Evaluation of an Ensemble Kalman Filter for WRF against a 3D-Var Assimilation System

Alain Caya

Jeff Anderson, Chris Snyder, Dale Barker

National Center for Atmospheric Research

Outline

- Motivations
- Brief introduction to data assimilation
- Description of the EnKF and 3D-Var methods
- Model, observations, experimental design
- Observation space diagnostics
- Imbalance introduced by the assimilation
- Conclusions

Motivations

- Develop an Ensemble Kalman Filter for the research community;
- Potential for a wide range of applications and scales;
- Ensemble gives an error estimate;
- Base for ensemble forecasting.

Data Assimilation Research Testbed

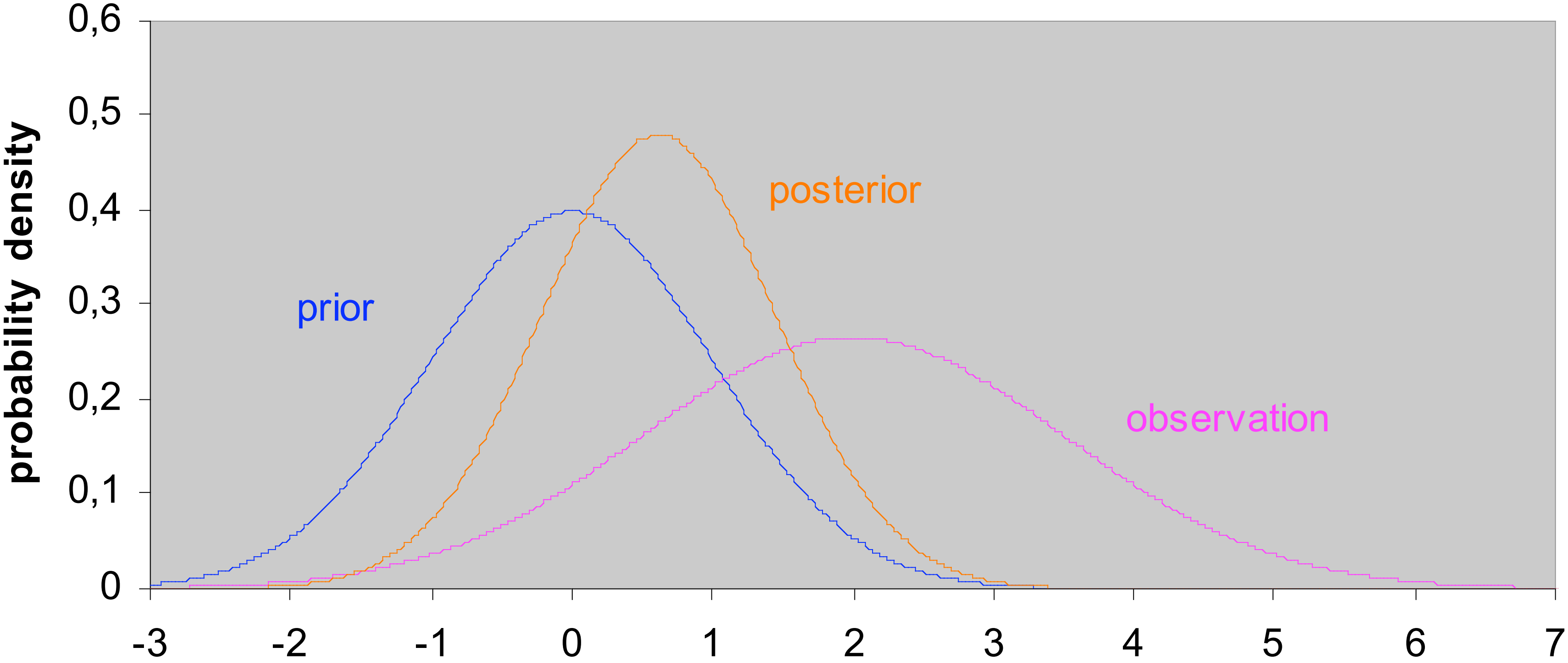
http://www.image.ucar.edu/DAI/DART

- Ensemble-based data assimilation schemes
- Compliant models:

1. Many low-order models (Lorenz63, L84, L96, ...)
2. Global 2-level PE model (from NOAA/CDC)
3. CGD's CAM 2.0 & 3.0 (global spectral model)
4. GFDL FMS B-grid GCM (global grid point model)
5. MIT GCM (from Jim Hansen)
6. Weather Research and Forecast model
7. NCEP GFS (assisted by NOAA/CDC)
8. GFDL MOM3/4 ocean model
9. ACD's ROSE model (upper atmosphere with chemistry)

Data Assimilation

$$\mathbf{p}\left(\mathbf{x}_t \mid \mathbf{Y}_t\right) = \frac{\mathbf{p}\left(\mathbf{y}_t \mid \mathbf{x}_t\right)\mathbf{p}\left(\mathbf{x}_t \mid \mathbf{Y}_{t-1}\right)}{\int \mathbf{p}\left(\mathbf{y}_t \mid x\right)\mathbf{p}\left(x \mid \mathbf{Y}_{t-1}\right)dx}$$



