Ensemble-based data assimilation: an essential component of a future mesoscale ensemble system

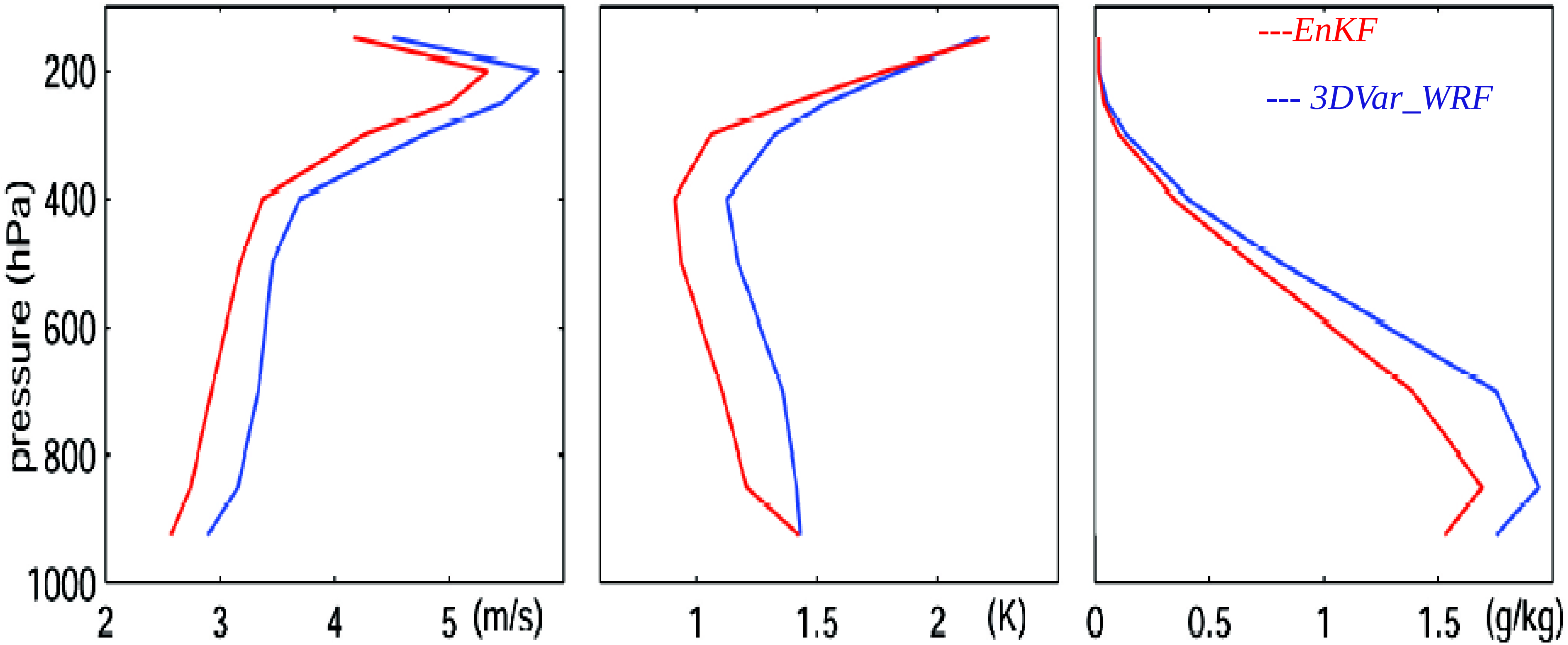
Fuqing Zhang
Penn State University

Mesoscale EnKF: Where are we at vs. VAR?

WRF EnKF vs. 3DVar assimilate soundings for BAMEX June 2003

WRF EnKF: 40-member, 30-km, multiphysics ensemble assimilating standard radiosondes every 12h quality controlled by WRF 3DVAR; domain is over most of the continental US with lateral boundary from a 90-km WRF ensemble (Meng and Zhang 2008a,b MWR)

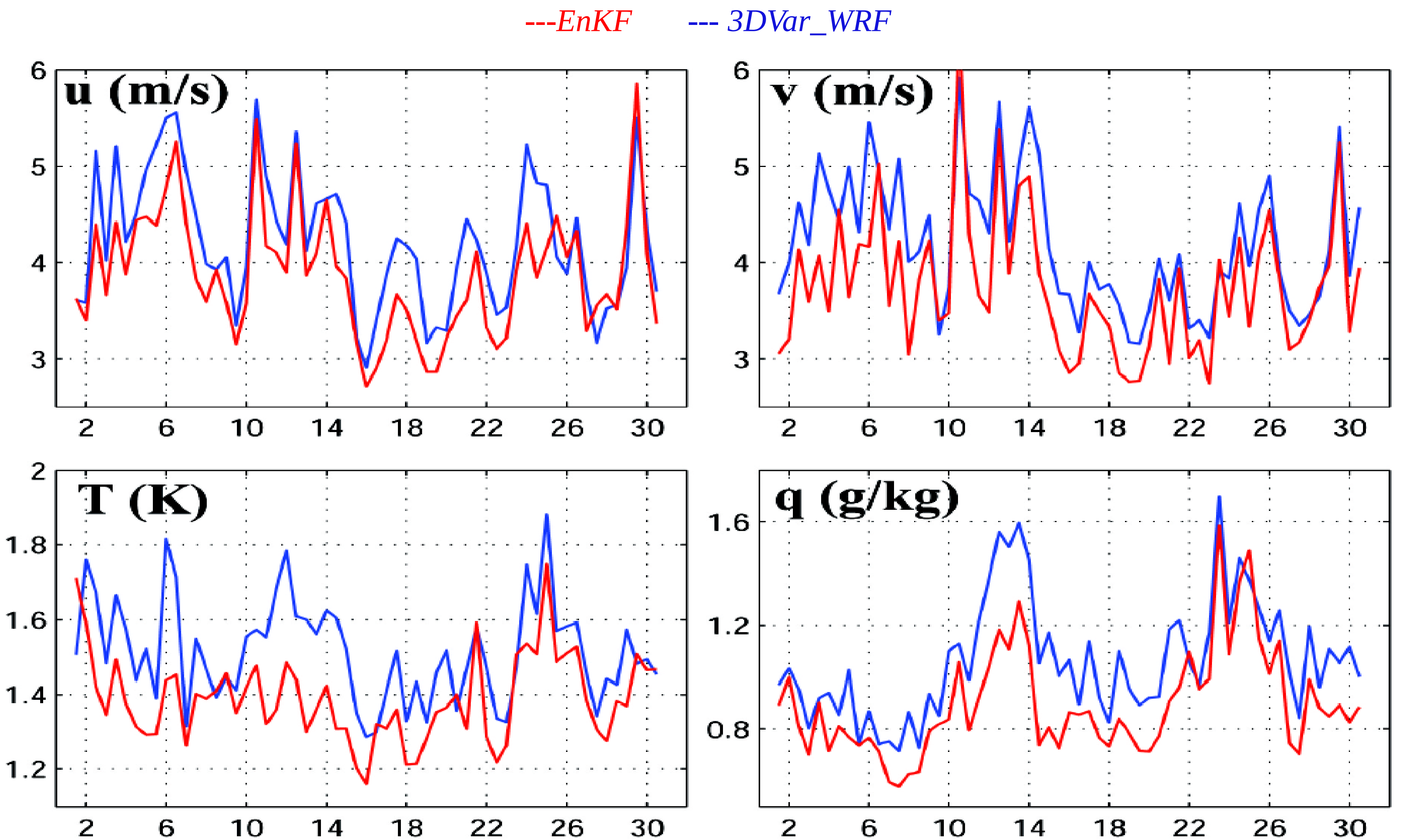
WRF 3DVAR: tuned for best performance, B from NMC method with forecast in May



***12-h forecast RMSE:** EnKF performs better than WRF 3DVar at almost every vertical level*

Mesoscale EnKF: Where are we at vs. VAR?

