En4D-Var: Combining Ensemble Forecast with 4D-Var and Experiments Using WRF

Qingnong Xiao[1], Chengsi Liu[1,2], and Bin Wang[2]

1. National Center for Atmospheric Research/MMM, Boulder, Colorado
2. Institute of Atmospheric Physics/LASG, Beijing, China

Outline

- Motivations
- En4D-Var scheme
- OSSEs with WRF En4D-Var
- Summary

Outline

- Motivations
- En4D-Var scheme
- OSSEs with WRF En4D-Var
- Summary

Advanced data assimilation

- 4D-Var
  - ✓ It is a non-sequential data assimilation technique, fitting observations in the whole assimilation window (optimal trajectory).
  - ✓ It is applied in many operational centers.
  - ✓ However, there are disadvantages compared with EnKF technique (TL and AD are difficult to code; background error covariance is evolved only within assimilation window and it is usually static at analysis time).
- Ensemble Kalman filter
  - ✓ It is a hot topic in recent years, and research shows promising results.
  - ✓ It is easy to design and code, and can include any physical process as needed.
  - ✓ One of the prominent advantages is its flow-dependent background error covariance.

Will EnKF replace 4D-Var in operational application?

- Although EnKF is promising in research, no evidence shows it can definitely outperform 4D-Var in operational. It has its own disadvantage, such as sampling errors.

- Variational data assimilation is well established in operational, it is difficult to be replaced, politically and technically.

How should we do?

- My view in the perspective of applications is
  - ✓ to include the flow-dependent background error covariance from ensemble forecast into 4D-Var, without significant change of the existing setup of operational 4D-Var system,
  - ✓ to use the ensemble perturbation matrix in the 4D-Var formulation and avoid tangent linear and adjoint model development in the 4D-Var setup.

Outline

- Motivations
- En4D-Var scheme
- OSSEs with WRF En4D-Var
- Summary

En3D-Var（Lorenc 2003）

$$J = \frac{1}{2}(x_a - x_b)^T B^{-1}(x_a - x_b) + \frac{1}{2}(\mathrm{H}(x_a) - x_y)^T O^{-1}(\mathrm{H}(x_a) - x_y)$$

3D-Var（incremental, precondition）

$$\delta\mathbf{x} = \mathbf{x}_a - \mathbf{x}_b \qquad \mathbf{B} = \mathbf{U}\mathbf{U}^T \qquad \delta\mathbf{x} = \mathbf{U}\mathbf{w}$$

$$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}(\mathbf{H}\mathbf{U}\mathbf{w} + \mathbf{d})^T \mathbf{O}^{-1}(\mathbf{H}\mathbf{U}\mathbf{w} + \mathbf{d})$$

$$\nabla_{\mathbf{w}} J = \mathbf{w} + \mathbf{U}^T\mathbf{H}^T\mathbf{O}^{-1}(\mathbf{H}\mathbf{U}\mathbf{w} + \mathbf{d})$$

$$w_a^N = w_a^{N-1} - \rho d$$

$\nabla J \neq 0$

$\nabla J = 0$

$$x_a = x_b + U w_a$$

En3D-Var

$$\delta\mathbf{x} = \mathbf{x}_a - \mathbf{x}_b \qquad \mathbf{B} = X_b' X_b'^T \qquad \delta\mathbf{x} = X_b' \mathbf{w}'$$

$$J(\mathbf{w}') = \frac{1}{2}\mathbf{w}'^T\mathbf{w}' + \frac{1}{2}(\mathbf{H}X_b'\mathbf{w}' + \mathbf{d})^T \mathbf{O}^{-1}(\mathbf{H}X_b'\mathbf{w}' + \mathbf{d})$$

$$\nabla_{\mathbf{w}'} J = \mathbf{w}' + X_b'^T\mathbf{H}^T\mathbf{O}^{-1}(\mathbf{H}X_b'\mathbf{w}' + \mathbf{d})$$

$$w_a^N = w_a^{N-1} - \rho d$$

$\nabla J \neq 0$

$\nabla J = 0$

$$x_a = x_b + X_b' w_a'$$

En4D-Var

$$J = \frac{1}{2}(x_a - x_b)^T B^{-1}(x_a - x_b) + \sum_{i=0}^{N} \frac{1}{2}(HM_{0\sim i}(x_a) - y_i)^T O^{-1}(HM_{0\sim i}(x_a) - y_i)$$

En4D-Var (generalized En3D-Var)

$$\mathbf{M}X_b' \approx \frac{1}{\sqrt{N-1}}(MX_{bi} - \overline{MX_b})$$

$$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}\sum_{i=0}^{I}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w} + \mathbf{d}_i)^T\mathbf{O}^{-1}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w} + \mathbf{d}_i)$$

$$\nabla_{\mathbf{w}} J = \mathbf{w} + \sum_{i=0}^{I} X_b'^T \mathbf{M}^T_{i\sim 0}\mathbf{H}^T\mathbf{O}^{-1}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w} + \mathbf{d}_i)$$

$$w_a^N = w_a^{N-1} - \rho d$$

$\nabla J \neq 0$

$\nabla J = 0$

$$x_a = x_b + X_b' w_a'$$

En4D-Var (Opt.2)

$$\mathbf{HM}X_b' \approx \frac{1}{\sqrt{N-1}}(HMX_{bi} - \overline{HMX_b})$$

$$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}\sum_{i=0}^{I}(\mathbf{HM}X_b'\mathbf{w} + \mathbf{d}_i)^T\mathbf{O}^{-1}(\mathbf{HM}X_b'\mathbf{w} + \mathbf{d}_i)$$

$$\nabla_{\mathbf{w}} J = \mathbf{w} + \sum_{i=0}^{N}(\mathbf{HM}_{0\sim i}X_b')^T\mathbf{O}^{-1}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w} + \mathbf{d}_i)$$

