Hybrid Data Assimilation System: Integrating 3D-VAR with Ensembles (ETKF)

Meral Demirtas

WRF-Var Tutorial Presentation

NCAR, Boulder, Colorado

Acknowledgements with my special thanks:

Dale Barker, Xuguang Wang, Chris Snyder, Josh Hacker, and Yongsheng Chen

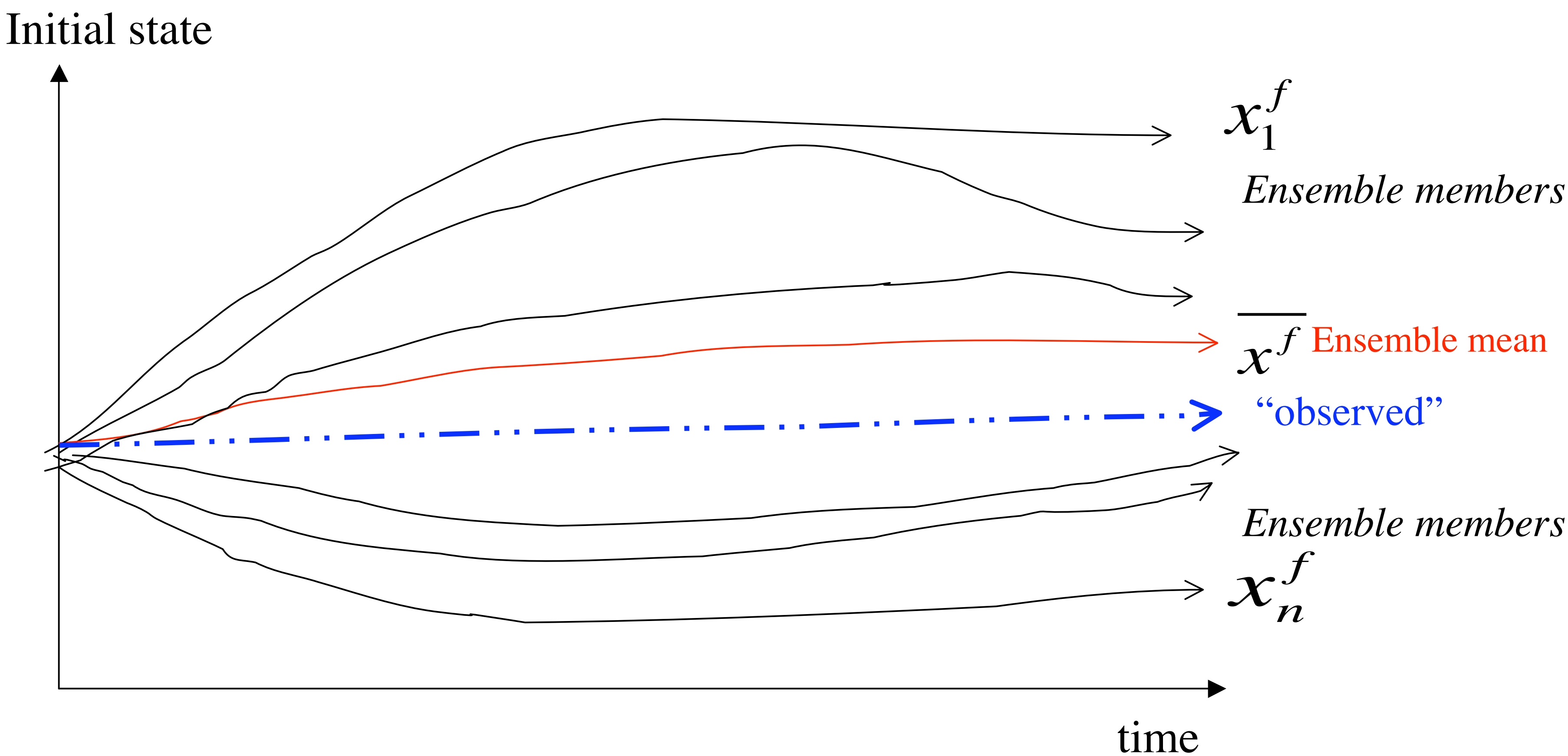
What is on the menu?

- Basic ingredients of a hybrid DA system
- Ensemble Transform Kalman Filter (ETKF)
- Hybrid system: Integrating ETKF into the 3D-VAR system
- Hybrid-ETKF implementation at Data Assimilation Testbed Center (DATC)
- Some examples from the latest NCAR applications
- the menu for the hybrid practice session

Basic ingredients of a hybrid system

1. WRF-Ensembles
2. Updating ensembles: Ensemble Transform Kalman Filter (ETKF)
3. Data assimilation system: 3D-VAR
4. Hybrid: Integrating ETKF into 3D-VAR system

Ensembles to address uncertainties in initial state



Ensembles in Formulas

Assume the following ensembles:

$$X^f = (x_1^f, x_2^f, x_3^f, ...x_N^f)$$

Ensemble mean: $\overline{x}^f = \frac{1}{N}\sum_{i=1}^{N} X_n^f$

Ensemble perturbations: $\delta X_n^f = x_n^f - \overline{x^f}$

Ensemble perturbations in vector form:

$$\delta X^f = (\delta x_1^f, \delta x_2^f, \delta x_3^f, ...,\delta_N^f)$$

Updating ensemble-based analyses

There are several approaches for ensemble-based data assimilation, we shall cover only Ensemble Transform Kalman Filter (ETKF) in this presentation. (For more details on DA ensemble techniques, I'd recommend Chris Snyder's talk at 9:10 on 4th Feb. 2009.)

ETKF technique produces ensemble members by re-scaling innovations with a transformation matrix. (Wang and Bishop 2003, Wang et. al. 2004, 2007.)

$$x^a = x^f \boldsymbol{T}$$

← ***T****ransformation matrix*
(solved by Kalman Filter Theory)

How does ETKF inflates ensemble analysis?

Need to adaptively inflate at time ***i*** by matching spread to innovation vectors. A scalar inflation factor has been introduced by Wang and Bishop 2003, $\prod$, :

$$\tilde{d}_i^T \tilde{d}_i = trace(\tilde{H} \alpha_i P_i^e \tilde{H}^T + \mathrm{I})$$

$$\tilde{d} = \left[y^o - \overline{H(x^f)} \right] / \sigma_o$$

$$\alpha_i = (\tilde{d}_i^T \tilde{d}_i - N) / (\sum_{i=1}^{k-1} \lambda_i)$$

$$\tilde{H} = H / \sigma_o$$

λ are the eigenvalues of $1/(N_k - 1)\tilde{H}P_i^e\tilde{H}$

$$\prod_i = \sqrt{\alpha_1 \alpha_2 \ldots \alpha_i} \qquad x_i = x_i^f T_i \prod_i$$

Pros and Cons of ETKF Technique

- Desirable aspects:
  - ETKF is fast (computations are done in model ensemble perturbation subspace).
  - It is suitable for generating ensemble initial conditions.
  - It updates initial condition perturbations.