WRF EnKF vs. 3DVar assimilate soundings for BAMEX June 2003

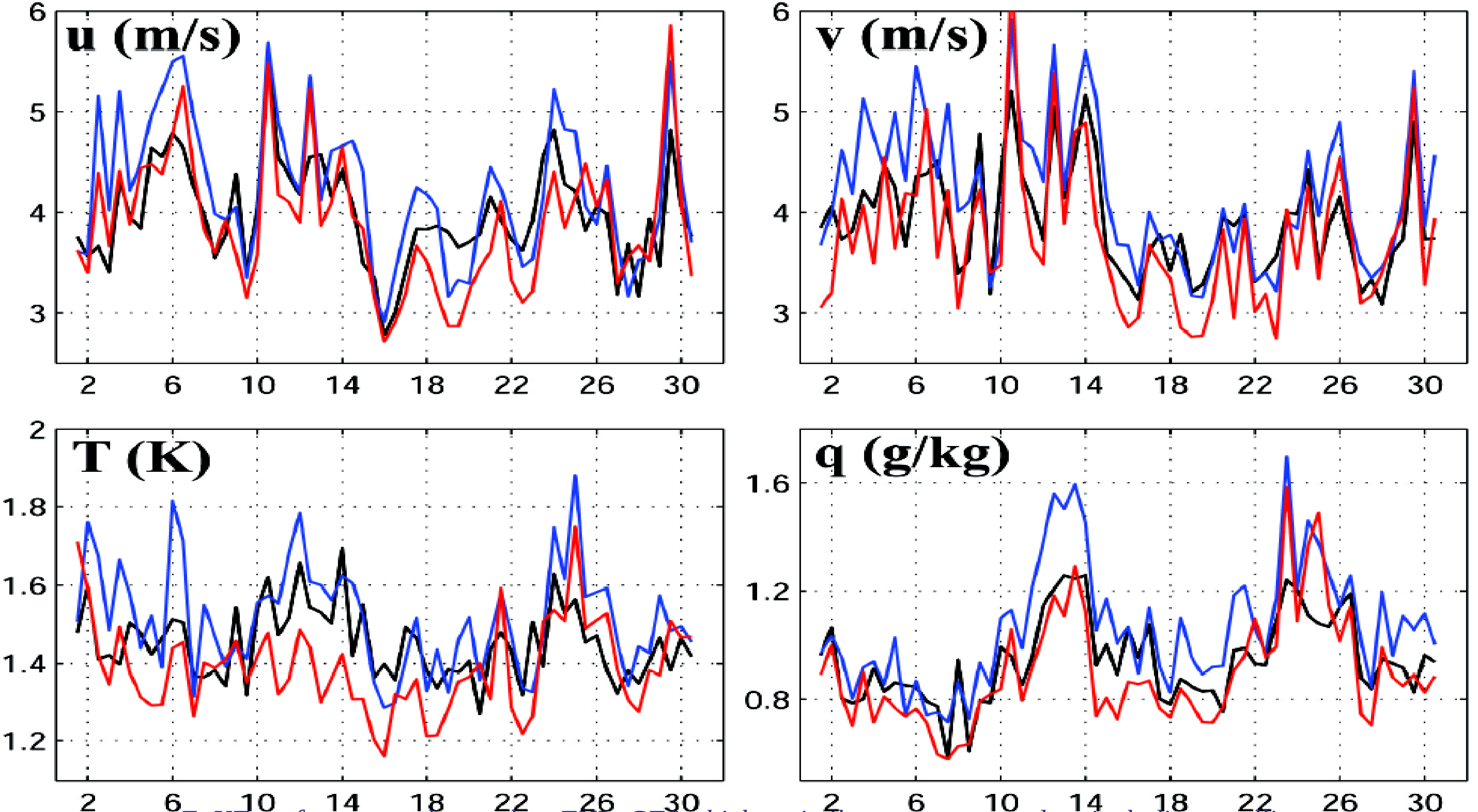


12-h forecast RMSE: *EnKF performs better than WRF 3DVar for the whole month of June 2003*

Mesoscale EnKF: A status update

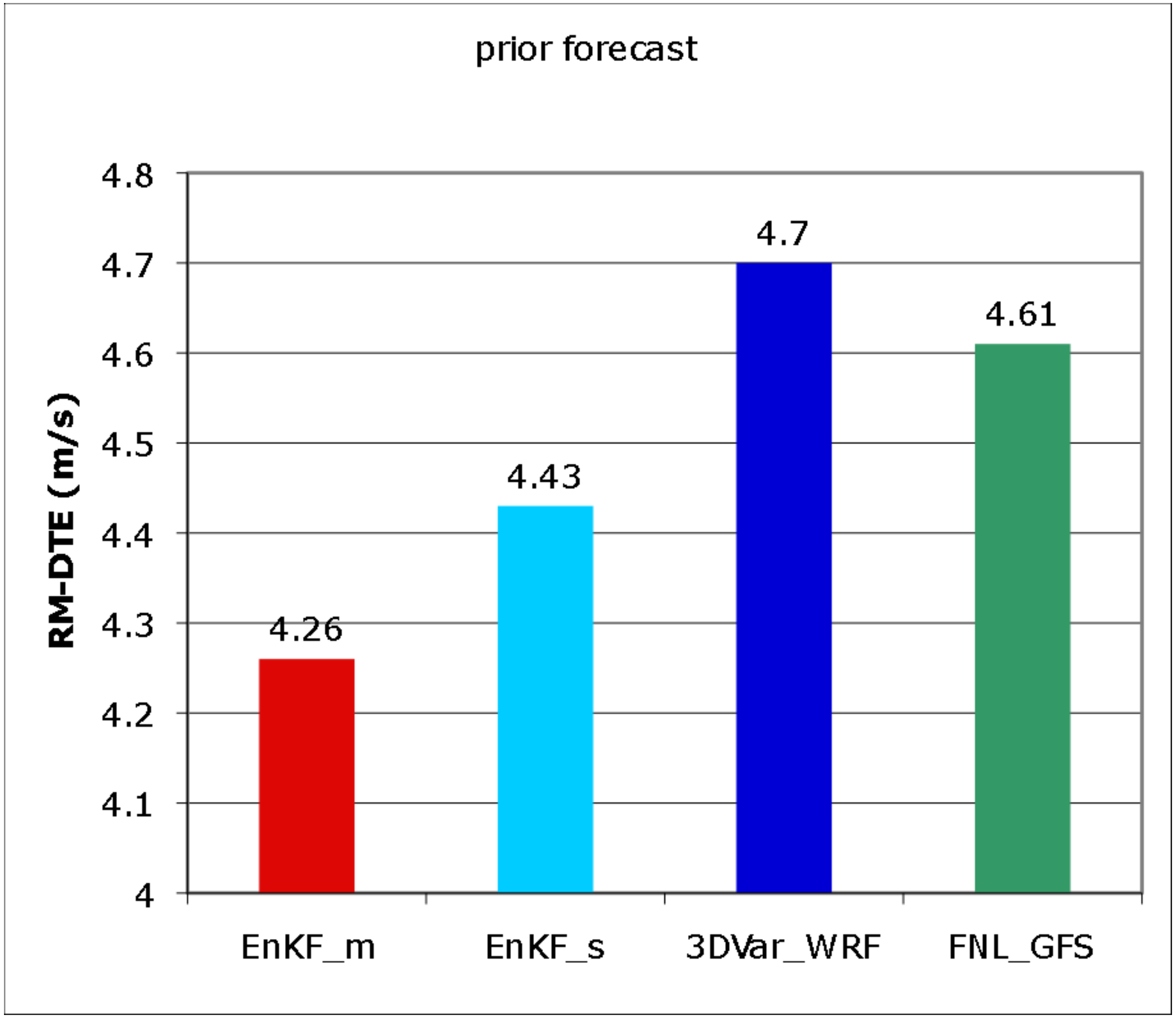
WRF EnKF vs. 3DVar assimilate soundings for June 2003:12h RMSE

--- EnKF *--- 3DVar_WRF* *--- WRF 12h fcst from FNL_GFS*



EnKF performs even better than FNL_GFS which assimilates many more data including satellite
FNL_GFS has a generally smaller 12-h forecast error than wrf-3DVar.

Monthly Averaged 12-h-Forecast RM-DTE for June 2003



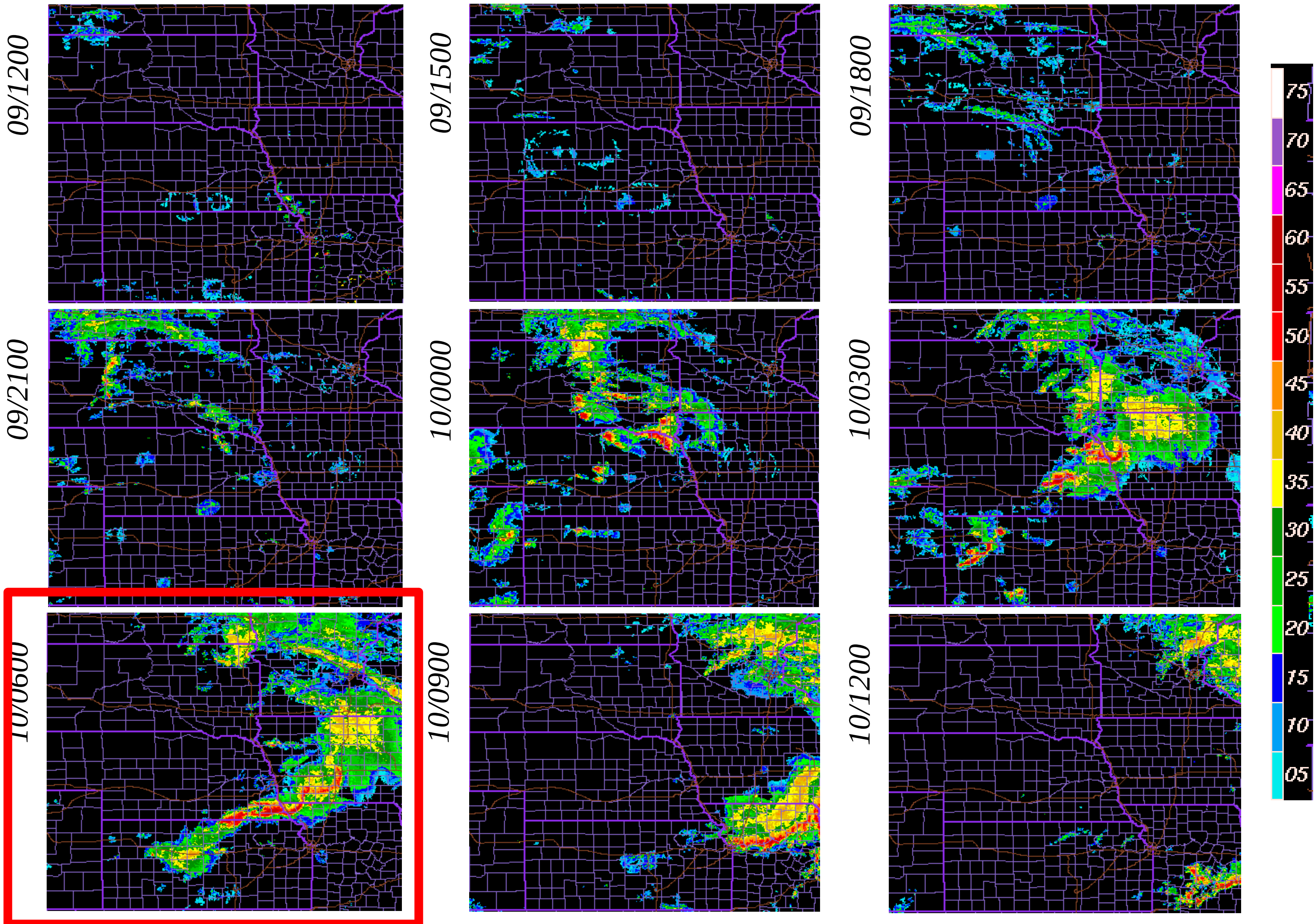
- ***EnKF has significantly smaller overall 12-h forecast error than both WRF-3DVar and FNL_GFS***
- ***FNL_GFS has smaller overall forecast error than WRF-3DVar***