Assuming normal statistics…

$$\mathbf{y} = \mathbf{Hx} + \mathbf{\mathring{a}}_r, \quad \mathbf{\mathring{a}}_r \sim \mathrm{N}(0, \mathbf{R})$$

$$\mathbf{p}(\mathbf{x}_t \mid \mathbf{Y}_t) \propto \mathrm{e}^{-(\mathbf{x}-\mathbf{x}_\mathrm{b})^\mathrm{T} \mathbf{P}^{b^{-1}} (\mathbf{x}-\mathbf{x}_\mathrm{b})} \mathrm{e}^{-(\mathbf{Hx}-\mathbf{y})^\mathrm{T} \mathbf{R}^{-1} (\mathbf{Hx}-\mathbf{y})}$$

$$\Downarrow$$

$$J(\mathbf{x}) = (\mathbf{x}-\mathbf{x}_\mathrm{b})^\mathrm{T} \mathbf{P}^{b^{-1}} (\mathbf{x}-\mathbf{x}_\mathrm{b}) + (\mathbf{Hx}-\mathbf{y})^\mathrm{T} \mathbf{R}^{-1} (\mathbf{Hx}-\mathbf{y})$$

True forecast-error statistics

- Are
  - Inhomogeneous, anisotropic, non-stationary
- Depend on
  - Model: error statistics, resolution, forecast range, …
  - Observations: error statistics, distribution, frequency, …

(Ensemble) Kalman filter

$$\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^b + \frac{\mathbf{P}^b\mathbf{H}^\mathrm{T}}{\mathbf{H}\mathbf{P}^b\mathbf{H}^\mathrm{T} + \mathbf{R}}\left(\mathbf{y} - \mathbf{H}\overline{\mathbf{x}}^b\right)$$

$$\mathbf{P}^b\mathbf{H}^\mathrm{T} \cong \frac{1}{n-1}\sum_{i=1}^{n} \mathbf{x}_i'^b \left(\mathbf{H}\mathbf{x}_i'^b\right)$$

$$\mathbf{H}\mathbf{P}^b\mathbf{H}^\mathrm{T} \cong \frac{1}{n-1}\sum_{i=1}^{n} \left(\mathbf{H}\mathbf{x}_i'^b\right)\left(\mathbf{H}\mathbf{x}_i'^b\right)$$

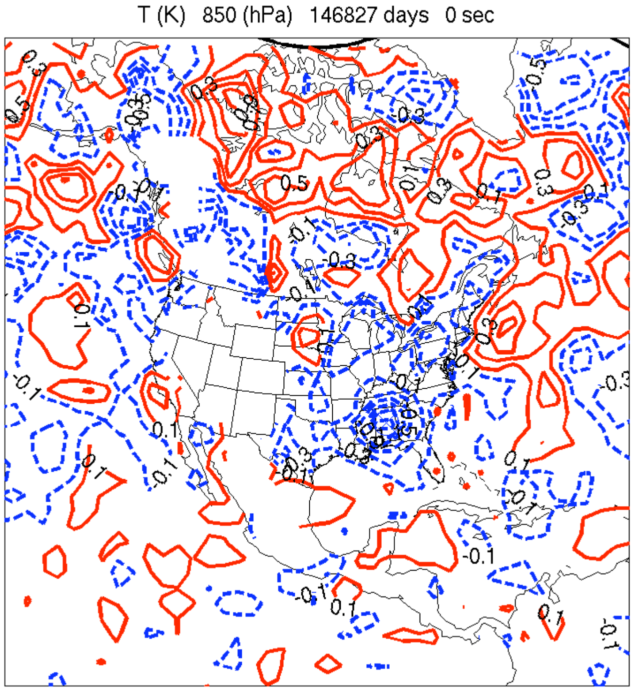
Ensemble Kalman filter

- $\mathbf{x}'^{a}$ are obtained using the square-root filter approach.
- 56 ensemble members.
- No inflation.
- Observations assimilated serially, one at a time.
- Horizontal localization half-width: ~ 2500 km.

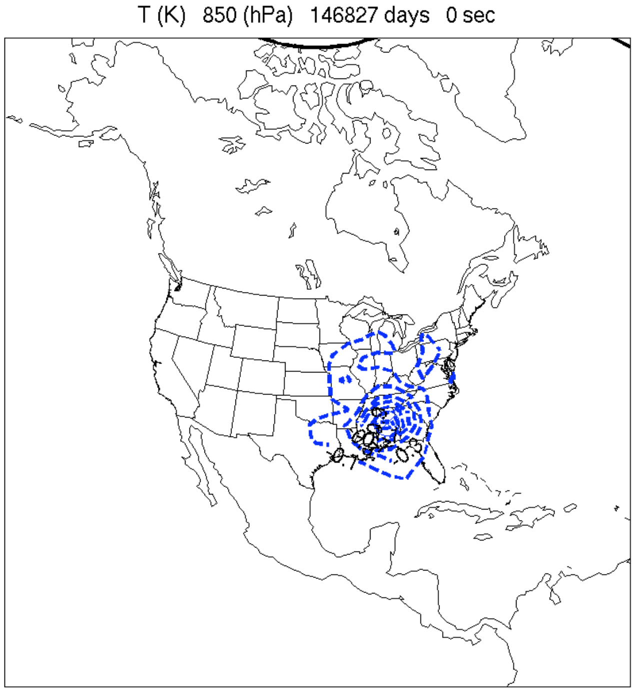
Temperature Increments

(T-obs at 850 hPa, initial time)