- Less desirable aspects:
  - ETKF does not localize, therefore it does not represent sampling error efficiently. It may need very high inflation factors.

Hybrid: 3D-VAR and Ensembles Integrated

- Flow-dependent covariance through ensembles.
- Coupling wind, temperature and moisture fields.
- Hybrid can be more robust for small size ensembles and/or model errors (Wang et al. 2007, 2008a).
- It can be adapted to an existing 3D-VAR system.
- It is less expensive compared to other ensemble filters.

Advantages of a Hybrid System

- Background errors are flow-dependent:
  - 3D-Var: uses static background error covariances
  - Ensemble DA: computes flow-dependent covariances
  - Hybrid: flow-dependent information from ensemble perturbation fed into the WRF-Var system.

  Flow-dependent ensemble covariance has the largest impact over and downstream of where observation is sparse (Wang et al. 2008b).

The theory behind hybrid DA….

Ensemble covariance is implemented into the 3DVAR cost function via control variables:

$$J(x_1^{'},\alpha) = \beta_1 \frac{1}{2} x_1^{'T} B^{-1} x_1^{'} + \beta_2 \frac{1}{2} \alpha^T C^{-1} \alpha + \frac{1}{2} (y^{o'} - Hx^{'})^T R^{-1} (y^{o'} - Hx^{'})$$

$$x^{'} = x_1^{'} + \sum_{k=1}^{K} (\alpha_k \circ x_k^e)$$

(Wang et. al. 2007, 2008a)

C: correlation matrix for ensemble covariance localization

$x_1^{'}$ 3D-VAR increment

$x^{'}$ Total increment including hybrid

α Extended control variable

β_1 Weighting coefficient for static 3D-VAR covariance

β_2 Weighting coefficient for ensemble covariance

The theory part continued….

Conserving total variance requires: $$\frac{1}{\beta_1} + \frac{1}{\beta_2} = 1$$

Ensemble covariance horizontal localization is done through recursive filters. Since extended control variables are constrained by horizontal correlation matrix, **C**, only horizontal localization is utilized. Preconditioning designed as:

$$x'_1 = U_1 v_1 \qquad U_1 \approx B^{1/2}$$

(Wang et. al. 2007, 2008a)

$$\alpha = U_2 v_2 \qquad U_2 \approx C^{1/2}$$

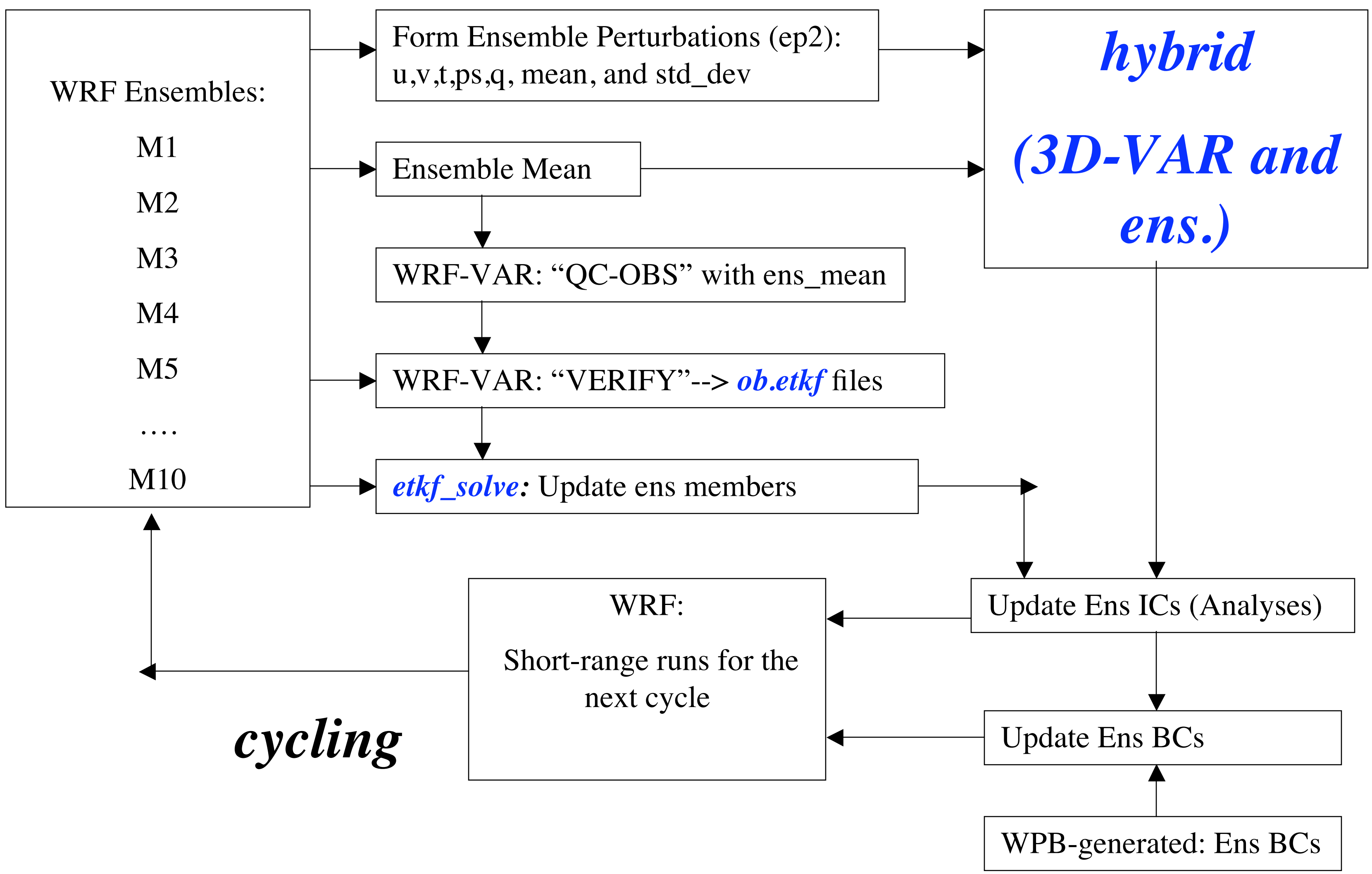
Some Hybrid Namelist Parameters

- alpha_corr_scale=1500km (default)
- je_factor (β_2)= 2.0
- jb_factor (β_1)= je_factor/(je_factor -1)
- alphacv_method =2 (ensemble perturbations on model space)
- ensdim_alpha= Ensemble size
- alpha_std_dev=1.0 (default)