(Meng and Zhang 2008a,b)

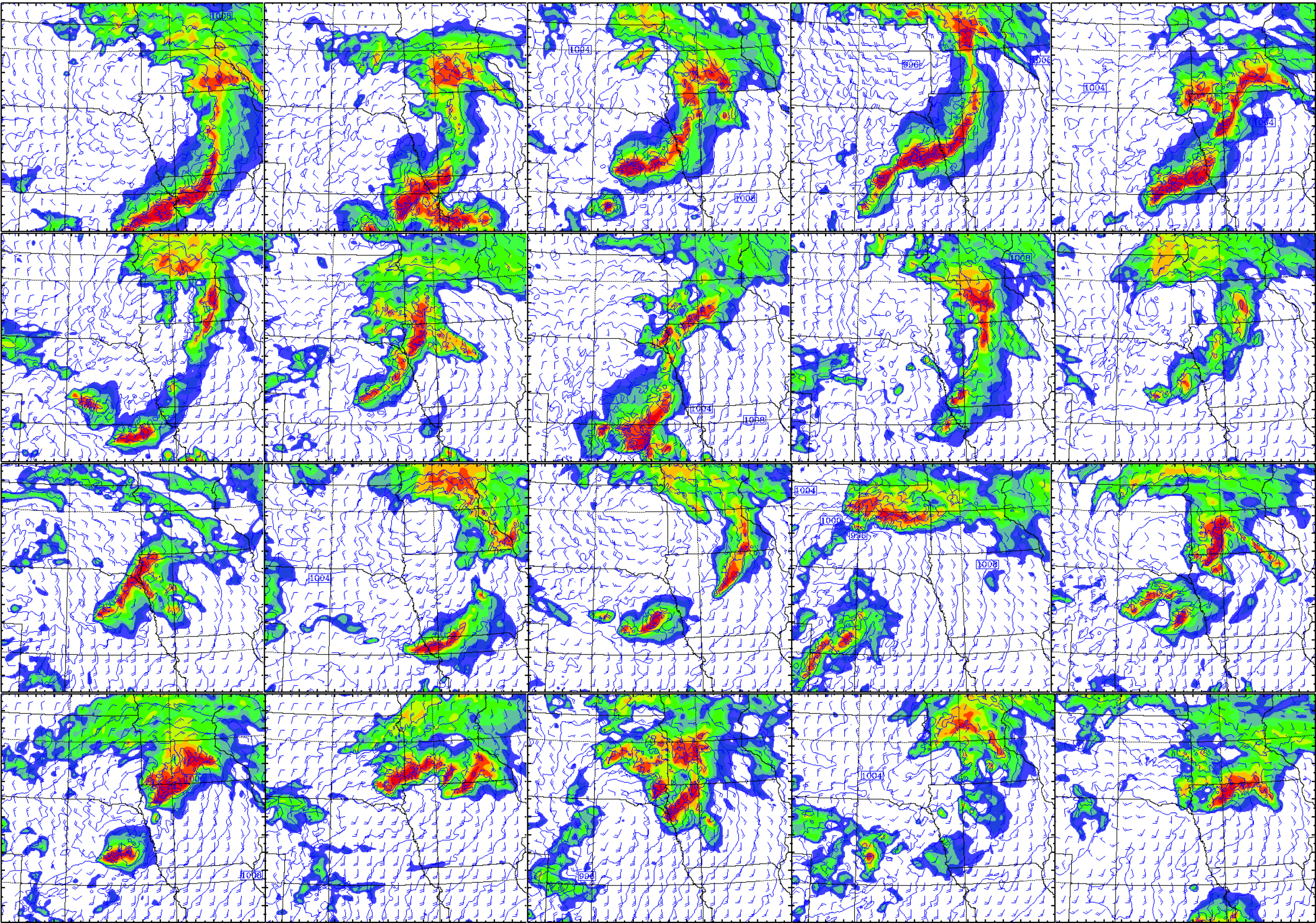
Issues Specific to Mesoscale EnKF

- Multi-scale in nature
  - Balance versus imbalance: need additional balance constraint?
  - Moist error growth dynamics at meso-/convective scales
  - Significance of model error, esp. in moist physics and boundary layer
  - Strong inhomogeneity in data coverage, lack of good thermodynamic obs
  - Localization challenge: moving beyond empirical tuning?
  - Ensemble initiation, startup vs. lead time, DFI windows
- Needs for lateral boundary conditions and nesting
  - Perturbation availability and consistency from global models
  - Multiple domain updating, one-way versus two-way nesting
  - Related: Unified model, dual resolution
- Satellite data assimilation for mesoscales
  - Bias correction
  - Model top
- Validation and inter-comparison with variational methods
  - Lack of common domains and metrics
  - Grid-point RMSE versus feature-based verifications
- Computing and parallelization

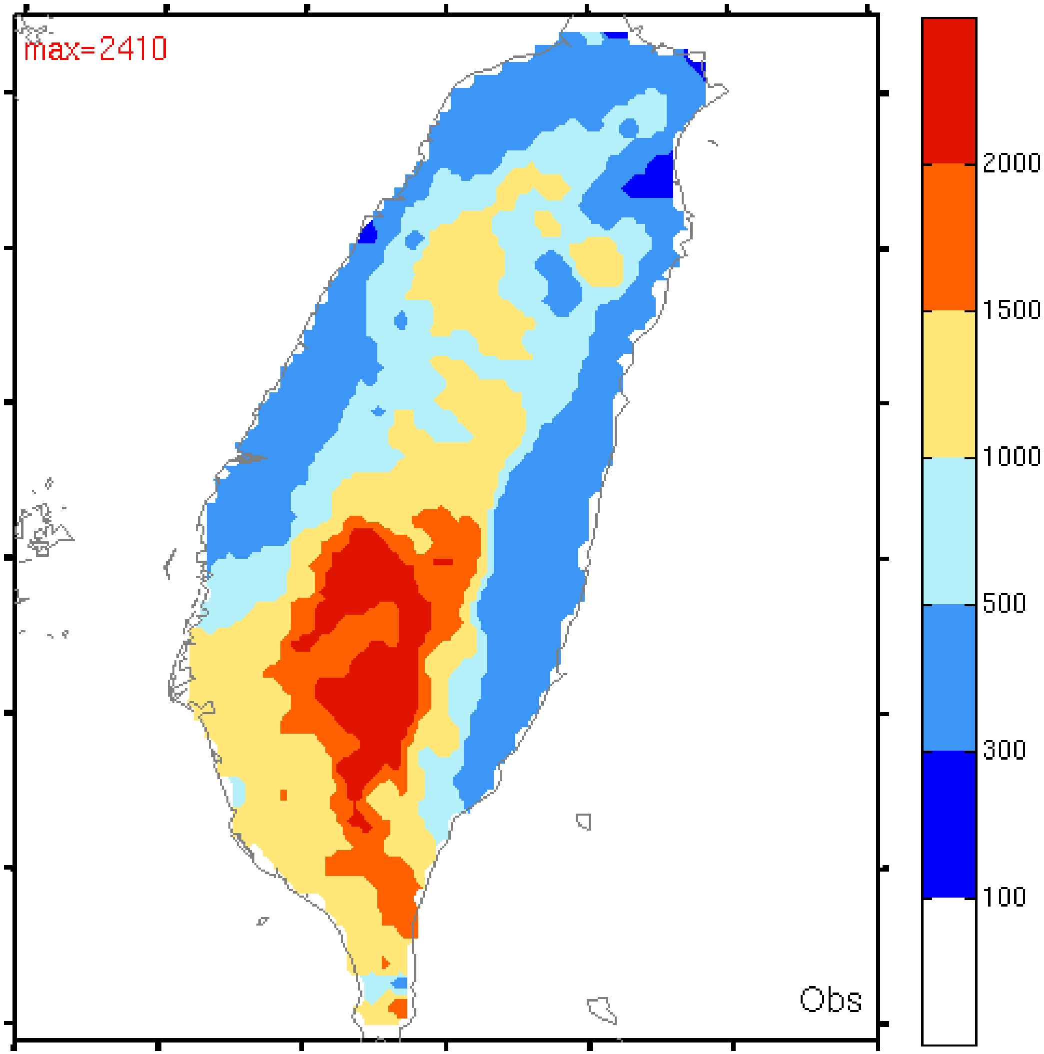
Cloud-resolving ensemble initialized with EnKF: IOP 7b