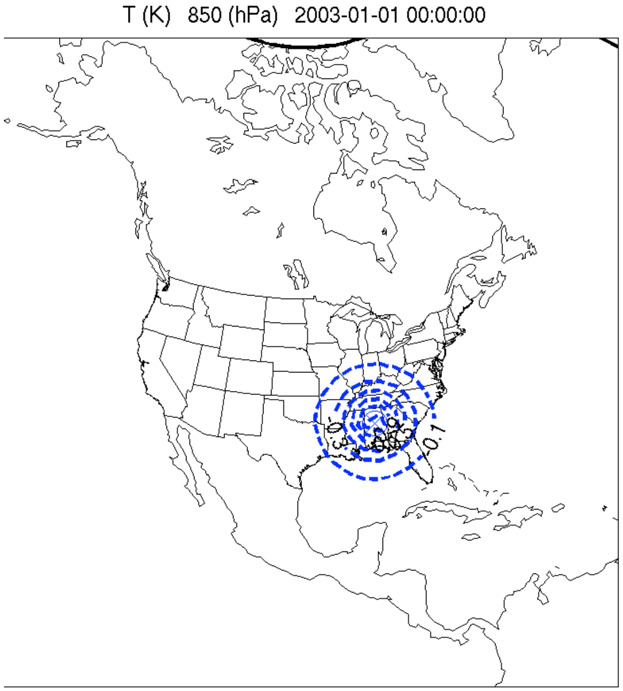
EnKF



With Localization



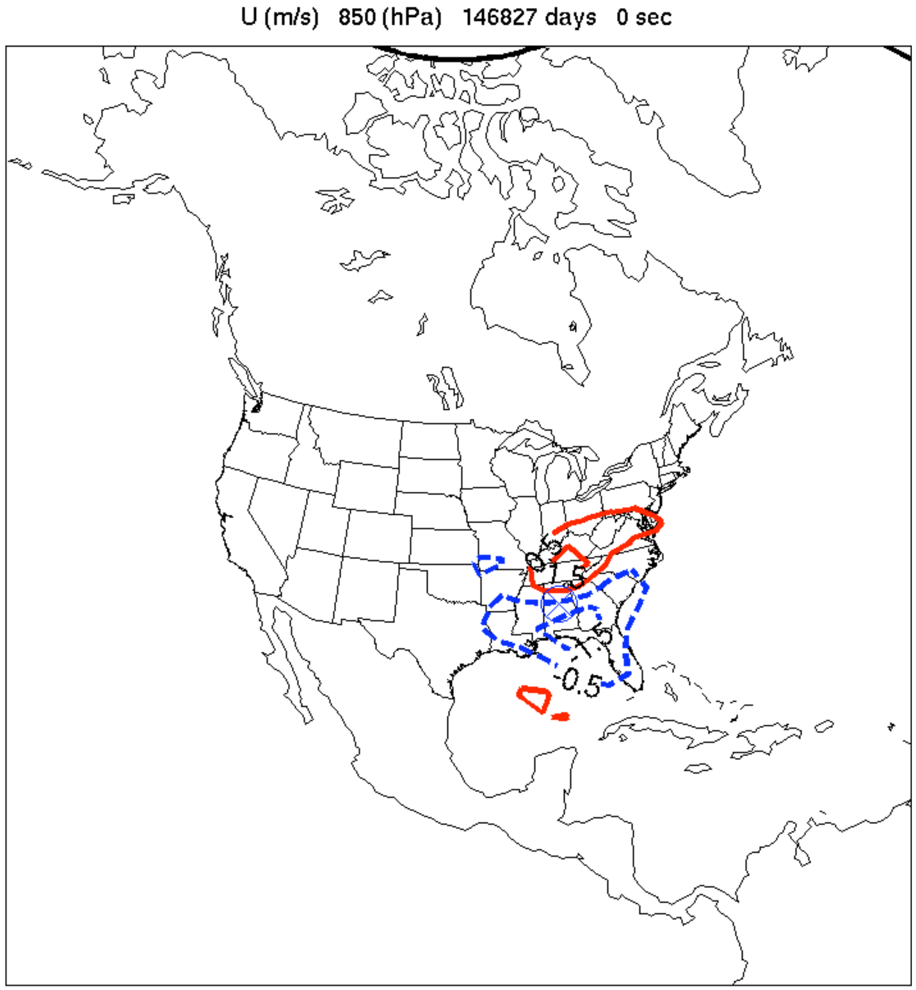
3D-Var



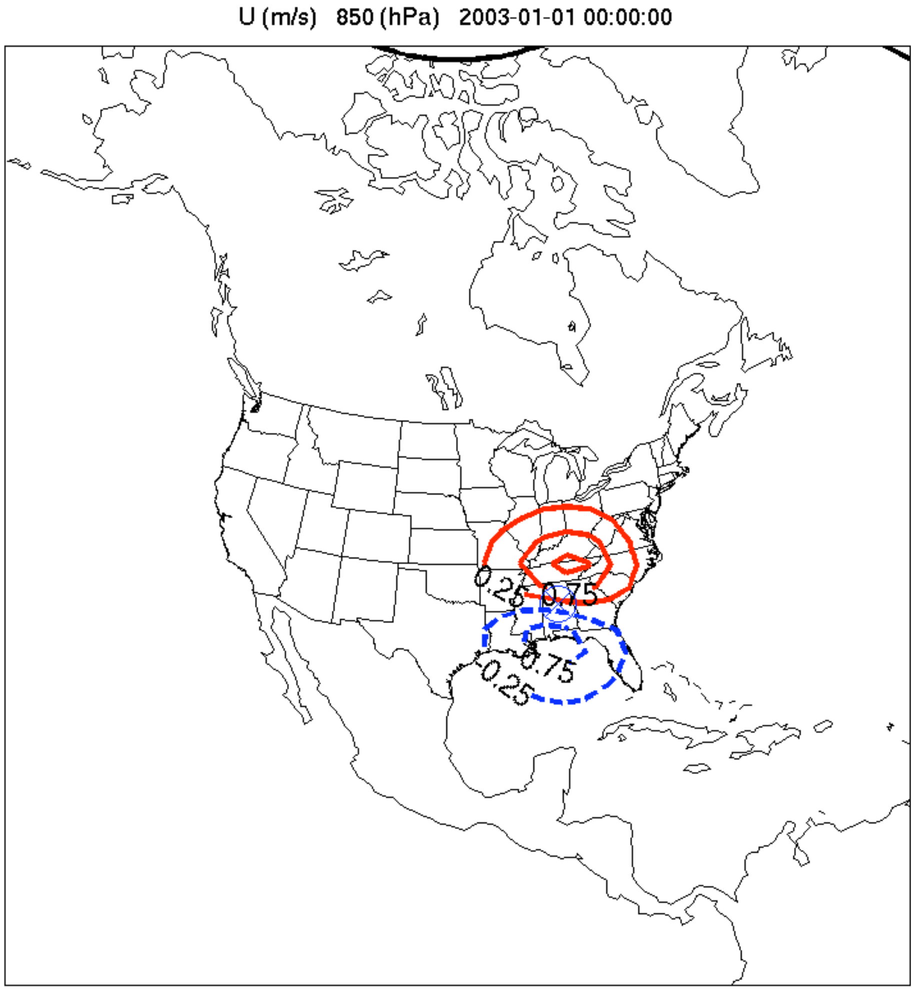
Zonal-wind Increments

(T-obs at 850 hPa, initial time)

EnKF



3D-Var



WRF 3D-VAR

http://box.mmm.ucar.edu/wrf/WG4/

- Background covariance model given by recursive filters (Wu et al., 2002, *MWR*)
- isotropic - inhomogeneous
- Control variables:
  - streamfunction (ψ),