Some Recent Work

- NCAR:
  - DATC applications (Meral Demirtas)
  - JME applications (Josh Hacker)
- The UK Met Office:
  - Global 4D-Var/Localized-ETKF (Dale Barker)
  - Adaptive localization within hybrid (with BOM, NRL)

WRF-VAR-ETKF Hybrid DA System Implementation at Data Assimilation Testbed Center (DATC)



DATC's t8_45km Domain Application

Ensemble size: 10

Test Period: 20070815-20070915

Cycle frequency: 3 hours

Ensemble analysis: ETKF technique

WRF-VAR technique: hybrid

WRF: Short-range (3hrs) runs for the next cycle

Observations: GTS

Initial and boundary conditions: GFS (via WPB)

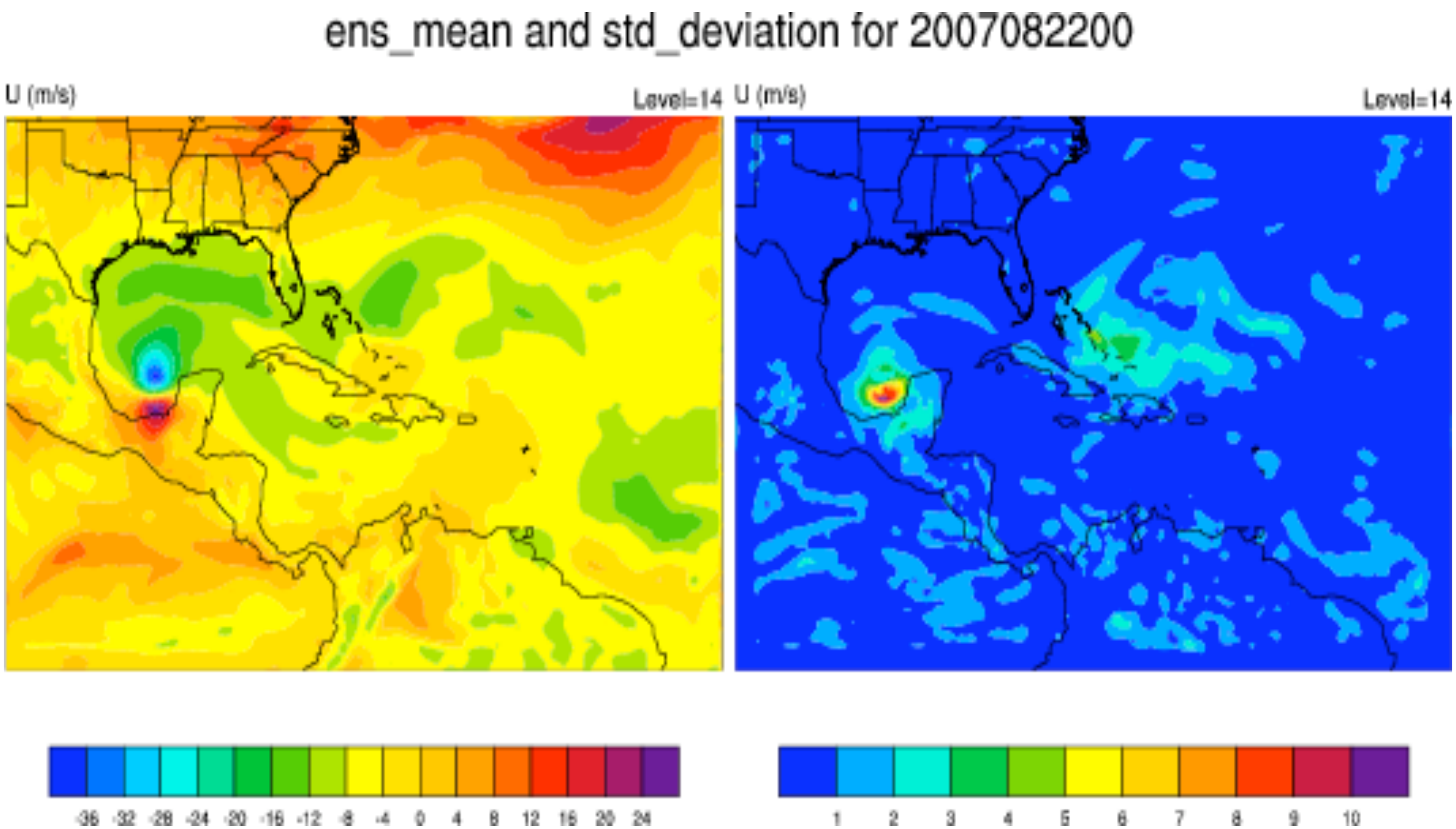
Horizontal resolution: 45km

Number of vertical levels: 57

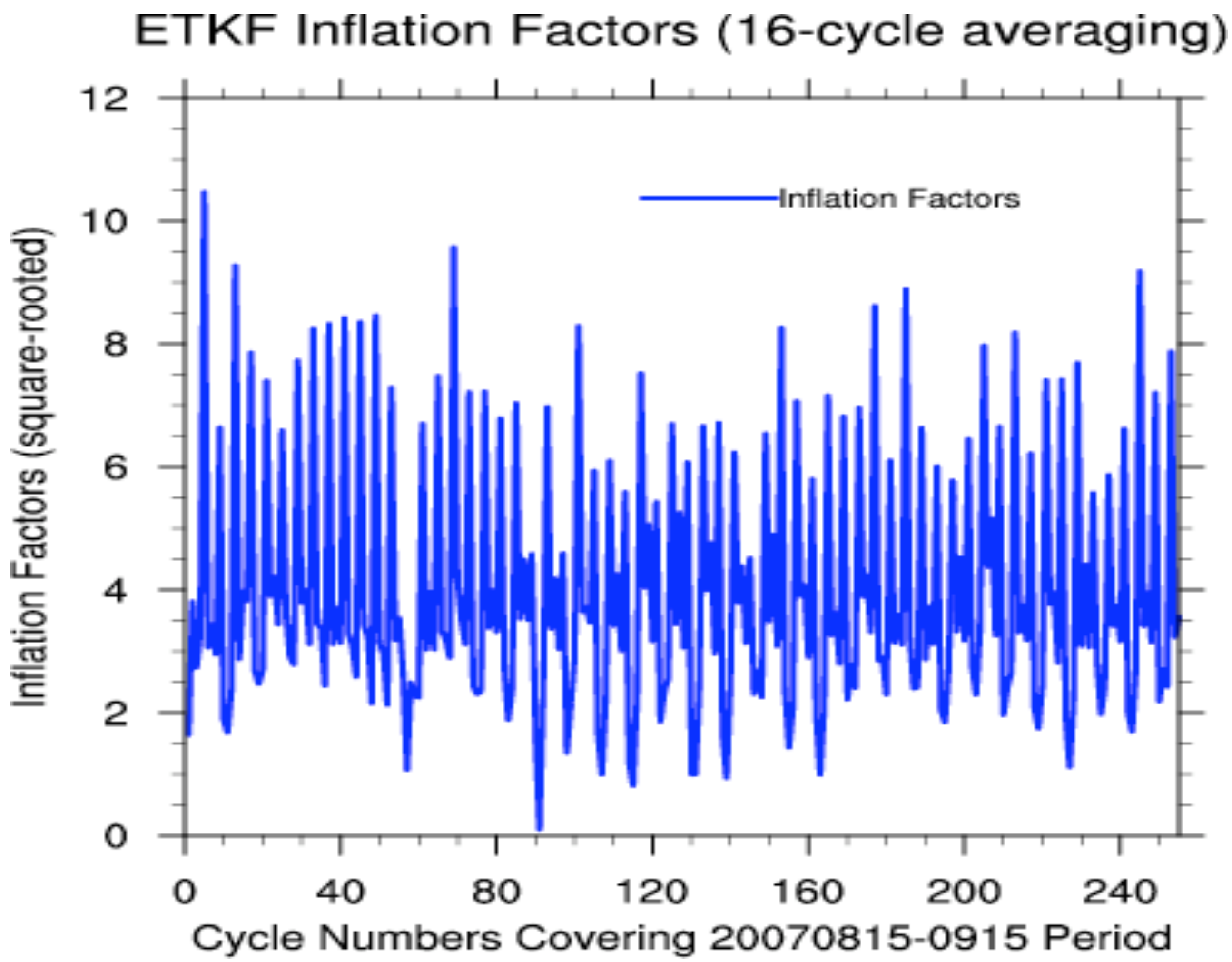
Preliminary results from DATC applications (snapshots)

Note that the work is still in progress…..

Ensemble Mean and Std. Deviation (spread)

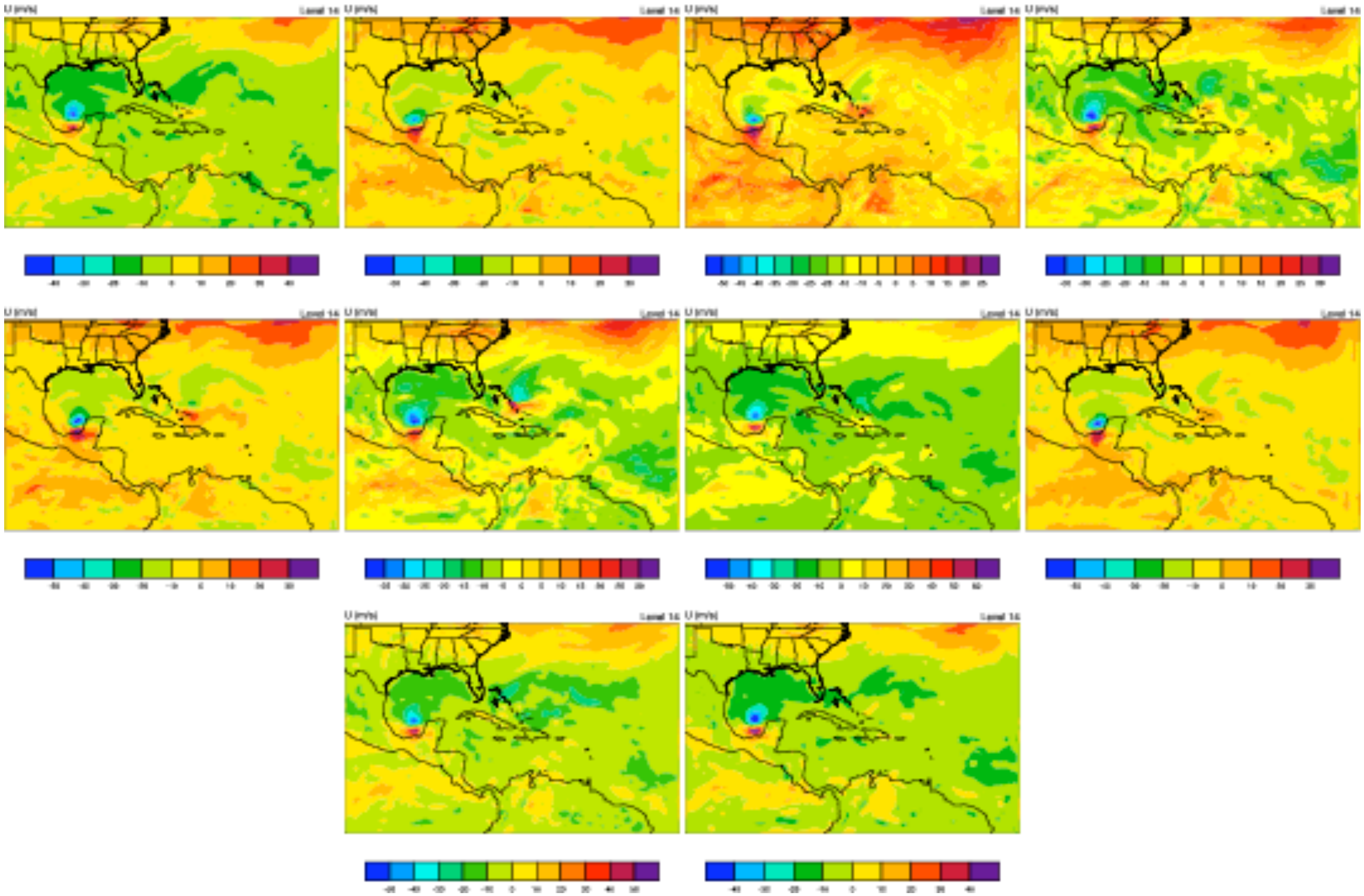


Inflation Factors (from 3-hourly cycling)