$$w_a^N = w_a^{N-1} - \rho d$$

$\nabla J \neq 0$

$\nabla J = 0$

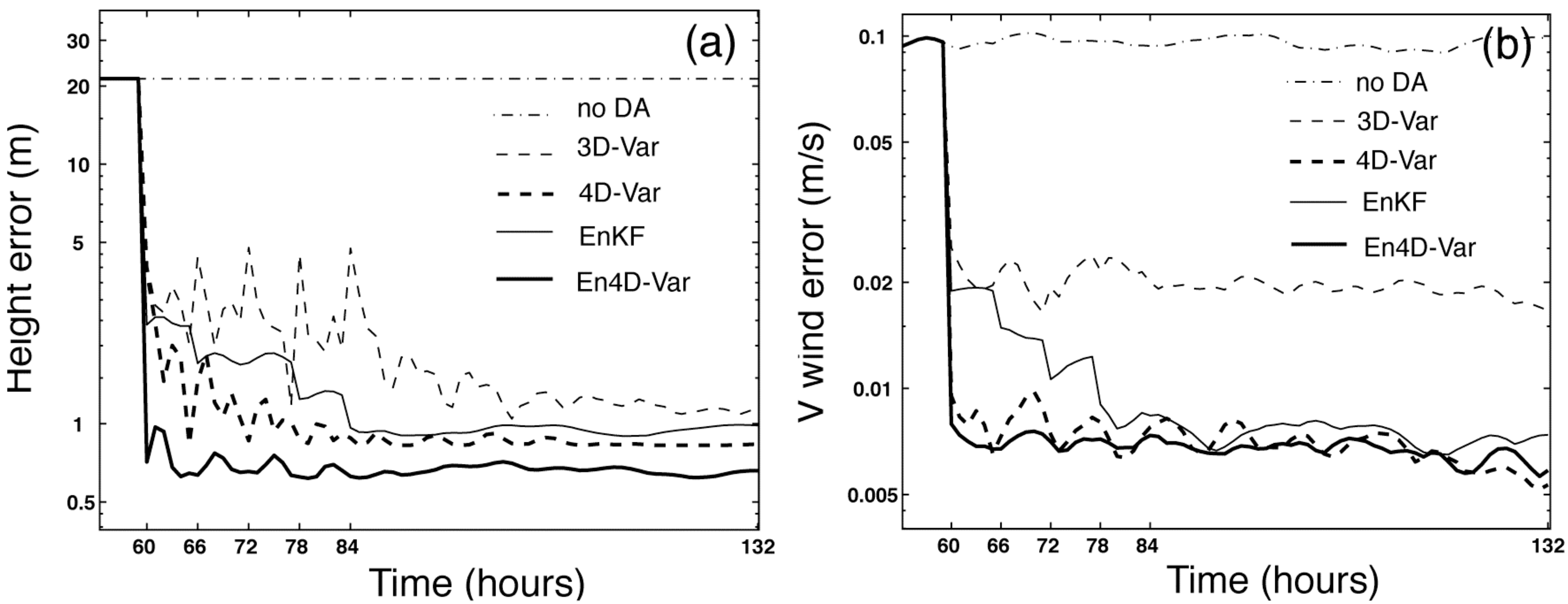
$$x_a = x_b + X_b' w_a'$$

adjoint Tangent linear

Some characteristics of En4D-Var

- En4D-Var uses the flow-dependent B matrix from ensemble forecast.
- It avoids tangent linear and adjoint models in its formulation (in Opt.2).
- It couples incremental approach with preconditioning using ensemble perturbation matrix.
- But sampling errors are introduced to En4D-Var (in Opt.2).

Proof-of-concept test with shallow water model



Evolution of domain-average RMSE

Outline

- Motivations
- En4D-Var scheme
- OSSEs with WRF En4D-Var
- Summary

WRF En4D-Var

- The success of En4D-Var with simple models gives us great motivations to implement the technique using WRF model.
- The biggest challenge for En4D-Var in real atmospheric model (e.g. WRF) is how to deal with sampling errors.

Localization in ensemble-based data assimilation

- Why
  - Imperfect ensemble => sampling errors => analysis increment noise
  - Ensemble dimension is far less than model dimension =>
    B matrix rank is restricted to the low-dimension sub-space =>
    deficient rank and underdetermined problem
- How
  - local truncation (Houtekamer and Mitchell 1998)
  - hybrid scheme (Hamill and Snyder 2001, Lorenc 2003)
  - Schur product (Houtekamer and Mitchell 2001, Lorenc 2003, Buehner 2005)

WRF En4D-Var

- We conduct horizontal and vertical localizations using Schur operator to deal with spatial sampling errors, similar to the method in EnKF localizations.
- We empirically put the analysis time at the mid of assimilation window to alleviate the temporal sampling errors.

Horizontal and vertical localization

- EOF decomposed correlation function operator

$$P' = [E_v \lambda_v^{1/2} \cdot (E_{h1} \lambda_{h1}^{1/2} \cdot X'_{b1}, \ldots, E_{h1} \lambda_{h1}^{1/2} \cdot X'_{bN}), \ldots, E_v \lambda_v^{1/2} \cdot (E_{hn} \lambda_{hn}^{1/2} \cdot X'_{b1}, \ldots, E_{hn} \lambda_{hn}^{1/2} \cdot X'_{bN})]$$

