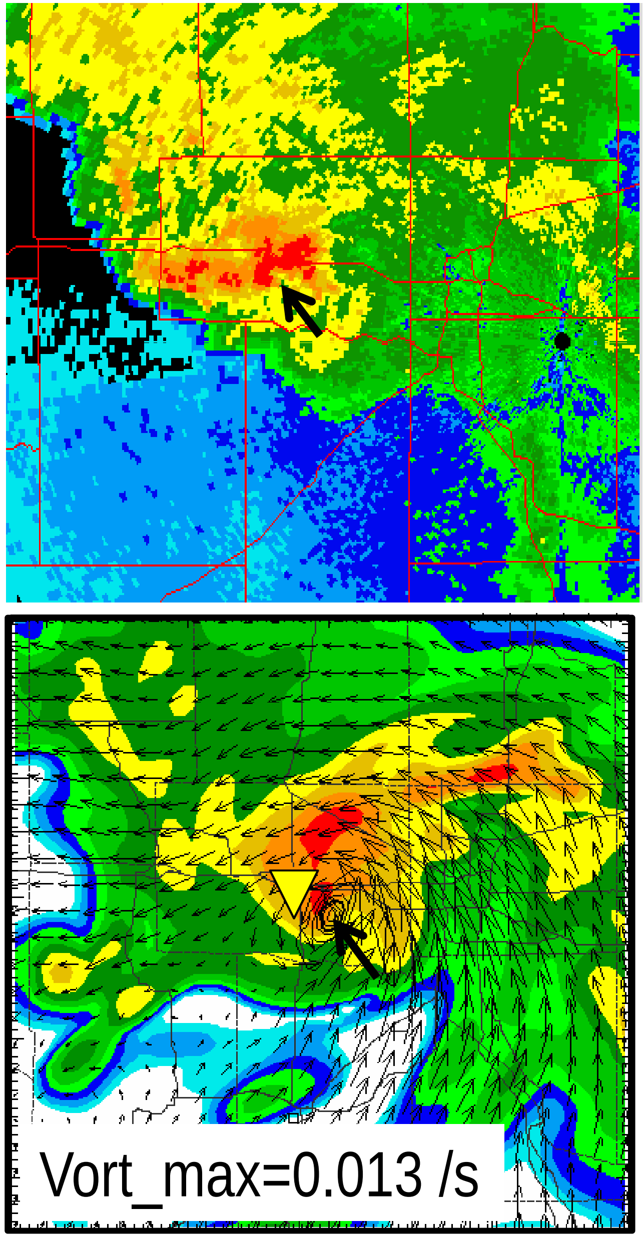
3DVar analysis run in 5-minute cycles for 1 hour at 2 km resolution

(Schenkman et al. MWR 2011a,b)

0350 UTC, 1h 50min forecast
Time of Minco Tornado



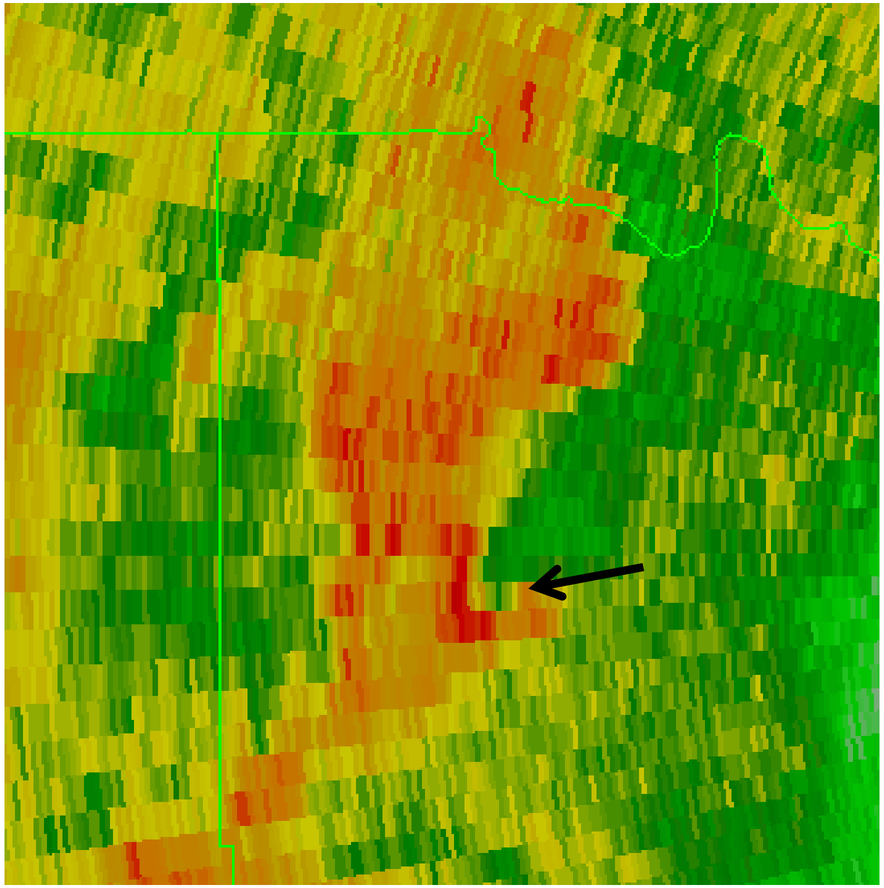
0440 UTC 2h 40m forecast
Time of El Reno Tornado



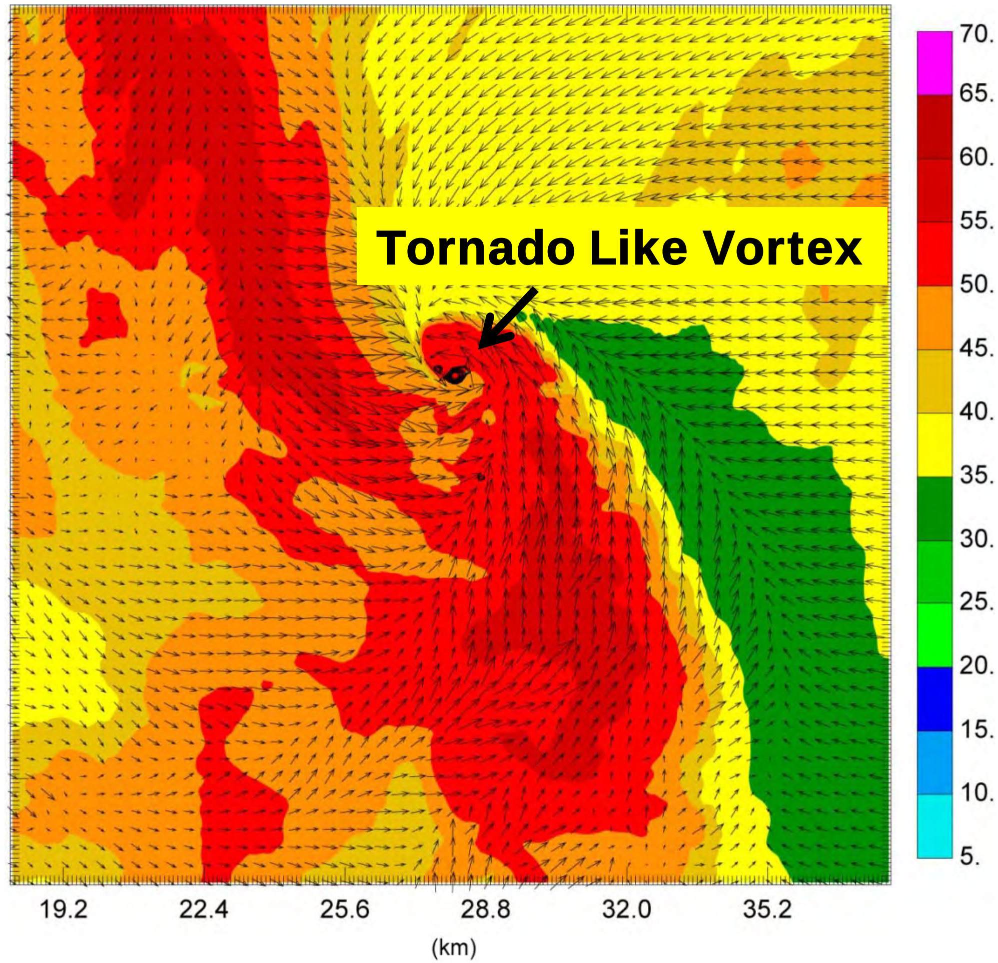
(Schenkman et al. MWR 2011a,b)

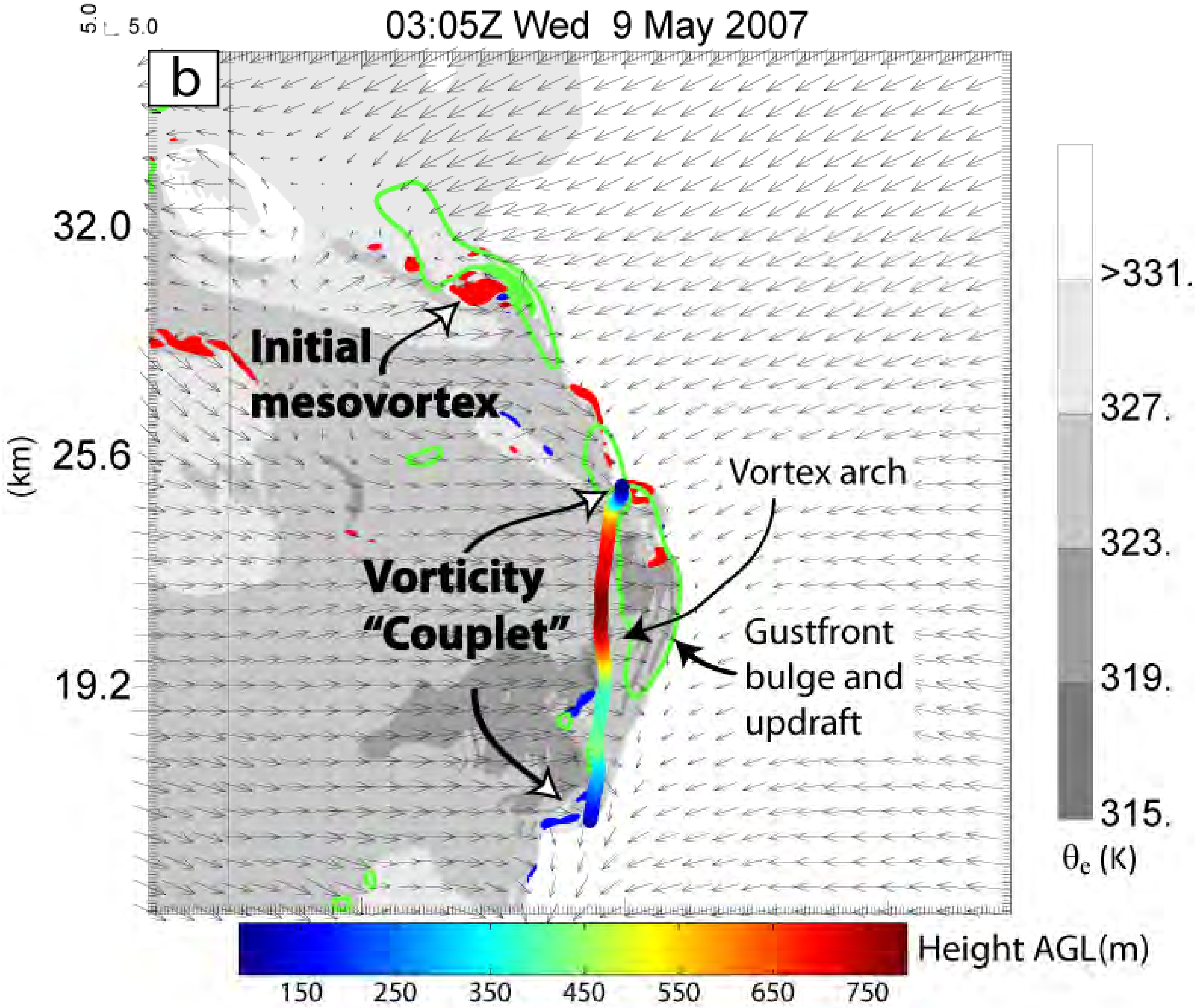
100 m nest captures TLV in Minco Mesovortex

Observed



100-m forecast

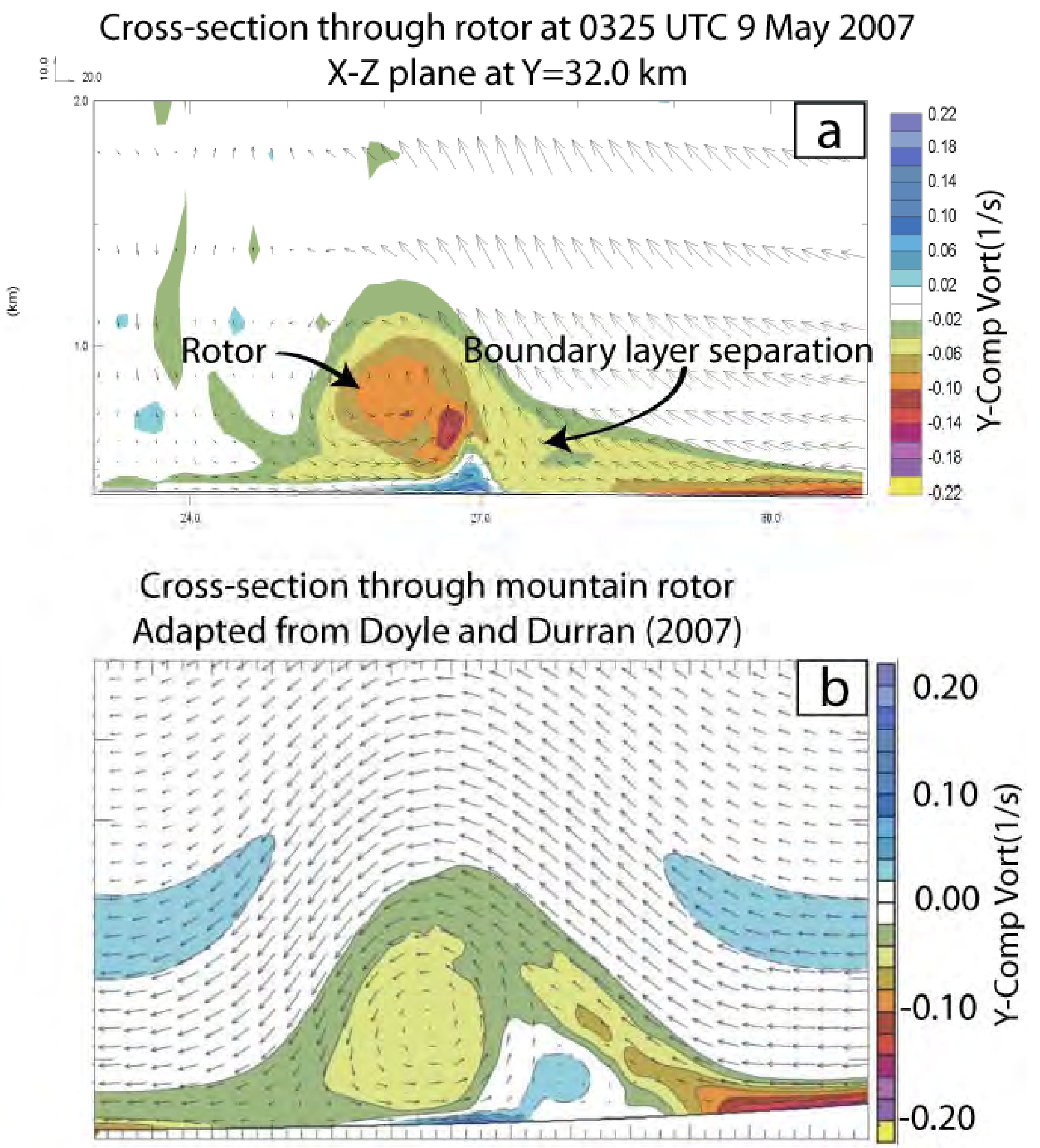


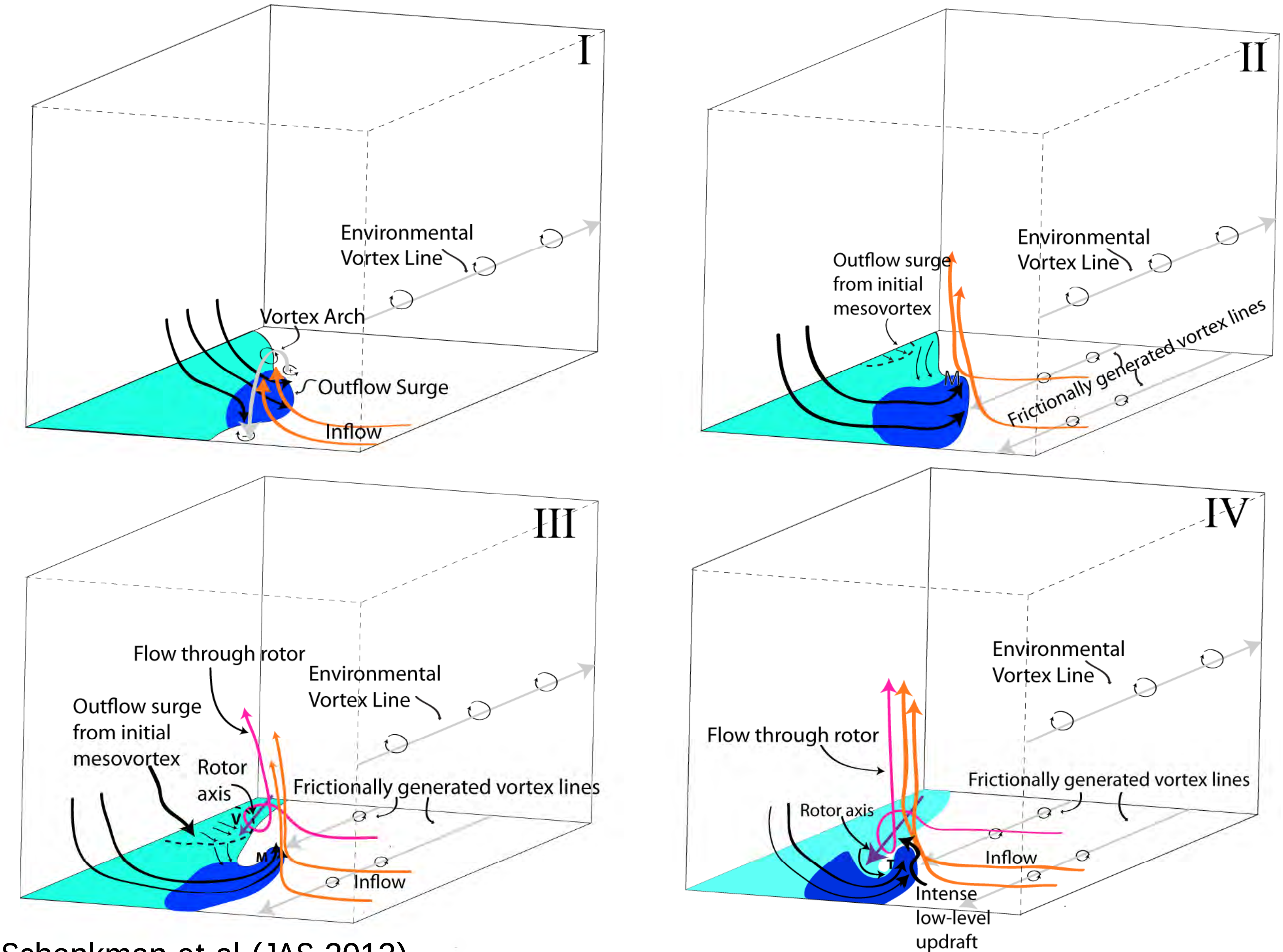


5.0
5.0
03:05Z Wed 9 May 2007
b
32.0
25.6
19.2
(km)
Initial mesovortex
Vortex arch
Vorticity "Couplet"
Gustfront bulge and updraft
>331.
327.
323.
319.
315.
θe (K)
150
250
350
450
550
650
750
Height AGL(m)

- Drag in model found to be responsible for horizontal rotor and strong updraft that concentrates vorticity to tornadic strength.

- Process found to be analogous to rotors that form in the lee of mountains.





Schenkman et al (JAS 2012)

EnSRF for (Polarimetric) Radar Data Assimilation

$$\underbrace{\begin{pmatrix} u^a \\ v^a \\ w^a \\ \theta^a \\ p^a \\ q_v^a \\ q_c^a \\ q_r^a \\ q_i^a \\ q_s^a \\ q_h^a \end{pmatrix}}_{N_x^a} = \underbrace{\begin{pmatrix} u^f \\ v^f \\ w^f \\ \theta^f \\ p^f \\ q_v^f \\ q_c^f \\ q_r^f \\ q_i^f \\ q_s^f \\ q_h^f \end{pmatrix}}_{N_x^f} + \alpha K \left[\begin{pmatrix} V_r \\ Z_{hh} \\ Z_{vv} \\ Z_{DR} \\ Z_{DP} \\ K_{DP} \end{pmatrix} - H \underbrace{\begin{pmatrix} u^f \\ v^f \\ w^f \\ \theta^f \\ p^f \\ q_v^f \\ q_c^f \\ q_r^f \\ q_i^f \\ q_s^f \\ q_h^f \end{pmatrix}}_{N_x^f} \right]$$

Radar data used:

V_r (Z_{hh} > 10 dBZ)

Z_{hh} (everywhere)

:

K_{DP} (K_{DP} > 0.3 deg./km)

$V_r = V_r(u, v, w)$

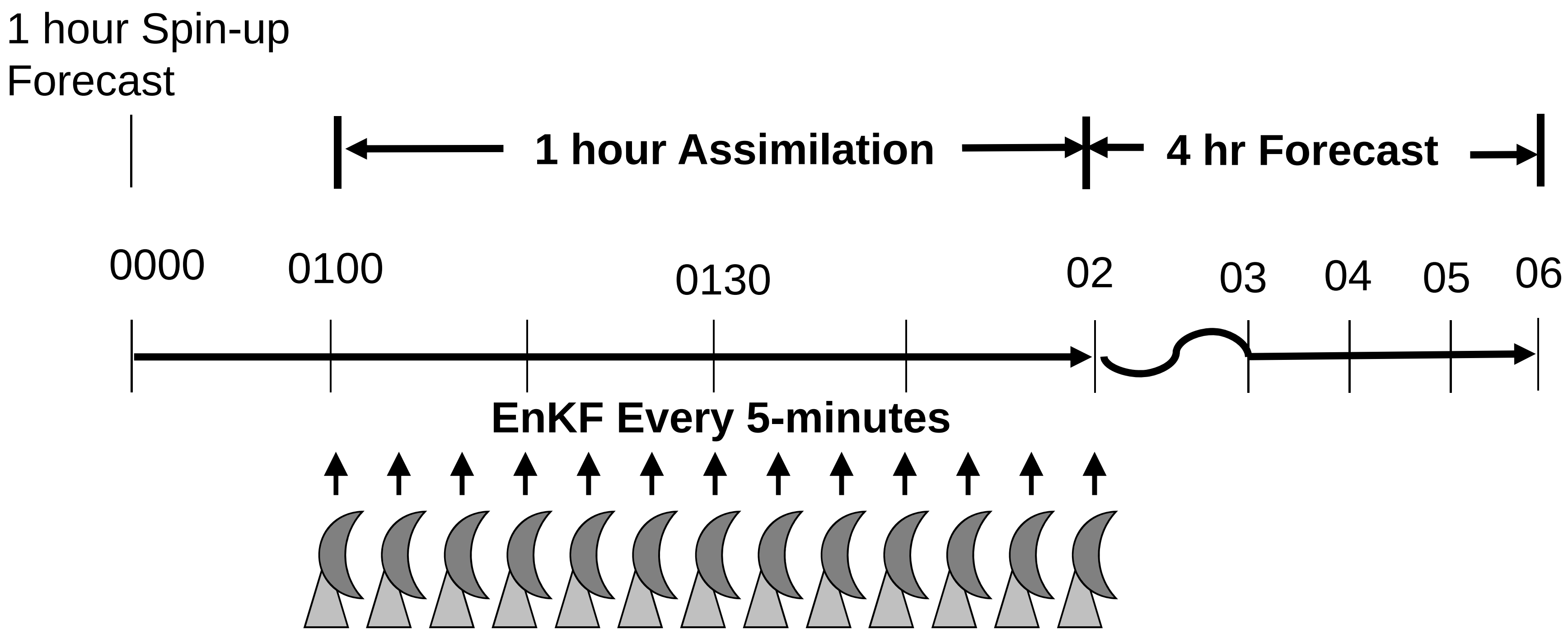
$Z_{hh} = Z_{hh}(q_r, q_s, q_h)$

:

$K_{DP} = K_{DP}(qr, qs)$

K is the Kalman Gain, a function of background and obs error covariances

ARPS EnKF Data Assimilation Experiments

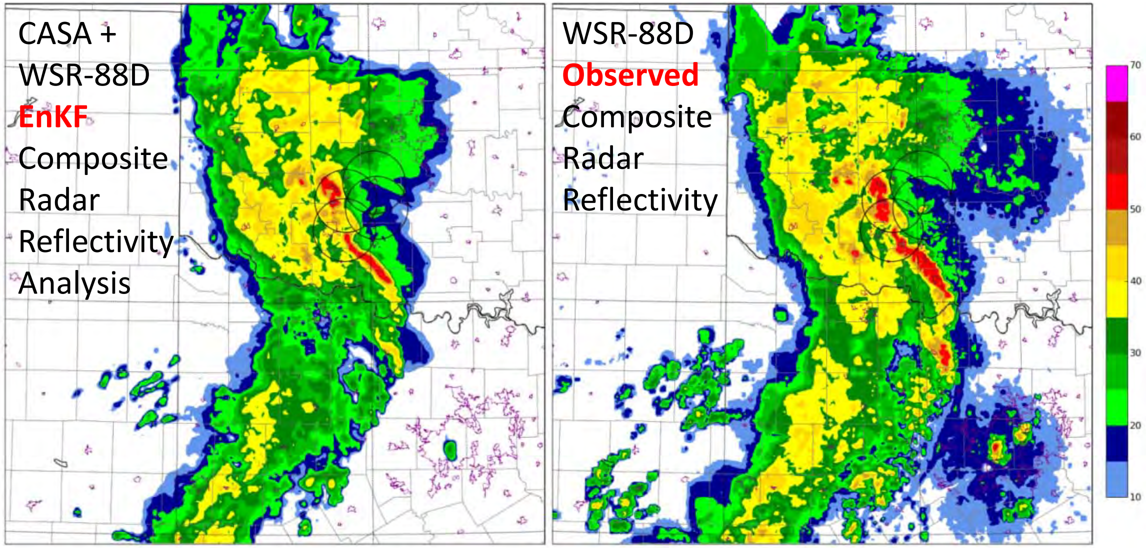


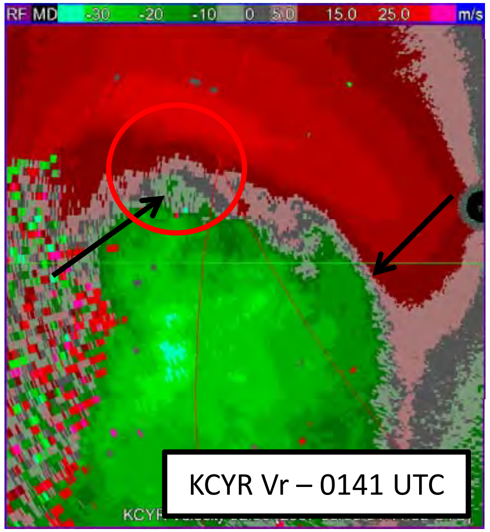
EnKF DA at 5-min intervals for 1 hour
at 2 km resolution