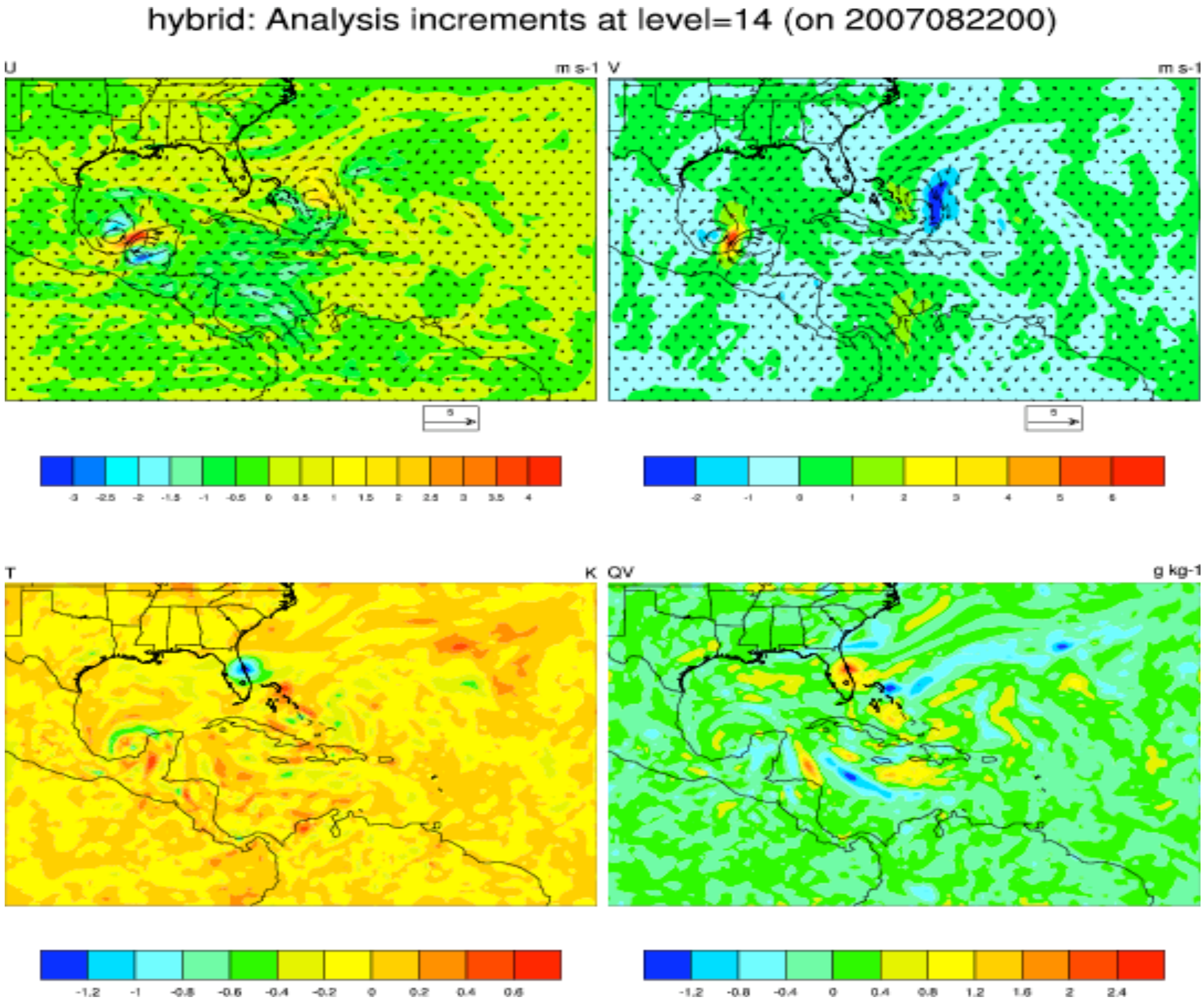


Stable, no-smoothing has been applied yet.