Analysis time tuning

- Why analysis time tuning

Assimilation window y_i

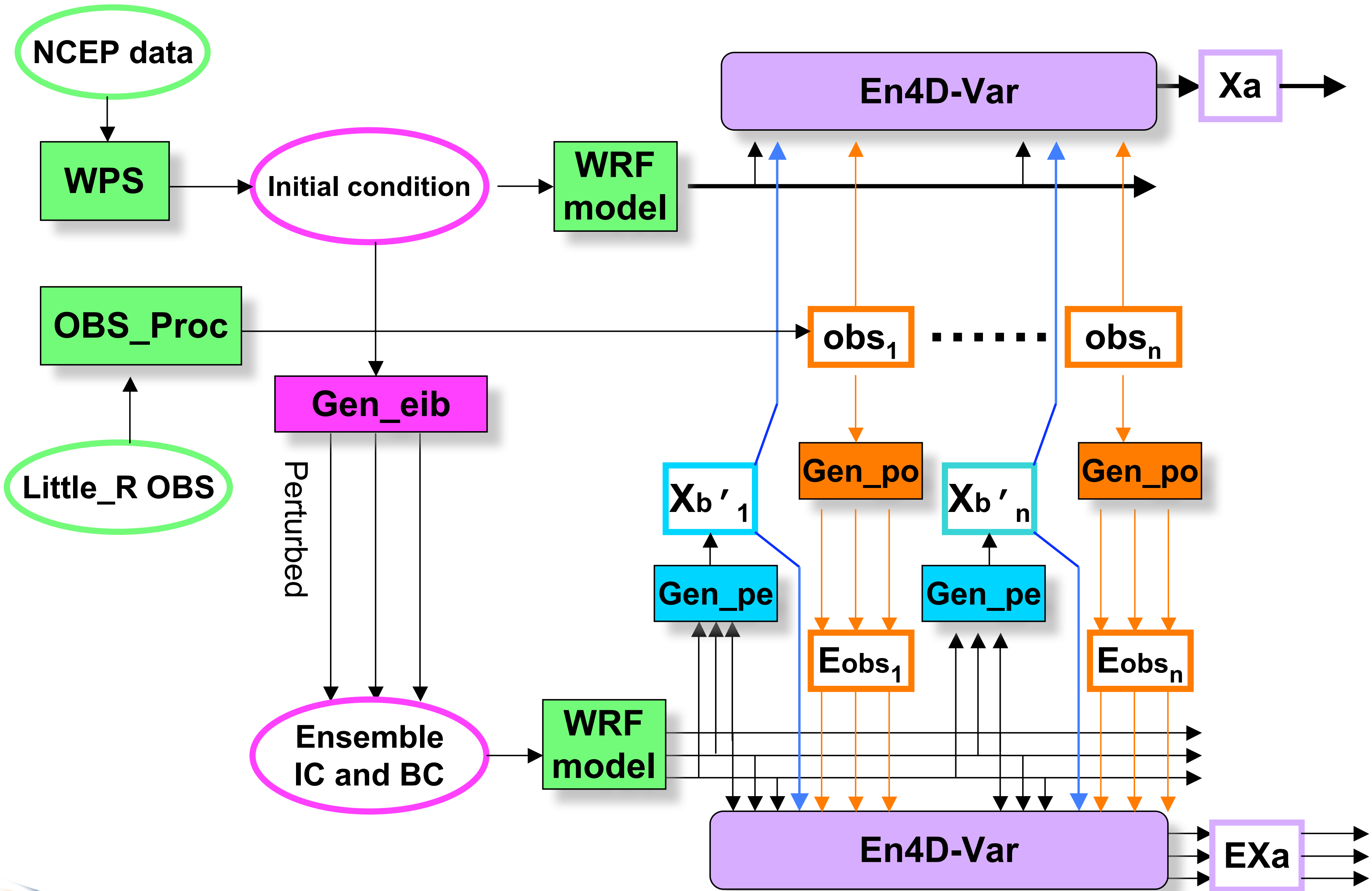
X_{a0}

$$X_{a0} = X_{b0} + X_{b0}'(\mathrm{H}X_{bi}')^{\mathrm{T}}(\mathrm{H}B_i\mathrm{H}^{\mathrm{T}} + O_i)^{-1}(y_i - HX_{bi})$$

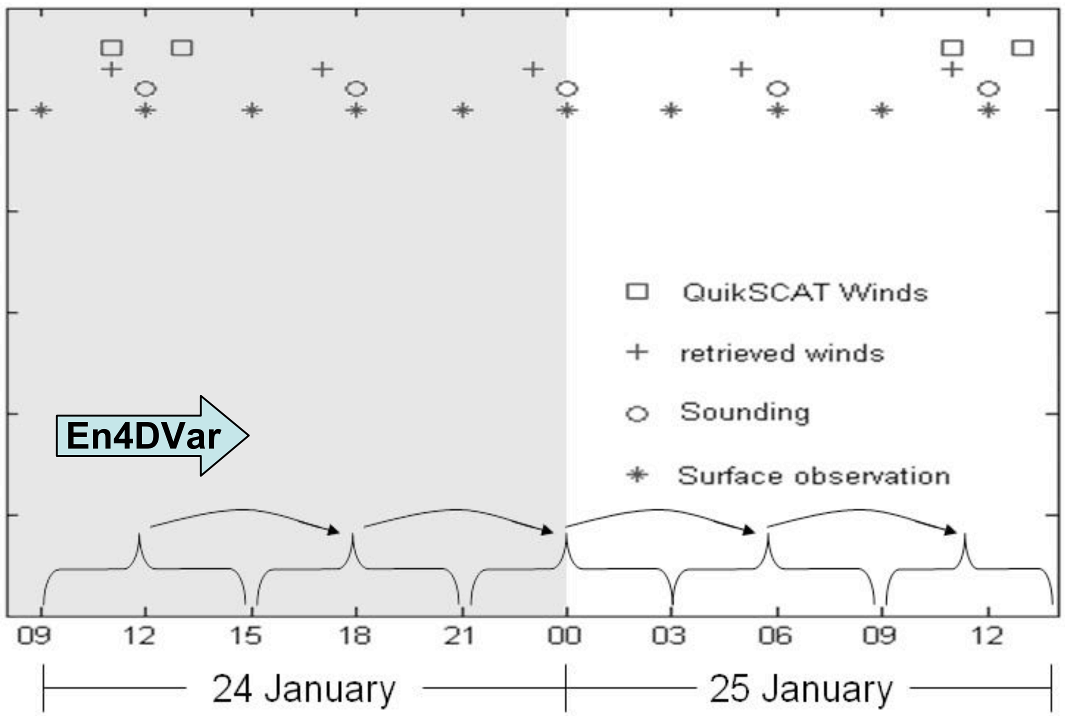
With perfect ensemble, if y_i is far enough from analysis time, $X_{b0}'(\mathrm{H}X_{bi}')^{\mathrm{T}}$ is close to zero.

Due to imperfect ensemble, $X_{b0}'(\mathrm{H}X_{bi}')^{\mathrm{T}}$ contains noise so that the analysis is contaminated by sampling errors.

Flow Chart for WRF-En4DVar



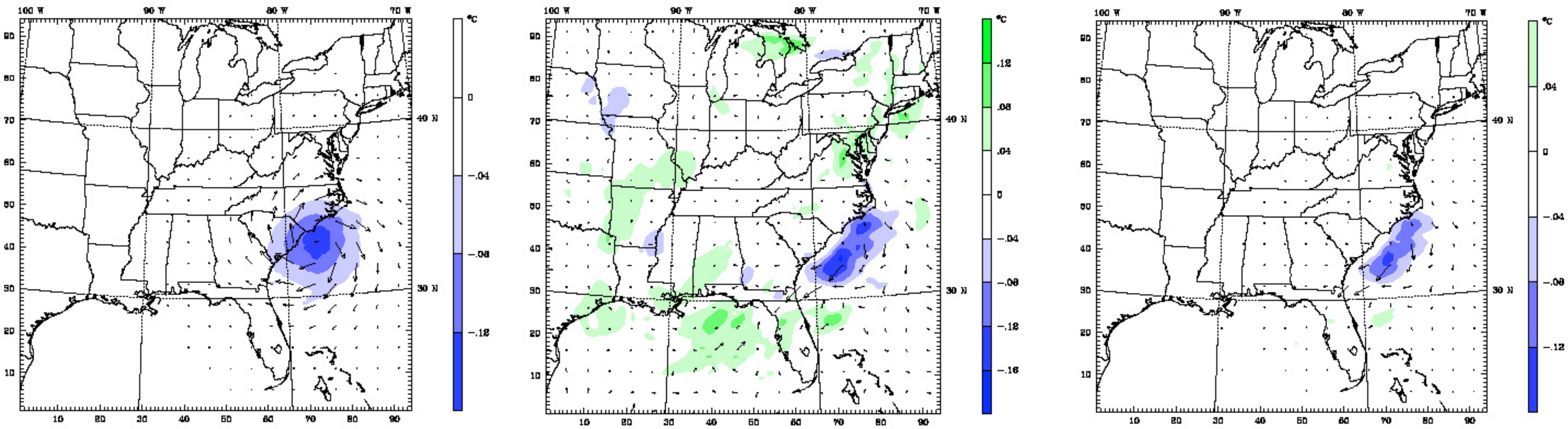
En4D-Var OSSE design



- Test with the “blizzard of 2000” case: 24-25 January 2000
- Assimilation window: 6 hours
- Cycling: From 0900 UCT 24 to 1500 UTC 25 Januare 2000
- Observations are simulated with real positions