(Snook et al. 2011, 2012, Putnam et al. 2013)

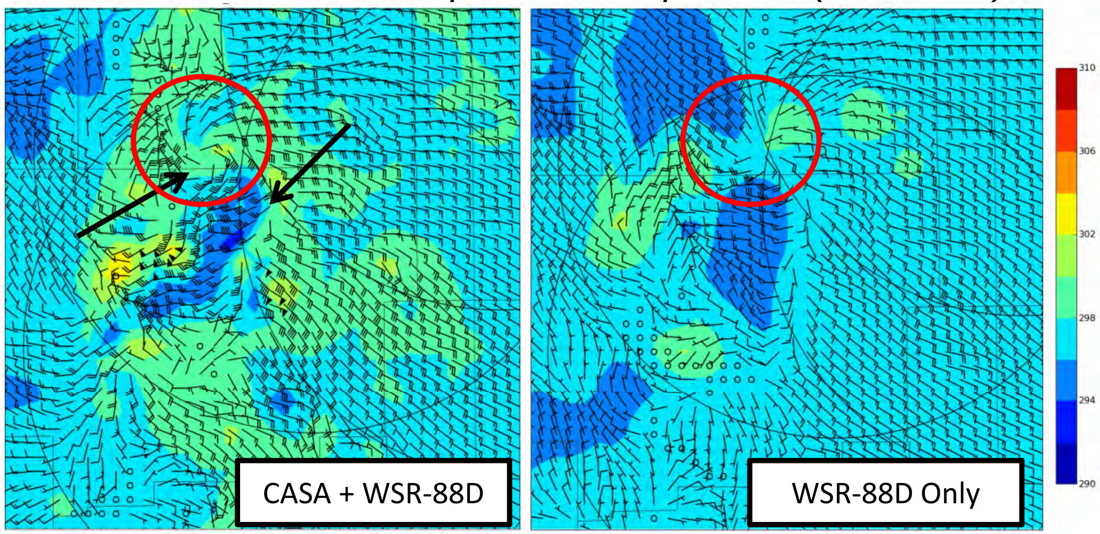
Assimilation of CASA and 88D radar data using mixed-microphysics ensemble: Results of EnKF Analysis using dx=2km

Final analysis (0200 UTC) Reflectivity





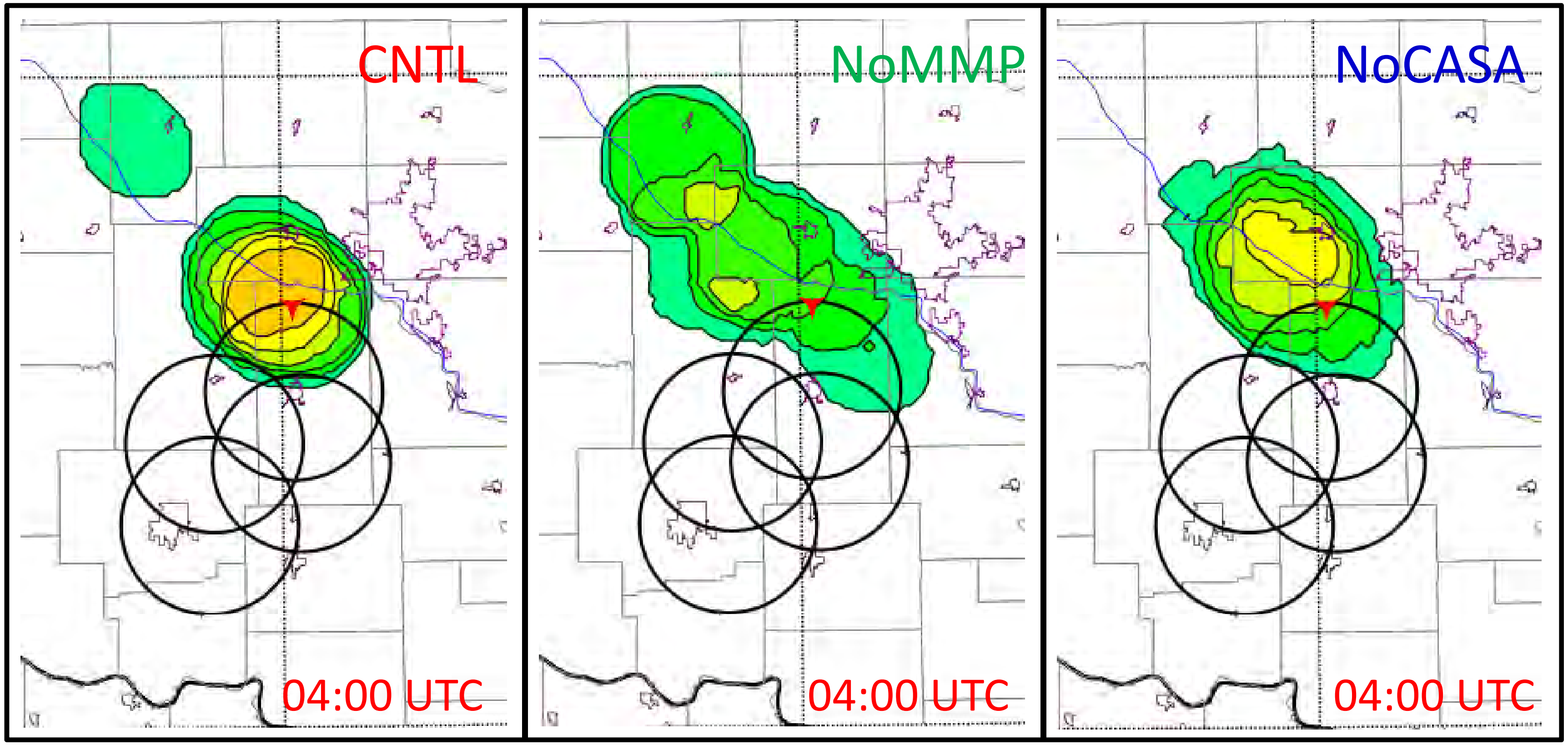
Near-surface winds and potential temperature (0140 UTC)



Snook et al (MWR 2011)

Effects of Assimilating CASA data and Mixed-microphysics Ensemble

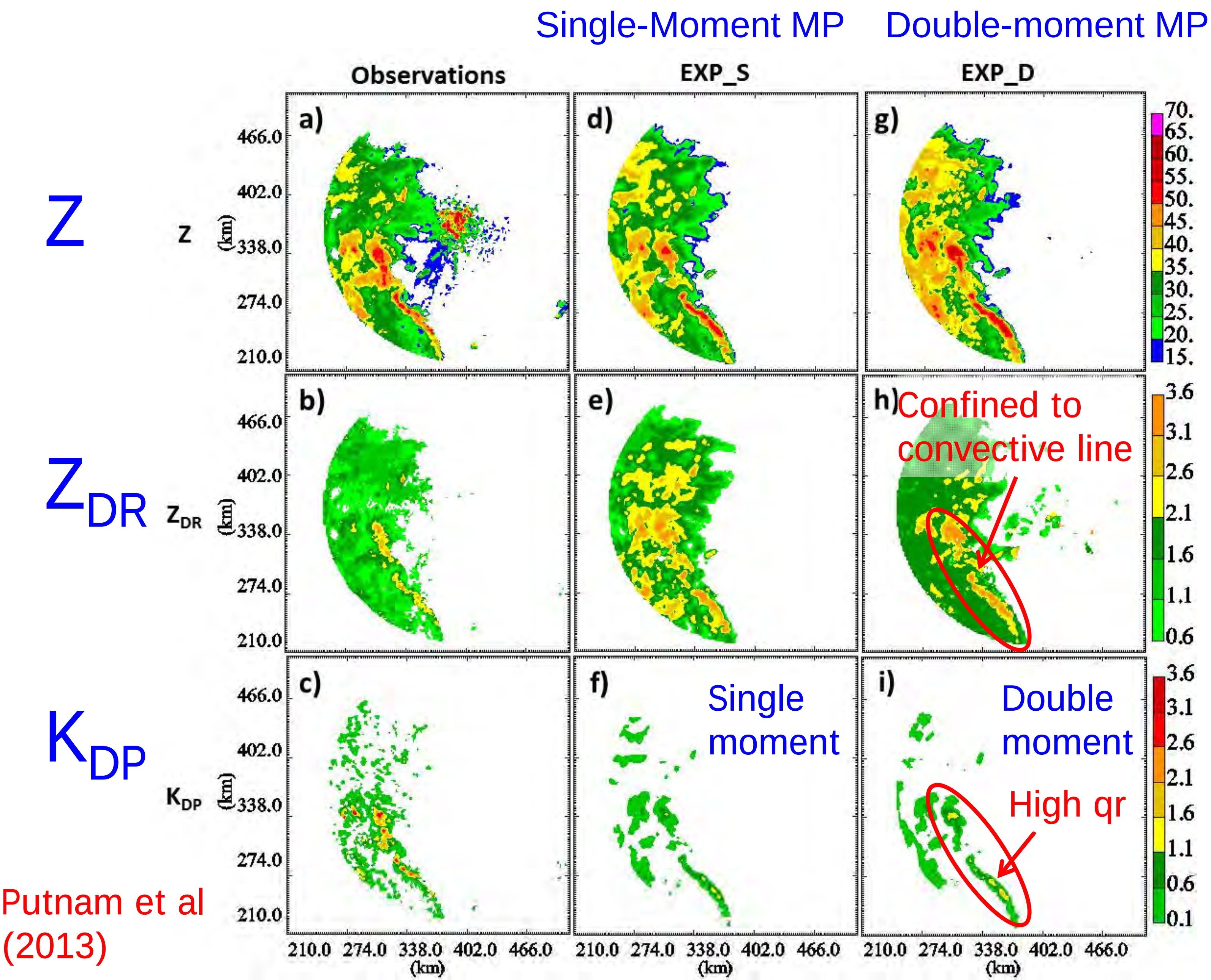
2h Probabilistic Forecasts of Near-surface Vortices



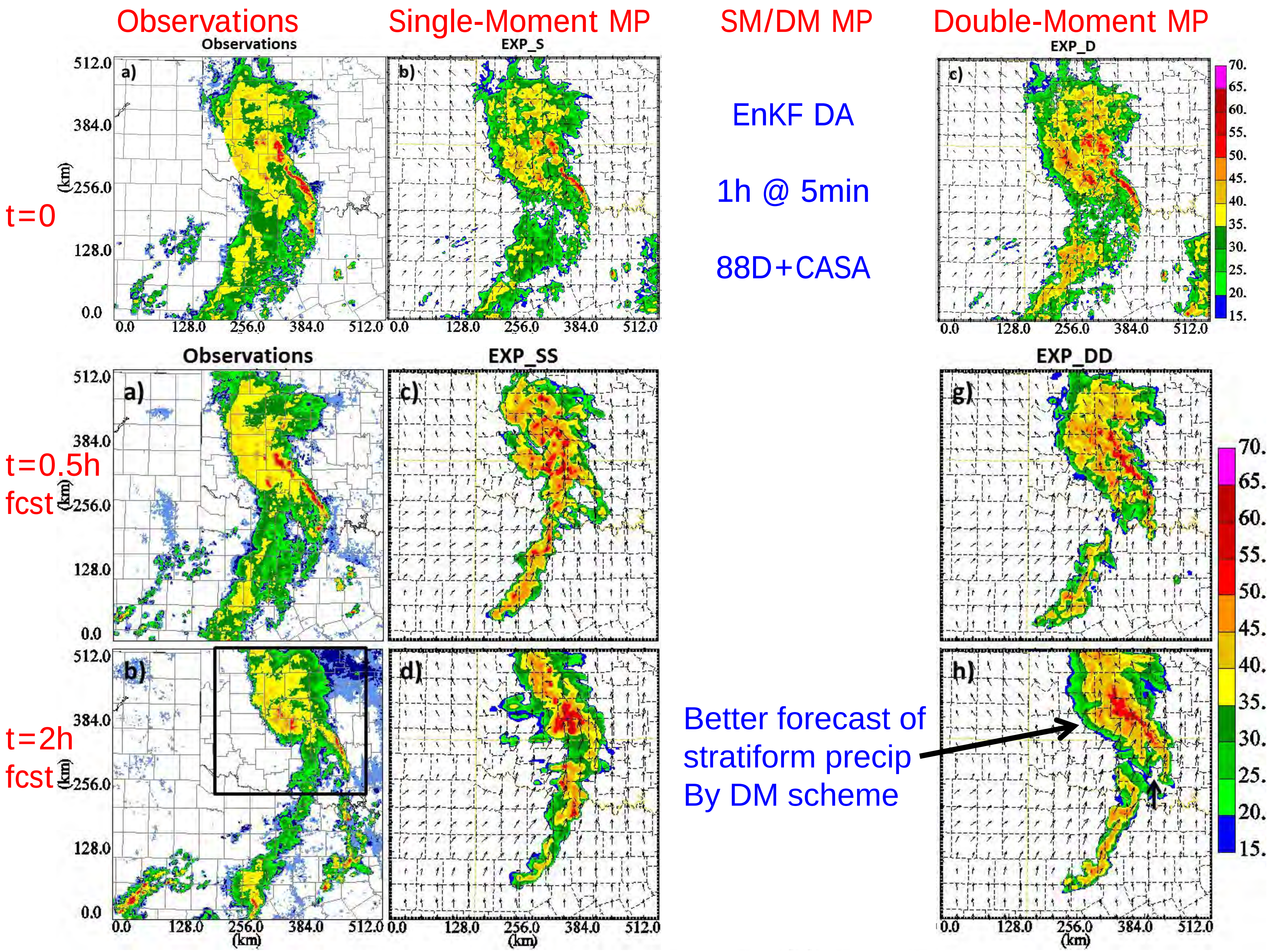
CNTL: 0.65 NoMMP: 0.35 NoCASA: 0.43

Snook et al (MWR 2012)

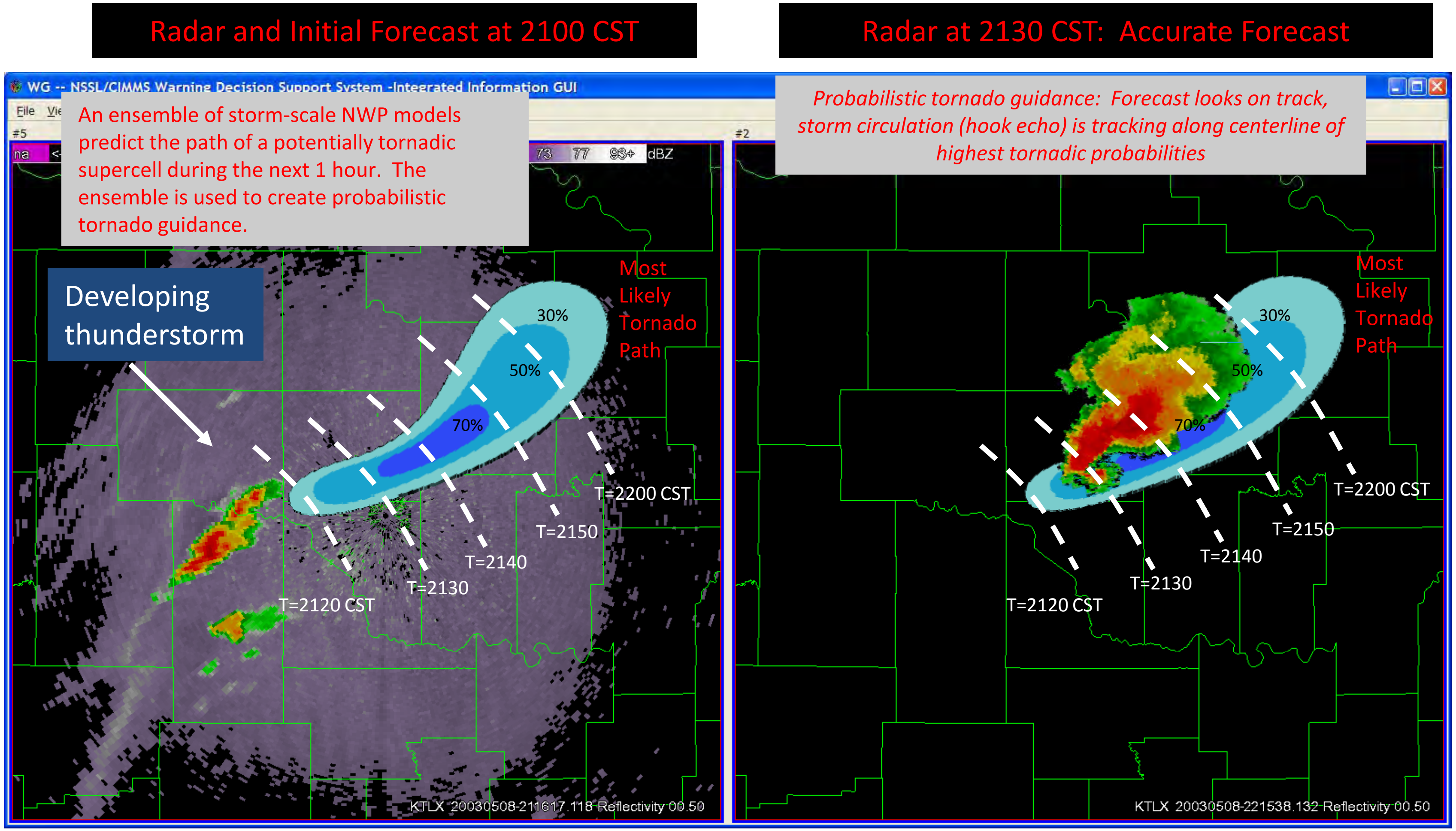
EnKF Analyses of Dual-Pol Variables at 0.5 tilt



Putnam et al (2013)


Observations
Single-Moment MP
SM/DM MP
Double-Moment MP
EnKF DA
1h @ 5min
88D+CASA
t=0
t=0.5h fcst
t=2h fcst
Observations
EXP_S
EXP_D
EXP_SS
EXP_DD
Better forecast of stratiform precip By DM scheme

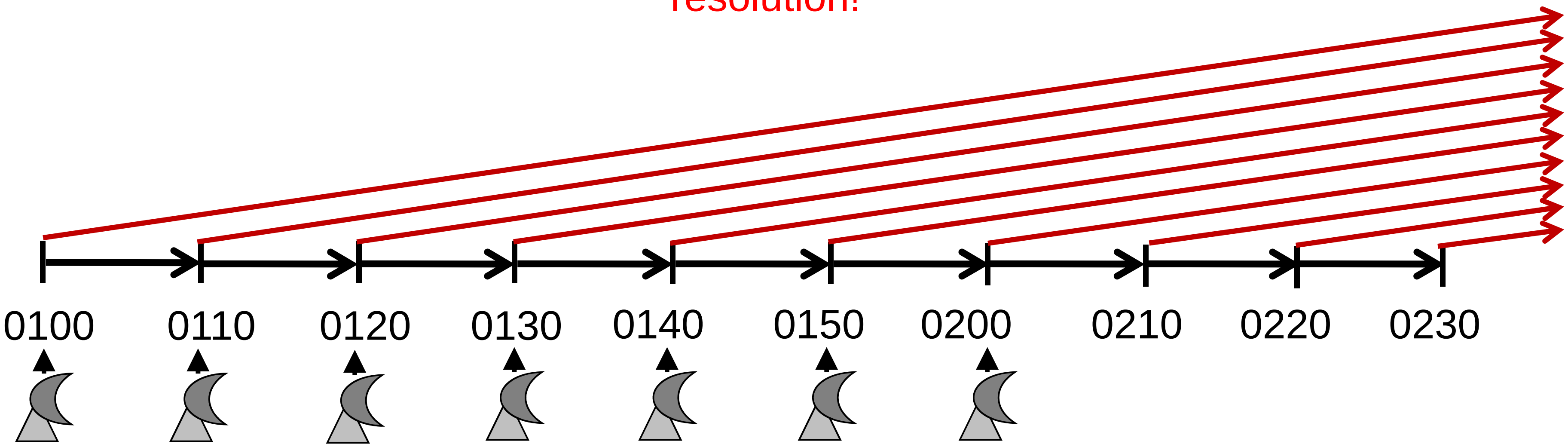
Probabilistic Warn-on-Forecast for Tornadoes
- The ultimate challenge – need ~100 m resolution



(Stensrud, Xue, et al. BAMS 2009)

CASA Realtime Forecast Configuration of Spring 2010

2 hour-long rapidly updated forecasts, every 10 minutes, at 1 km resolution!



88D + CASA Vr and Z data

350 x 320 x 53 grid using 800 cores

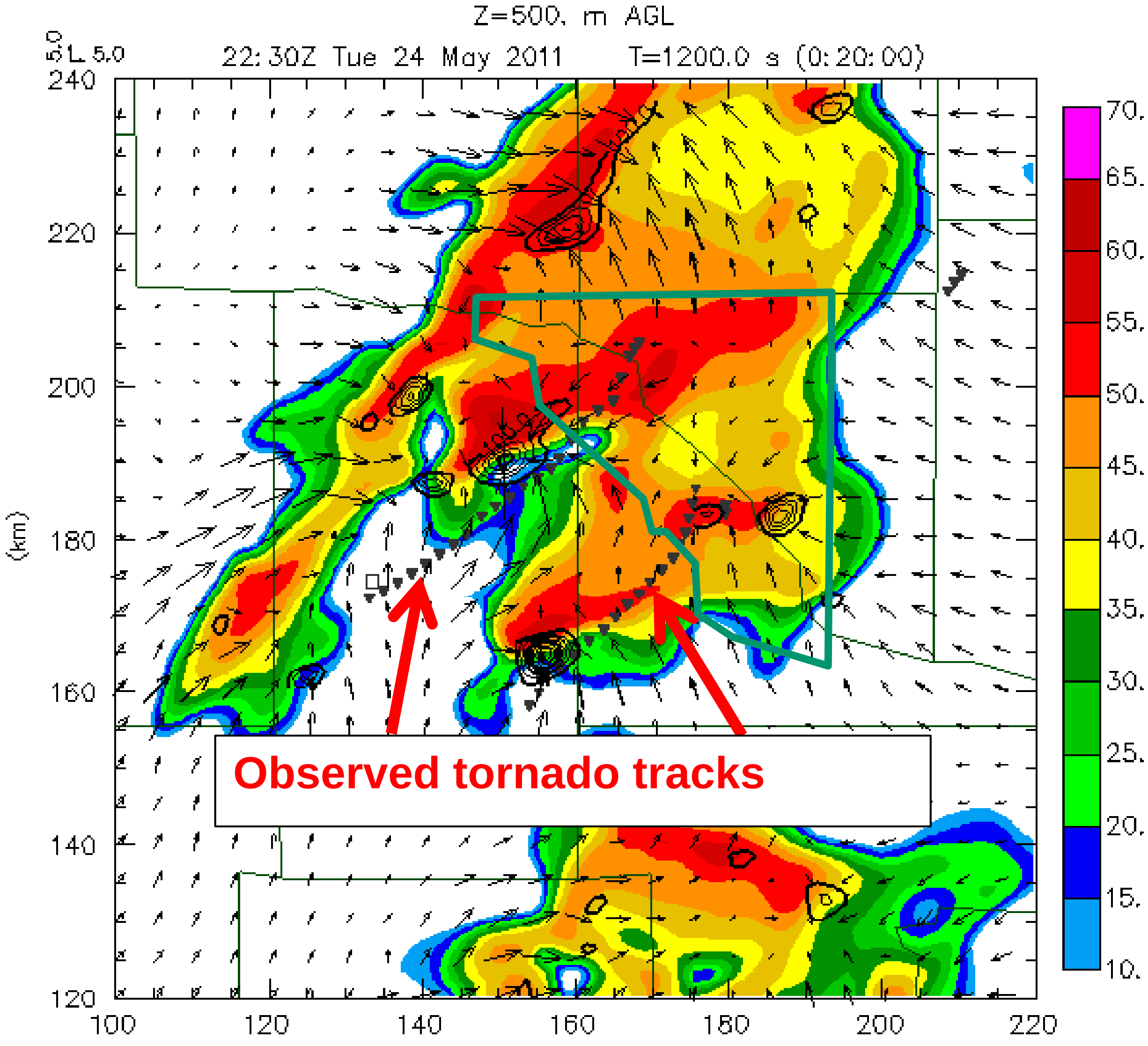
Data ingest, QC, pre-processing, DA, forecast, post-processing, product generation all completed within 10 minutes!

Every single step is parallelized using MPI.