Bow echo forecasts by 3.3-km cloud-resolving ensemble



Record rainfall by Typhoon Morakot over Taiwan



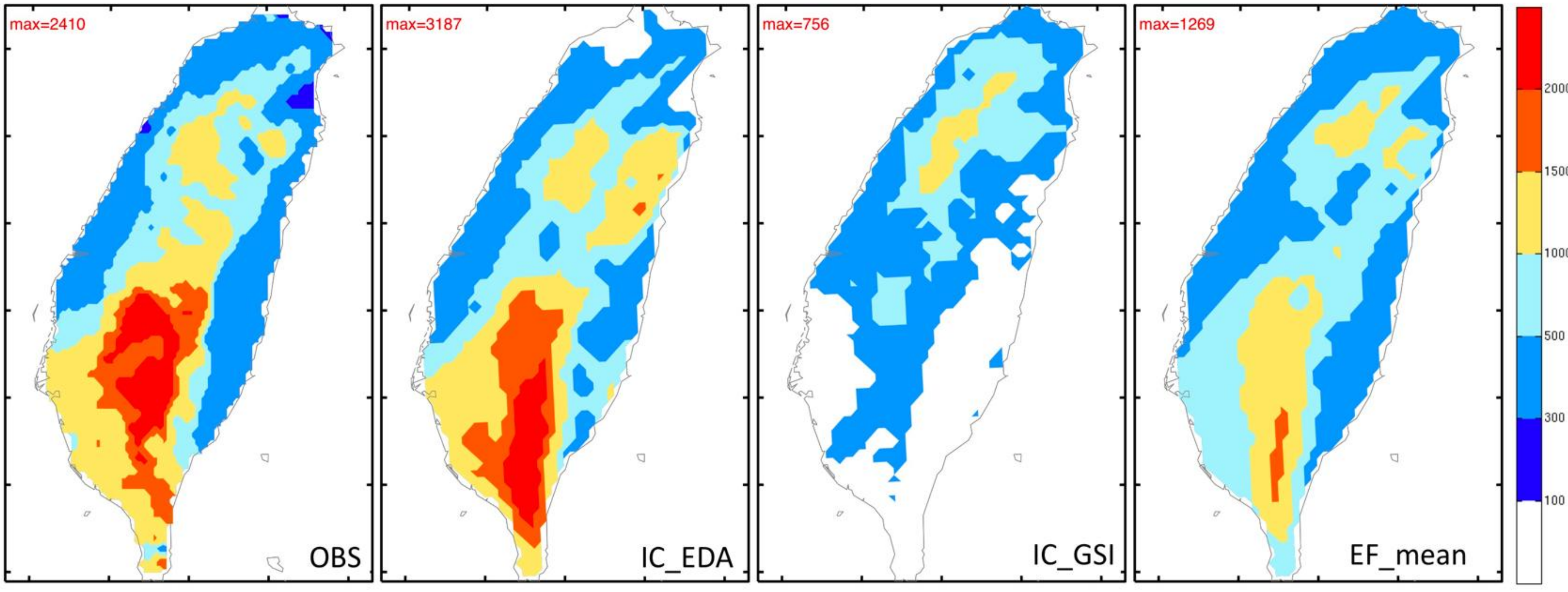
Courtesy of Bill Kuo

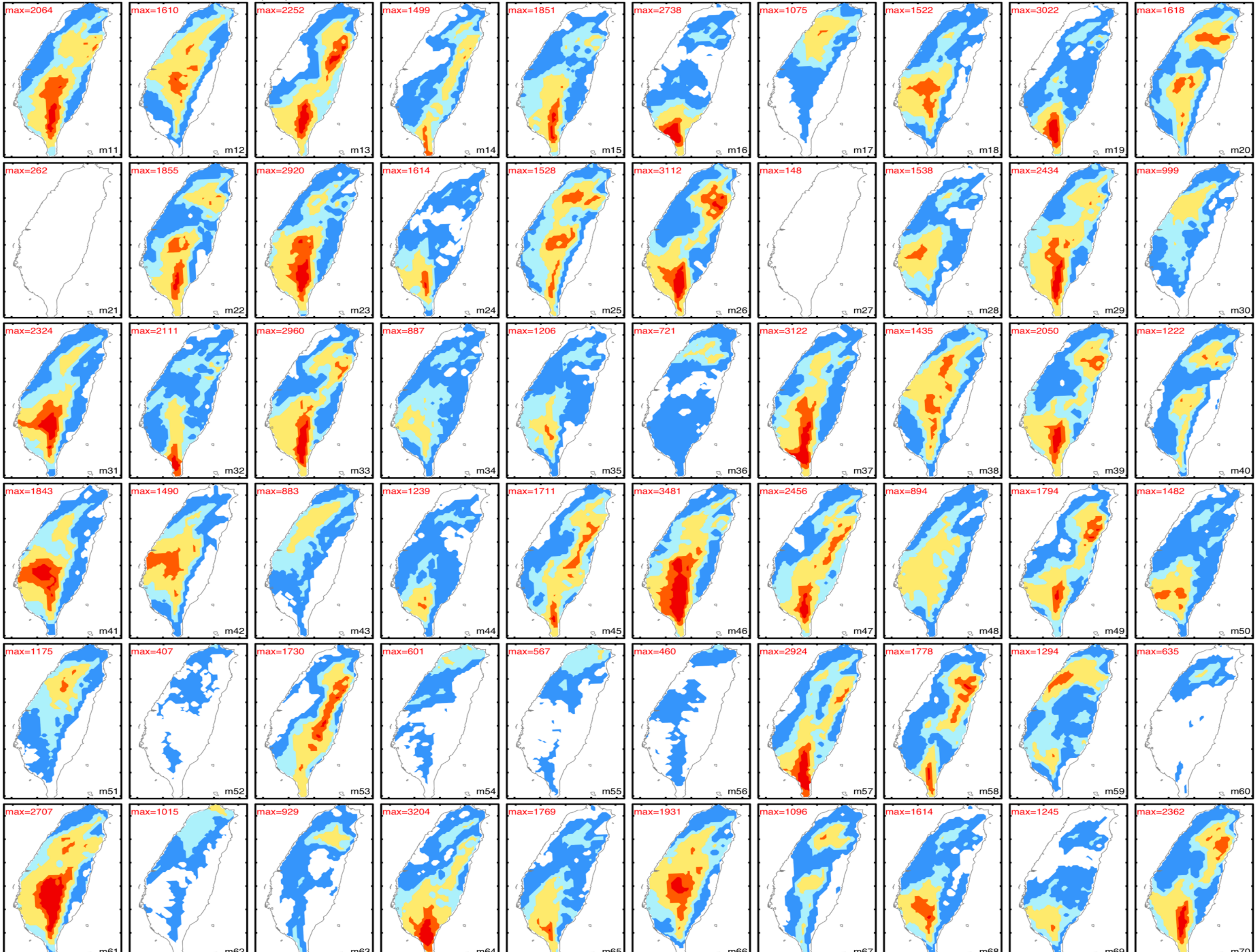
Cloud-resolving ensemble prediction of Morakot

WRF ensemble: two domains with 13.5/4.5km grid spacing, 60-member initialized with GFS EnKF perturbations running at TACC ranger cluster as part of HFIP

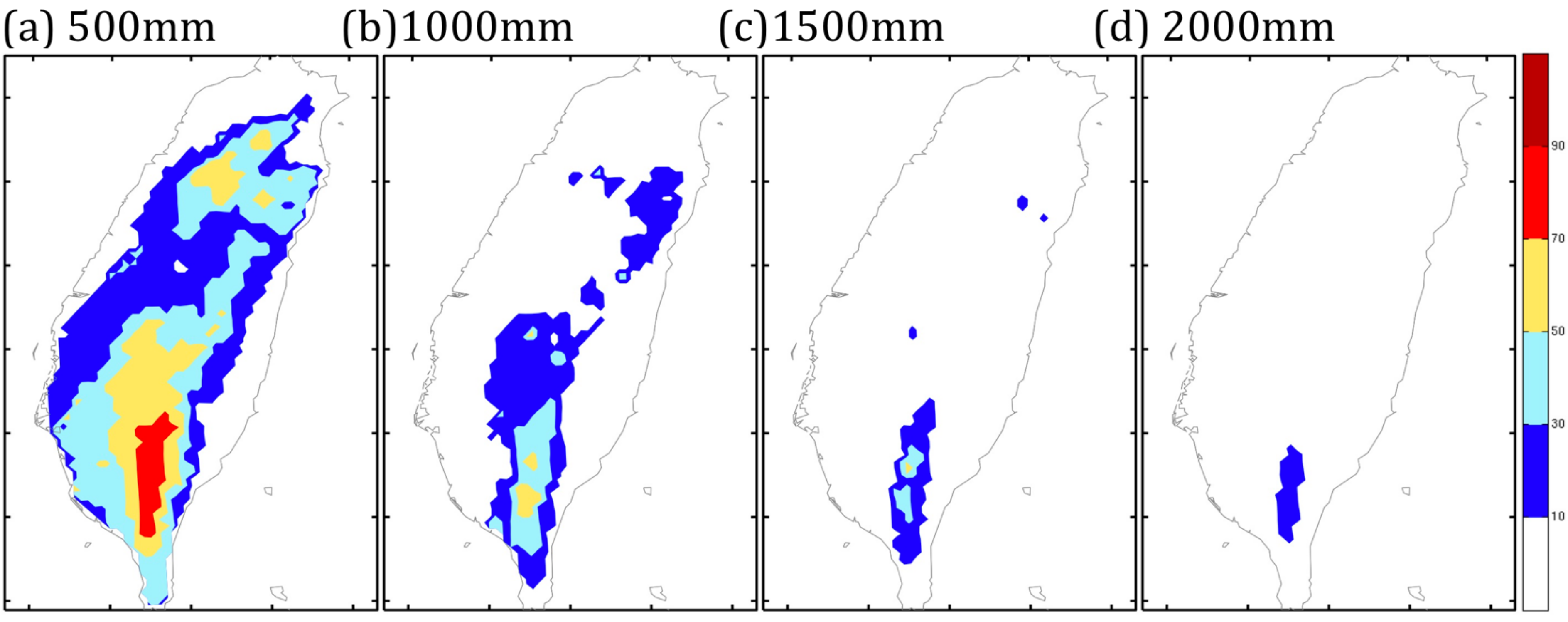
Two deterministic WRF forecasts: one initialized with mean GFS EnKF analysis (IC_EDA) and the other with operational GFS GSI analysis (IC_GSI)