ETKF Output Images for U-wind 2007082200

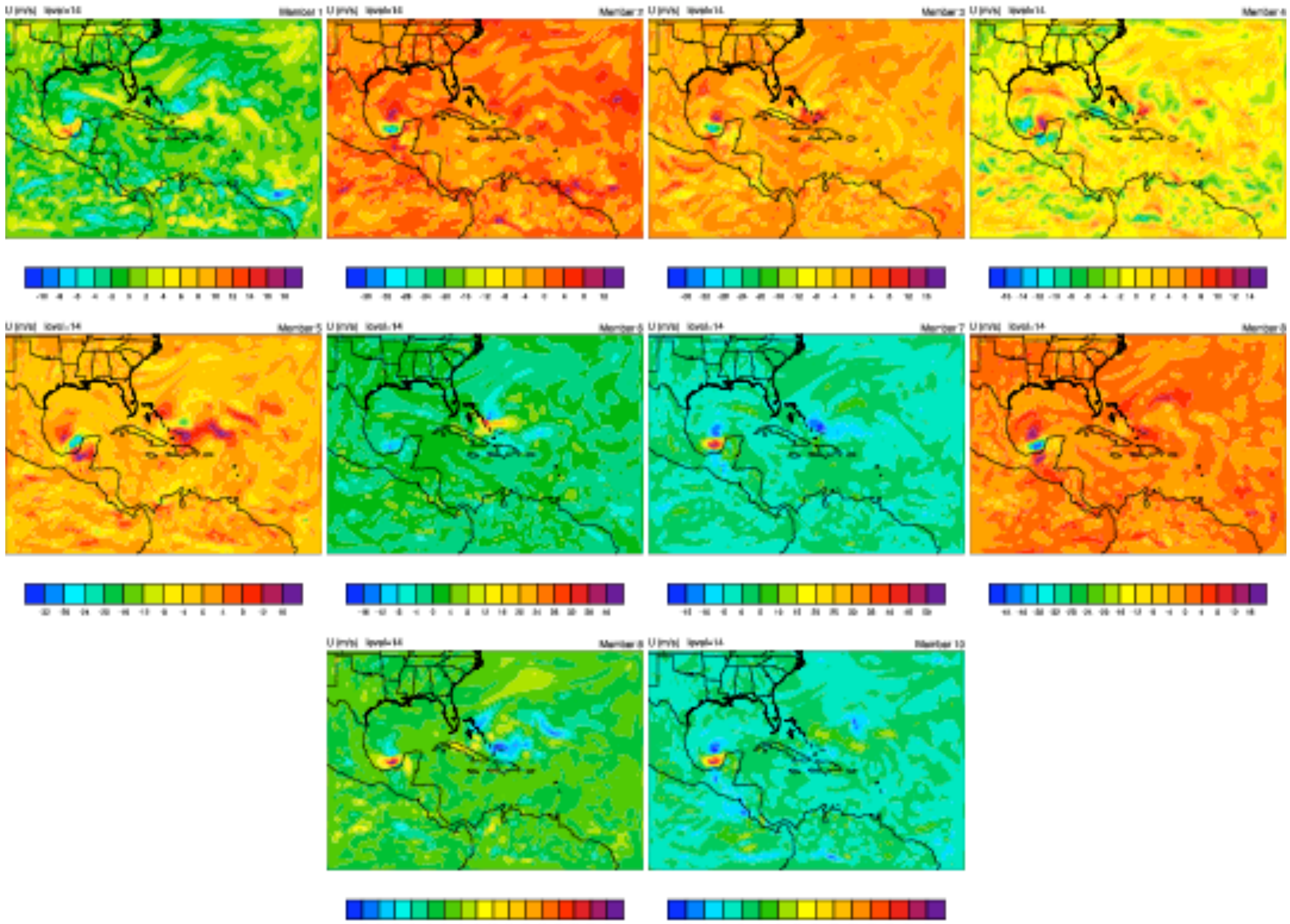


A snapshot of 10 ensemble members after ETKF procedure.