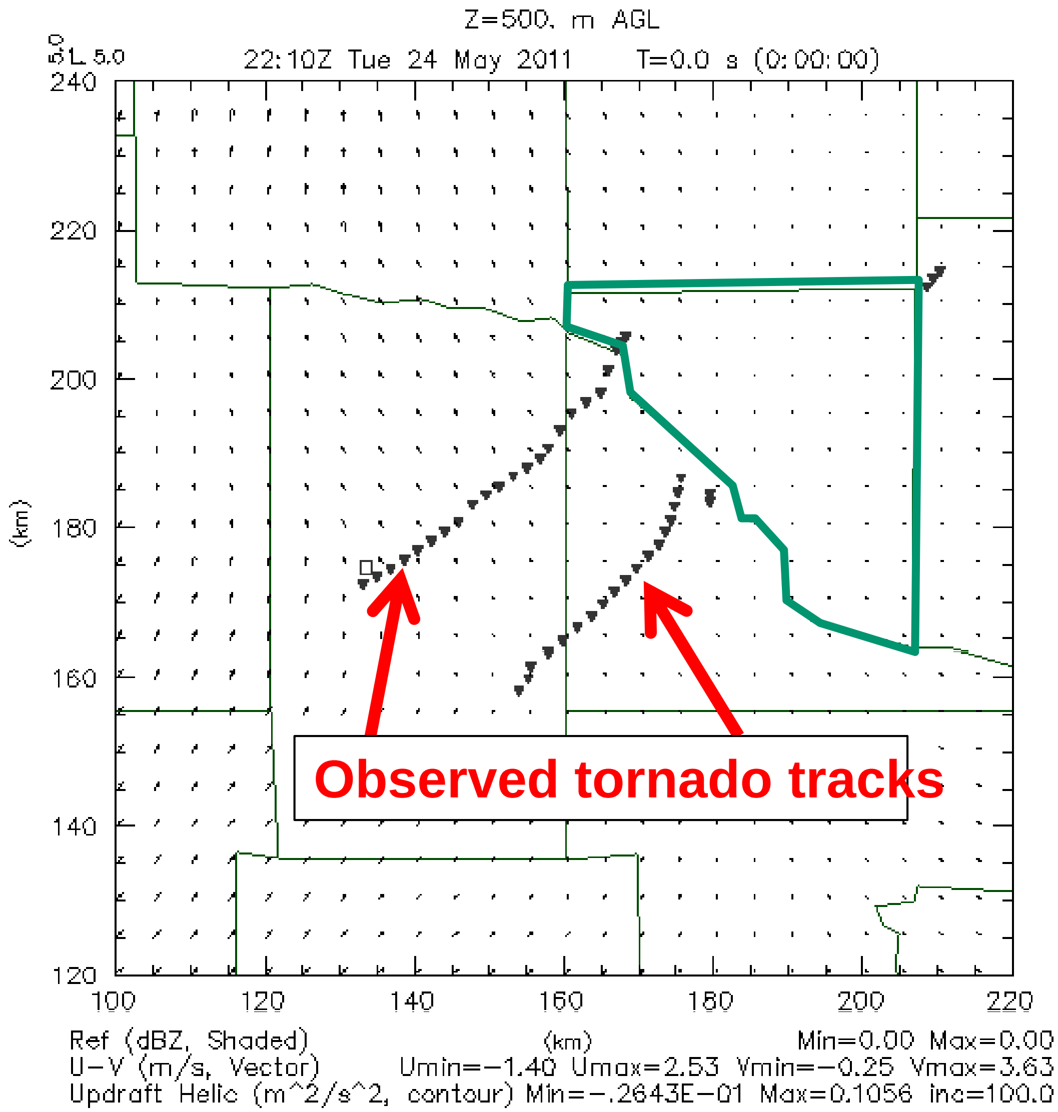
20 min forecast of low-level reflectivity and vertical vorticity for 24 May 2011 tornado case (dx=1km)

ARPS forecast model

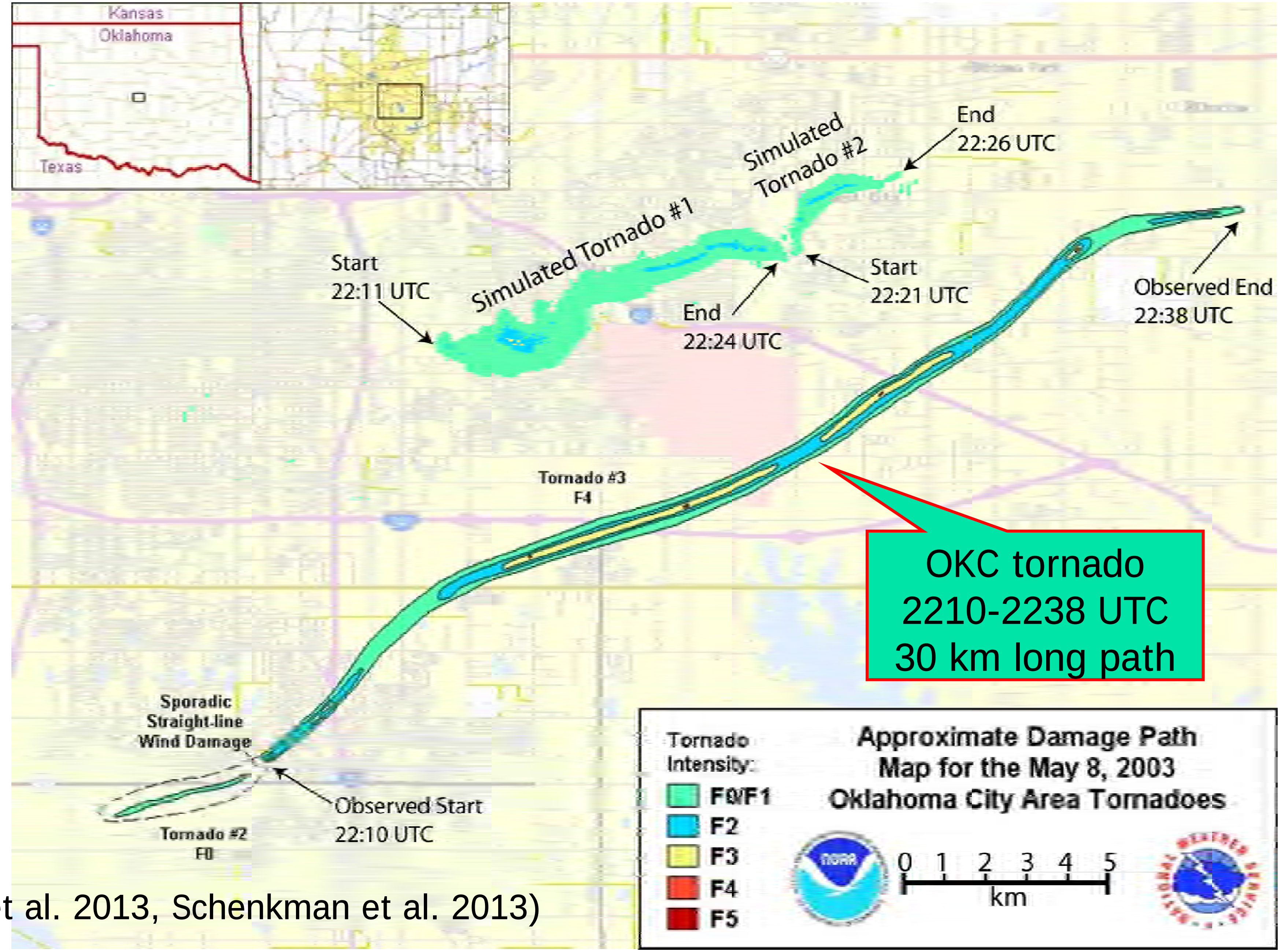
ARPS 3DVAR Cloud analysis



70 min forecast of low-level Z and vertical vorticity for 24 May 2011 tornado case (animation)

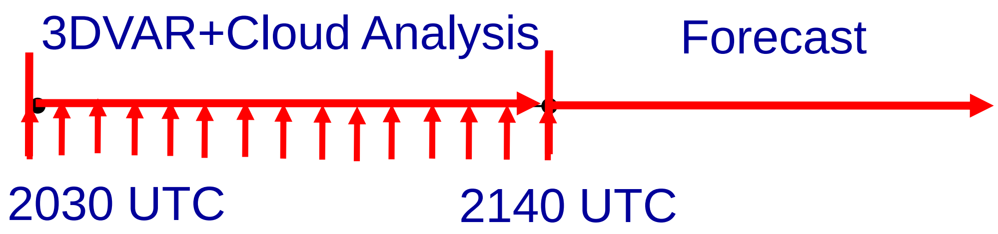


May 8th, 2003 OKC tornado

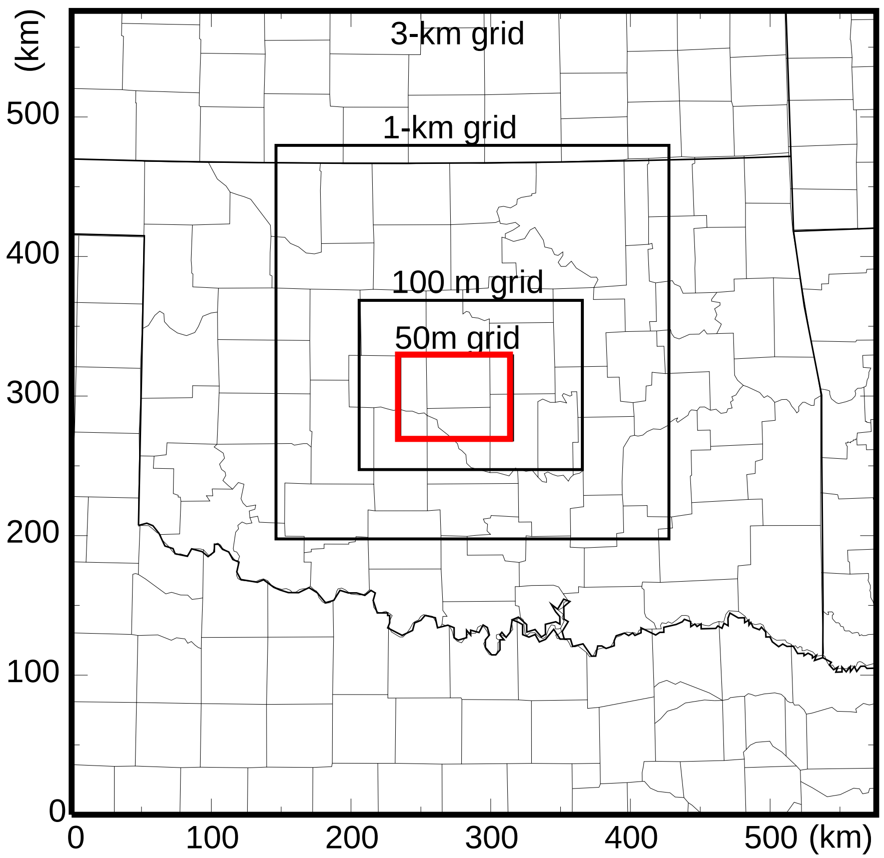


(Xue et al. 2013, Schenkman et al. 2013)

DA cycles on 1-km Grid

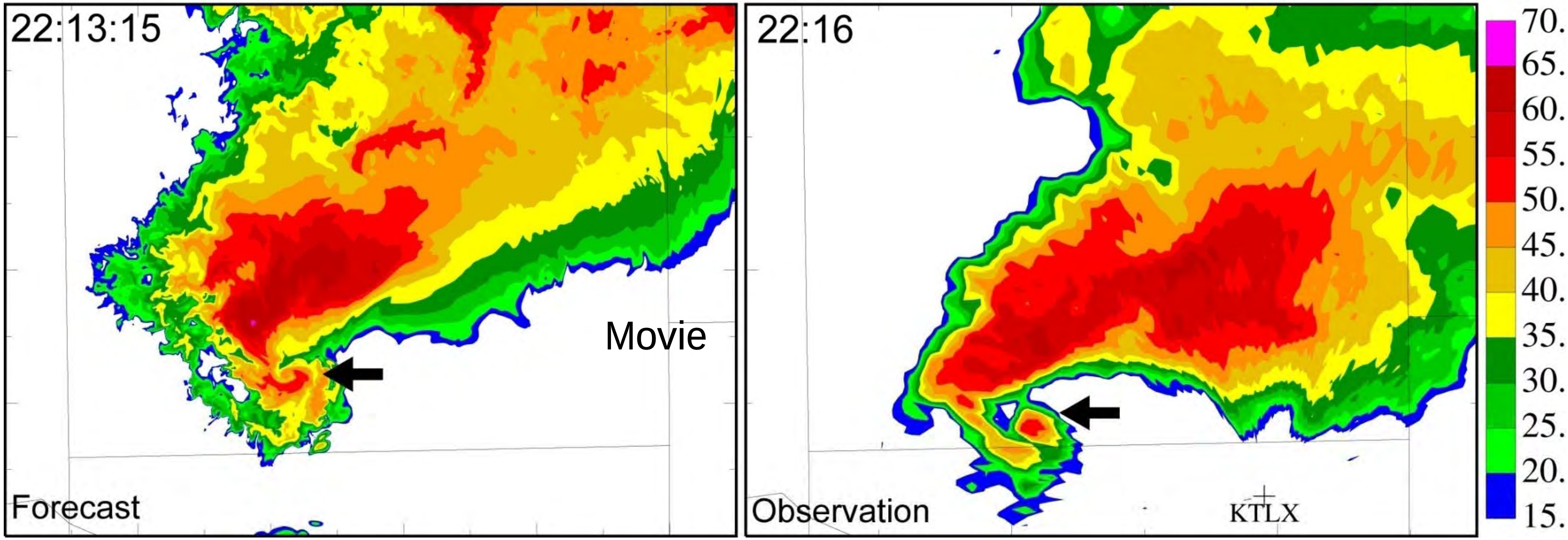


4 nested grids



Short-Range Radar Initialized Prediction of Thunderstorms, Strong Winds, Gust Fronts, Downbursts, and Tornadoes using NWP Model

50-m Grid Forecast v.s. Observation (Movie)

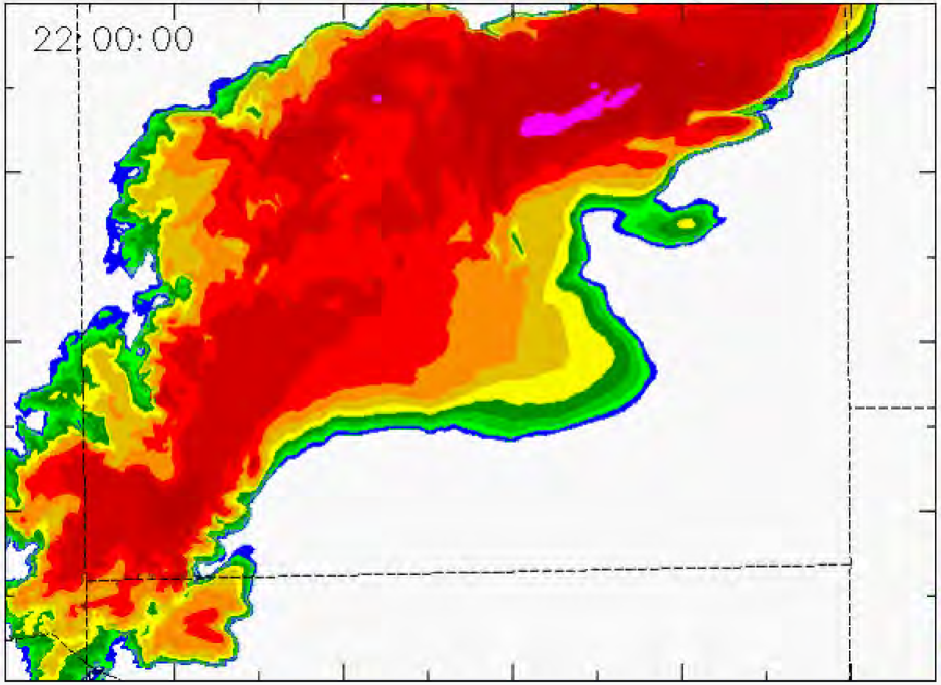


Forecast Low-level Reflectivity

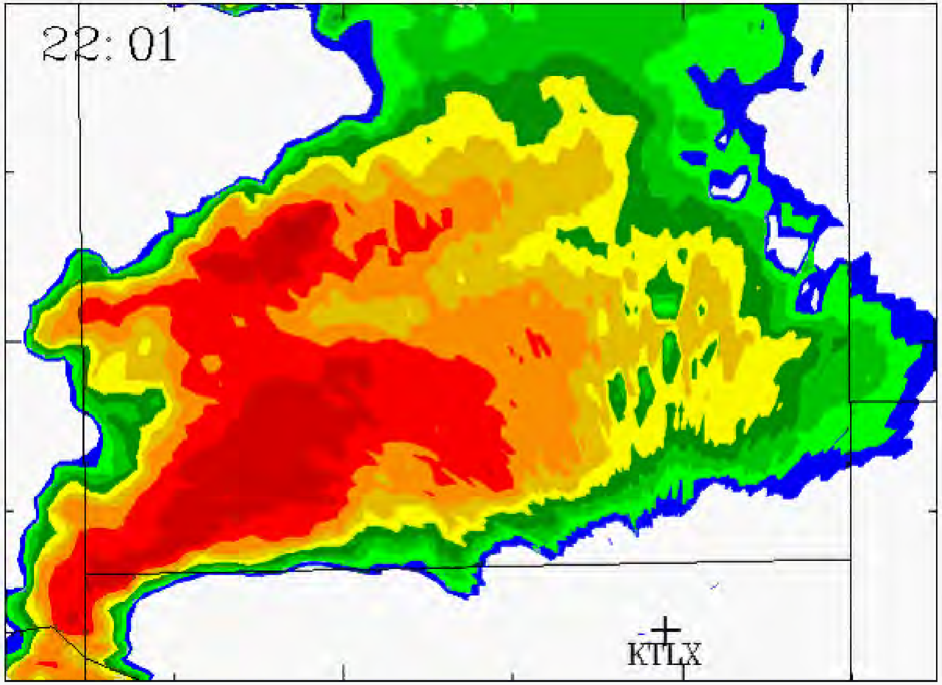
Observed Low-level Reflectivity

43 minute forecast

50-m Grid Forecast v.s. Observation

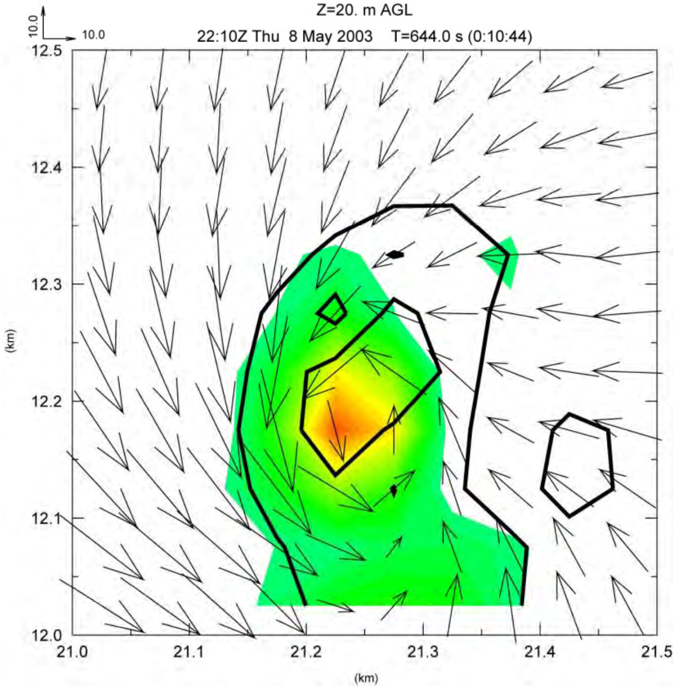


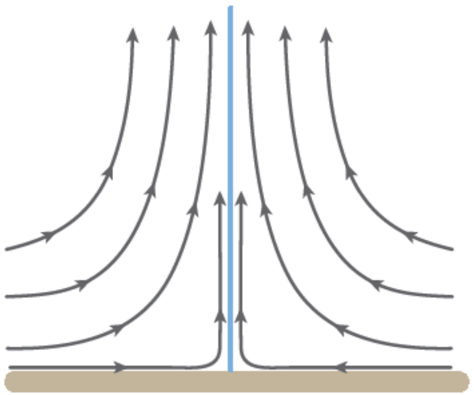
Forecast Low-level Reflectivity



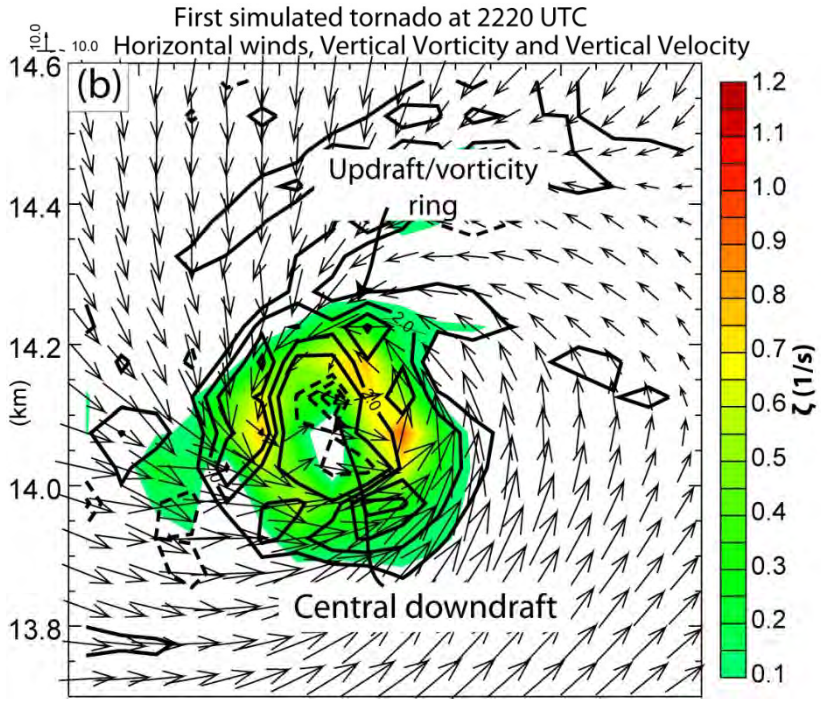
Observed Low-level Reflectivity

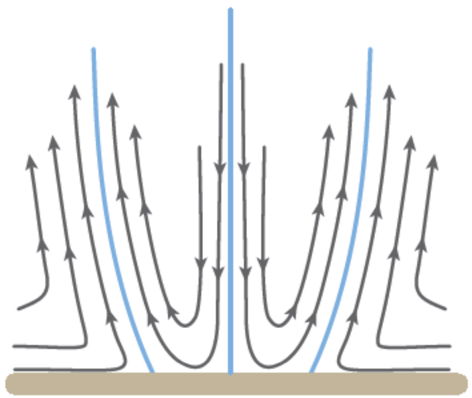
One Cell



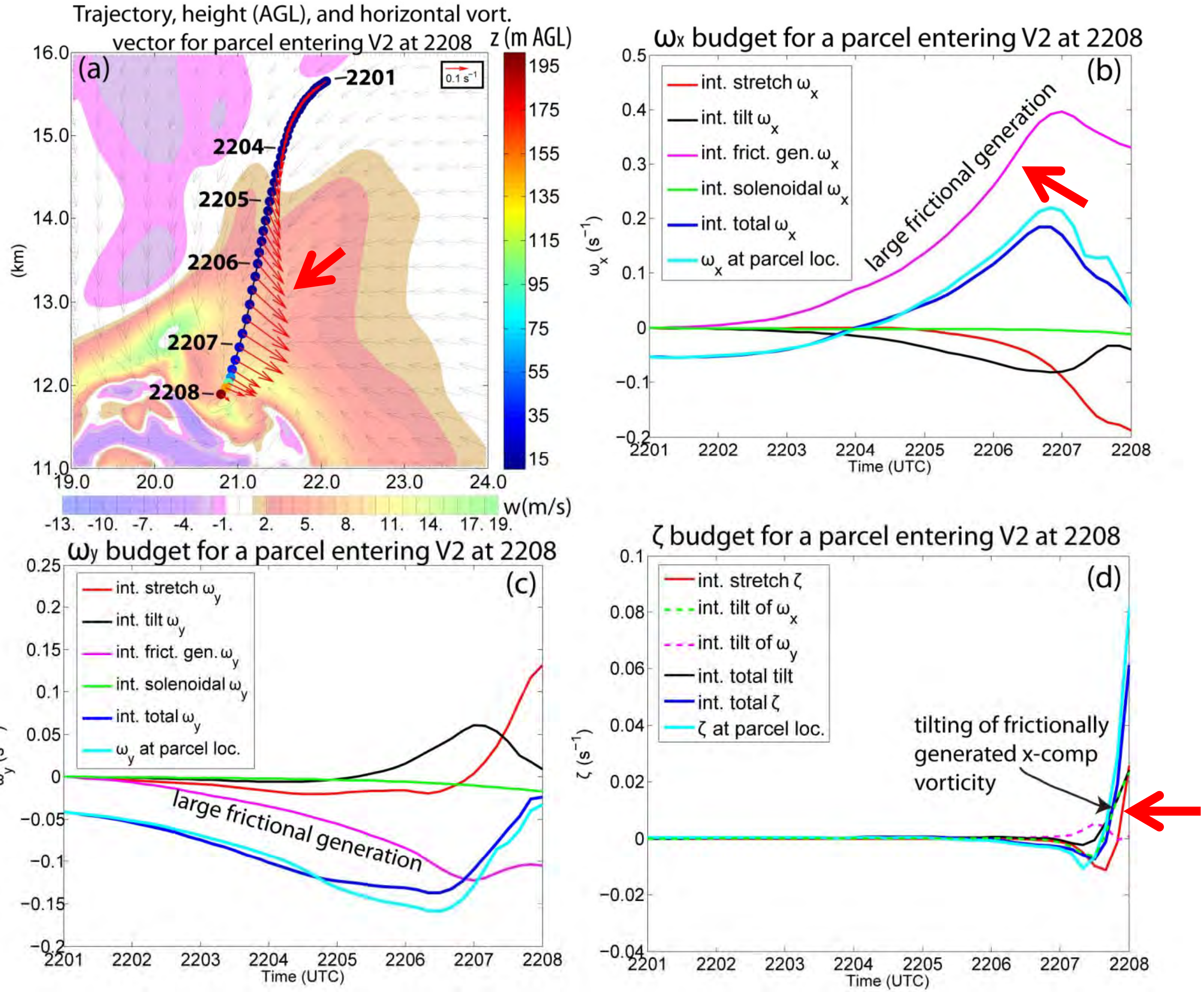


Two Cell

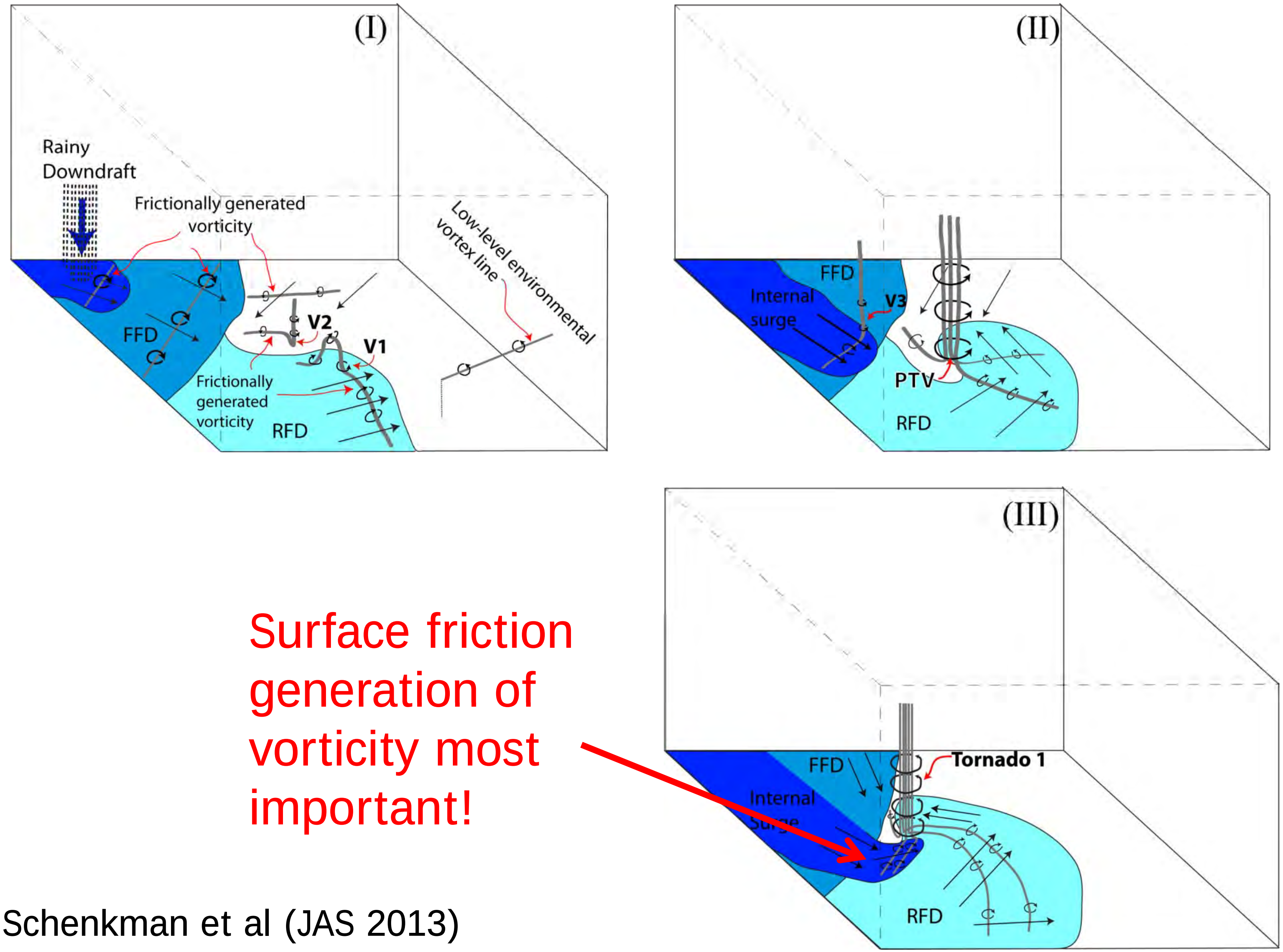




Trajectory analysis into V2



3-stage Conceptual model of Tornadogenesis



Surface friction generation of vorticity most important!