GFS EnKF: real-time analysis since July 1 at the operational resolution assimilating the same data as in NCEP operational GFS GSI system, also at TACC (Jeff Whitaker)




max=2064
m11
max=1610
m12
max=2252
m13
max=1499
m14
max=1851
m15
max=2738
m16
max=1075
m17
max=1522
m18
max=3022
m19
max=1618
m20
max=262
m21
max=1855
m22
max=2920
m23
max=1614
m24
max=1528
m25
max=3112
m26
max=148
m27
max=1538
m28
max=2434
m29
max=999
m30
max=2324
m31
max=2111
m32
max=2960
m33
max=887
m34
max=1206
m35
max=721
m36
max=3122
m37
max=1435
m38
max=2050
m39
max=1222
m40
max=1843
m41
max=1490
m42
max=883
m43
max=1239
m44
max=1711
m45
max=3481
m46
max=2456
m47
max=894
m48
max=1794
m49
max=1482
m50
max=1175
m51
max=407
m52
max=1730
m53
max=601
m54
max=567
m55
max=460
m56
max=2924
m57
max=1778
m58
max=1294
m59
max=635
m60
max=2707
max=1015
max=929
max=3204
max=1769
max=1931
max=1096
max=1614
max=1245
max=2362

Probabilistic Forecast of Morakot Rainfall by the Cloud-resolving ensemble prediction



Convective-scale EnKF Assimilation for Hurricanes

(Zhang et al. 2009 MWR)

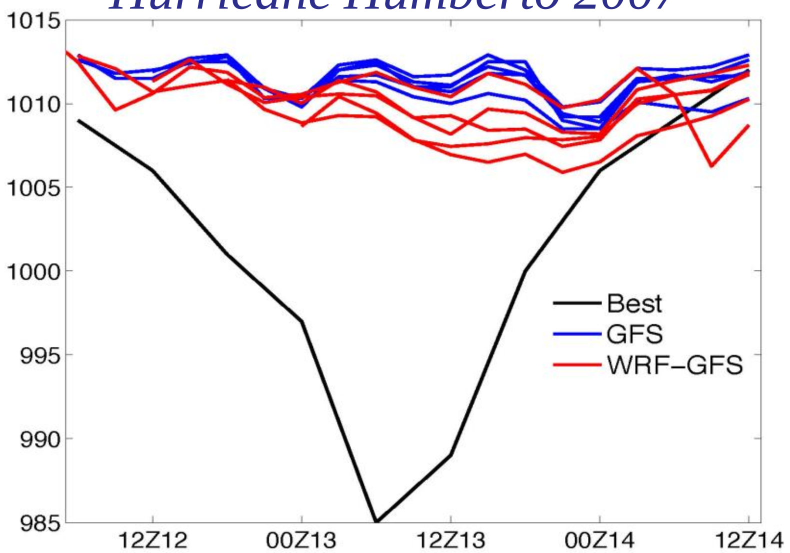
- WRF domains: D1-D2-D3 with grid sizes of 40.5, 13.5 and 4.5km
  - Physics: WSM 6-class microphysics; YSU PBL; Grell-Devenyi CPS

- EnKF system (Meng & Zhang 2008b): but 30-member no multiphysics

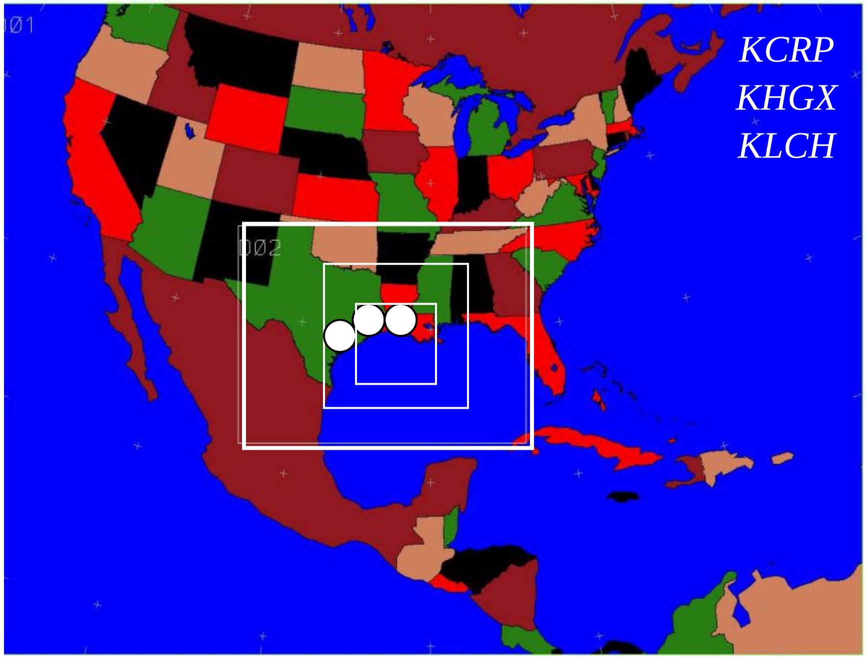
  - Initialized at 00Z 12 using 3DVar background uncertainty with FNL analysis; GFS forecast used for boundary condition in forecasts

- Data assimilated: WSR88D Vr at KCRP, KHGX & from KLCH 09Z12 to 12Z 13 Sept 2007; Successive covariance localization; obs err 3m/s

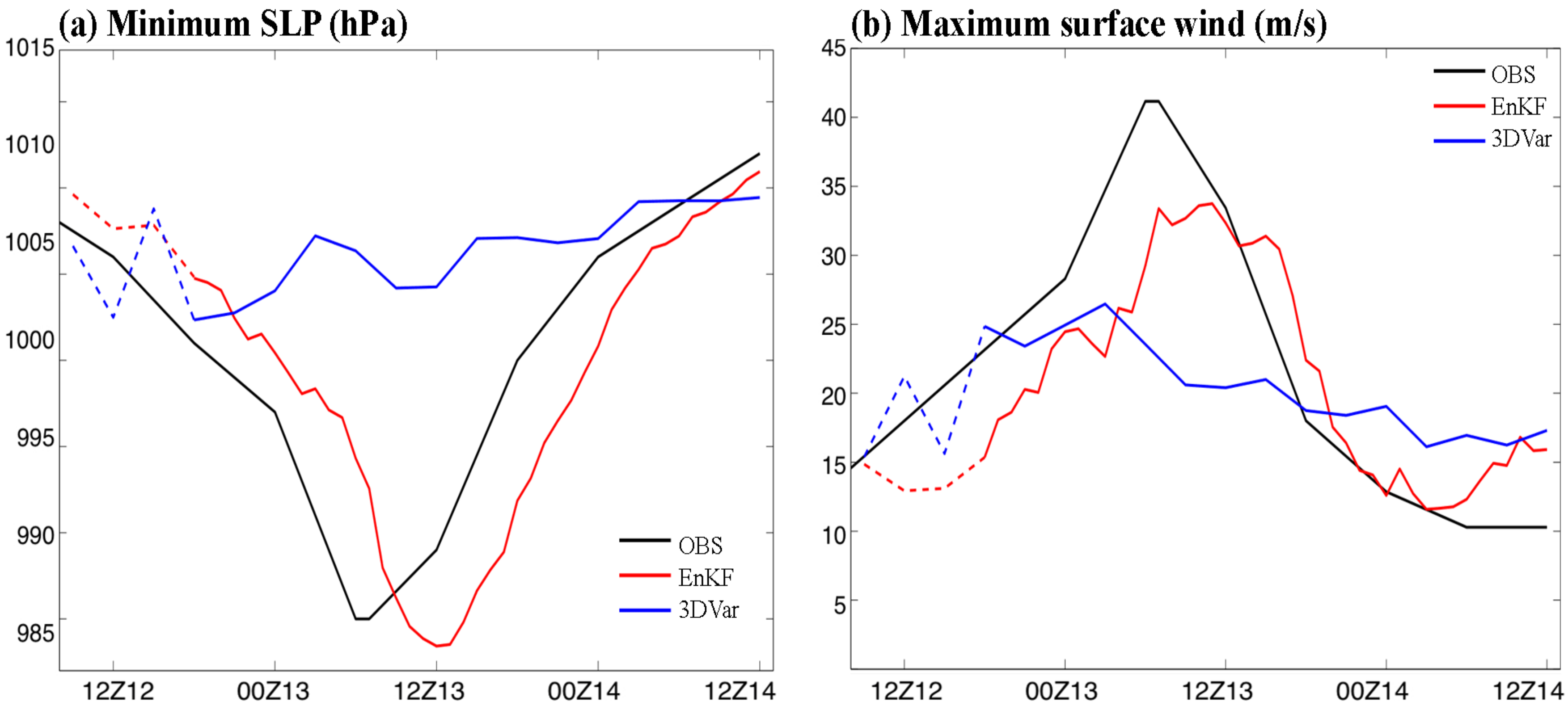
Hurricane Humberto 2007



(GFS ops run and WRF run from GFS)