Single observation test

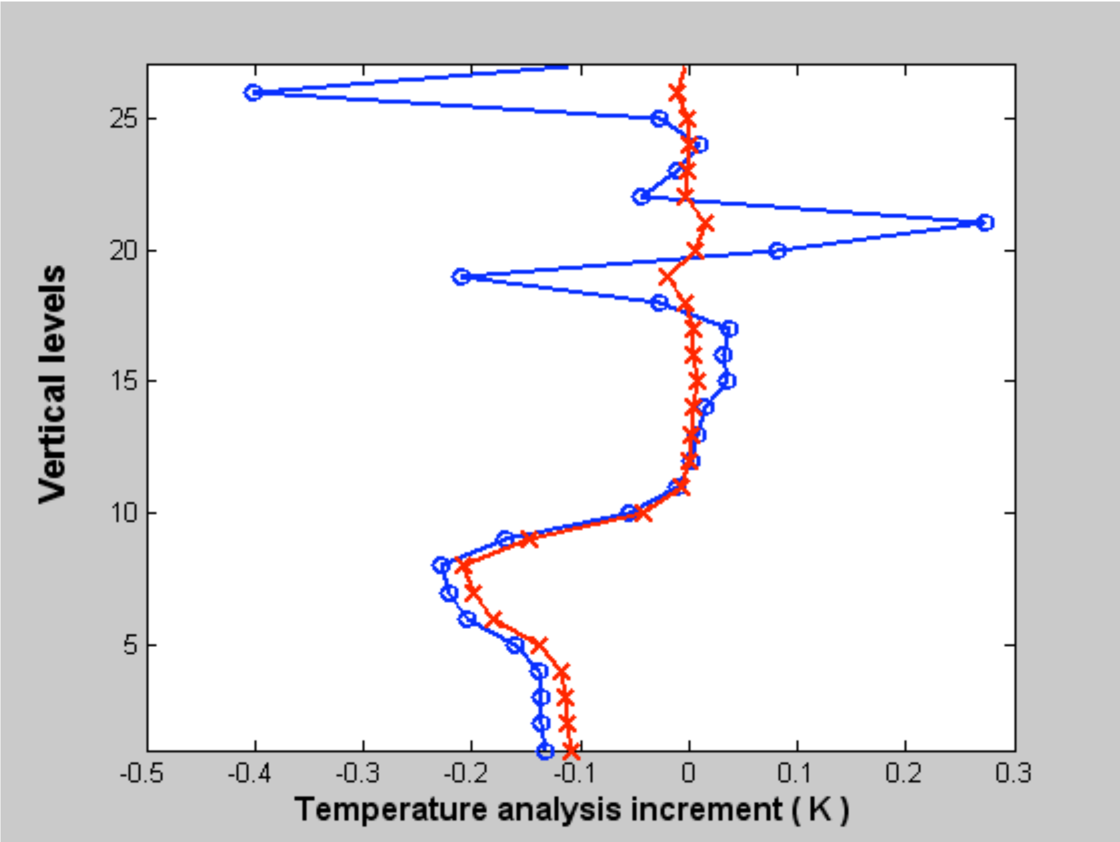
(single T observation at 850hpa at 24-12Z Jan.)



WRF-Var **En4D-Var without localization** **En4D-Var with localization**

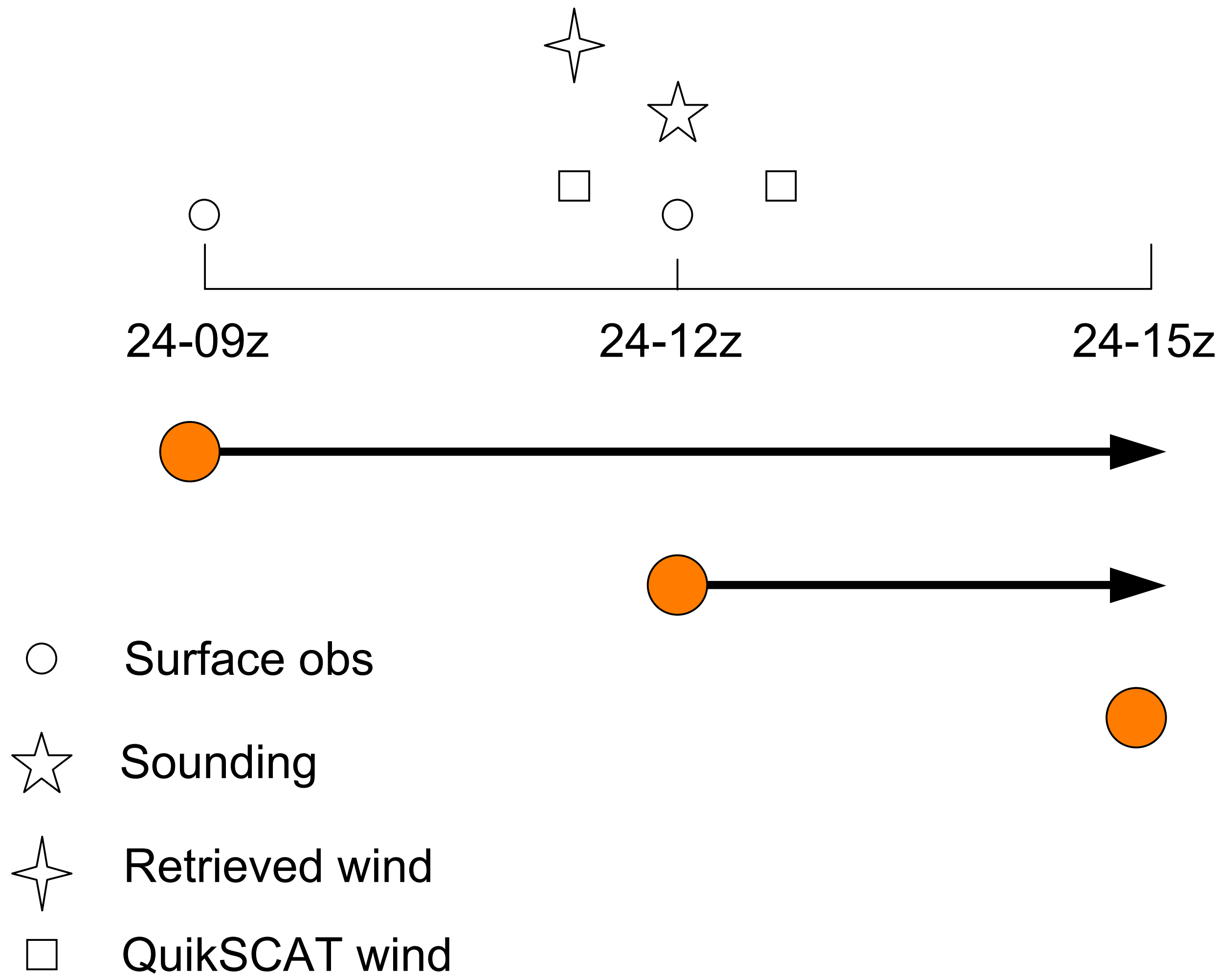
Increments of wind vector and temperature at 1000hpa