Schenkman et al (JAS 2013)

Storm-Scale Ensemble Forecasting (SSEF) for NOAA Hazardous Weather Testbed (HWT)

- CAPS/OU has been producing realtime convection-resolving storm-scale ensemble forecasts during springs for the NOAA Hazardous Weather Testbed (HWT) organized by NSSL and SPC since 2007.
- Spring 2011 configuration
  - 51-member 4-km and one 1-km forecasts.
  - 36 hours.
  - Full CONUS domain.
  - 4 mesoscale models (WRF ARW, NMM, ARPS, COAMPS)
  - Multiple physics combinations
- Assimilation of level-2 Vr and Z data from all WSR-88D radars, and sfc obs, in the CONUS-sized domain using ARPS 3DVAR/Cloud analysis to initialize WRF ARW, NMM and ARPS since 2008, at 4, 2 and 1 km resolutions.

The Computers Used

NICS, Kraken (~99K cores)

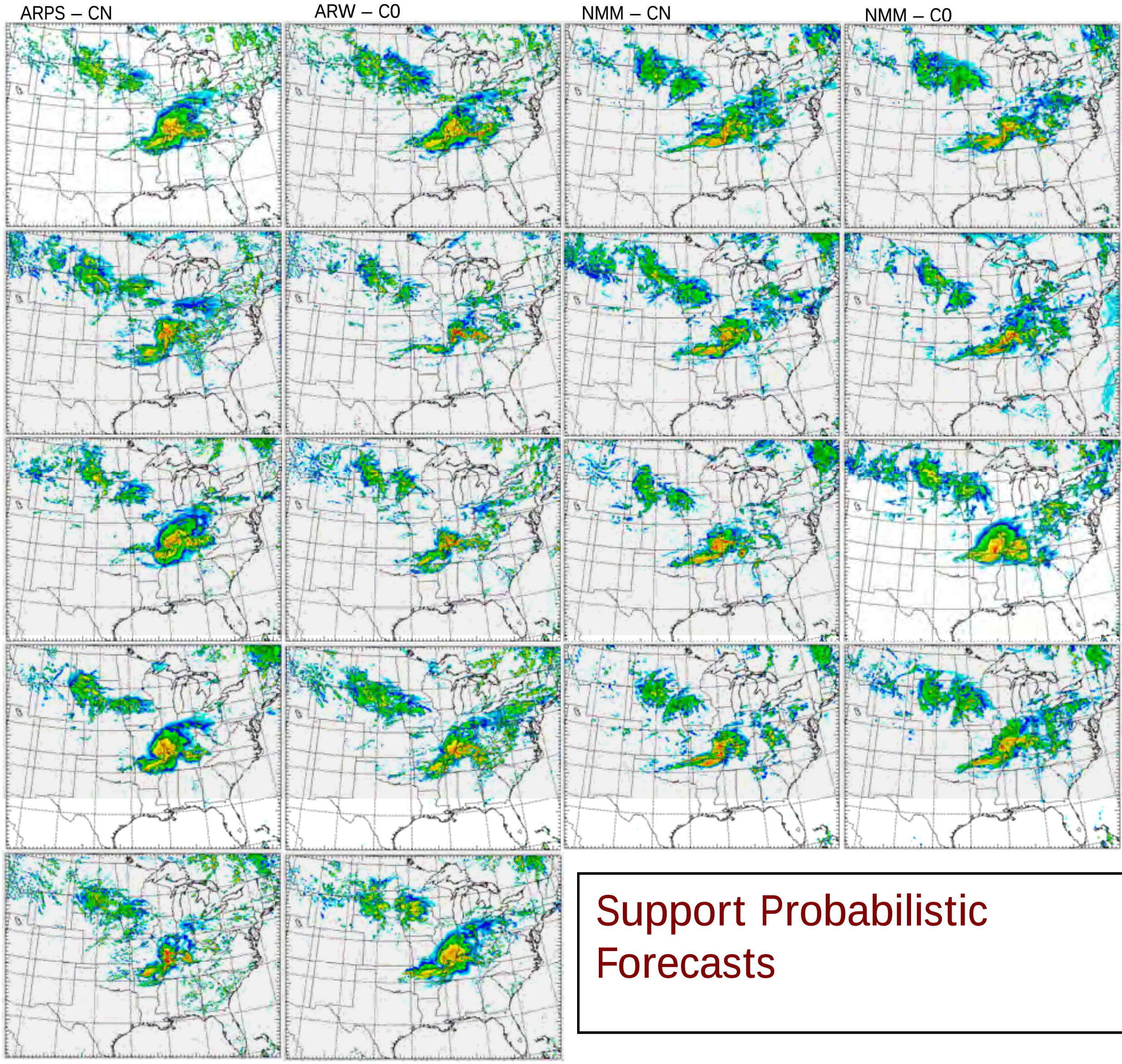
For 2010: Exclusive use of a 18,000-core Cray XT-4 at
NICS/University of Tennessee 6 hours a day

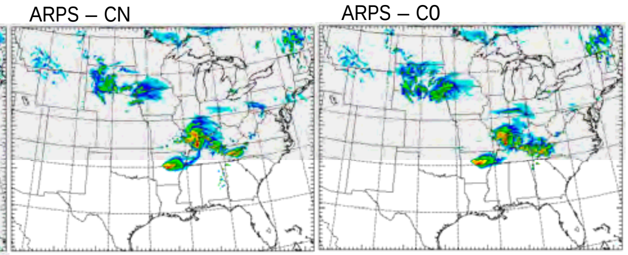
NOAA Hazardous Weather Testbed (HWT)

- HWT is a facility jointly managed by NSSL, SPC, and NWS Norman WFO
- To accelerate the transition of promising new technologies into forecasting and warning for hazardous weather.
- HWT organizes annual Spring Experiment that attracts about 100 researchers and forecasters each year.
- Provides forecasters with a first-hand look at the latest research concepts and potential future products, and immerses researchers in the challenges, needs, and constraints of forecasters.

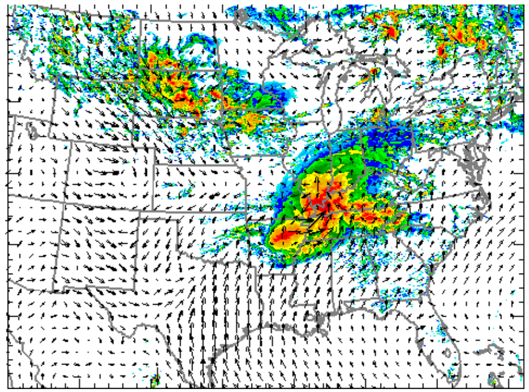
4 km ensemble forecasts produced by CAPS

Ensembles: Explore & Define Uncertainty

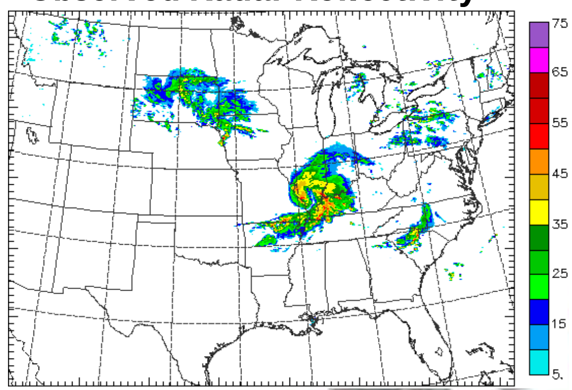




18 hr Forecast - 1 km "Radar"



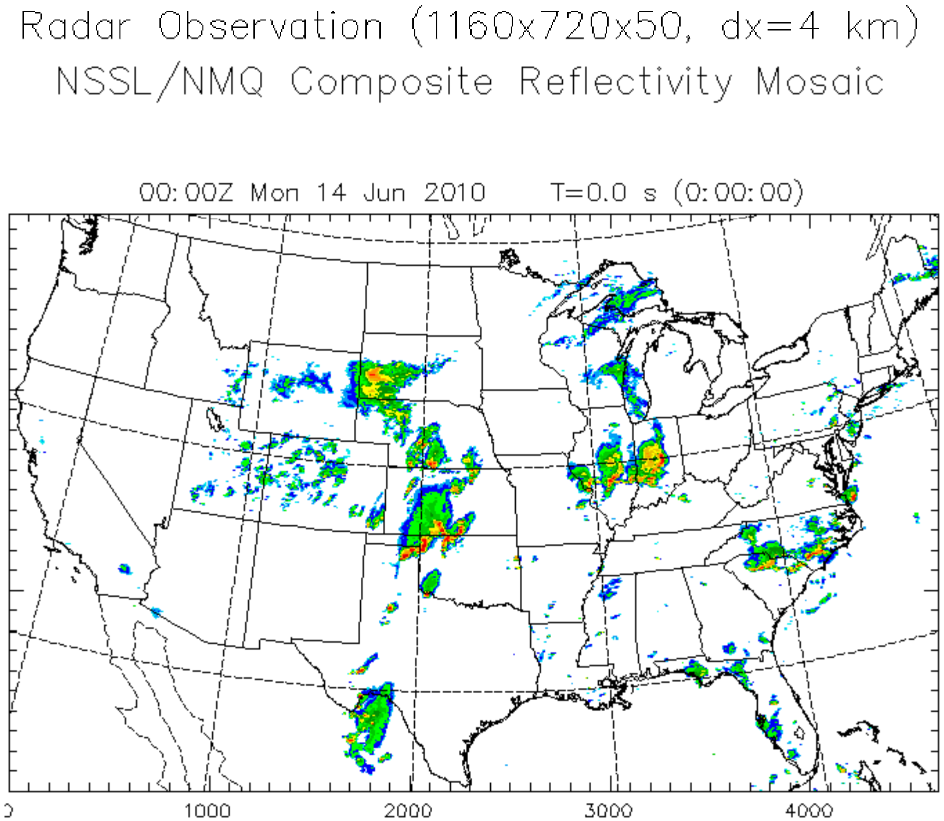
Observed Radar Reflectivity

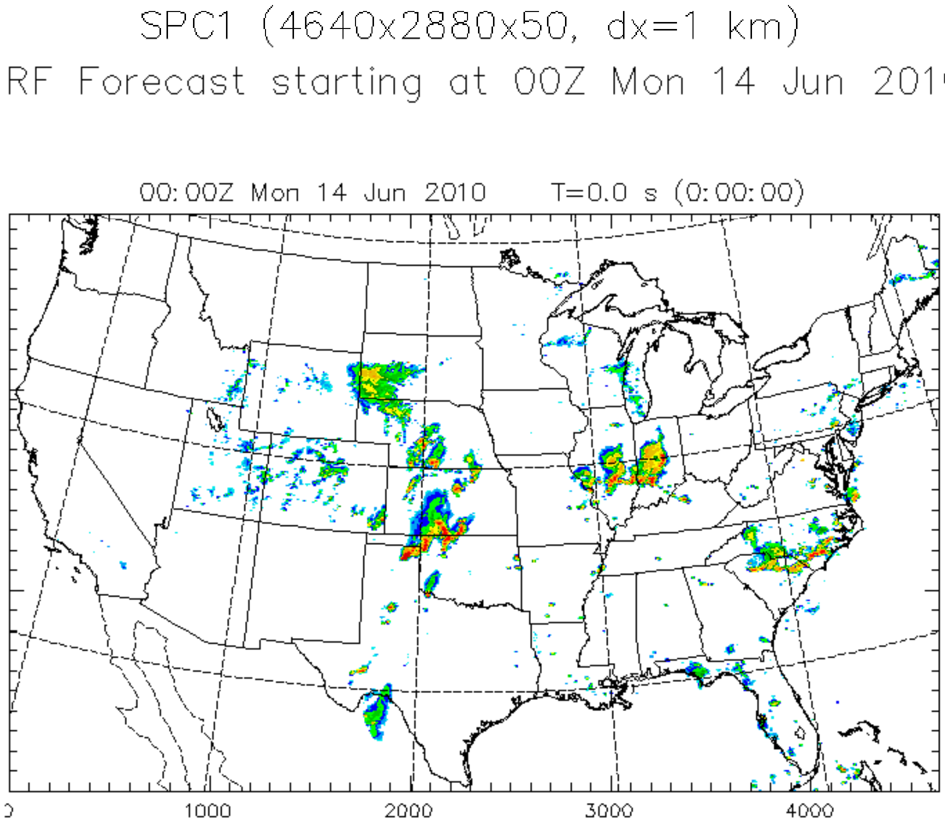


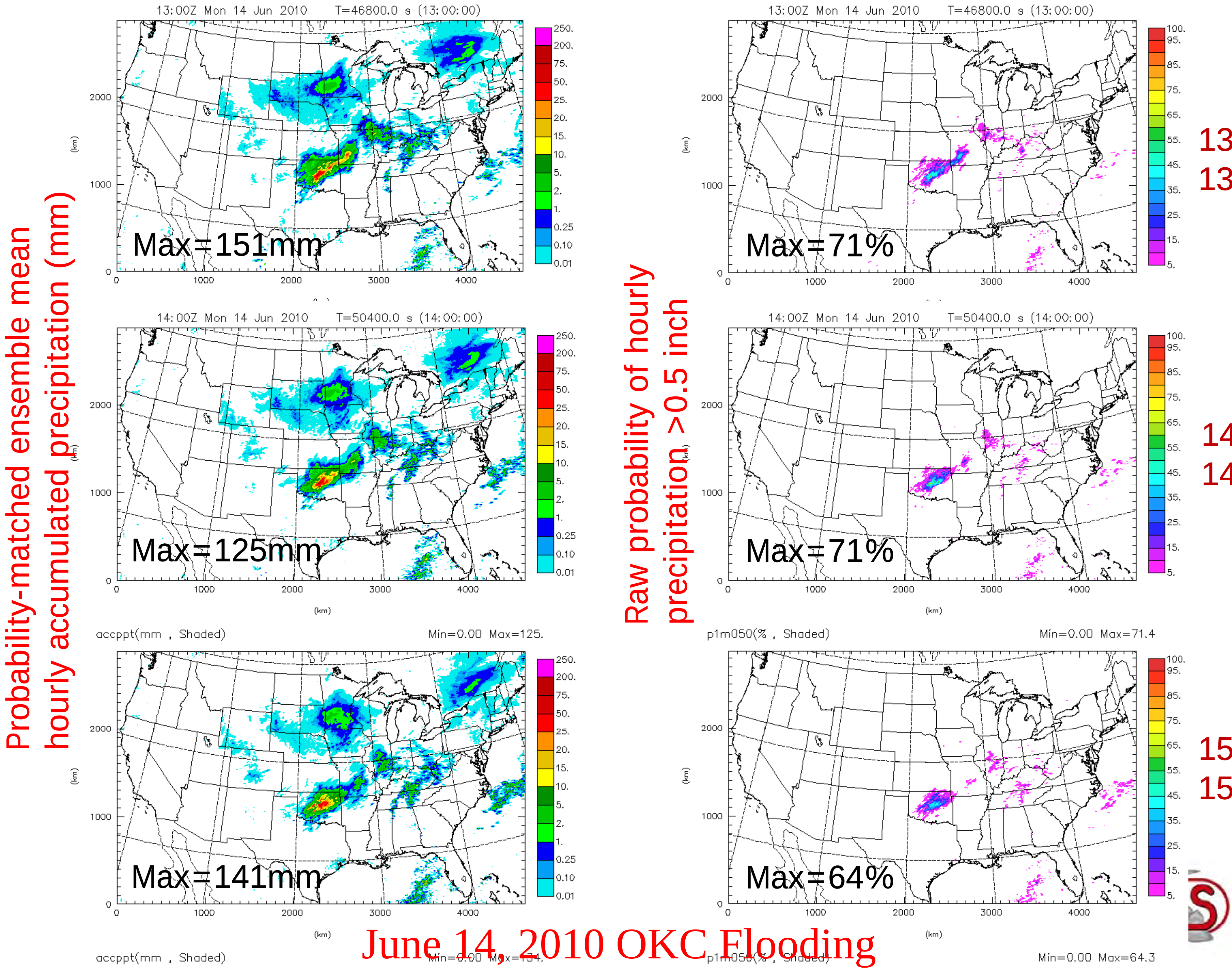
Support Probabilistic Forecasts

June 14, 2010 – Oklahoma City Flooding
~ 10 inch rain fell in part of OKC

Observed echo 1 km grid forecast



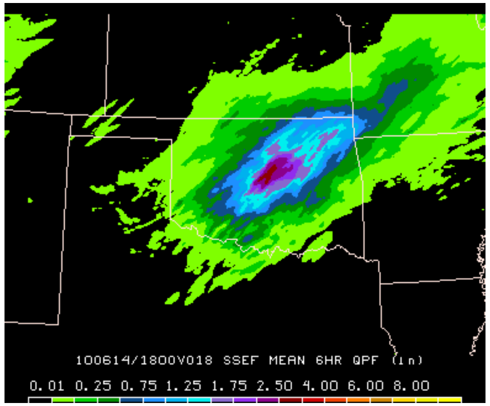




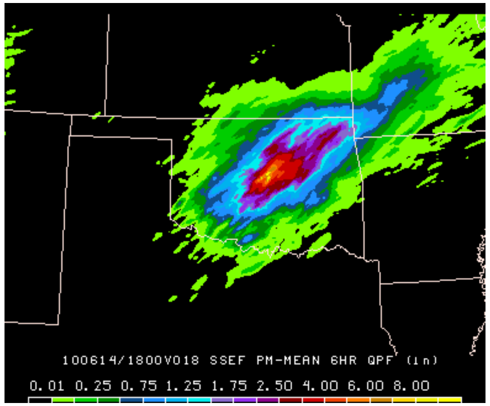
June 14, 2010 OKC Flooding

12–18Z accumulated precipitation: 18h (June 14, 2010 – OKC Flood Day)

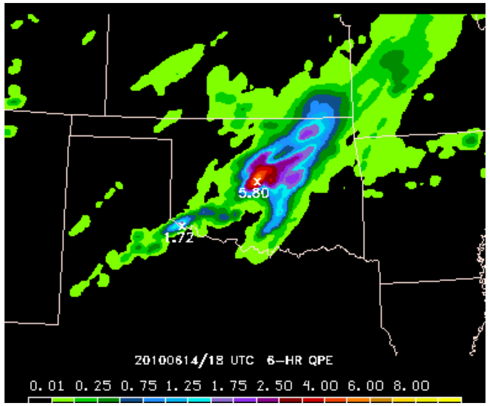
SSEF mean



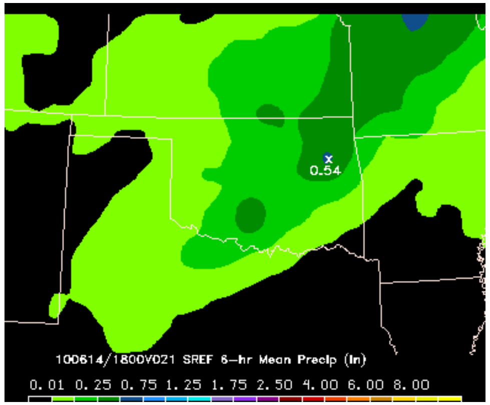
SSEF Prob match



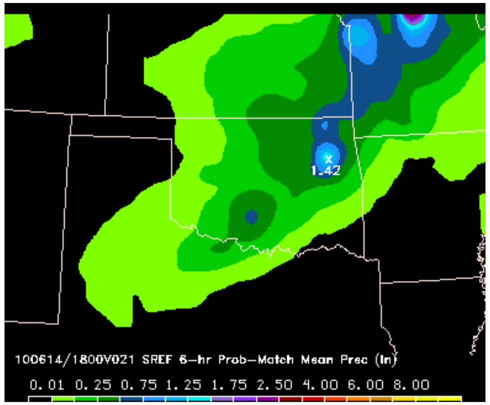
QPE



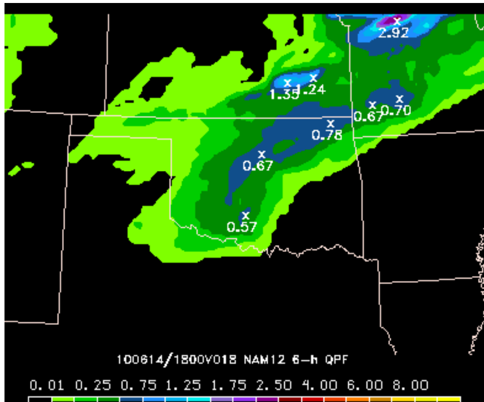
SREF mean



SREF Prob match



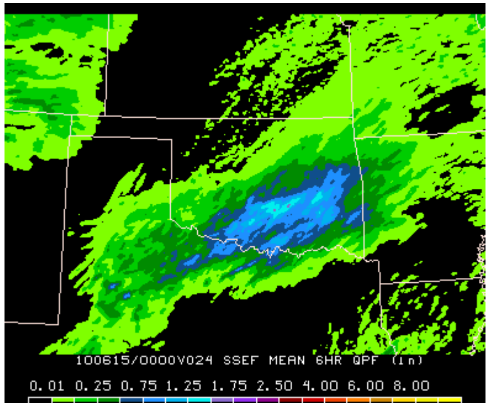
NAM



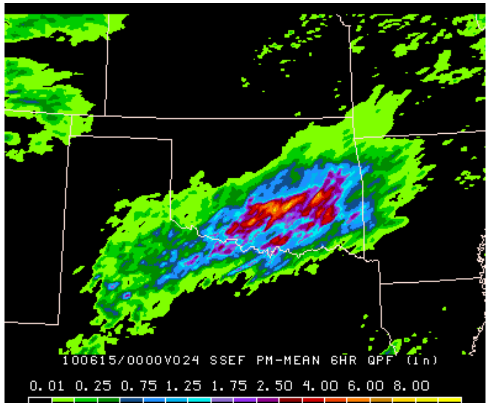
HWT images

18–0Z accumulated precipitation: 24h (June 14, 2010 – OKC Flood Day)