  - unbalanced velocity potential (χ),
  - unbalanced temperature (*T*),
  - unbalanced surface pressure (p_s),
  - pseudo-relative humidity.
- Linear balance equation

WRF 3D-VAR Background Error Covariance

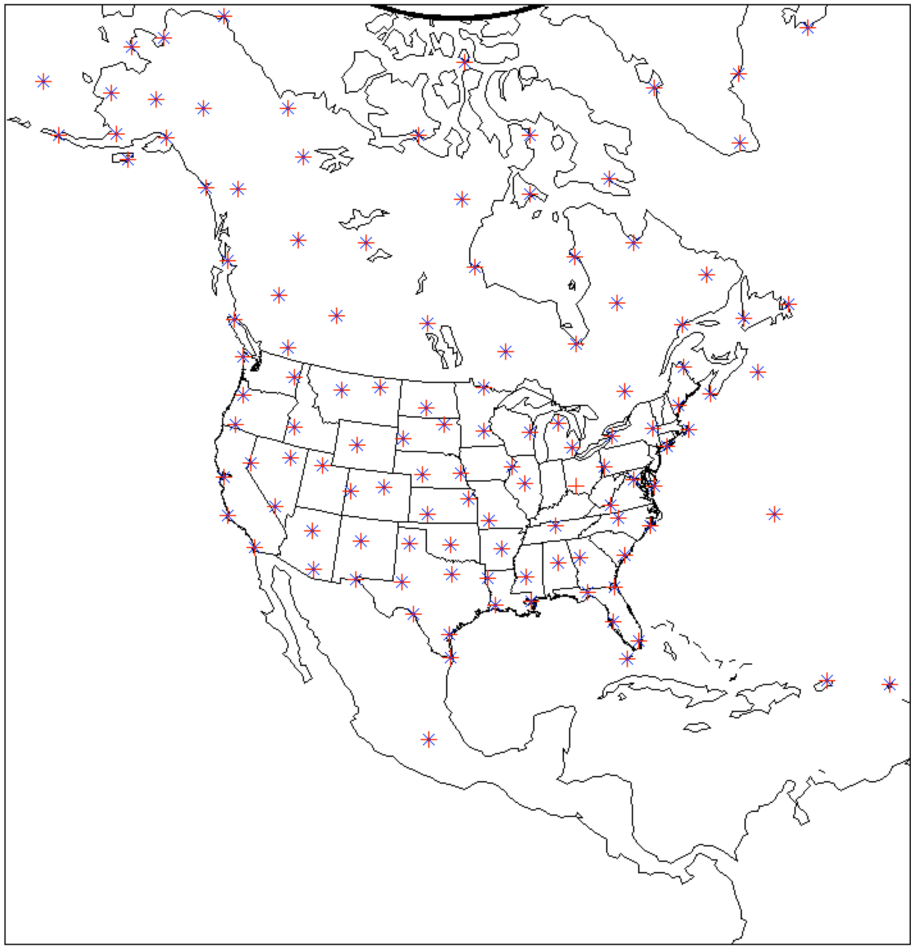
- “NMC” method.
- 48 – 24-hour GFS forecasts with T170 resolution valid at the same time for 357 cases distributed over one year.
- Amplitudes and scales have been tuned for WRF.