WRF/EnKF Forecast vs. Observations vs. 3DVAR

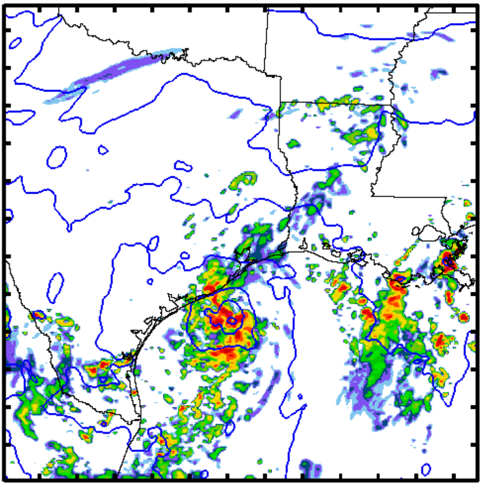


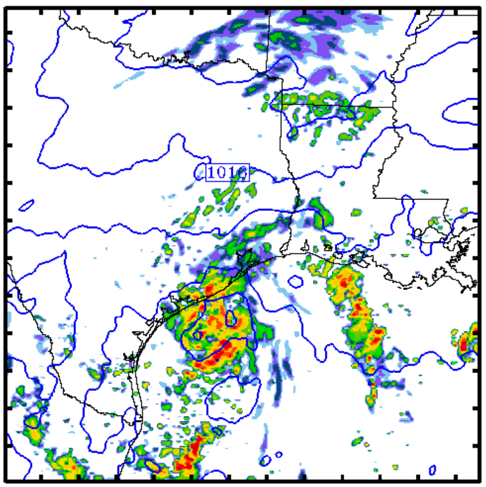
The WRF/3DVAR (as a surrogate of operational algorithm) assimilates the same radar data but without flow-dependent background error covariance, its forecast failed to develop the storm despite fit to the best-track observation better initially

(Zhang et al. 2009 MWR)

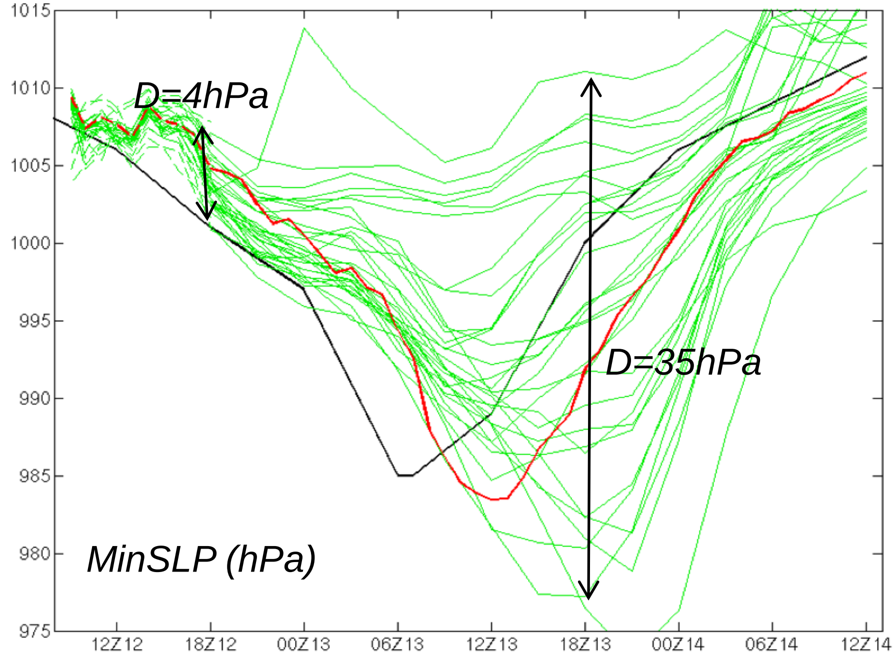
Cloud-resolving Ensemble Forecast and Predictability

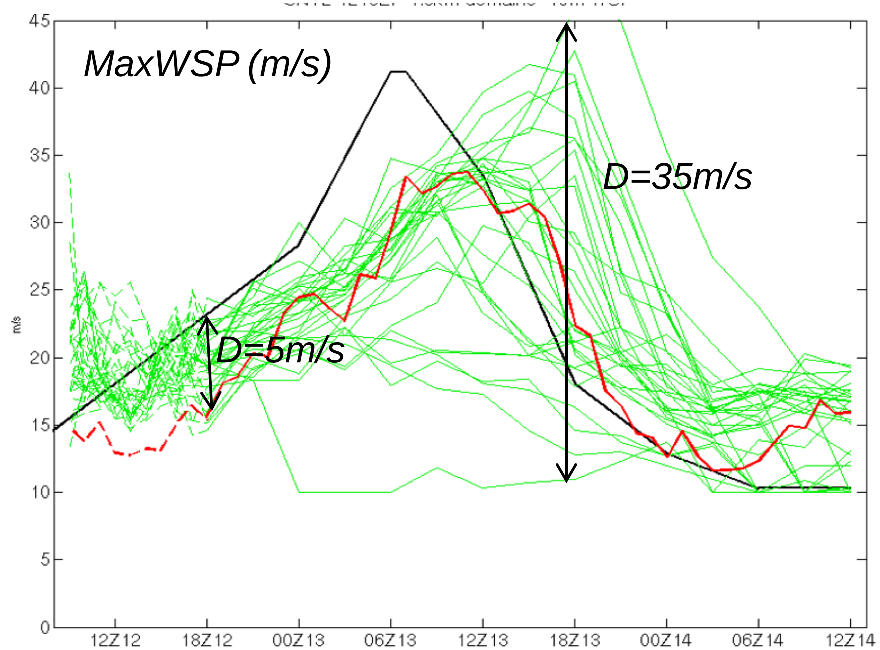
Strong member



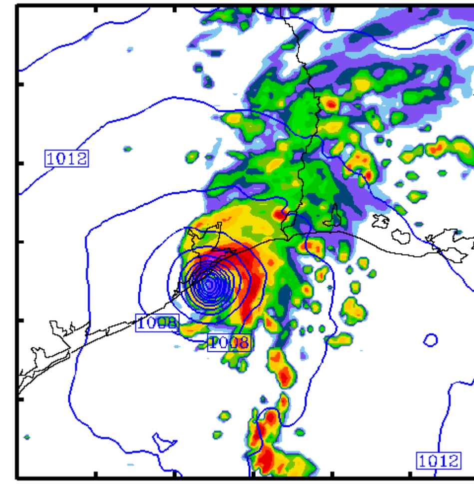


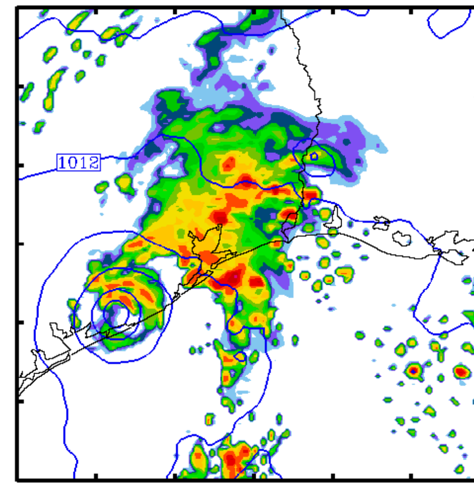
Weak member





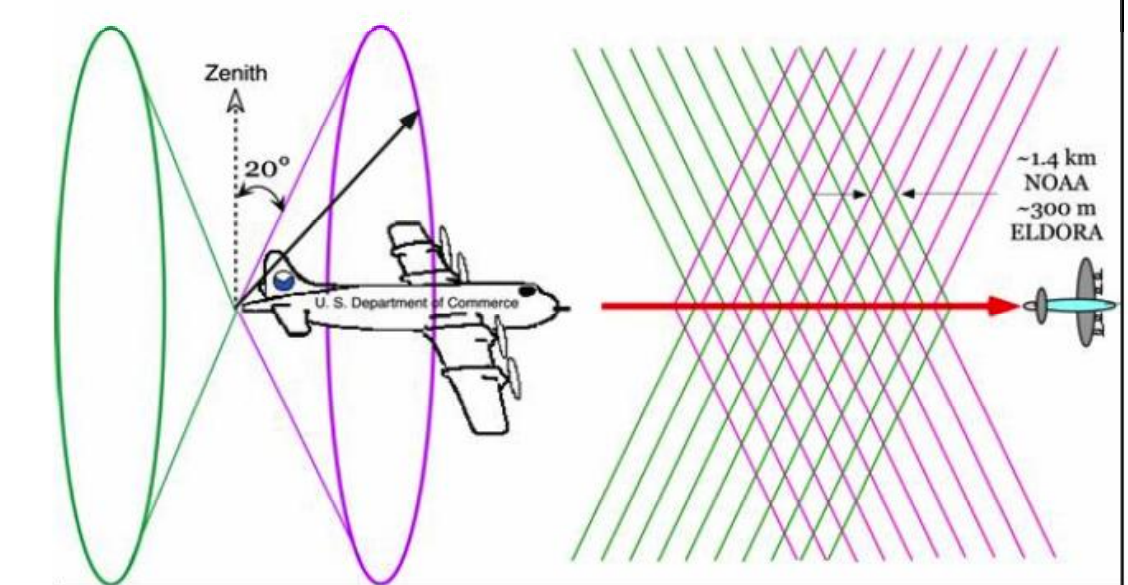
Strong member





Weak member

(Sippel and Zhang 2009 JAS)