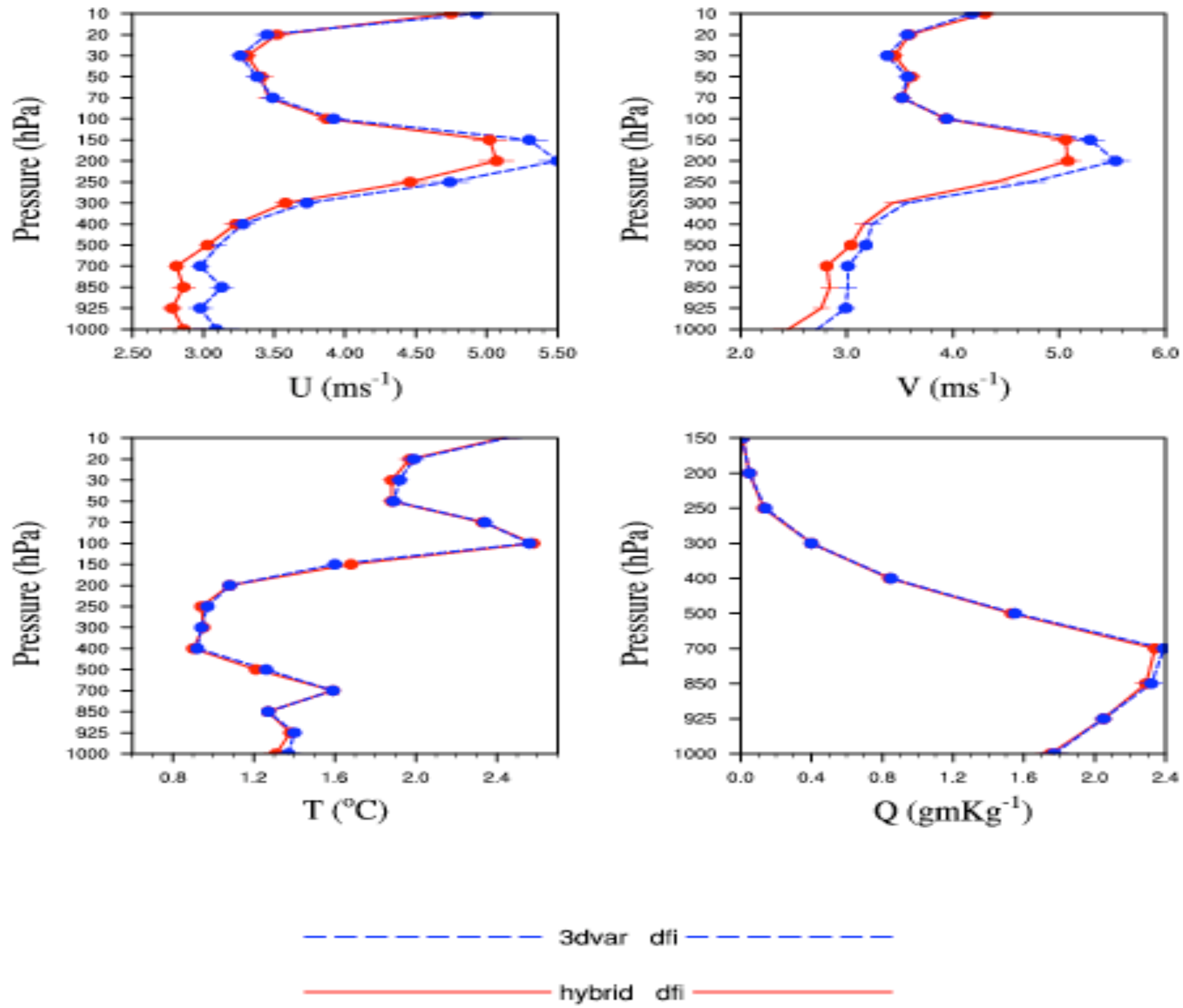


Flow-dependent increments from the hybrid system

Hybrid: Increments for Ensemble Members for 2007082200

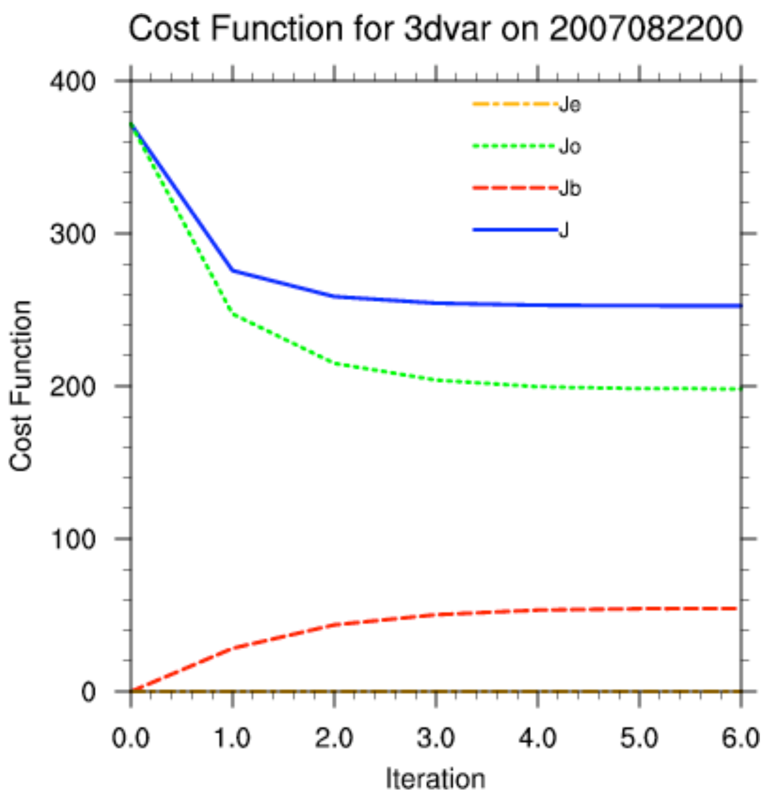


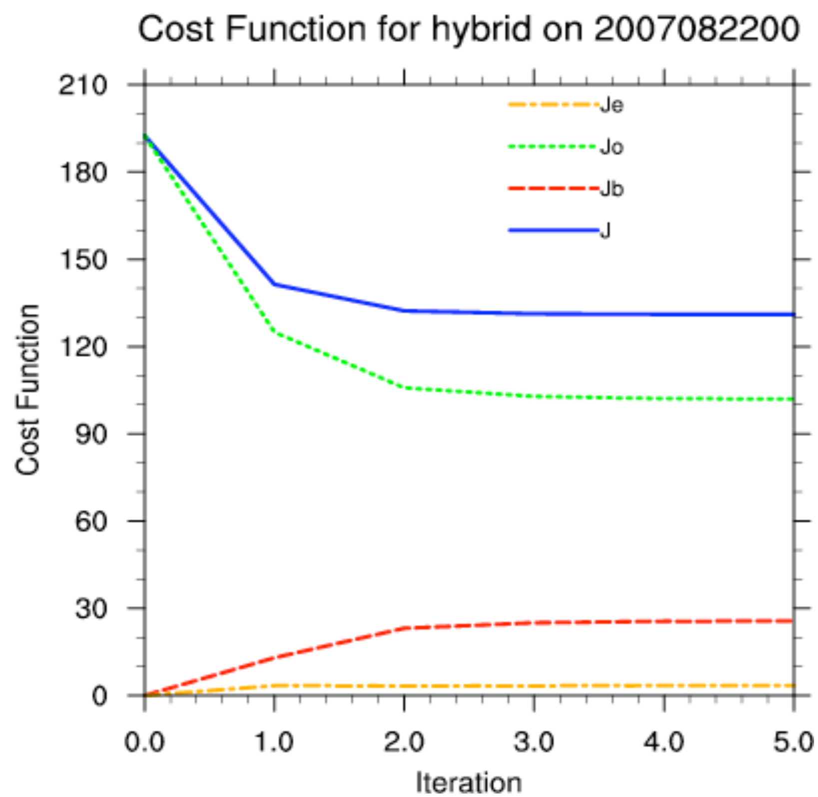
RMSE Profiles for t8_45km: 15th August-15th September 2007 (t+24)



Hybrid gives low RMSE scores for wind compared to 3D-Var.

Cost function: 3DVAR and Hybrid





Cost function is smaller in hybrid.

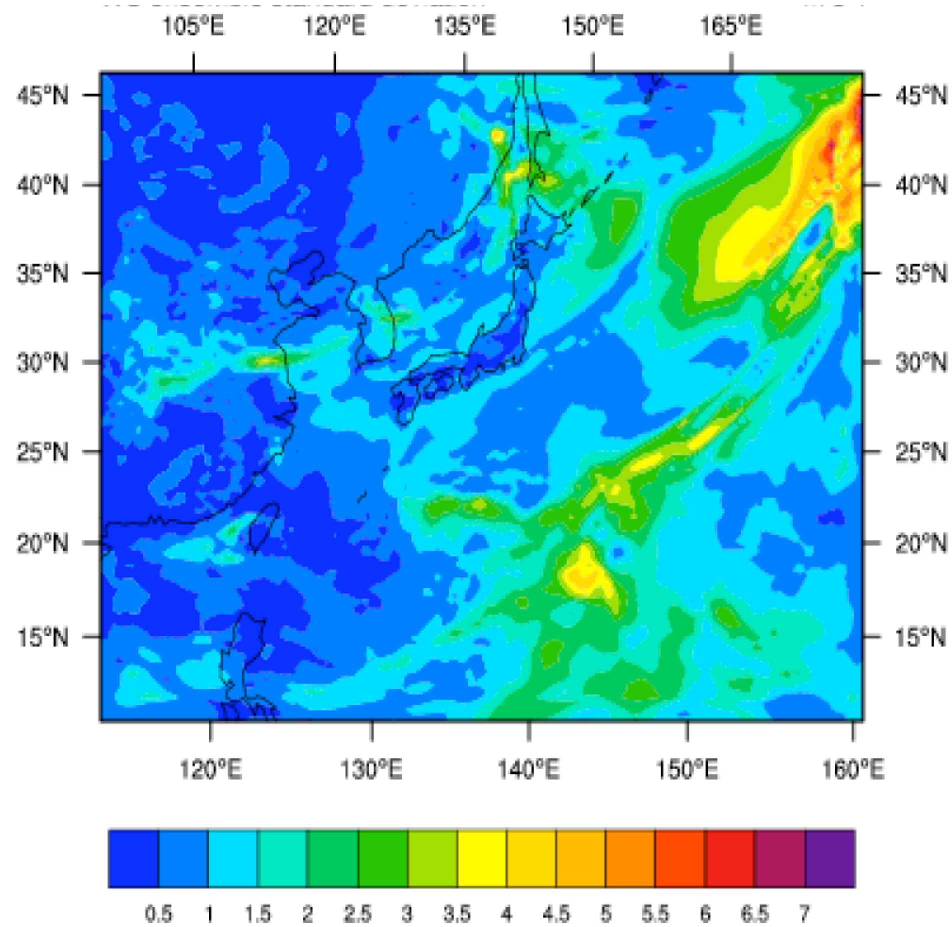
Preliminary results from JME applications (snapshots)

Note that the work is still in progress…..