WRF THE WEATHER RESEARCH & FORECASTING MODEL

http://wrf-model.org

- Regional, atmospheric model
- Non-hydrostatic, fully compressible
- Multiple nested domains
- Movable grids

Observation Network

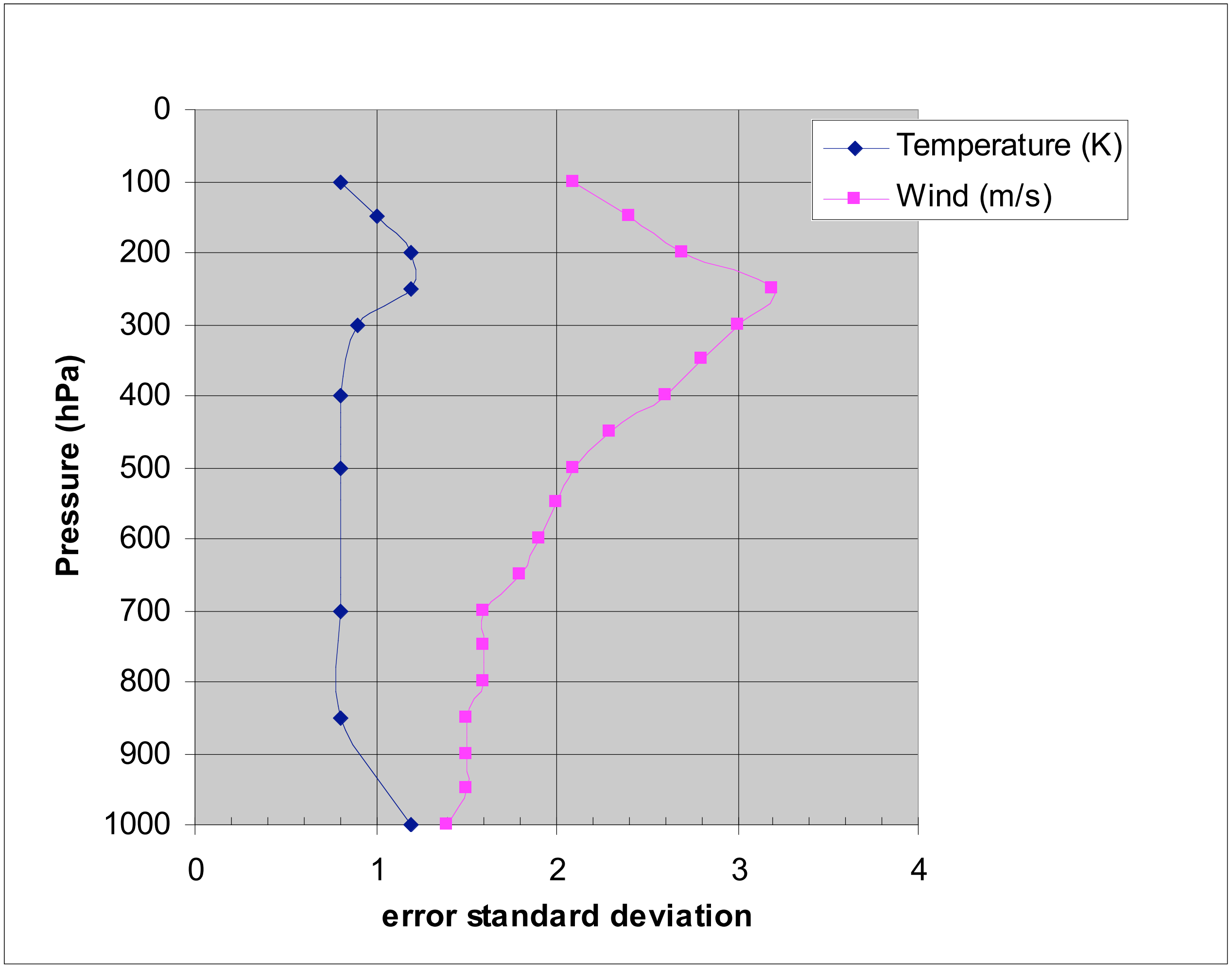


Real observations:

- NCEP radiosonde data set
- U, V wind components
- Temperature
- No background check