Cross-section of temperature increment

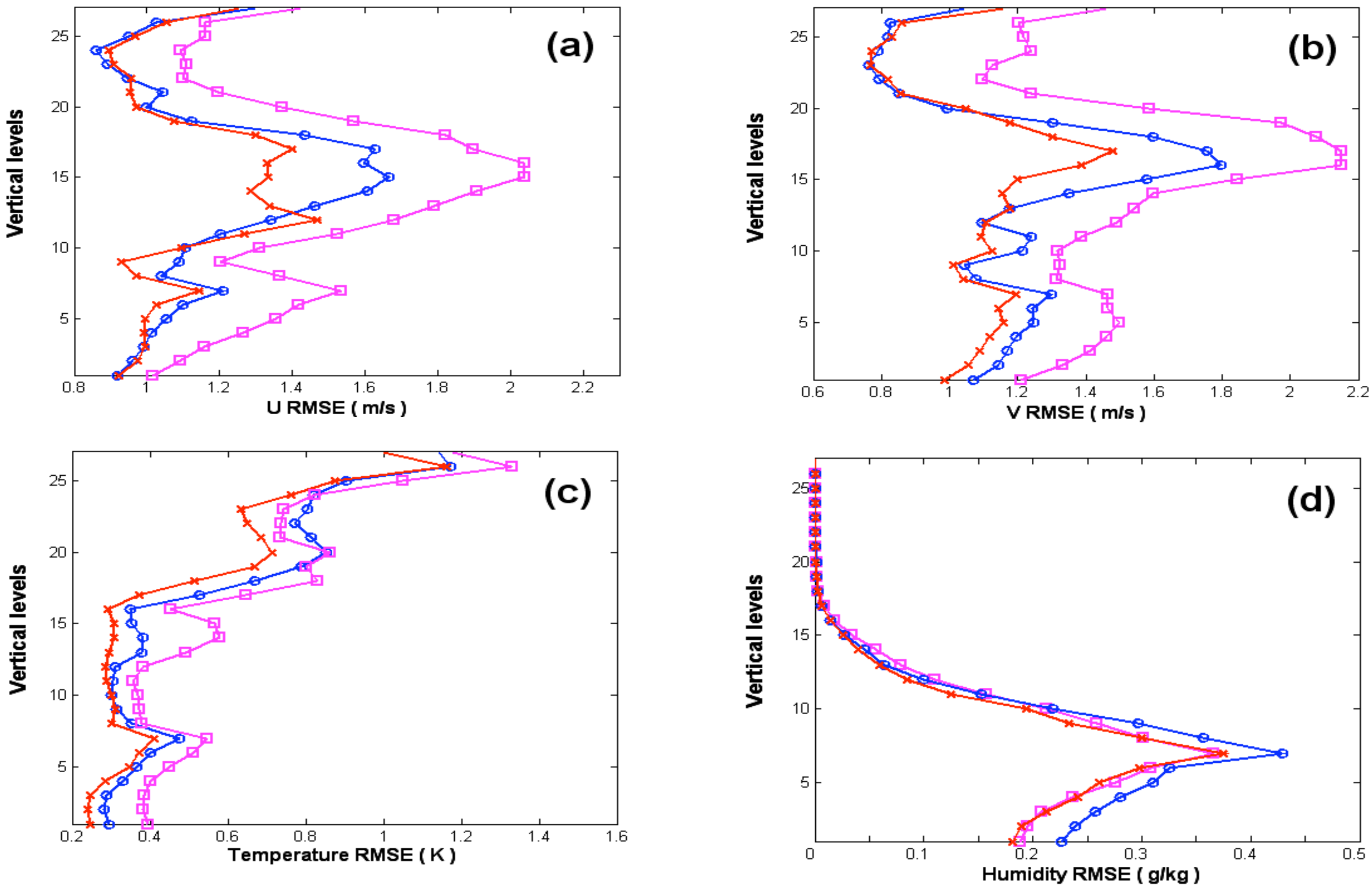


Blue Circle-line : analysis increment without localization
Red Cross-line : analysis increment with localization