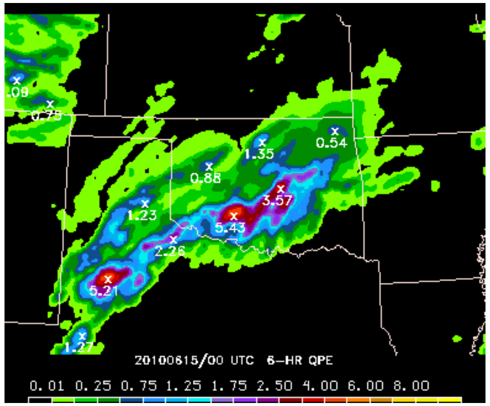
SSEF mean



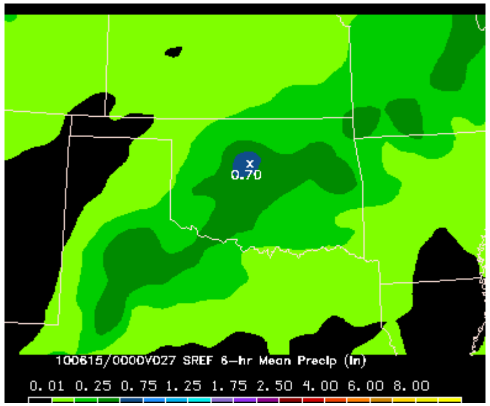
SSEF Prob match



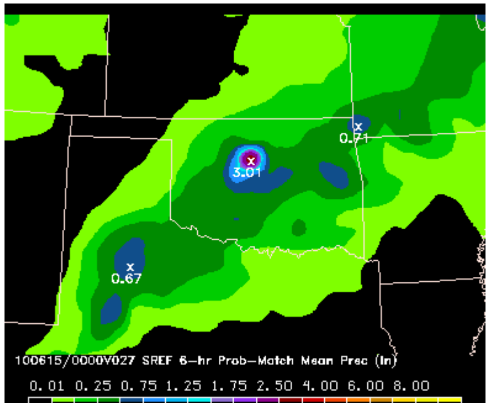
QPE



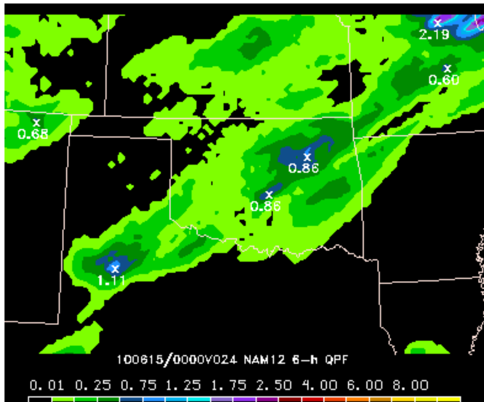
SREF mean



SREF Prob match

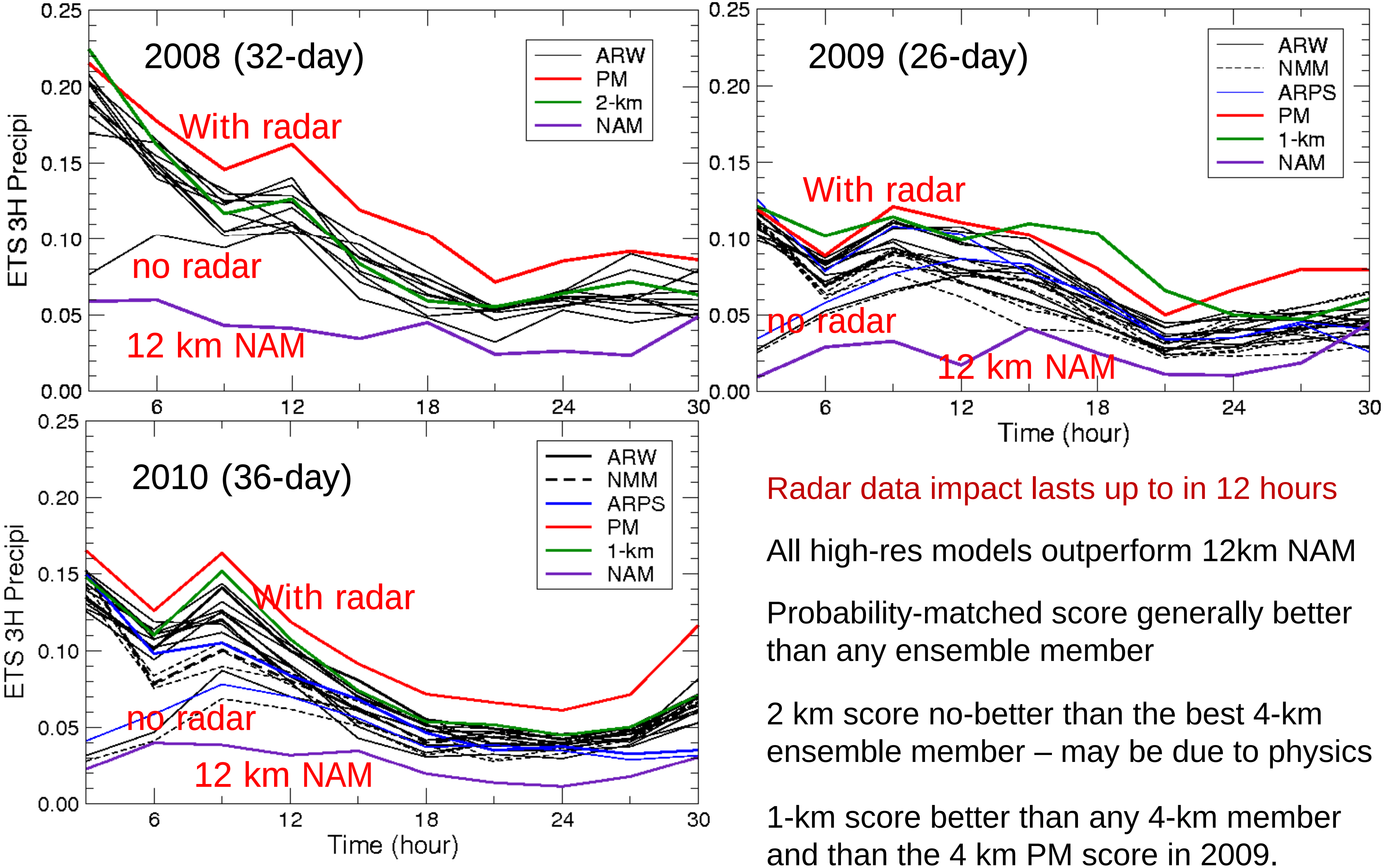


NAM



HWT images

Euitable Threat Scores for 3-hourly Precip. ≥ 0.5 in



Radar data impact lasts up to in 12 hours

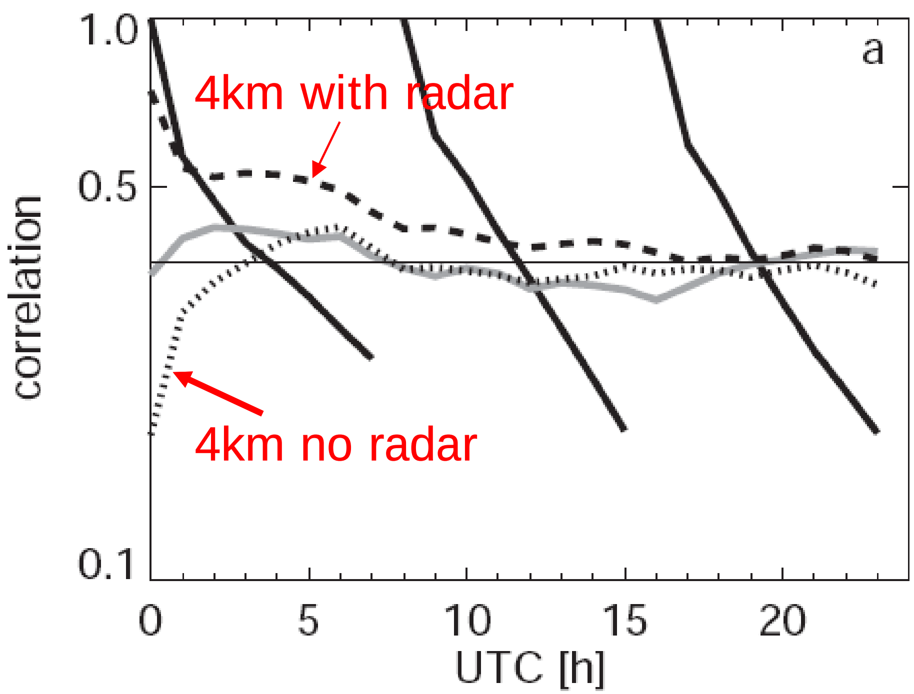
All high-res models outperform 12km NAM

Probability-matched score generally better than any ensemble member

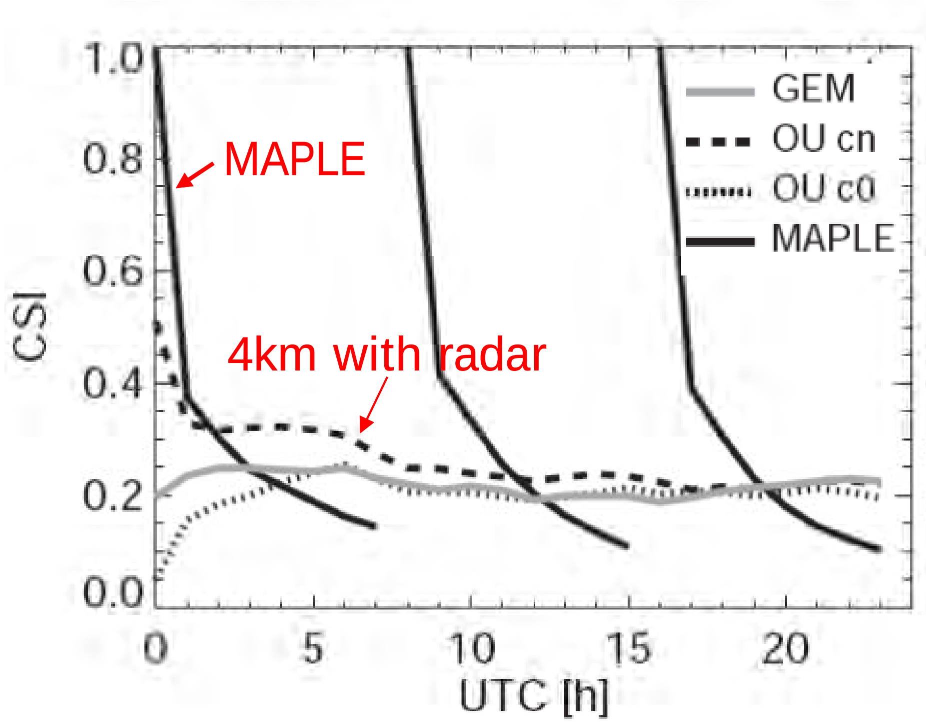
2 km score no-better than the best 4-km ensemble member – may be due to physics

1-km score better than any 4-km member and than the 4 km PM score in 2009.

Comparison of CAPS 4 km Cn/C0 2008 Forecasts with McGill 2-km MAPLE Nowcasting System and Canadian 15-km GEM Model



Correlation for reflectivity

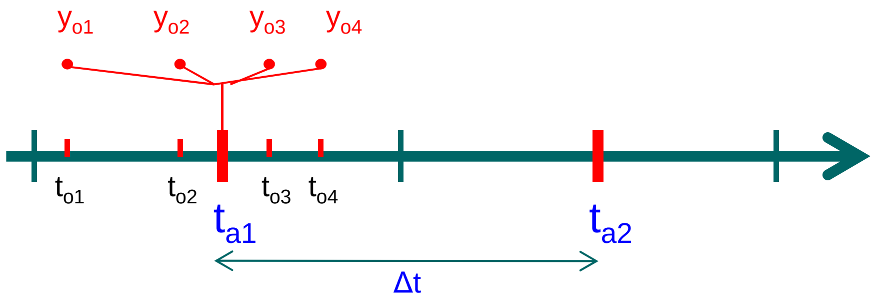


CSI for 0.2 mm/h

Courtesy of Madalina Surcel of McGill U. (Surcel et al. 2009 Radar Conf.)

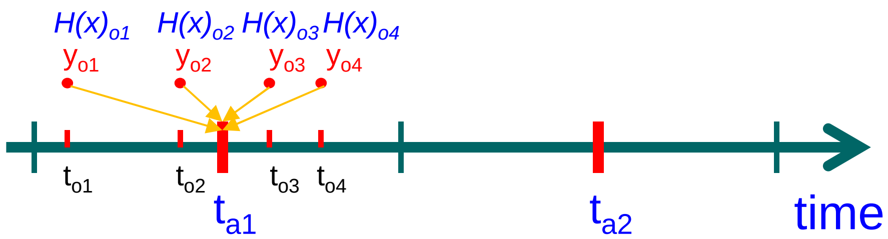
4DEnKF vs. Regular EnKF

EnKF: Observation are assumed to be taken at the assimilation time, like 3DVAR.



t_a: time of analysis, Δt: analysis time interval,
t_o: time of observation

4DEnKF: Observation can be assimilated at the times they are taken – no timing error, like 4DVAR

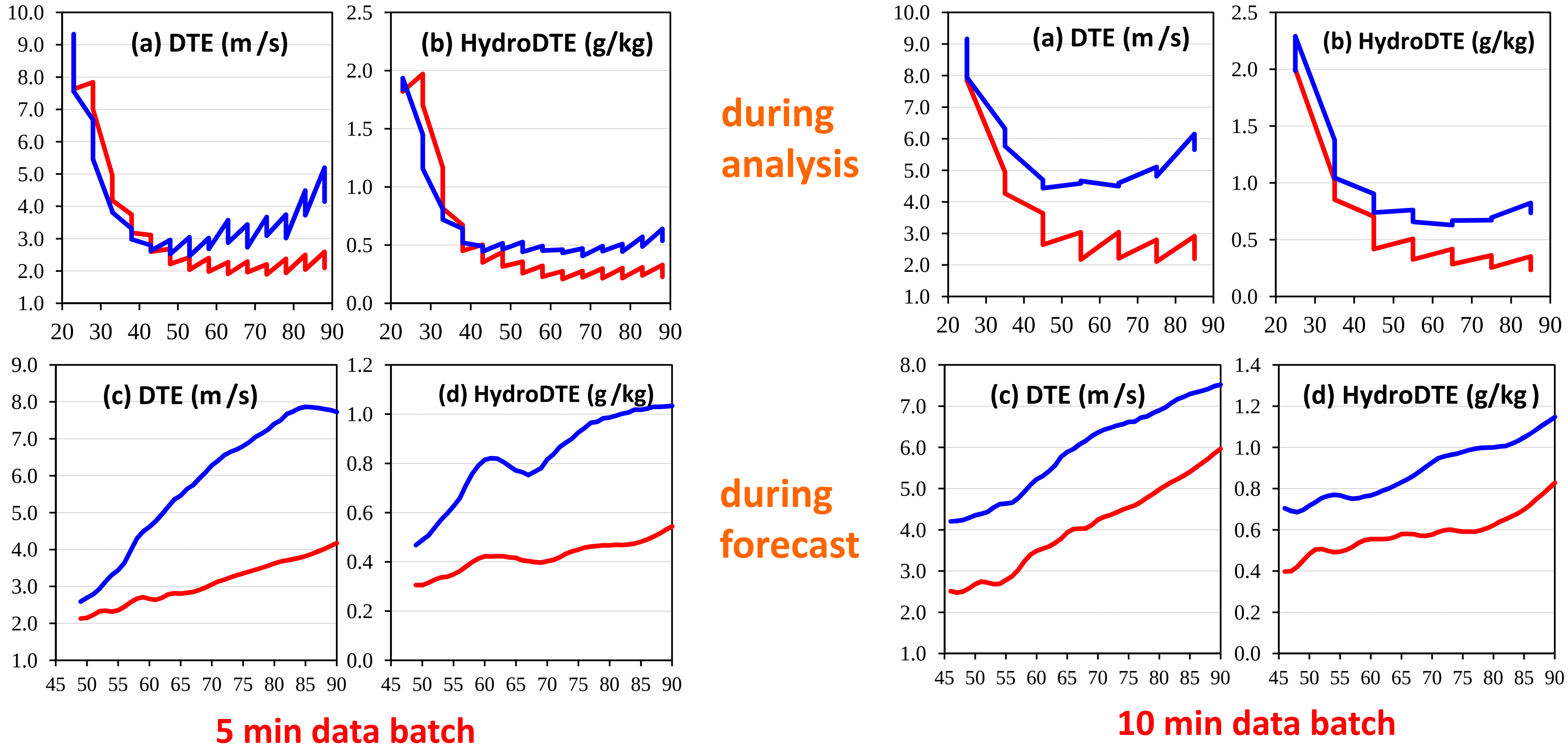


- Existing 4D ensemble algorithms include 4D-LETKF, En4DVAR.
- 4D EnSRF algorithm requires special treatment due to its serial nature.
- 4DEnSRF computes and update observation priors at the observations times.
- 4DEnSRF has been tested with simulated radar data for both WRF and ARPS, and with real data and ARPS.

H(x) can be calculated within the numerical model – added to ARPS.

EnSRF and 4DEnSRF Analysis and Forecast Errors: OSSEs for a Supercell with WRF

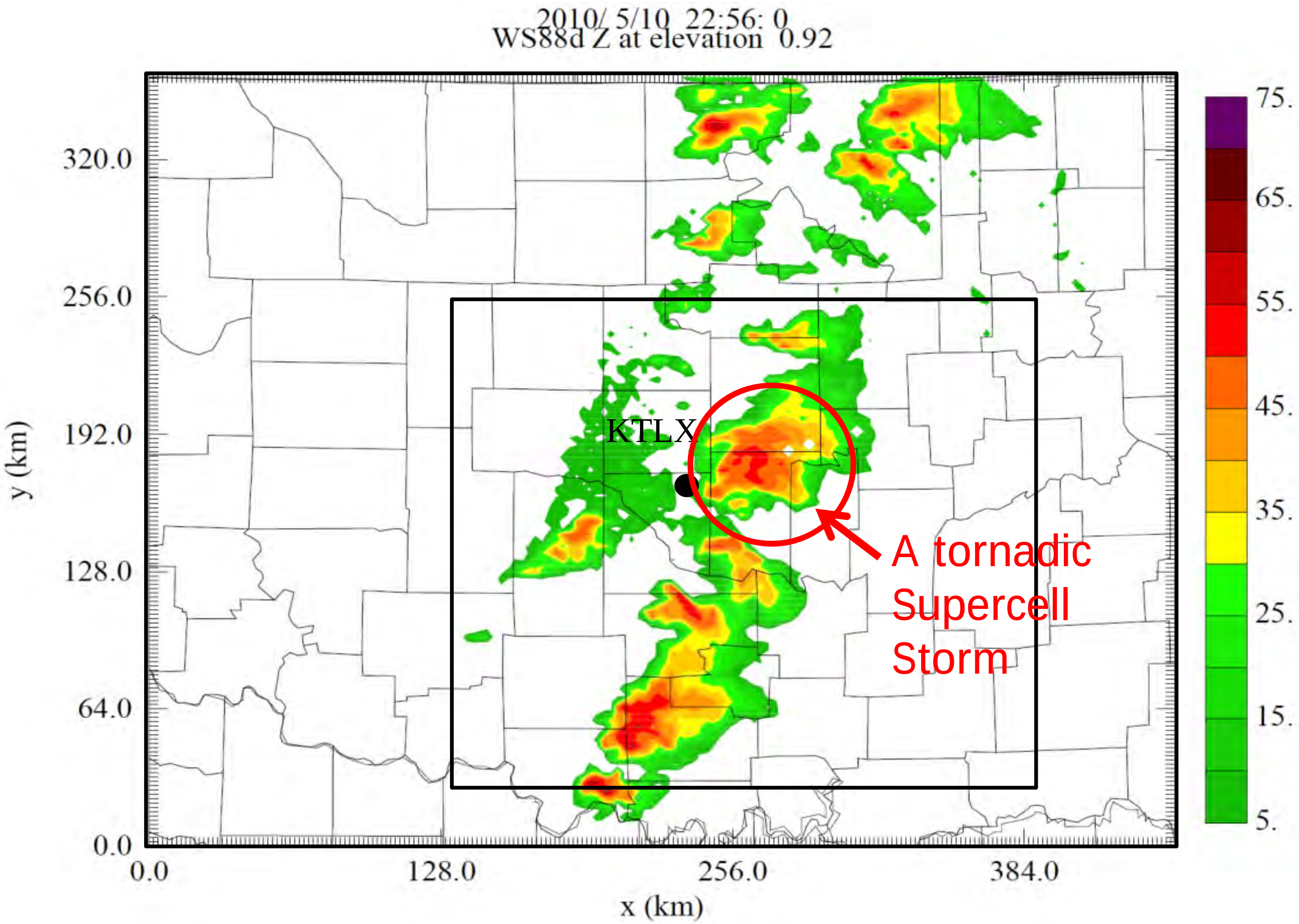
Blue: EnSRF Red: 4DEnSRF



- The RMS errors of 4DEnSRF are lower than those of EnSRF.
- The advantage of 4DEnSRF is more when analysis interval is longer (10 min or longer).
- Forecast error with 4DEnSRF grow faster for 10 min data batches than the 5 min case, while the growth rates for EnSRF are fast in both cases.

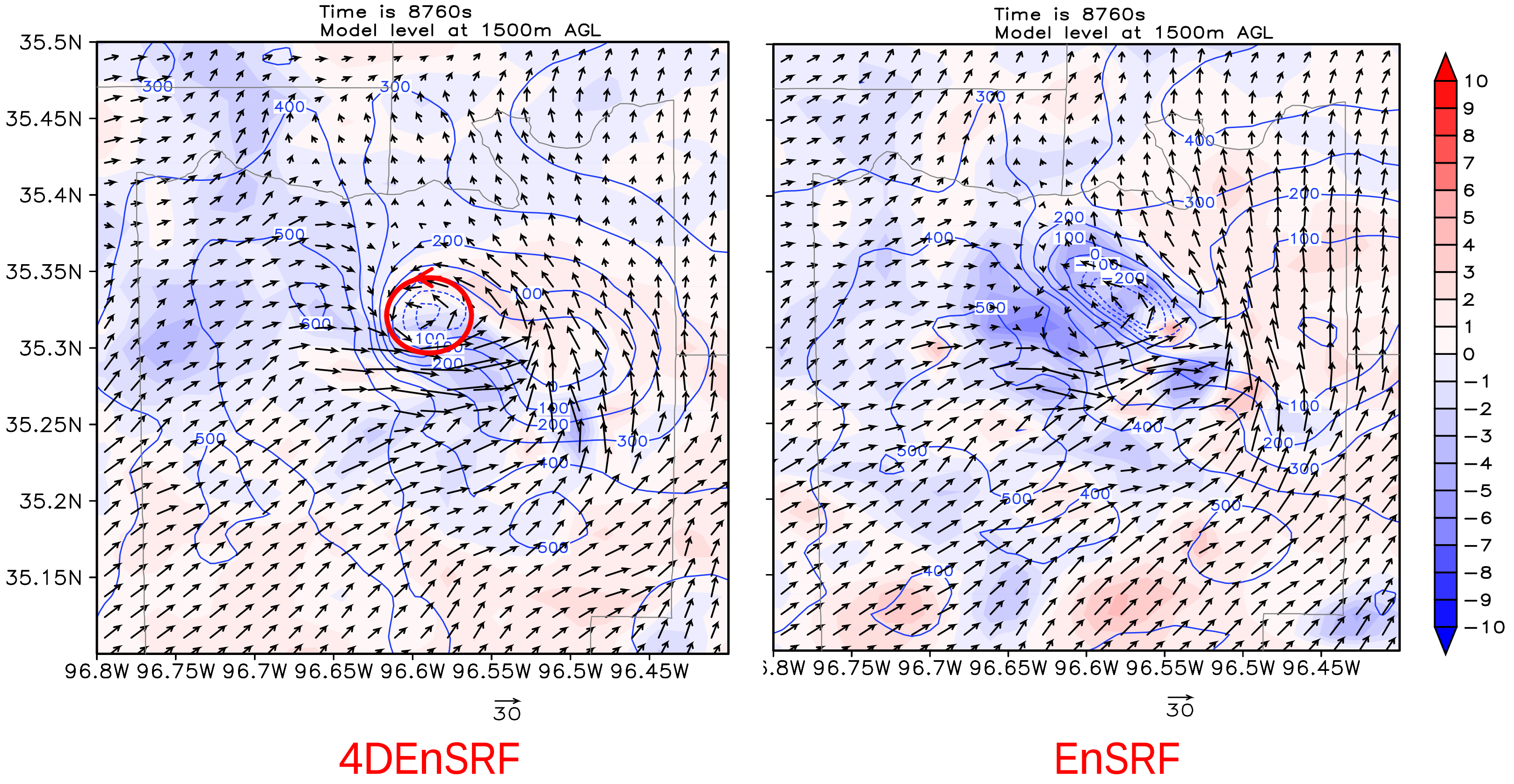
(Wang et al. 2012 QJRMS)

May 10, 2010 supercell case using EnSRF and 4DEnSRF



1.5 km grid
nested in
a 3 km grid

Analyzed T Perturbations (shaded), Pressure (contours) and Wind Vectors at 1.5 km AGL

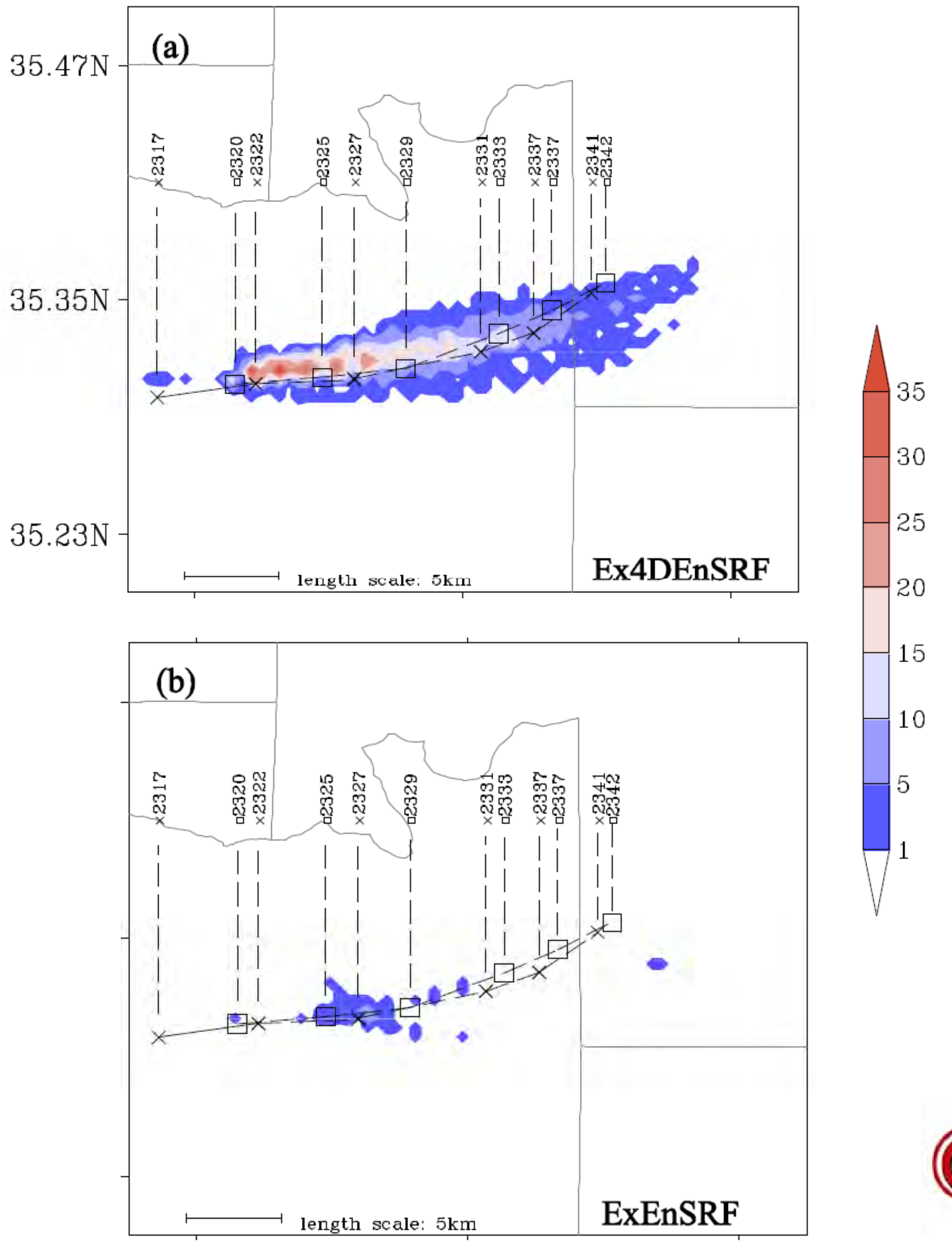


- Much better organization of low-level mesocyclone with 4DEnSRF
- Better dynamic consistency between wind, pressure and temperature fields