Observational Errors

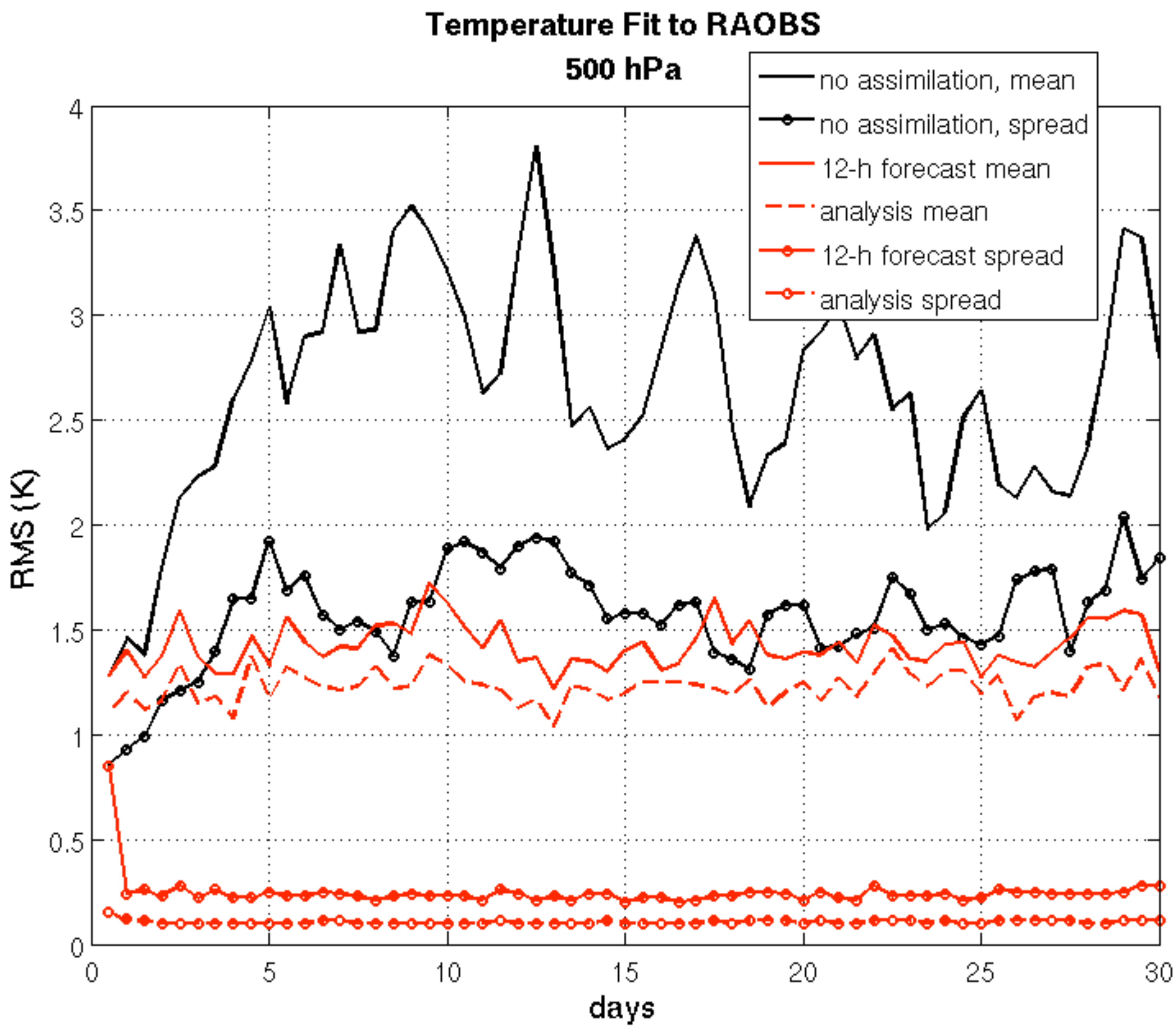
No observational-error covariances



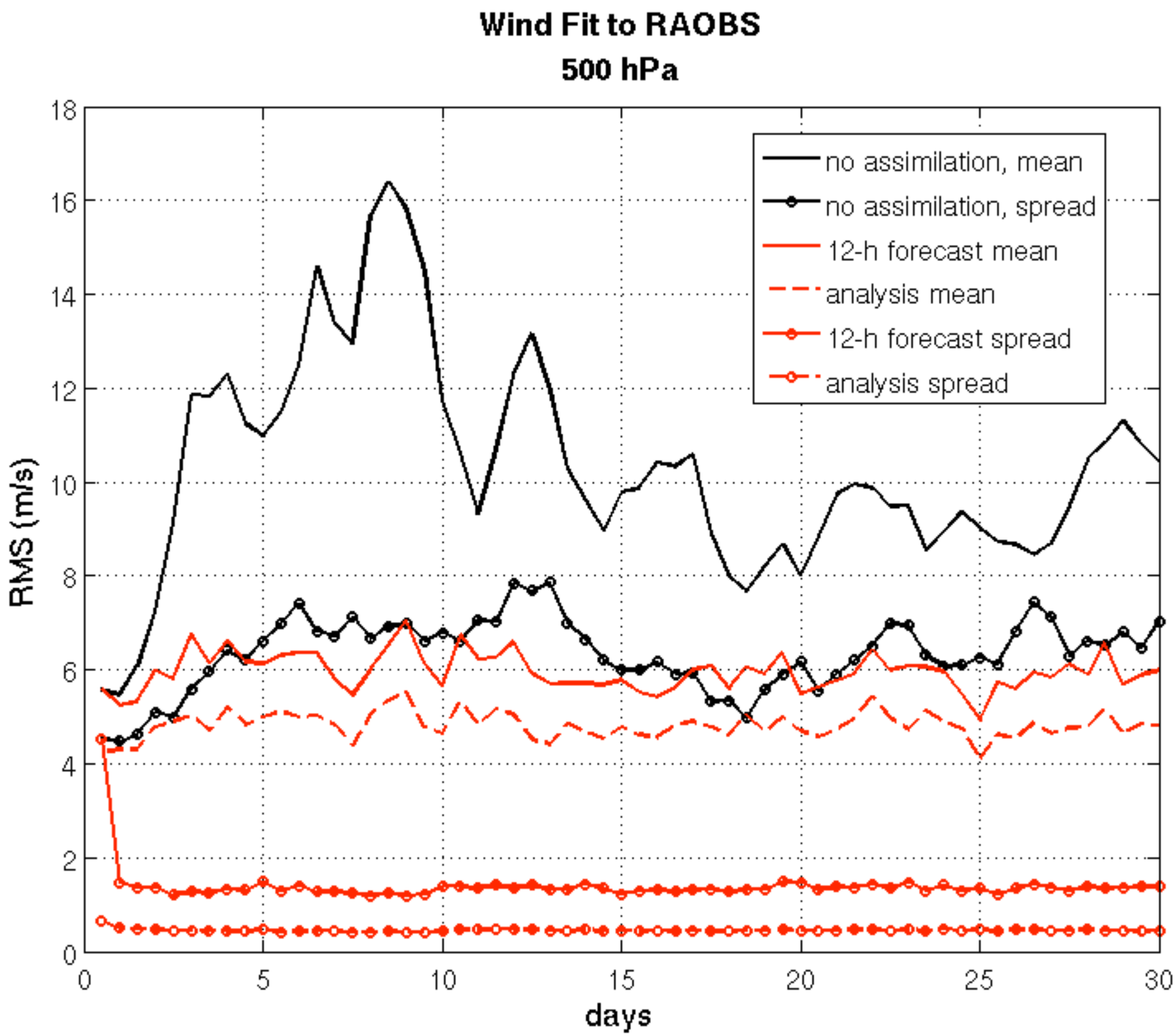
Experimental Setup

- North America, 45x45 grid, 200 km resolution,
- Real observations available every 12 h,
- 3D-Var uses NCEP/FNL global analyses as ICs (1 Jan 03) and LBCs,
- EnKF uses same for initial ensemble mean and ensemble mean LBCs,
- EnKF ICs and LBCs perturbations are drawn from 3D-Var background-error statistics.