Experiments on analysis time



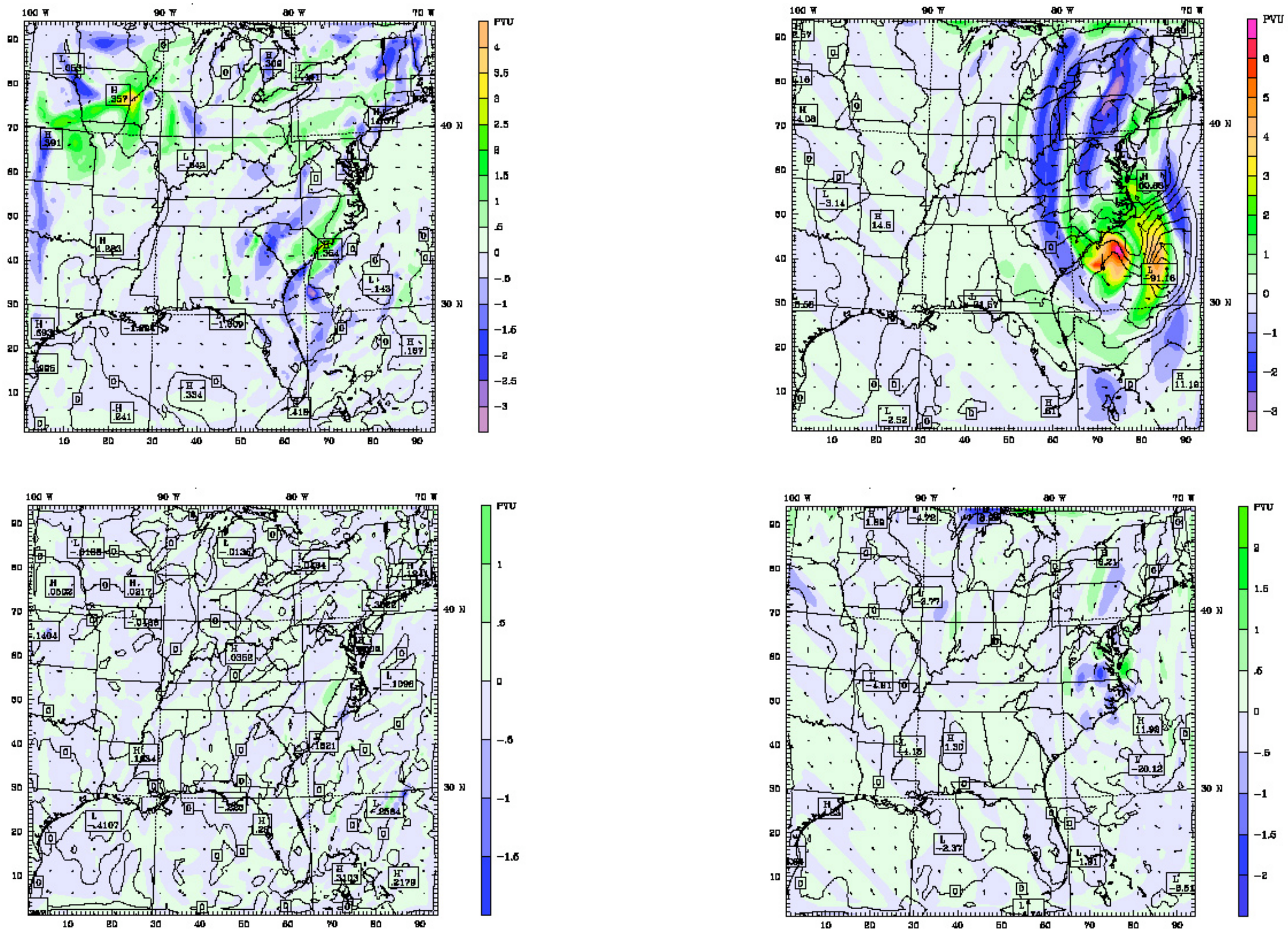
RMSE at different analysis time



Analysis at the beginning (pink), mid (red), and end (blue) of assimilation window