Probability of sfc vorticity forecast exceeding a threshold (dx=500m)

4DEnSRF

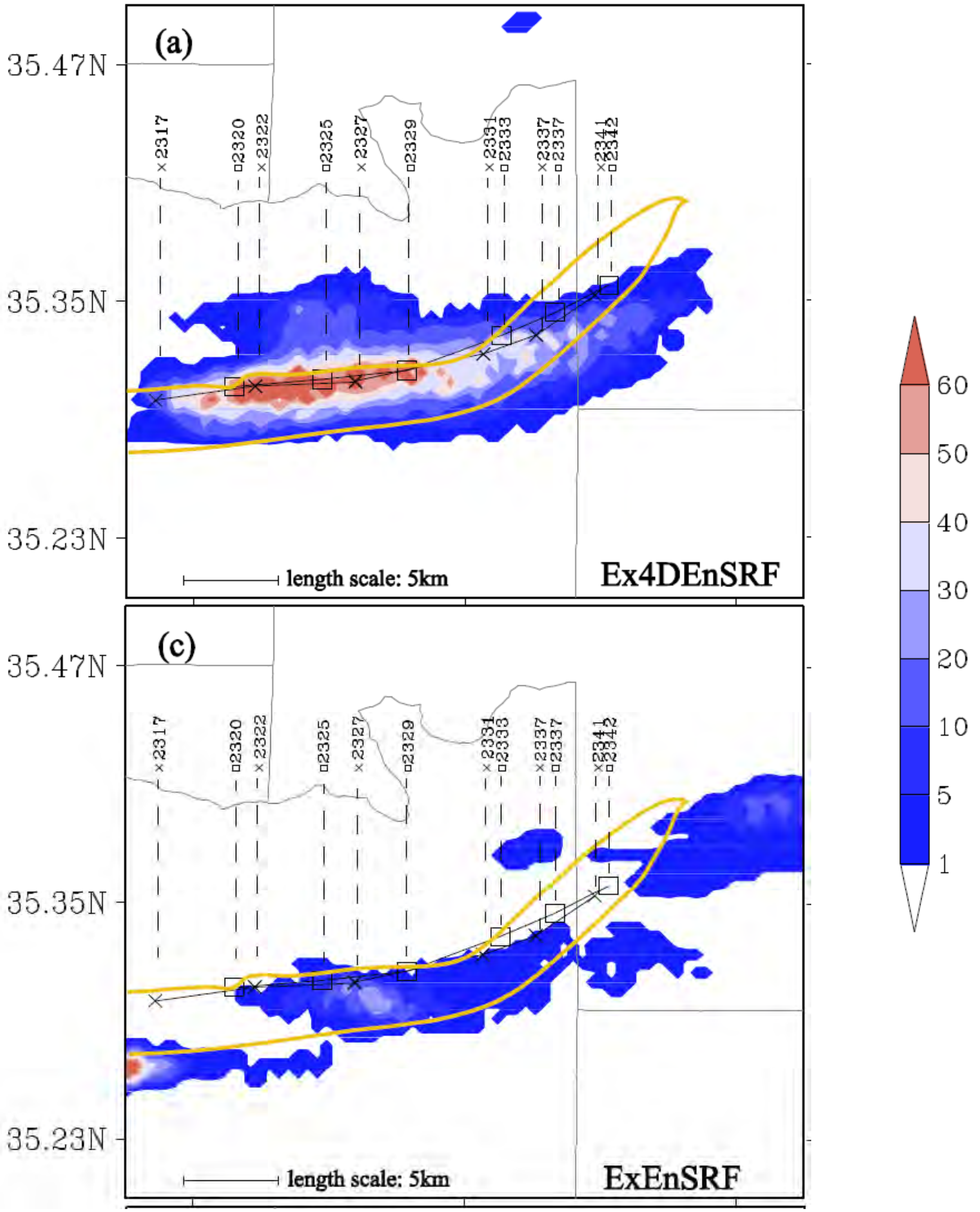
EnSRF



Probabilities of sfc wind speed forecast exceeding 29 ms^{-1} (EF0)

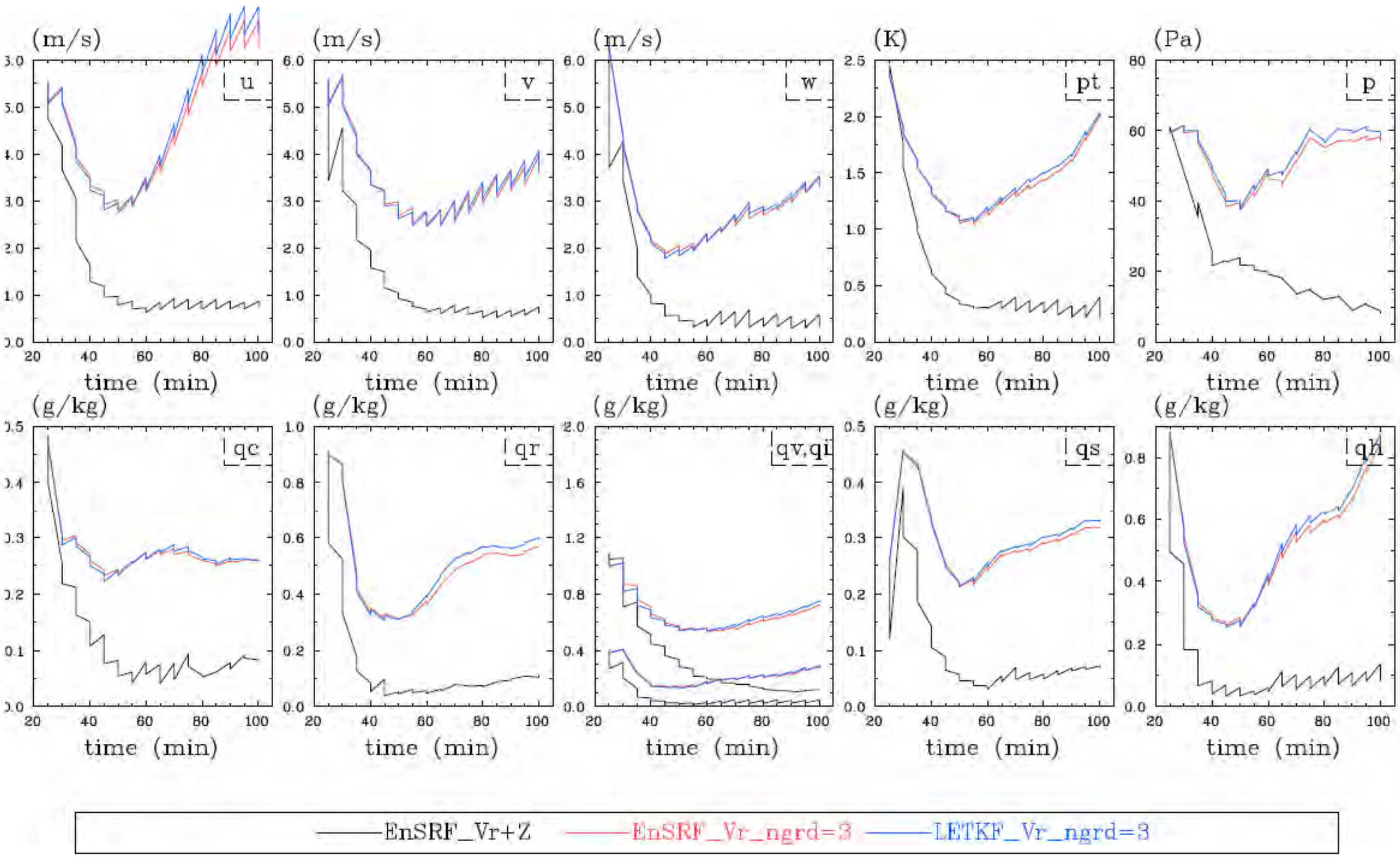
4DEnSRF

EnSRF



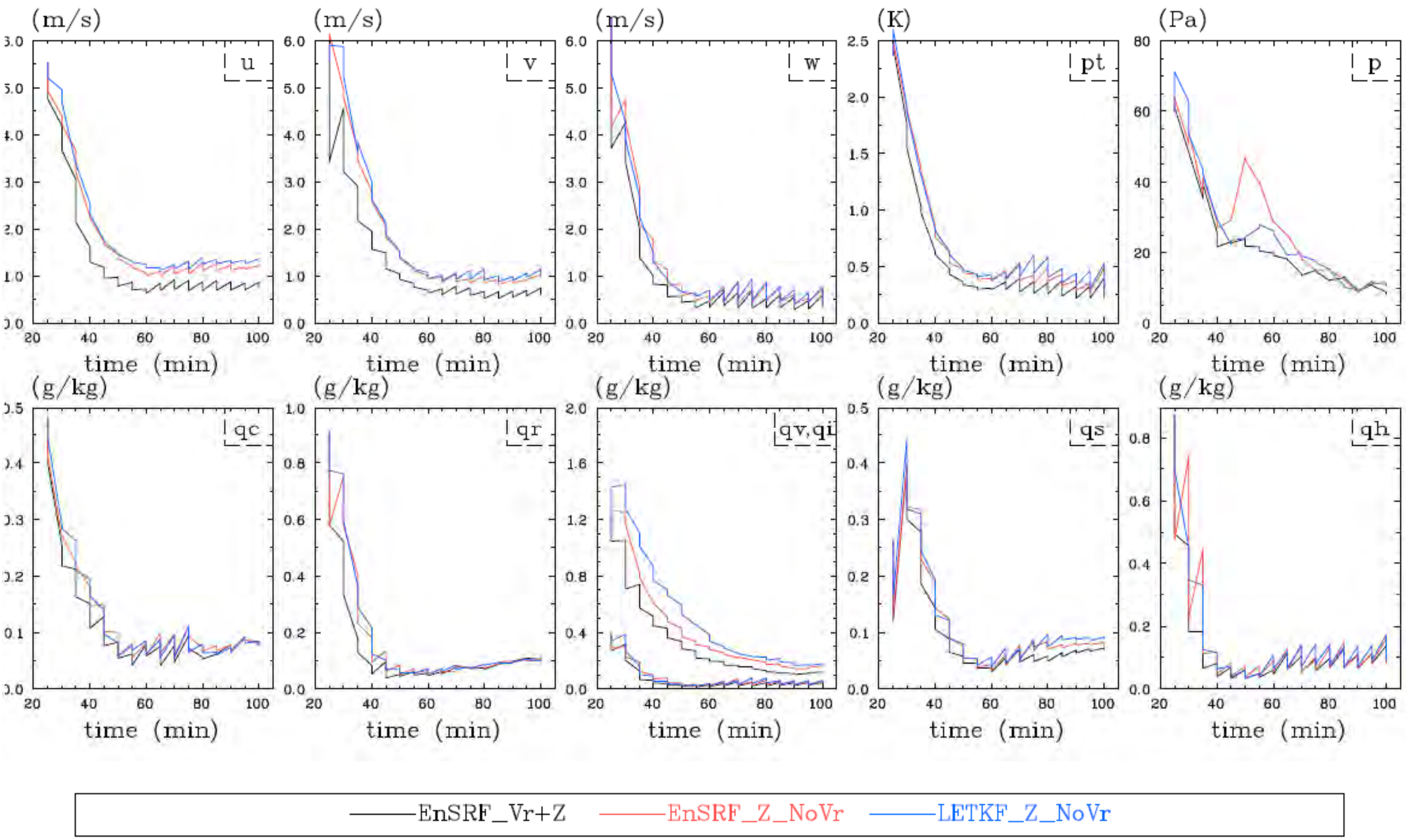
Comparison of EnSRF and LETKF Assimilation of Radar Data using OSSEs

EnSRF v.s. LETKF assimilation of Vr



- Little difference between EnSRF and LETKF when assimilating Vr data

EnSRF v.s. LETKF Assimilation of Z

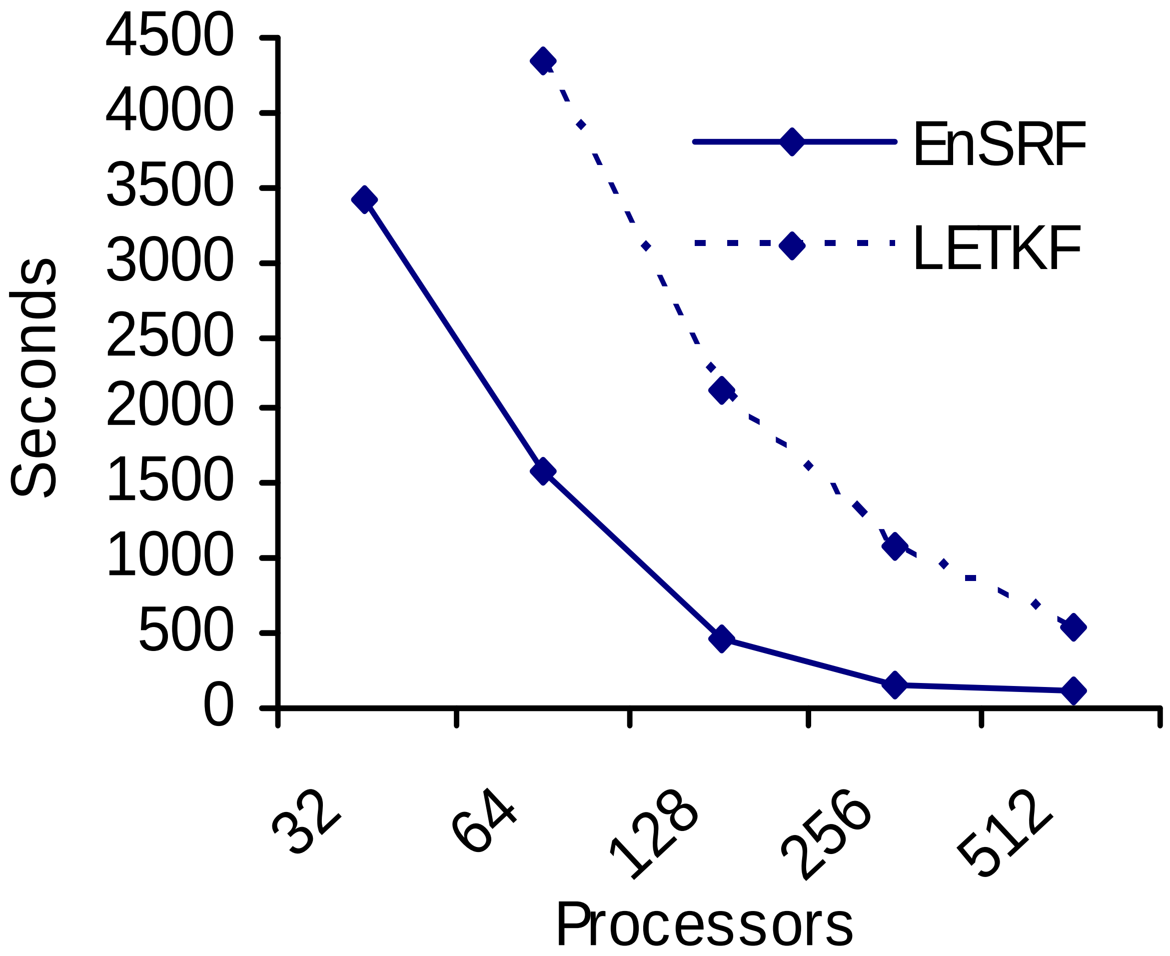


- More difference between EnSRF and LETKF when assimilating Z data
- LETKF slightly worse during initial cycles
- Poor initial estimated state and nonlinear obs operator of Z are believed to be the main cause of differences.

Computational Challenges:
Need to parallelize everything, efficiently

Timing statistics of parallel EnSRF and LETKF algorithms for a global model test problem of moderate resolution

(courtesy of Jeffrey Whitaker).



Here, the EnSRF parallelization is at the state vector level – observations are still assimilated on after another.

CAPS's Multi-scale Parallel EnKF system

- Hybrid OpenMP/MPI parallelization
- Multiple data types, especially high-density radar data
- 4D asynchronous filter (4DEnSRF), and iterative procedure (iEnSRF) for faster convergence/spin up
- Interface with both ARPS and WRF
- Support for 2-moment microphysics, perturbed physics and multi-physics ensemble, build-in attenuation correction
- Microphysical parameter estimation
- Polarimetric radar data assimilation
- Adaptive covariance inflation, successive spatial localization
- EnKF-3DVAR hybrid (testing) and LETKF algorithms
- En4DVAR (being implemented)
- Observation operators build into ARPS (for 4DEnKF).
- Next step: Running model and EnKF completely in memory?

Parallel EnSRF algorithm suitable for dense radar data via domain decomposition

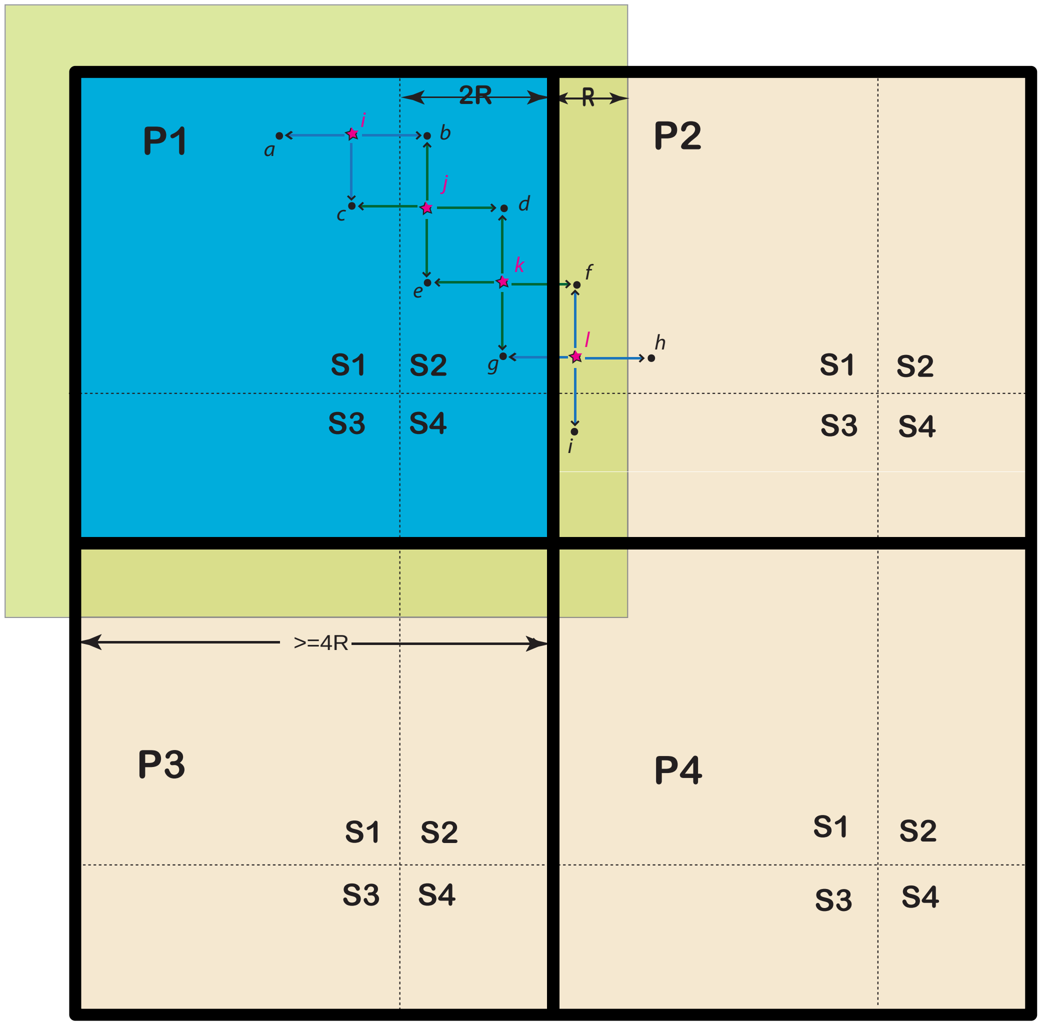
R is maximum localization radius

Grid is decomposed into sub-domains, each domain is assigned to a shared-memory node

Each sub-domain is further divided into S1-S4 sub-patches

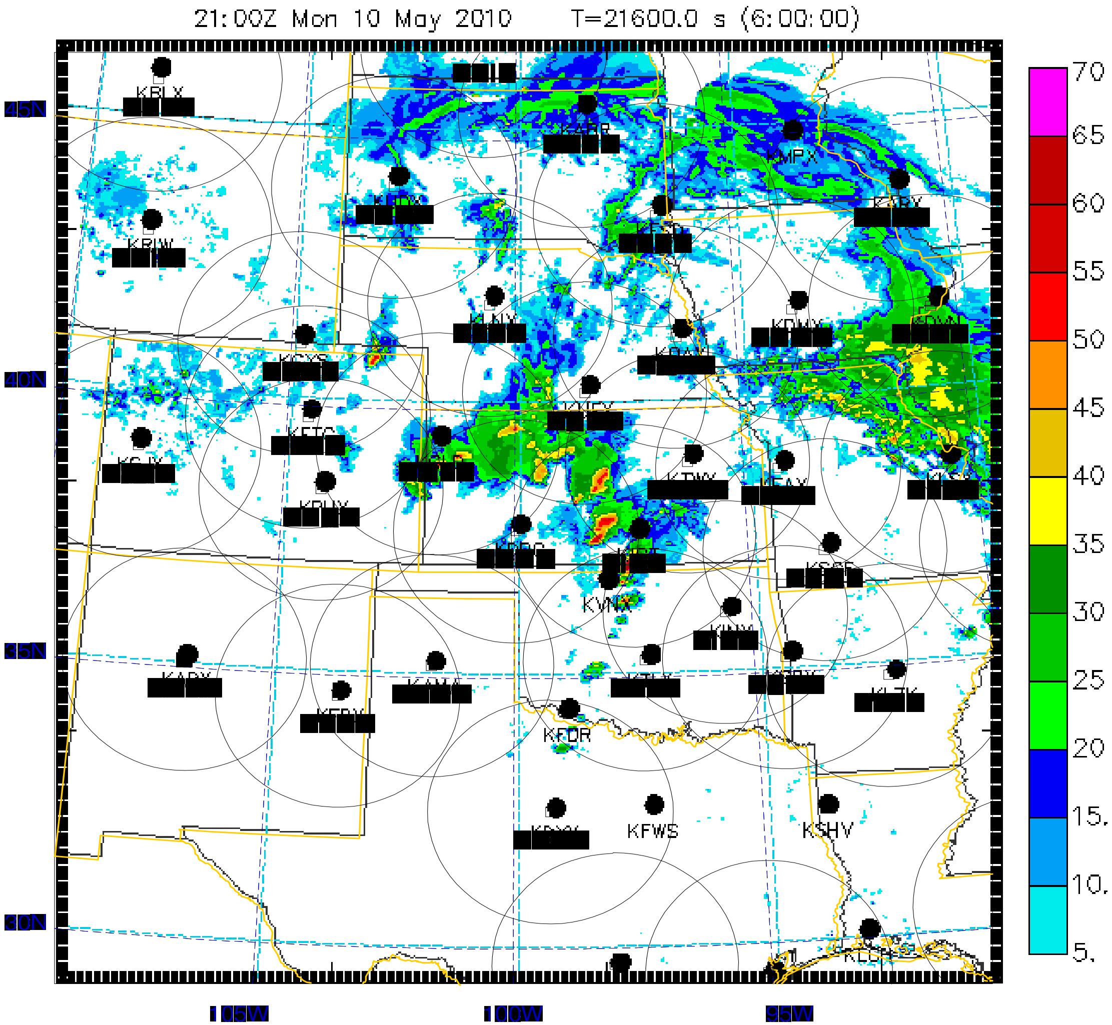
Sub-patches with the same label must be at least 2R apart so that obs within them don't affect common grid points and can be processed simultaneous

Loop through S1-S4 to complete obs assimilation



OpenMP/MPI hybrid Parallel EnKF Algorithm (Wang et al. 2013 JTech)

10 May 2010 OK-KS tornado outbreak



443x483x53 grid

1760 x 1920 km

~40 radars:

Vr and Z

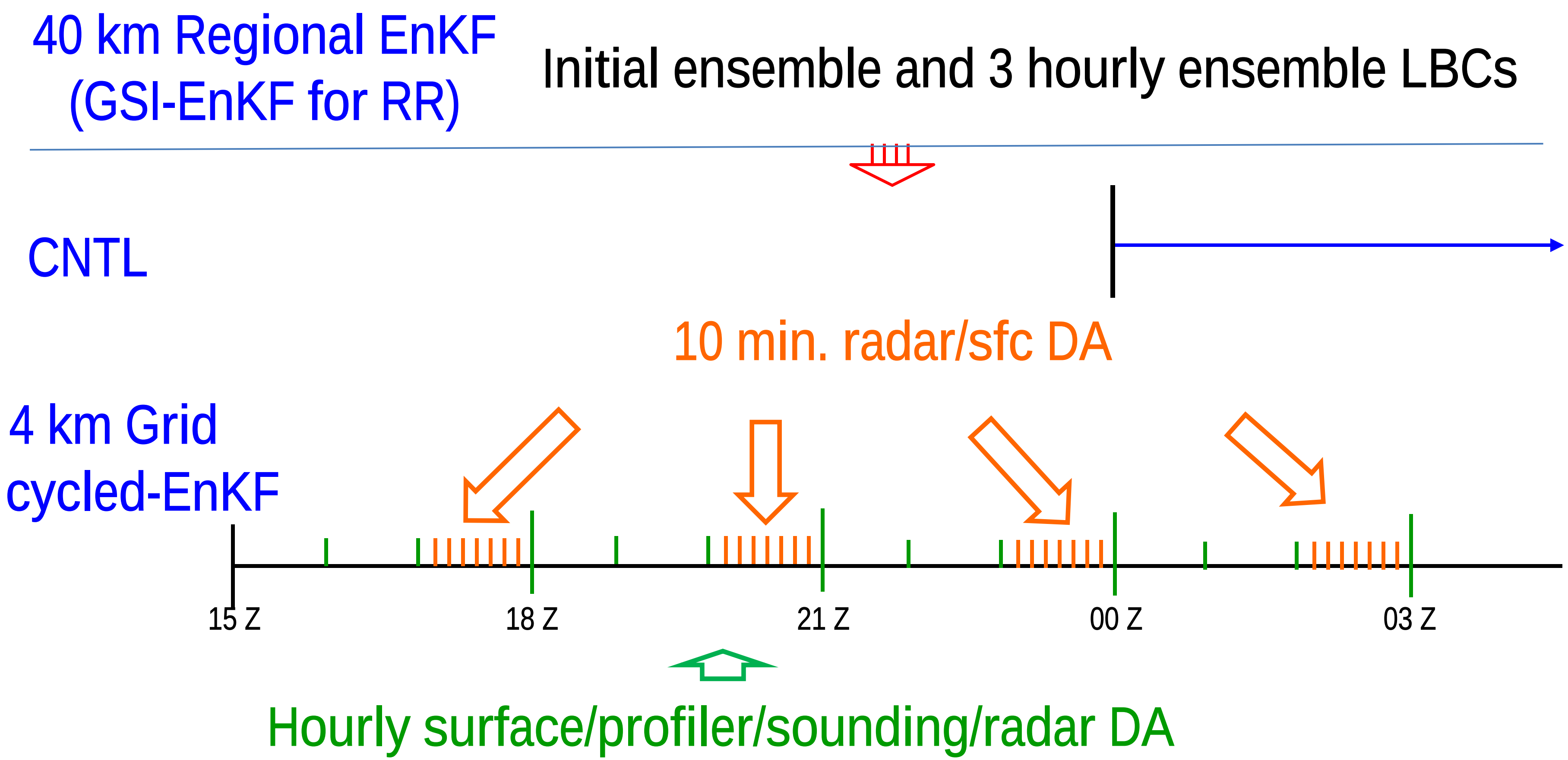
Surface obs,

Soundings

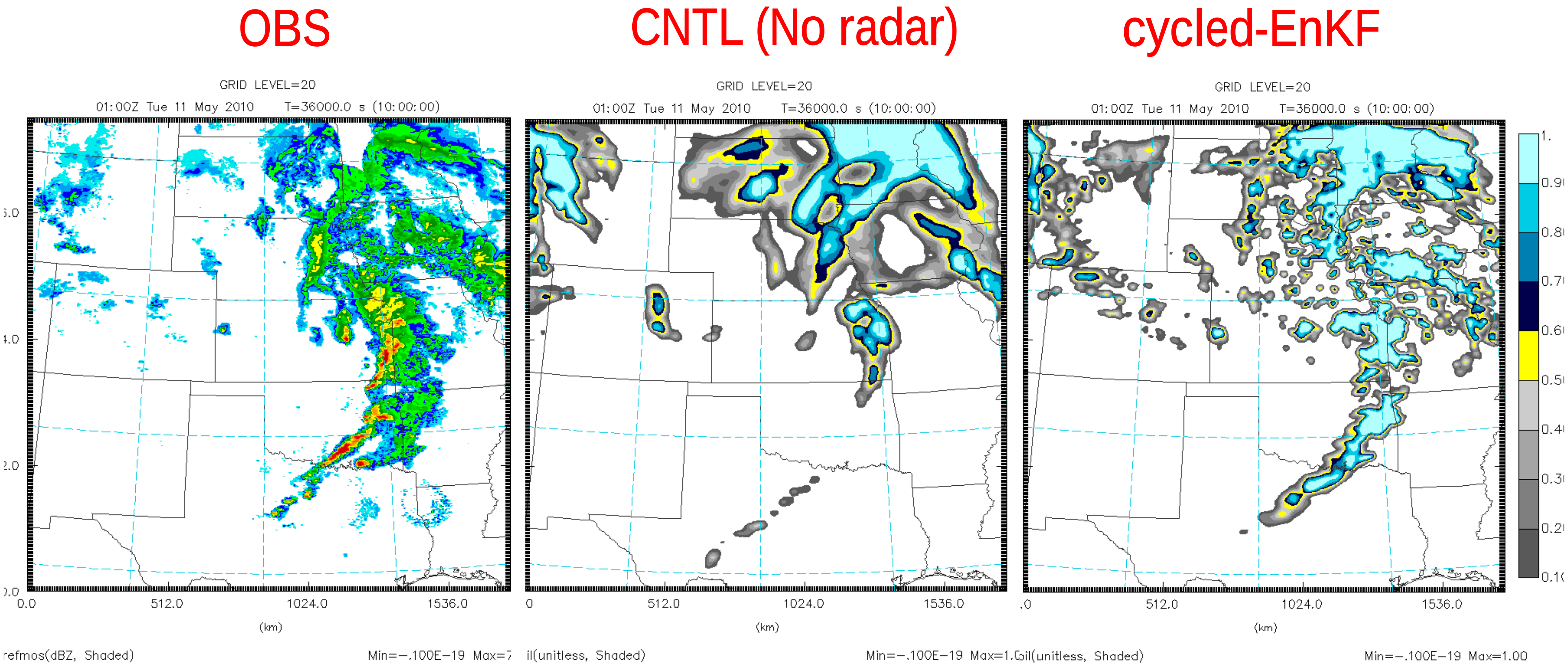
Profilers

Parallel multi-scale EnKF Algorithm (Wang et al. 2012)

Experiment setup for cycled-EnKF

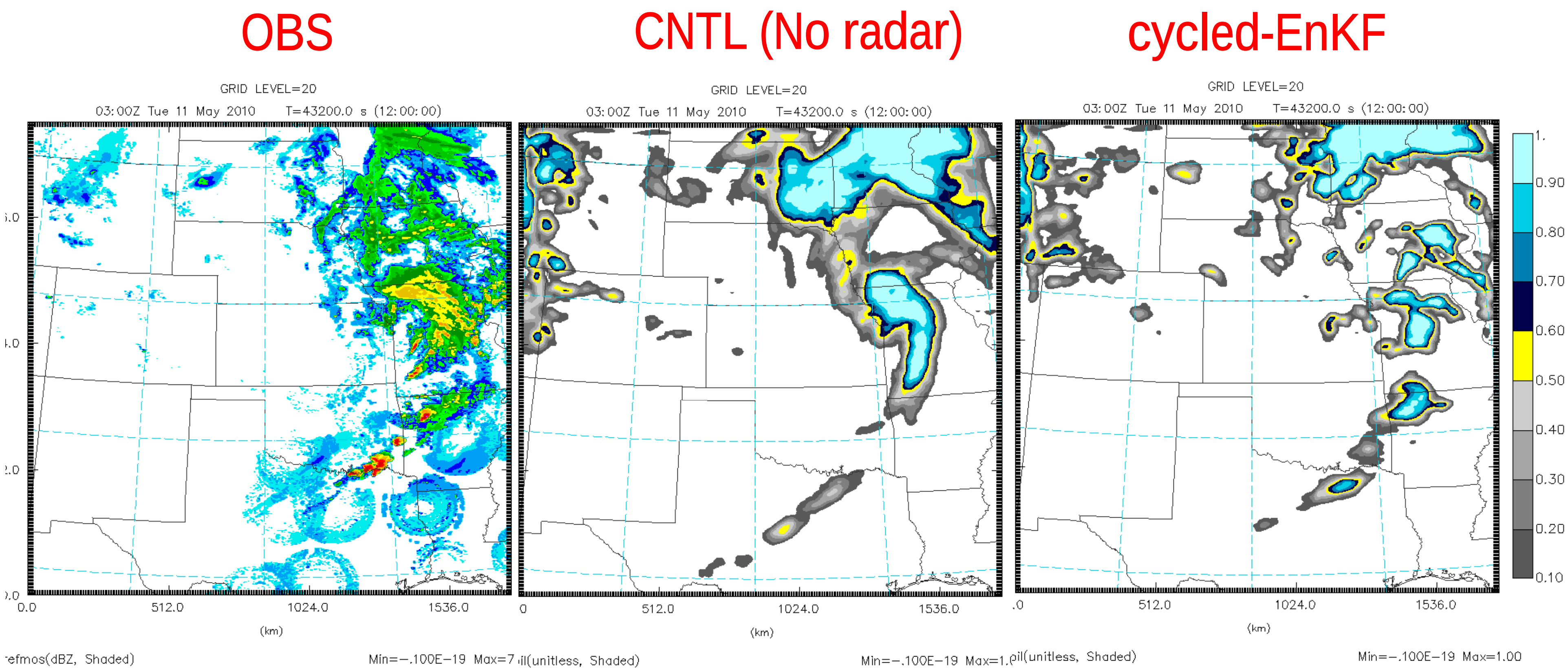


1-hour Probabilistic Forecast



Probability of Z_H > 15 dBZ (shade)

3-hour Probabilistic Forecast



Probability of Z_H > 15 dBZ (shade)

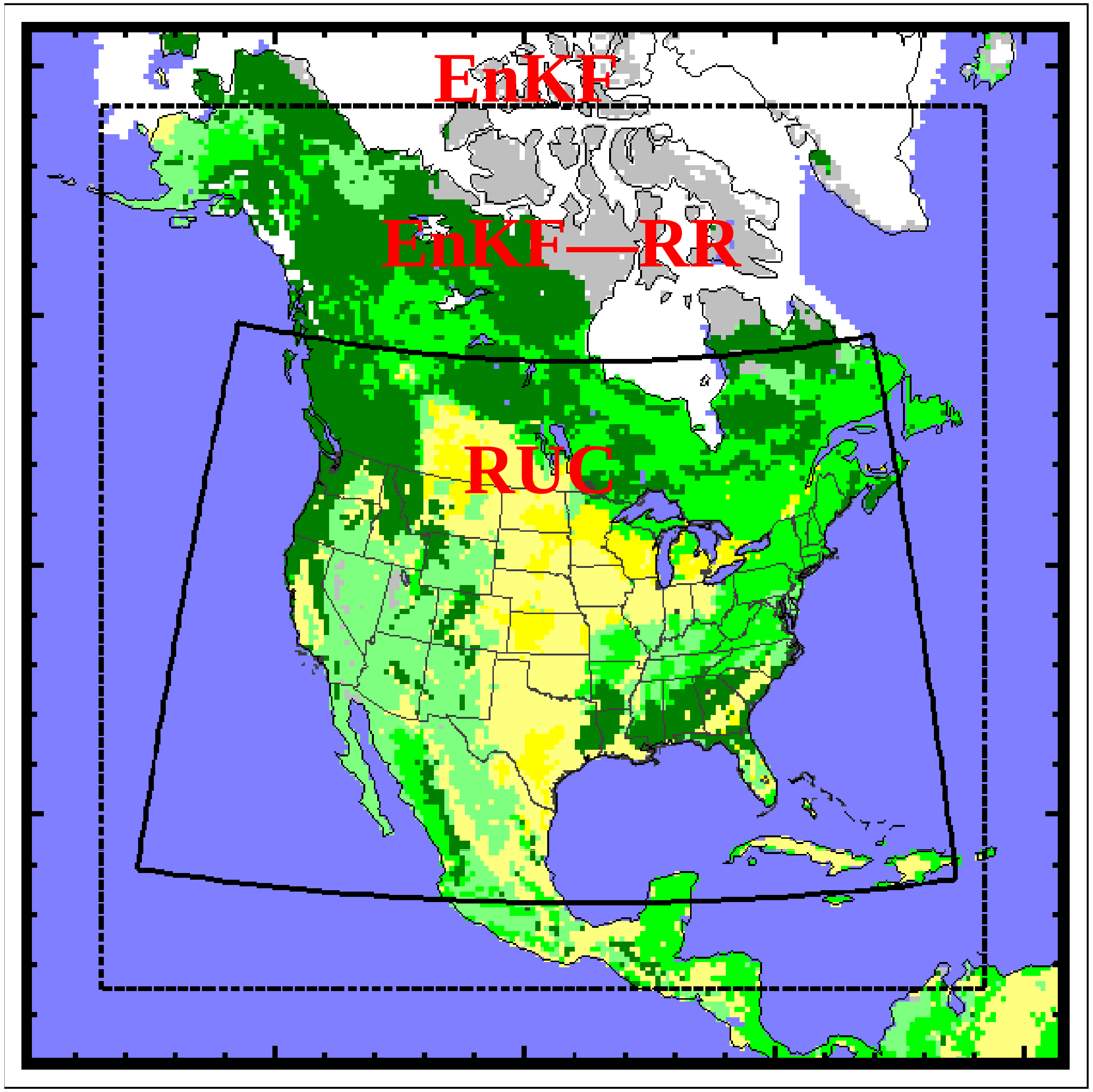
Development and Testing of GSI-Based EnKF and EnKF/3DVAR Hybrid Data Assimilation Systems for the Operational Rapid Refresh System

Kefeng Zhu[1], Yujie Pan[1], Ming Xue, Xuguang Wang[1,2], Jeffrey S. Whitaker[3], Stanley G. Benjamin[3], Stephen S. Weygandt[3] and Ming Hu[3]

Center for Analysis and Prediction of Storms[1] and School of Meteorology[2]
University of Oklahoma, Norman Oklahoma 73072
NOAA Earth System Research Laboratory[3], Boulder, Colorado

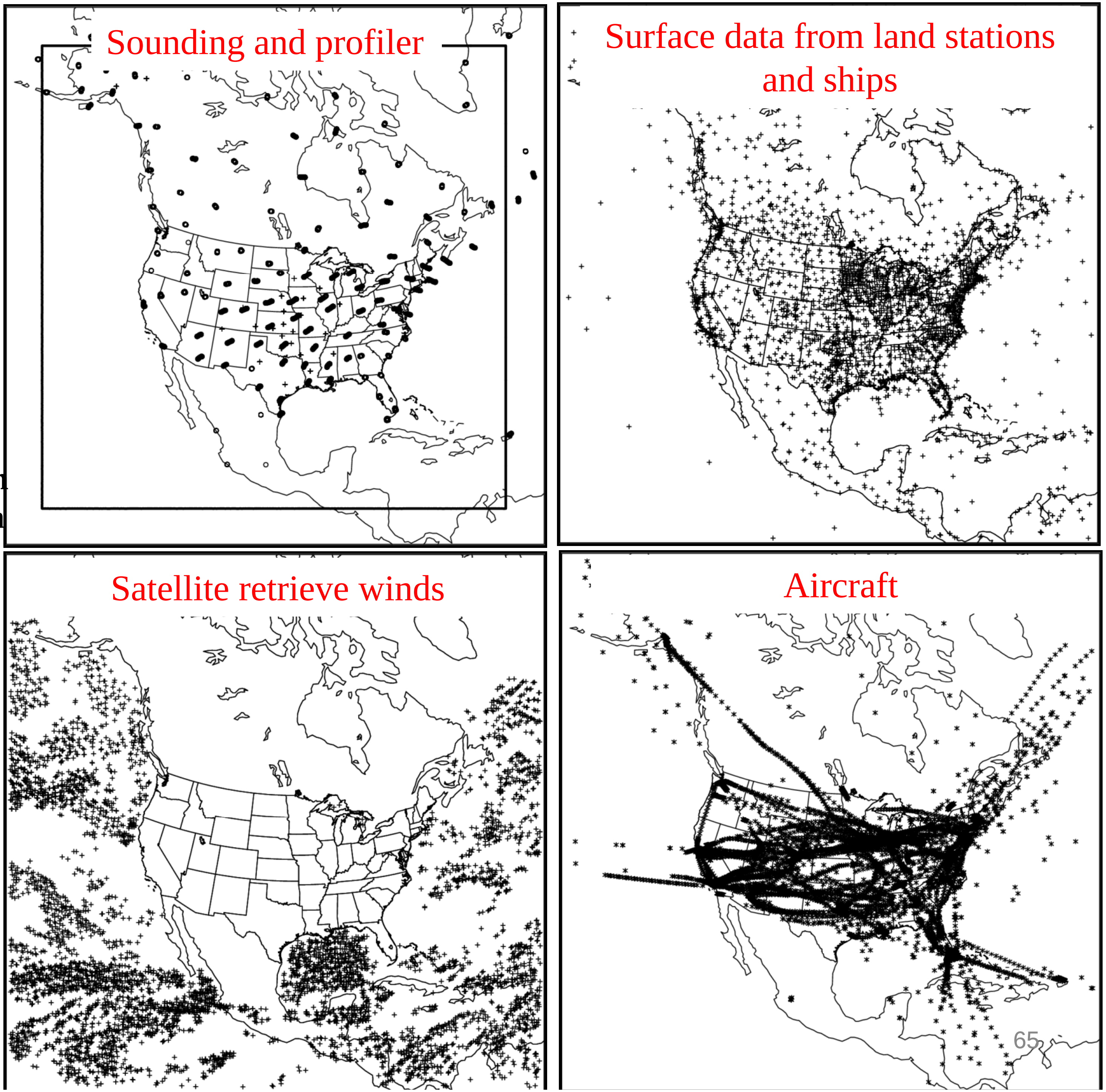
(Zhu et al. MWR 2013; Pan et al. MWR 2013)

Domains

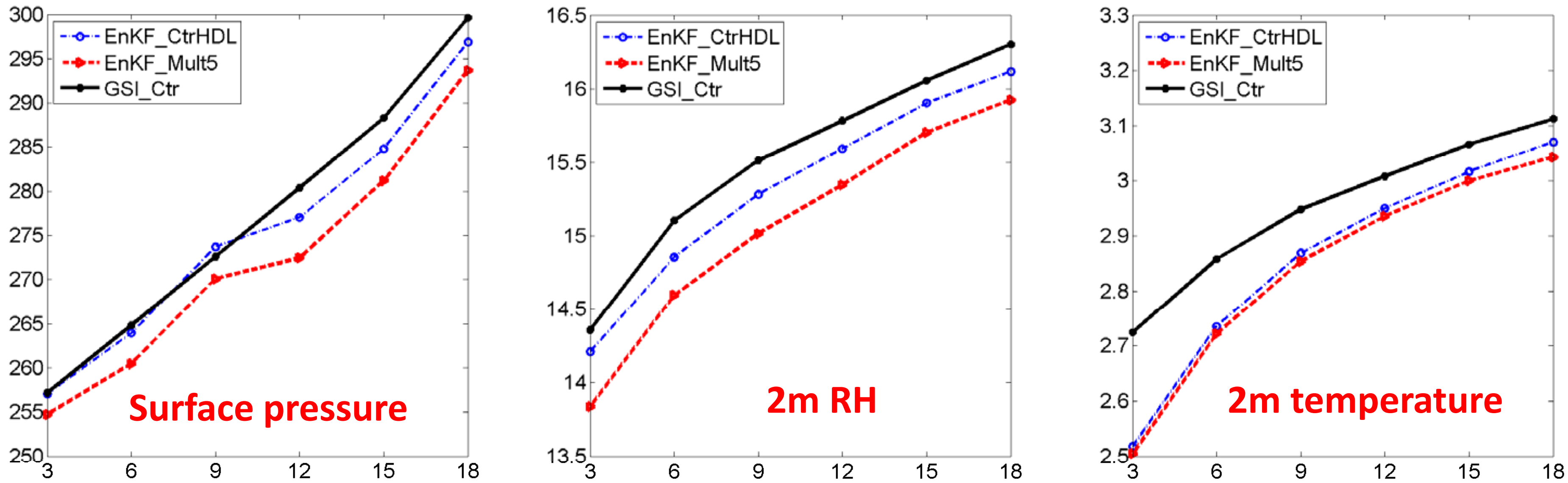


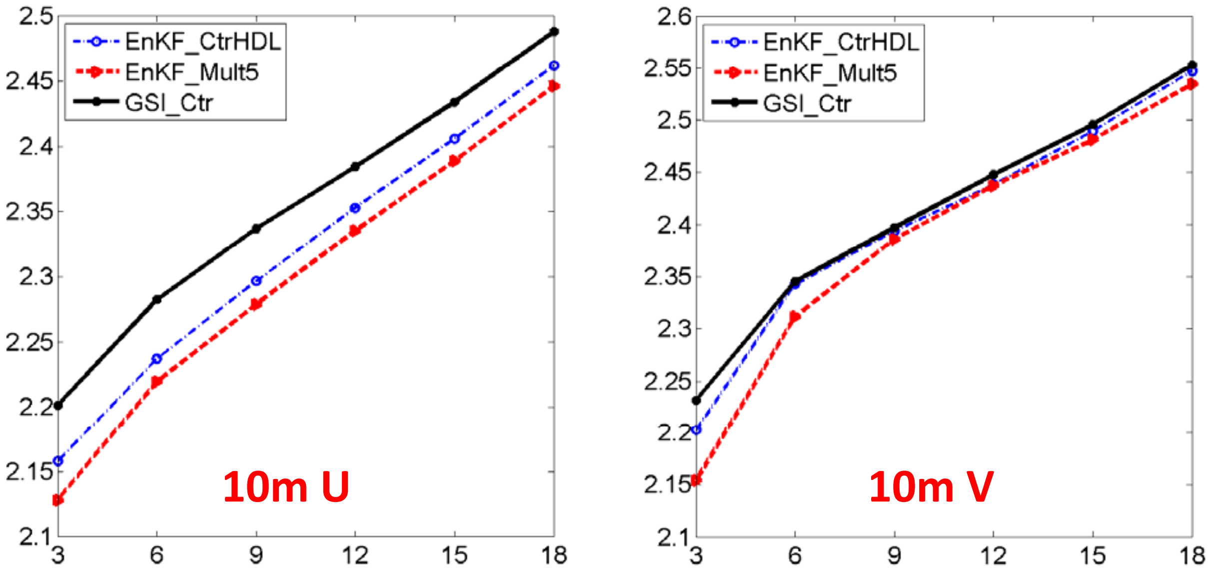
EnKF Test Domain
207x207 grid points
~40 km, 51 levels
The 13 km RR-like forecast Domain
532x532 grid points
~13 km, 51 levels
Current RUC Domain as indicated

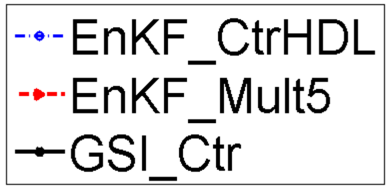
Horizontal distribution of selected observation kinds