EnKF

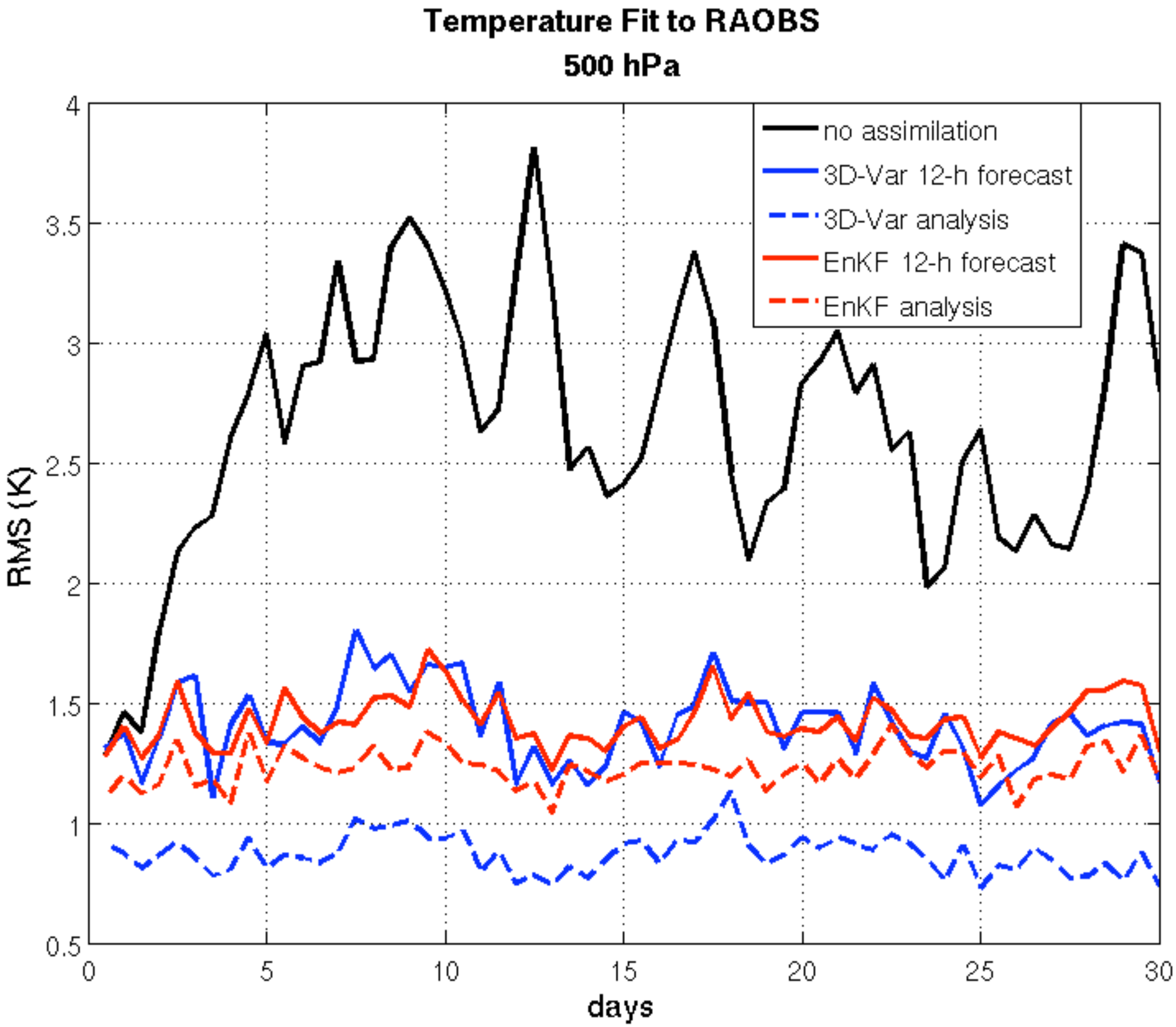

Temperature Fit to RAOBS
500 hPa
no assimilation, mean
no assimilation, spread
12-h forecast mean
analysis mean
12-h forecast spread
analysis spread
RMS (K)
days
0
0.5
1
1.5
2
2.5
3
3.5
4
0
5
10
15
20
25