Joint Mesoscale Ensemble (JME) Applications

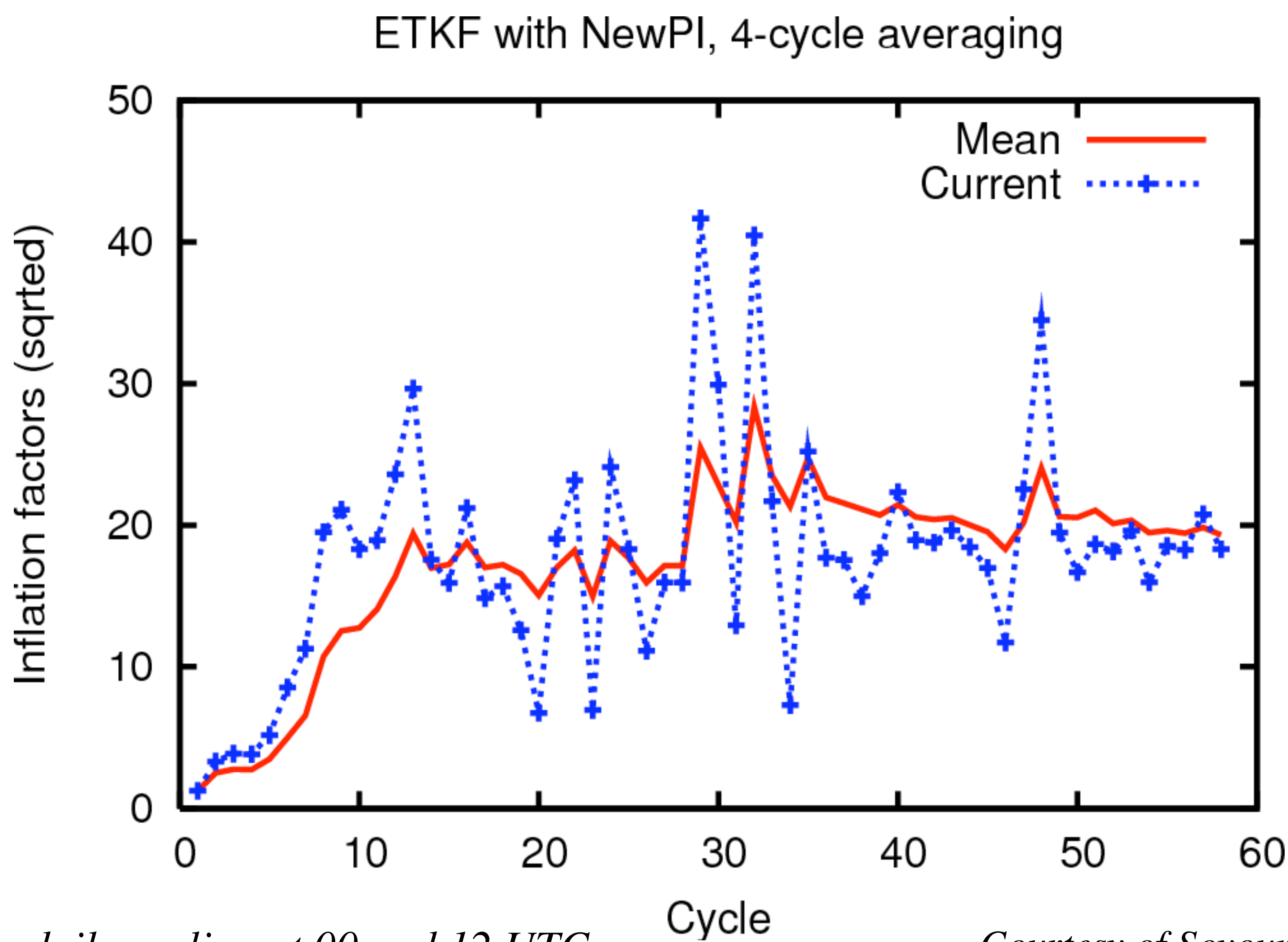
- 10 to 20 WRF physics configurations
- Capability for WRF-Var to update mean and/or individual members
- Capability for ETKF perturbations
- Lateral boundary conditions from global ensemble (GFS)
- Research on multi-parameter and stochastic approaches
- WRF-Var used to compute innovations.

10-m wind speed ensemble standard deviation (spread)



Courtesy of Josh Hacker

Inflation Factors Generated by JME



Twice daily cycling at 00 and 12 UTC

Courtesy of Soyoung Ha

Thanks for attending…..

What is on the menu for practice session

- **Computation:**
  - Computing ensemble mean.
  - Extracting ensemble perturbations (EP).
  - Running WRF-VAR for “hybrid”.
  - Displaying results for: ens_mean, std_dev, ensemble perturbations, hybrid increments, cost function and, etc.
  - If time permits, tailor your own test by changing hybrid settings; testing different values of “je_factor” and “alpha_corr_scale” parameters.

- **Scripts to use:**
  - Some NCL scripts to display results.

Brief information for the chosen case

Ensemble size: 10

Domain info:

- time_step=240,
- e_we=122,
- e_sn=110,
- e_vert=42,
- dx=45000,
- dy=45000,

Input data provided (courtesy of JME Group):

- WRF ensemble forecasts valid at 2006102800
- Observation data (ob.ascii) for 2006102800
- 3D-VAR “be.dat” file

Referred references

Wang, X., and C. H. Bishop, 2003: A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *J. Atmos. Sci.*, **60**, 1140-1158.

Wang, X., T. M. Hamill, J. S. Whitaker and C. H. Bishop, 2007: A comparison of hybrid ensemble transform Kalmanfilter-OI and ensemble square-root filter analysis schemes. *Mon. Wea. Rev.*, **135**, 1055-1076.

Wang, X., D. Barker, C. Snyder, T. M. Hamill, 2008a: A hybrid ETKF-3DVAR data assimilation scheme for the WRF model. Part I: observing system simulation experiment. *Mon. Wea. Rev.*, in press.

Wang, X., D. Barker, C. Snyder, T. M. Hamill, 2008b: A hybrid ETKF-3DVAR data assimilation scheme for the WRF model. Part II: real observation experiments. *Mon. Wea. Rev.*, in press.

Thanks for attending…..