Surface variables verification---against ADPSFC

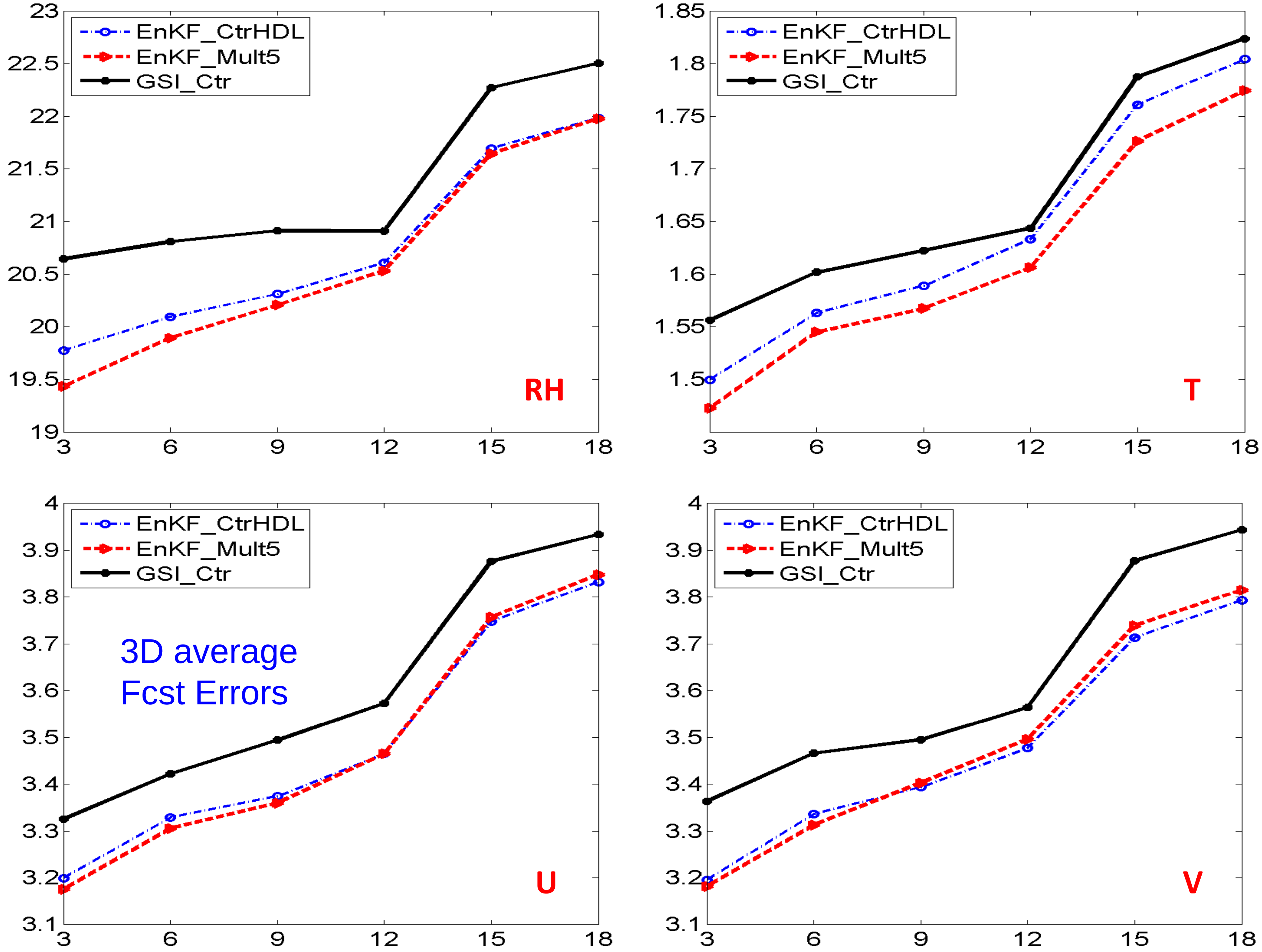


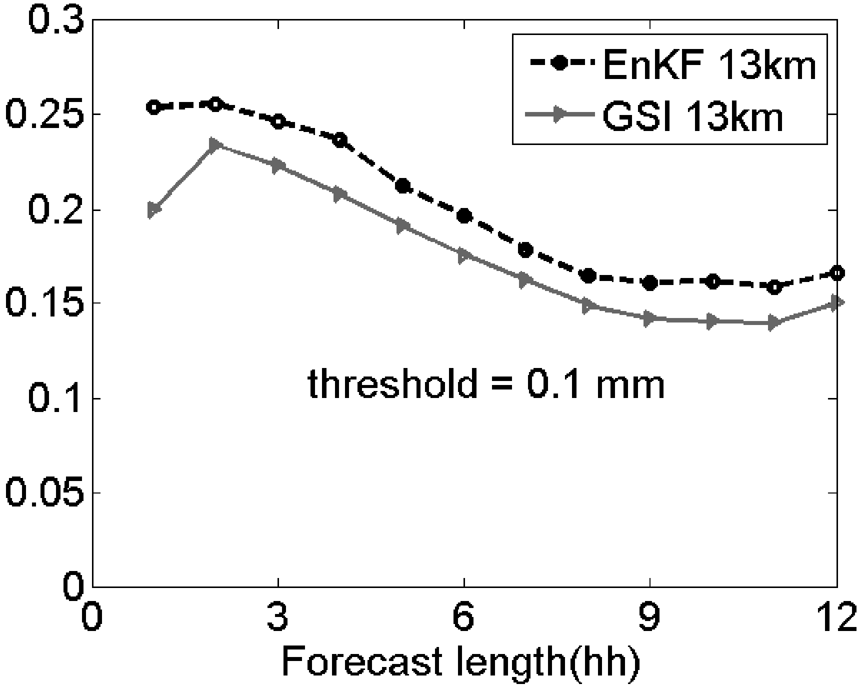


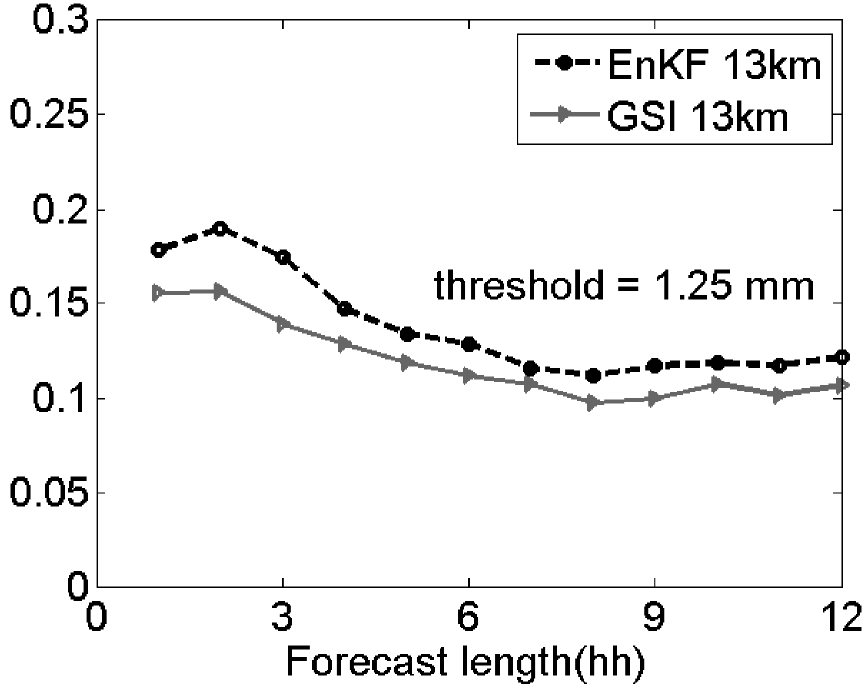


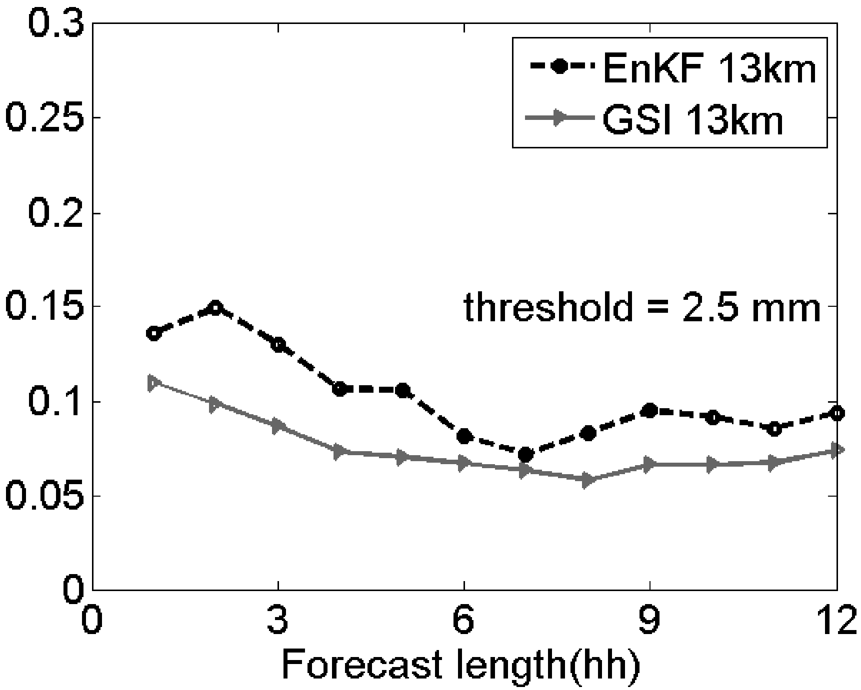
When multiple-physics is employed, EnKF consistently outperforms GSI in surface variable verification

Height-dependent localization radii are used.


EnKF_CtrHDL
EnKF_Mult5
GSI_Ctr
RH
T
U
V
3D average
Fcst Errors
3
6
9
12
15
18







EnKF is consistently better than GSI for the precipitation forecast

GSI-Hybrid: Method

Extended control variable method in 3D GSI hybrid (Wang 2010, MWR):

$$J\left(\mathbf{x}_1^{'},\boldsymbol{\alpha}\right)=\beta_1 J_1+\beta_2 J_e+J_o$$

$$=\beta_1\frac{1}{2}\mathbf{x}_1^{'T}\mathbf{B}^{-1}\mathbf{x}_1^{'}+\underbrace{\beta_2\frac{1}{2}\boldsymbol{\alpha}^T\mathbf{C}^{-1}\boldsymbol{\alpha}}+\frac{1}{2}\left(\mathbf{y}^{o'}-\mathbf{H}\mathbf{x}^{'}\right)^T\mathbf{R}^{-1}\left(\mathbf{y}^{o'}-\mathbf{H}\mathbf{x}^{'}\right)$$

Extra term associated with extended control variable

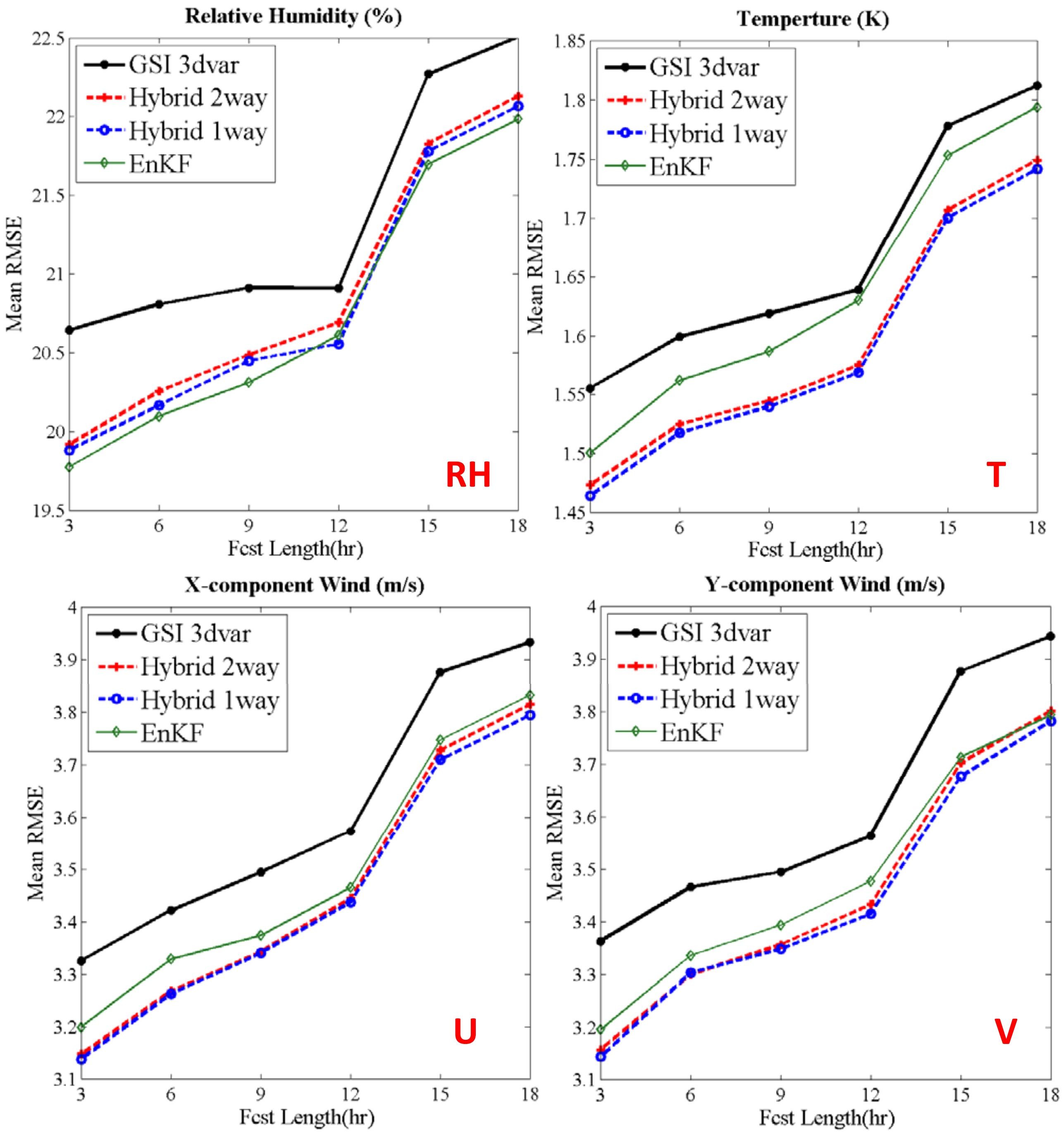
$$\mathbf{x}^{'}=\mathbf{x}_1^{'}+\underbrace{\sum_{k=1}^{K}\left(\boldsymbol{\alpha}_k\circ\mathbf{x}_k^e\right)}$$

Extra increment associated with ensemble

$\mathbf{B}$ 3DVAR static covariance; $\mathbf{R}$ observation error covariance; K ensemble size;
$\mathbf{C}$ correlation matrix for ensemble covariance localization; $\mathbf{x}_k^e$ kth ensemble perturbation;
$\mathbf{x}_1^{'}$ 3DVAR increment; $\mathbf{x}^{'}$ total (hybrid) increment; $\mathbf{y}^{o'}$ innovation vector;
$\mathbf{H}$ linearized observation operator; β_1 weighting coefficient for static covariance;
β_2 weighting coefficient for ensemble covariance; $\boldsymbol{\alpha}$ extended control variable.

Upper air verification against soundings

3D average Fcst Errors

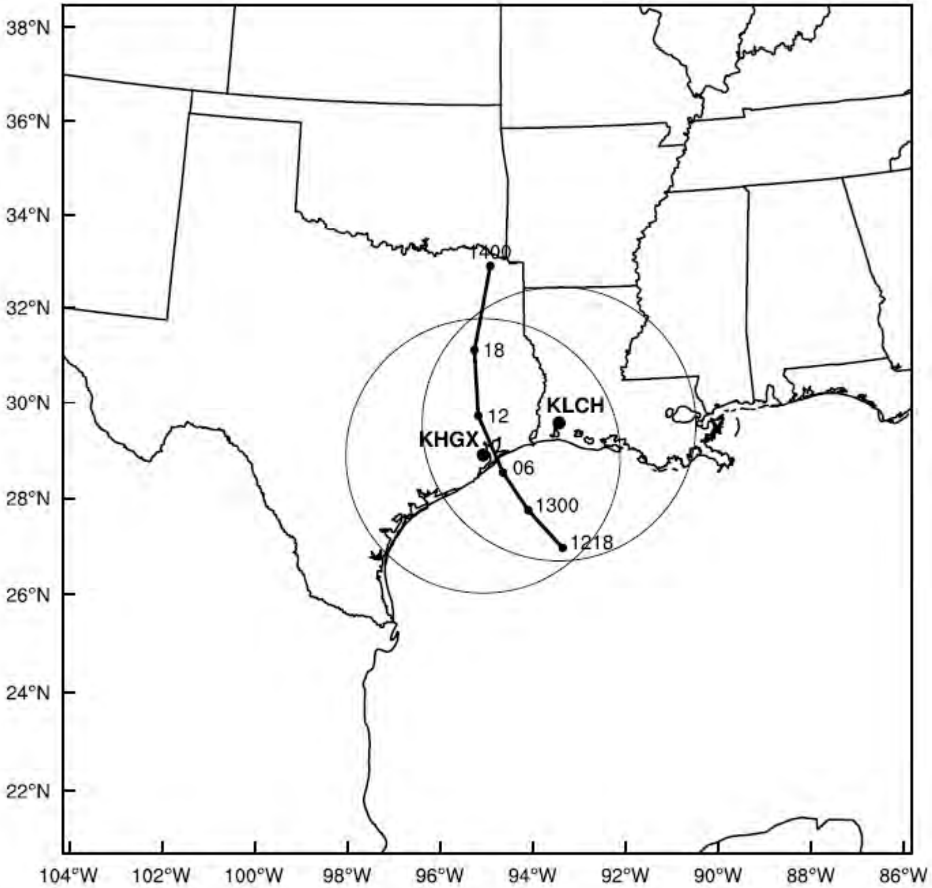


In 3D, hybrid better than EnKF in T, U, V fcst, slightly worse for RH

Hurricane Radar DA

- EnKF assimilation for Ike (2008) (Dong and Xue QJ2012)
- Hybrid-3DVAR DA for Ike (2008) (Li, Wang and Xue MWR 2012)
- Assimilation of GBVTD wind retrieval for Typhoon Shaomei (2006) (Zhao et al.)
- 3DVAR assimilation of airborne radar data for Ike (Du et al)
- 3DVAR+Cloud Analysis of ground-based radars for Ike (2008), Meranti (2010), Jangmi (2008), Marokot (2009)

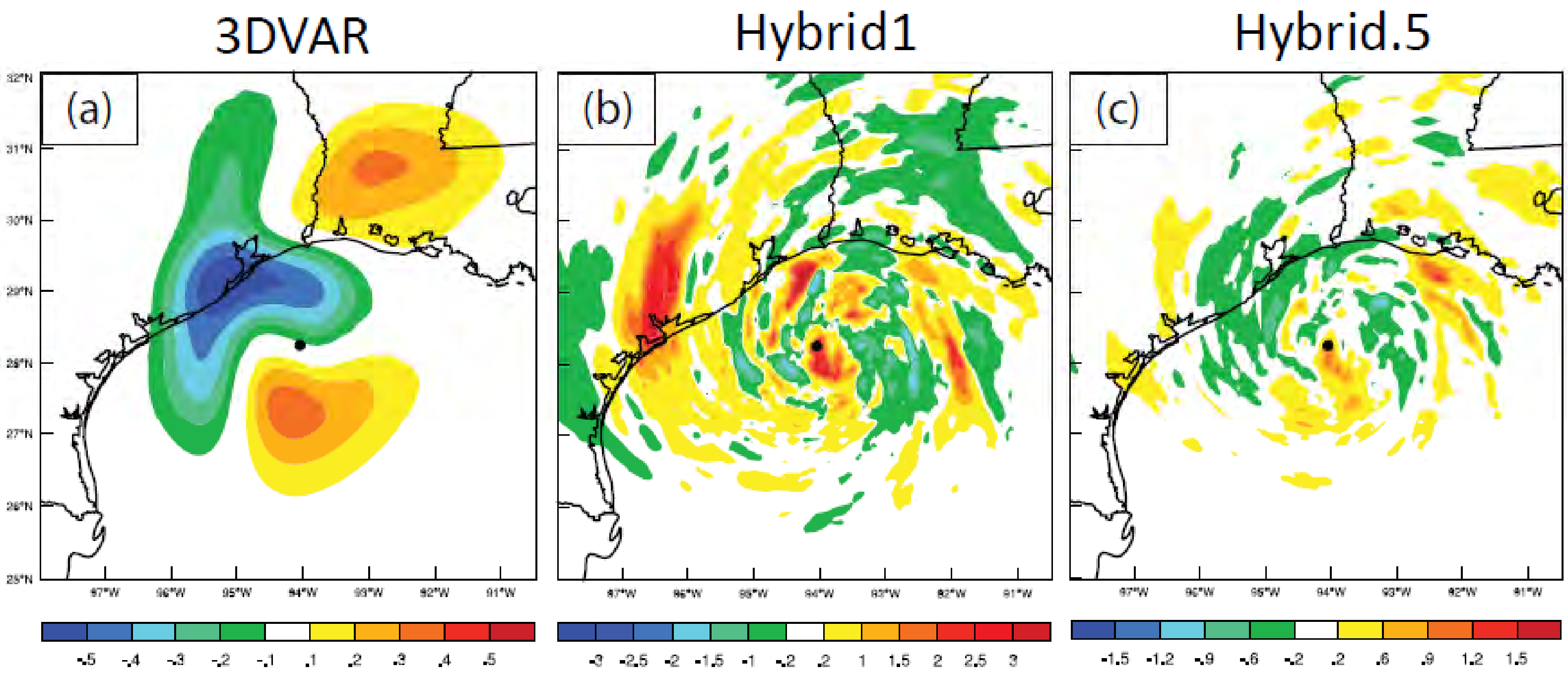
Coastal Radar DA using WRF 3DVAR/Hybrid for Ike (2008)



- WRF: Δx=5km
- Ensemble size: 40 members
- DA and forecast:
  - Cold start 6h ensemble initialized at 18ZSep12
  - Radar DA starting at 00ZSep13, every 30 min, for 3 hours
  - 21h deterministic forecast from 03ZSep13
- Hybrid DA system: WRFDA based hybrid (Wang et al. 2008a)
- Ensemble generation: perturbed observation
- Observations: radial velocity from two WSR88D radars (KHGX, KLCH)

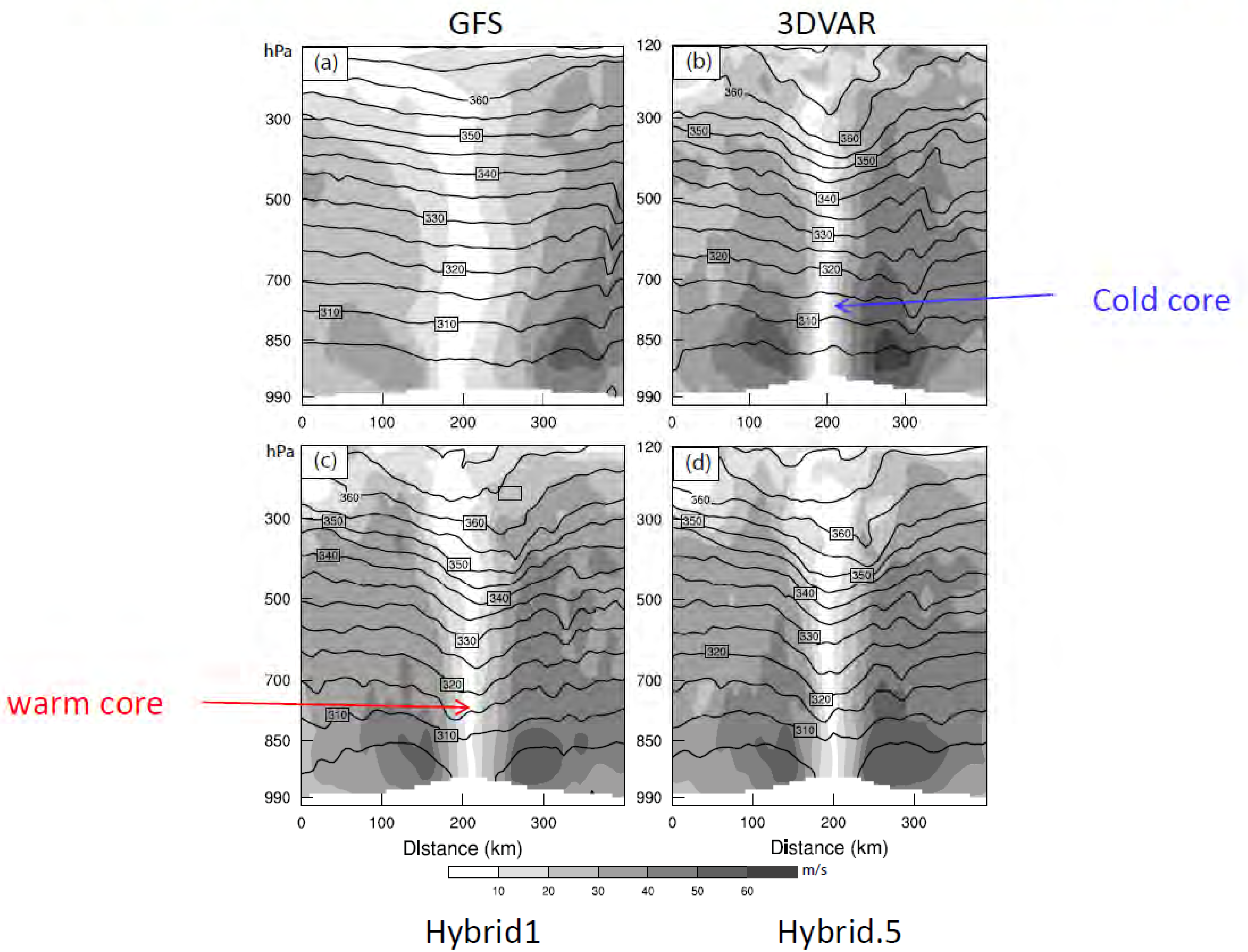
(Li, Wang and Xue 2012)

Temperature increment

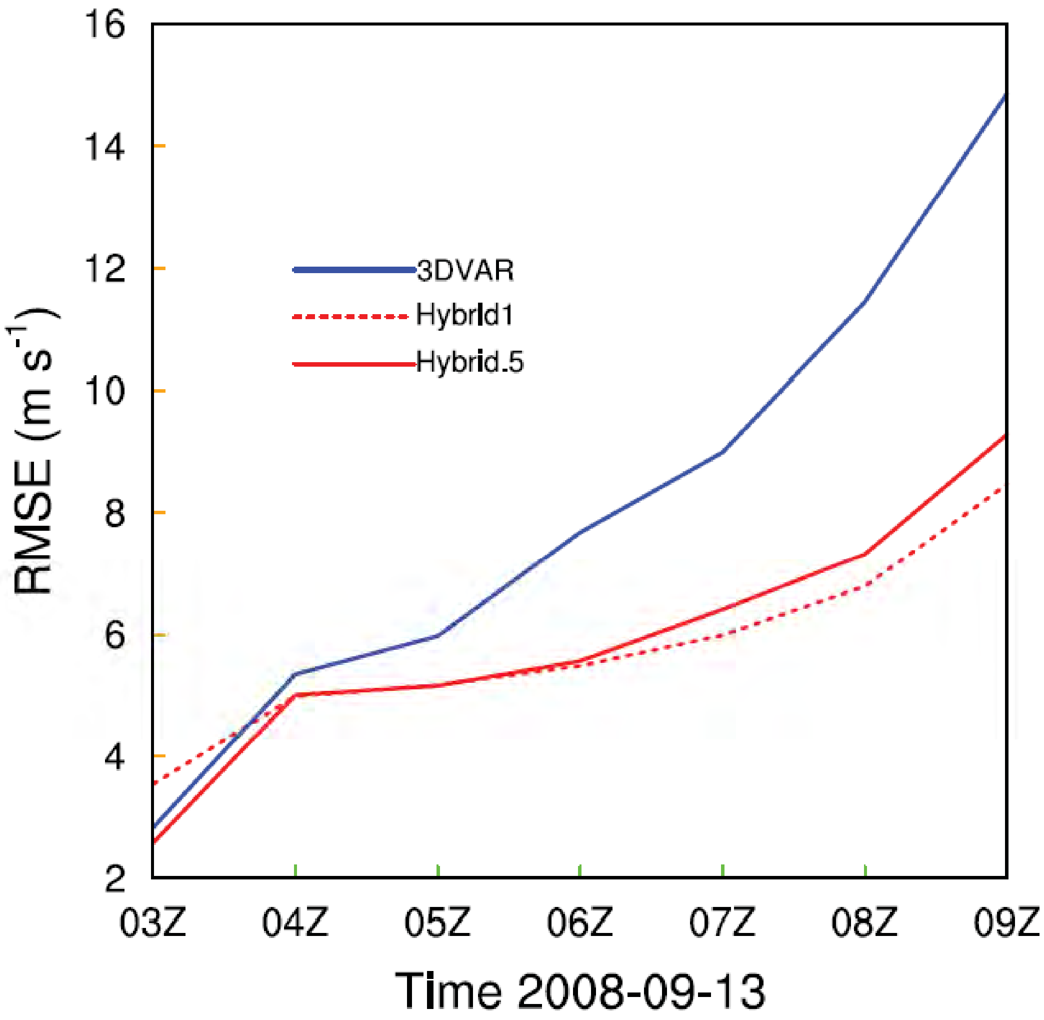


- Cross variable increment by the 3DVAR and hybrid were different, with the temperature increment by the hybrid reflected rain band structure.

Horizontal wind speed and potential temp. analyses



Forecasts verified against radial velocity



- Forecast initialized by the hybrid fit the radial velocity better than the 3DVAR.

- inclusion of static covariance in the hybrid did not help except at early lead time.

Thanks!

感謝!