Analysis error at 300hpa

	25-00z	25-12z
CTL PV Wind Height		
En4Dvar PV Wind Height		

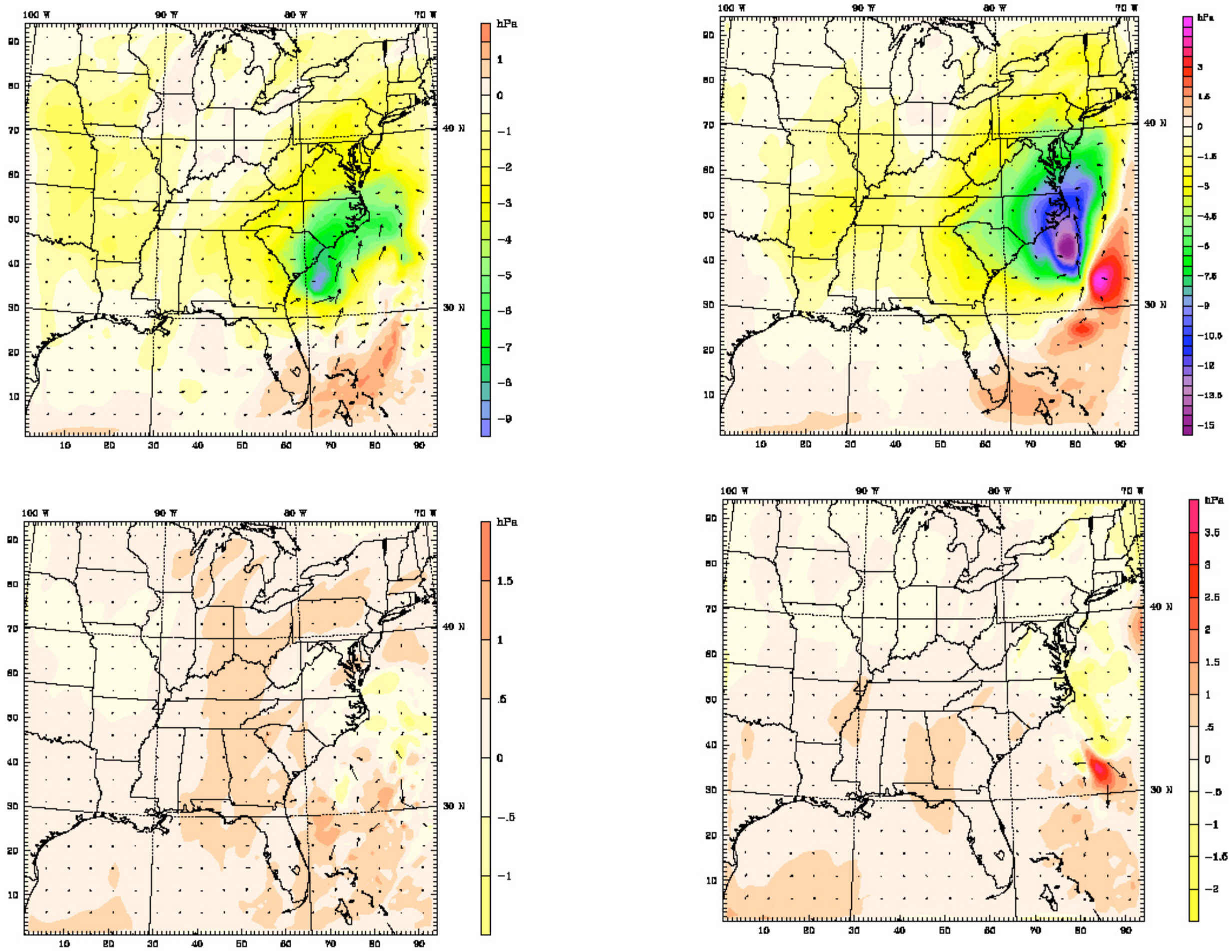


Analysis error at 1000hpa

25-00z　　25-12z

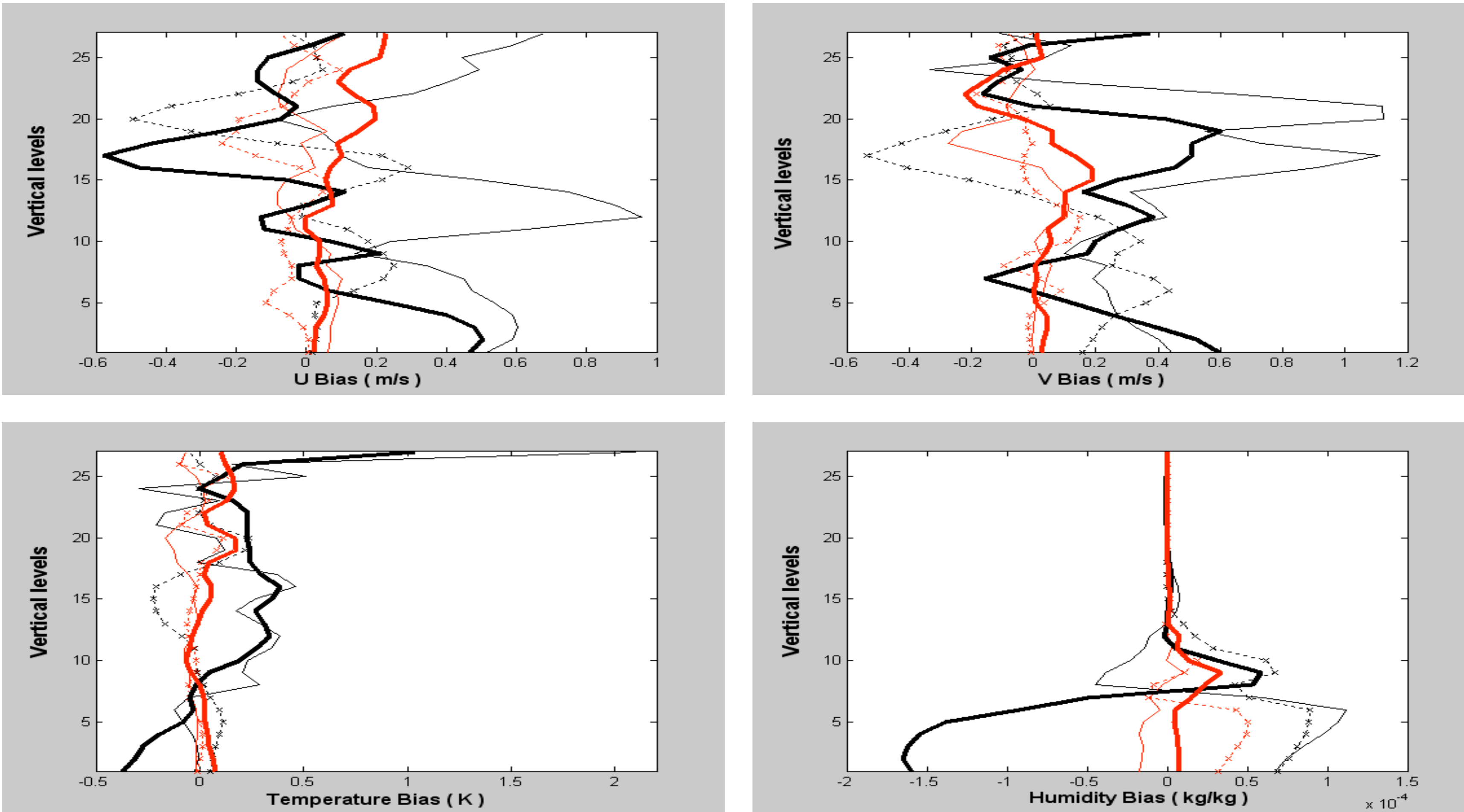
CTL
SLP
Wind

En4DVar
SLP
Wind



On 05/22/2009 at the 8th Adjoint Workshop in Tannersville, Pennsylvania

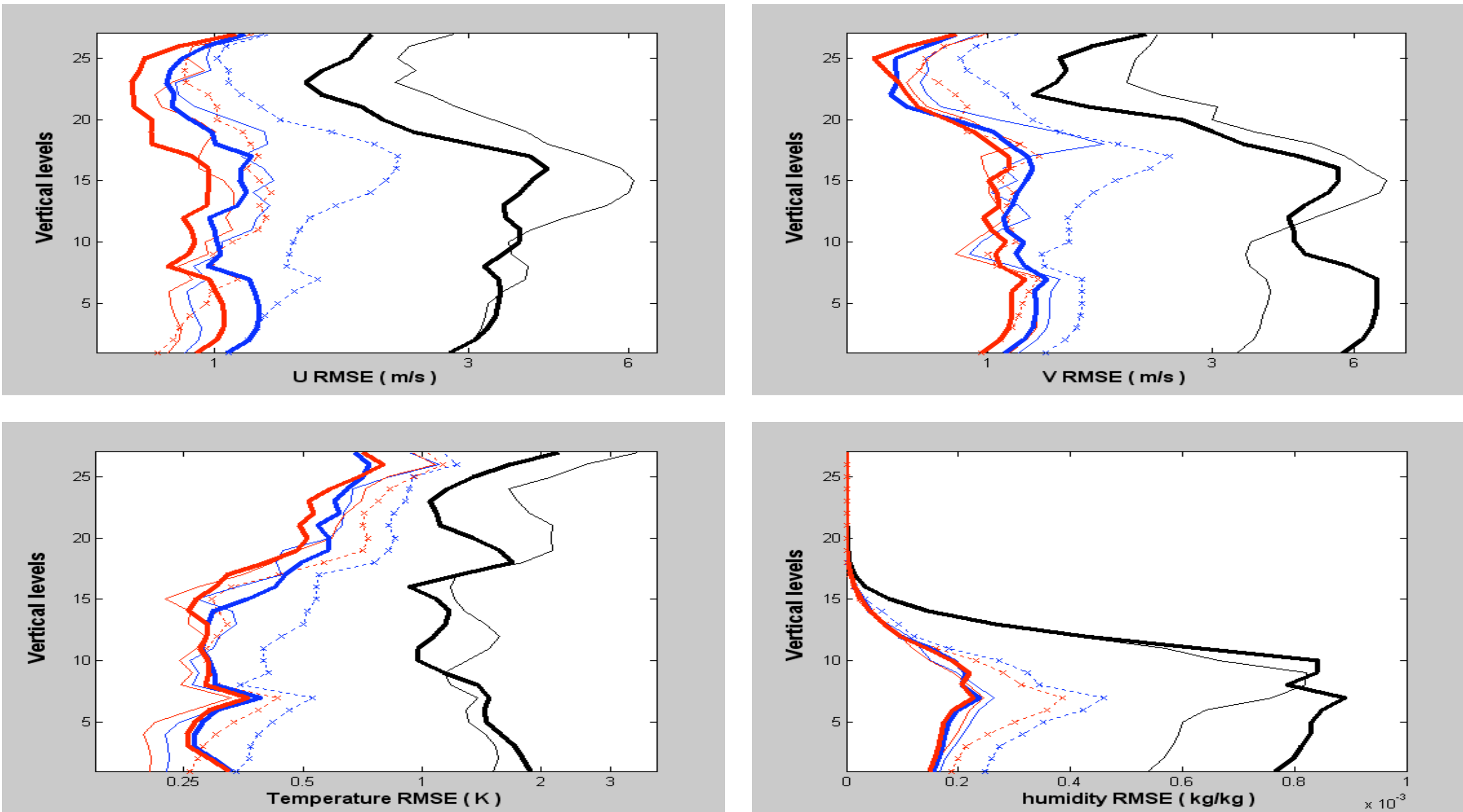
Vertical bias at 24-12Z/25-00Z/25-12Z



Red: analysis **Black: CTL**

Dot-cross: 24-12Z thin line: 25-00Z thick line: 25-12Z

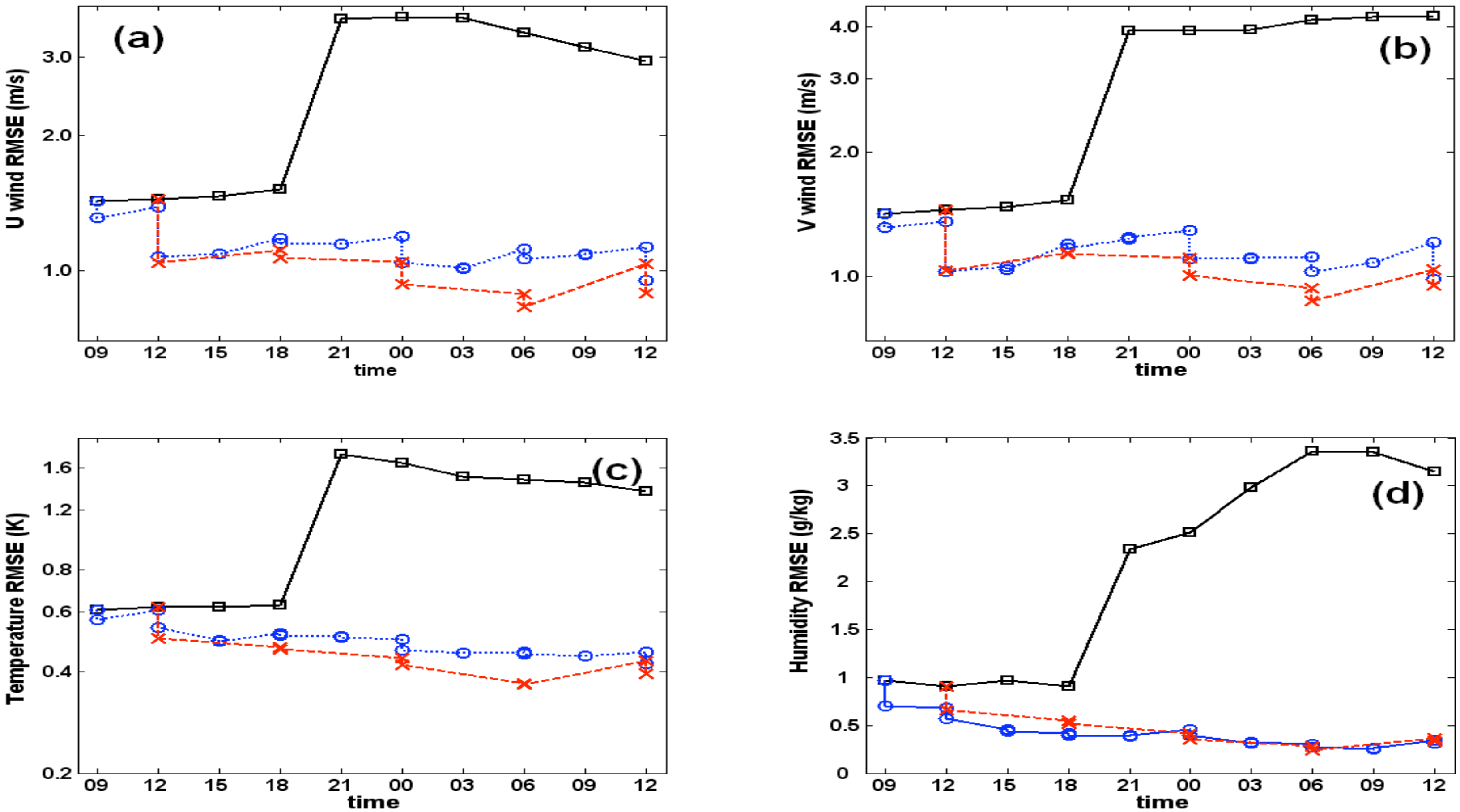
Vertical RMSE at 24-12Z/25-00Z/25-12Z



Red: En4D-Var analysis **Blue: forecast** **Black: CTL**

Dot-cross: 24-12Z thin line: 25-00z thick line: 25-12Z

Domain average RMSE in cycling



Black: CTL Blue: En3DVar Red: En4DVar

Summary

- WRF En4D-Var shows flow-dependant structure in its analysis increments.
- The localization with Schur operator can greatly reduce the analysis noise.
- The WRF En4D-Var optimal analysis time is at the middle (instead of the beginning) of assimilation window.
- OSSEs indicate that the analysis error using WRF En4D-Var is much less than that of control experiment.
- WRF En4D-Var gets a better analysis comparing with En3D-Var cycling.
- Comparison of WRF En4D-Var with WRF 4D-Var is under way.

Related publications:

Liu, C., Q. Xiao, and B. Wang, 2008: An ensemble-based four-dimensional variational data assimilation scheme: Part I: Technical formulation and preliminary test. *Mon. Wea. Rev.*, **136**, 3363-3373.

Liu, C., Q. Xiao, and B. Wang, 2009: An ensemble-based four-dimensional variational data assimilation scheme: Part II: Observing system simulation experiments with Advanced Research WRF (ARW). *Mon. Wea. Rev.*, **137**, 1687-1704.

Thank you !

Questions and comments
are welcome.