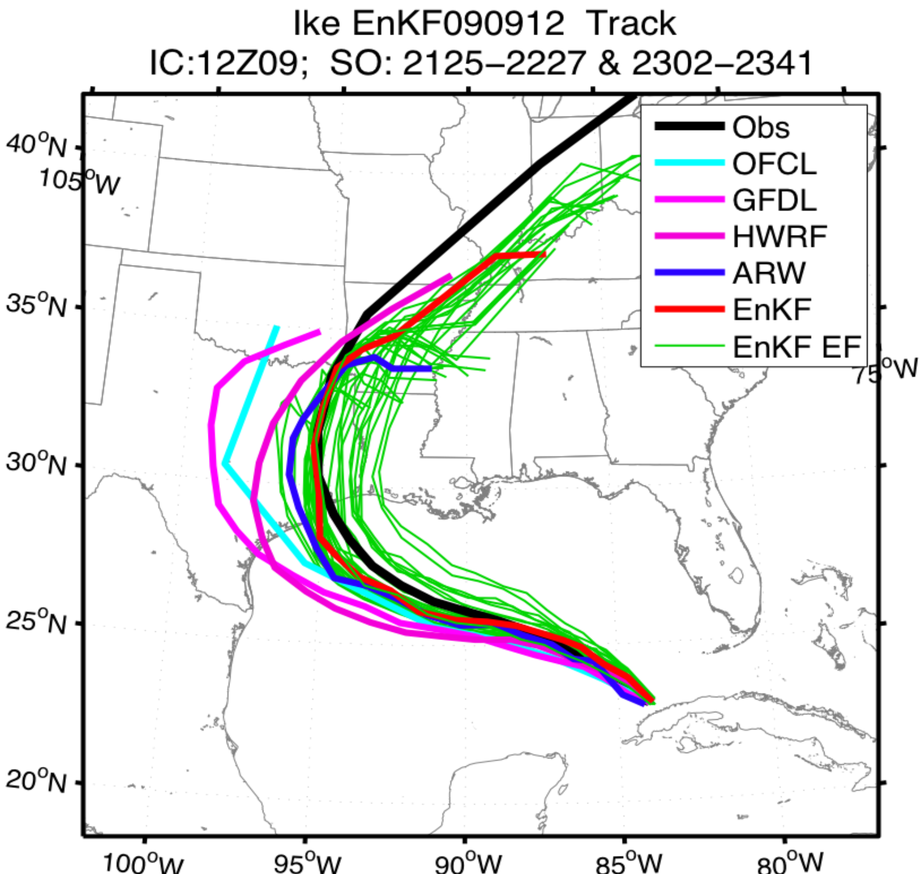


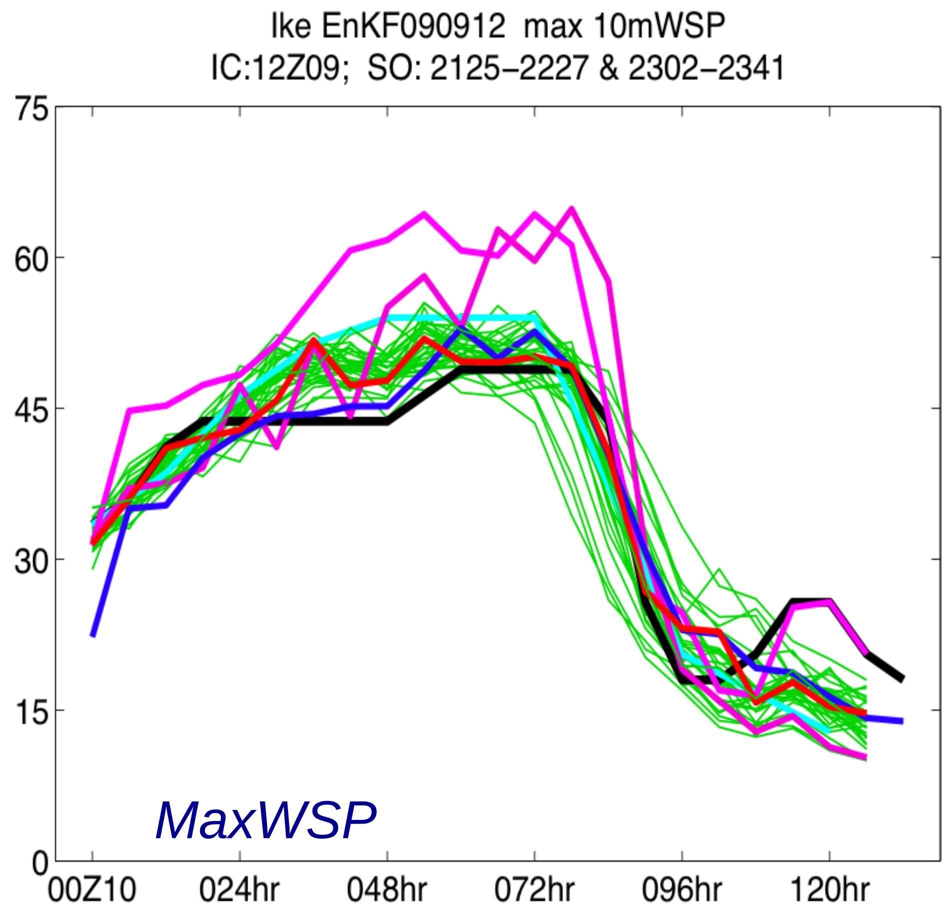
Hurricane Ike (2008)

103 deaths, $19.3 billion in estimated damage

Hurricane IKE (2008): Realtime WRF EnKF and ensemble forecast with assimilation of airborne Doppler winds

30,000 TACC linux cluster cores
30-member 5-d 1.5-km ensemble; 4 h after last observation time
30-member 5-d 1.5-km ensemble; 7 h after last observation time

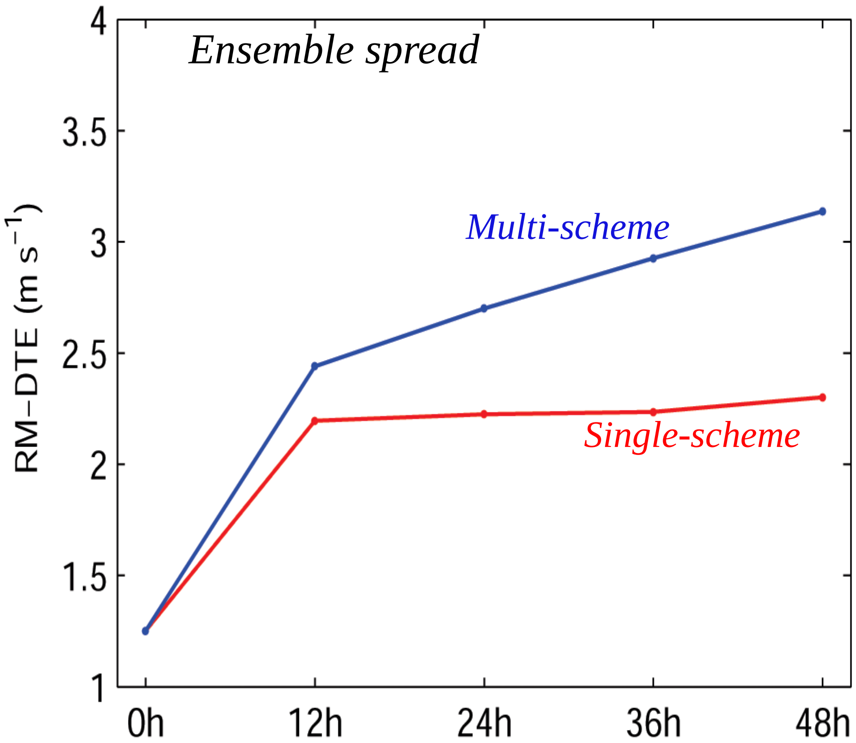


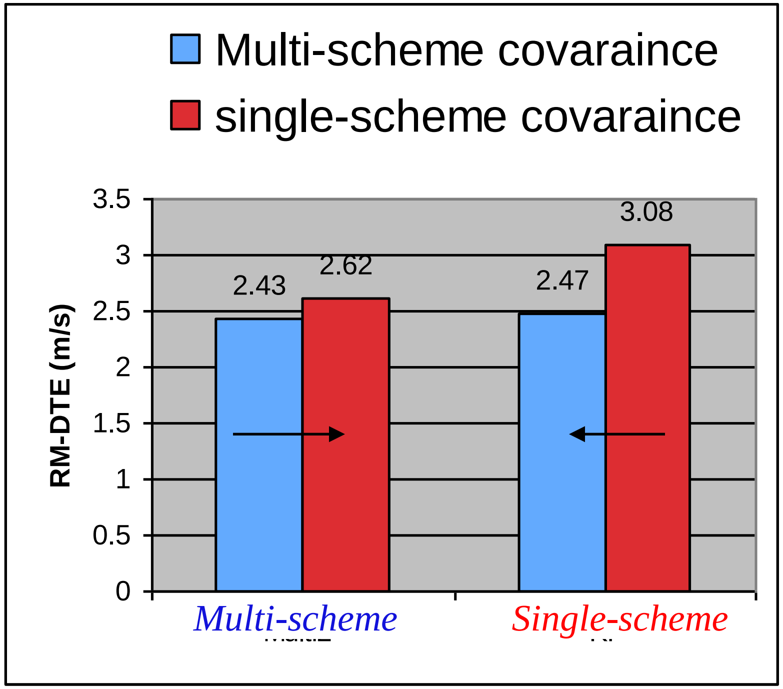


Concluding remarks

- *Ensemble forecast may coupled seamlessly with EnDA at the meso/convective scales to provides flow-dependent analysis and forecast uncertainty*
- *With advanced computing such as clusters at TACC, cloud-resolving ensemble analysis and forecasting for regional scales can be accomplished now, with apparent benefits in case studies*
- *Mesoscale ensemble forecast shares many of the same issues as mesoscale EnKF: model error, sampling, spinup, LBC, etc.*
- *Future directions:*
    - *Coupling EnKF with 4Dvar to improve analysis*
    - *Multi-model multi-physics to improve ensemble*
    - *Coupling with global ensemble analysis and forecasting*

Why multi-scheme ensemble: better mean or better covariance?





Exchange covariance between ...

Multi-scheme has a better prior forecast

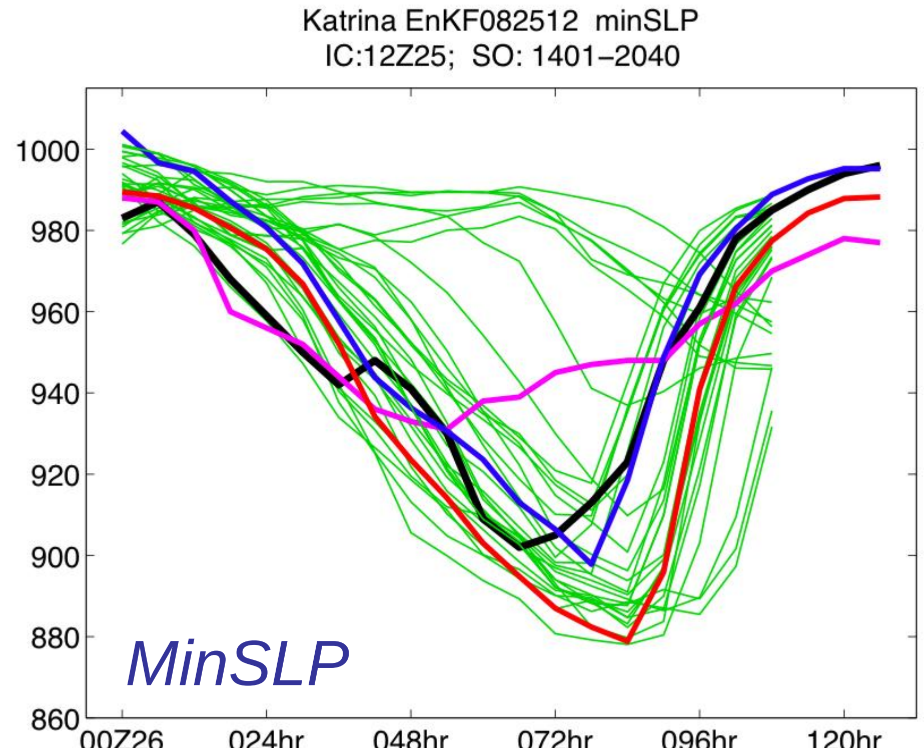
Multi-scheme is less vulnerable to filter divergence due to larger ensemble spread

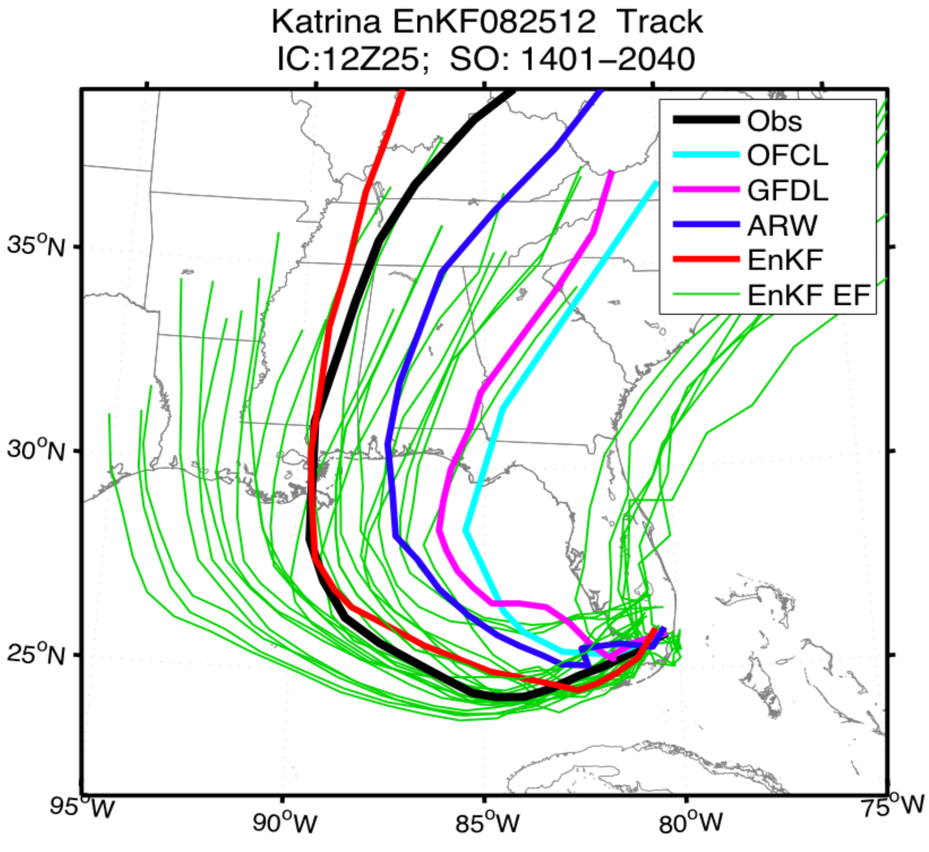
Multi-scheme has a better background error covariance structure

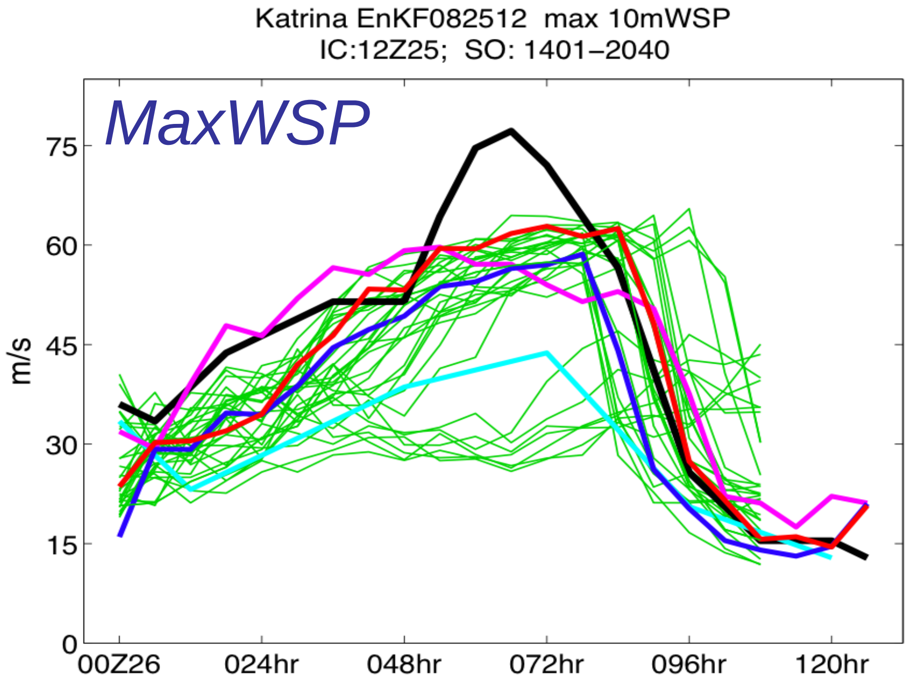
(Meng and Zhang 2007 MWR)

WRF/EnKF Performance (before Florida Landfall)

30-member ensemble forecast from EnKF posterior uncertainty







WRF/EnKF Performance (after Florida Landfall)