30

EnKF

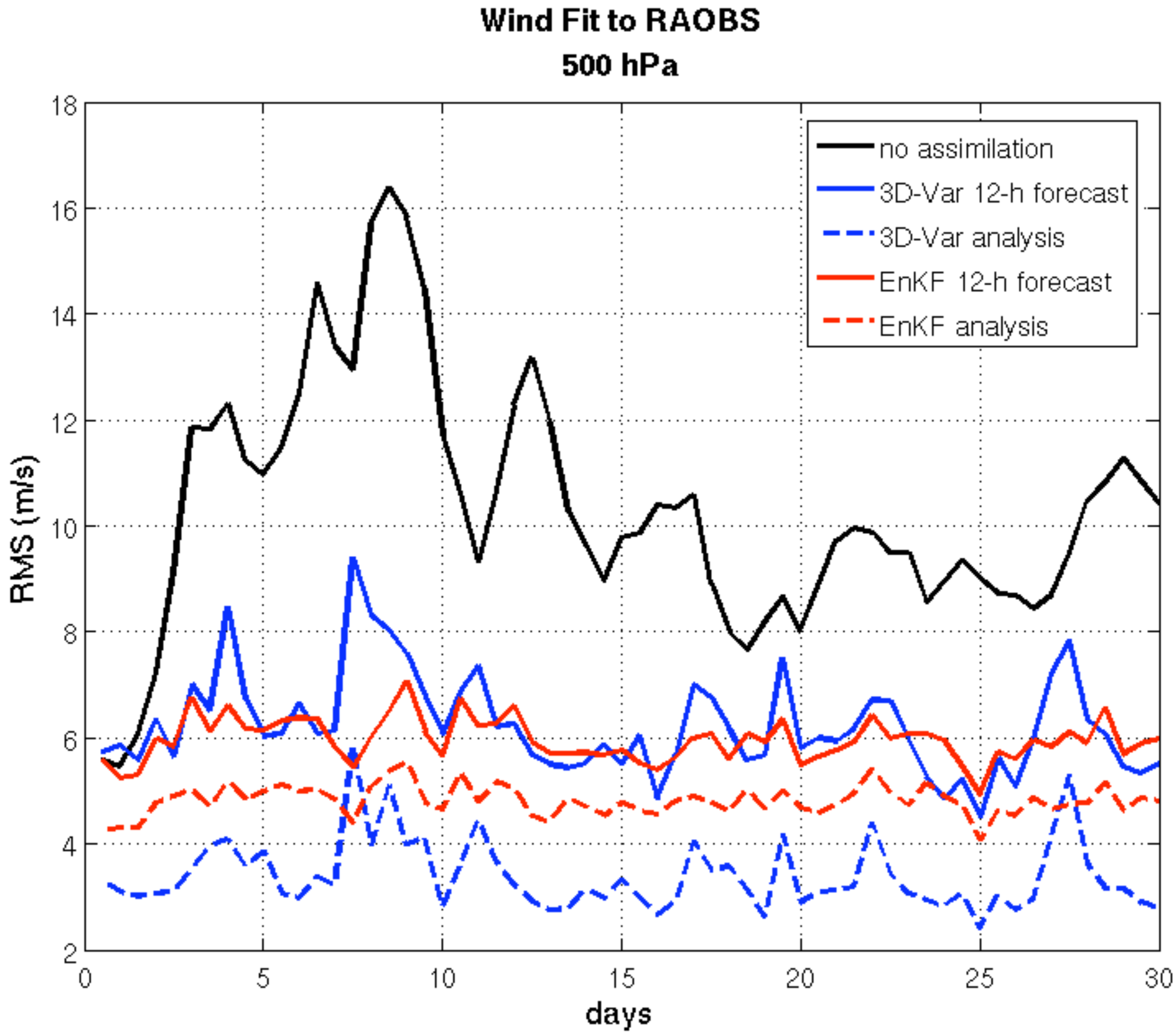


Wind Fit to RAOBS
500 hPa
RMS (m/s)
days
0
2
4
6
8
10
12
14
16
18
0
5
10
15
20
25
30
no assimilation, mean
no assimilation, spread
12-h forecast mean
analysis mean
12-h forecast spread
analysis spread

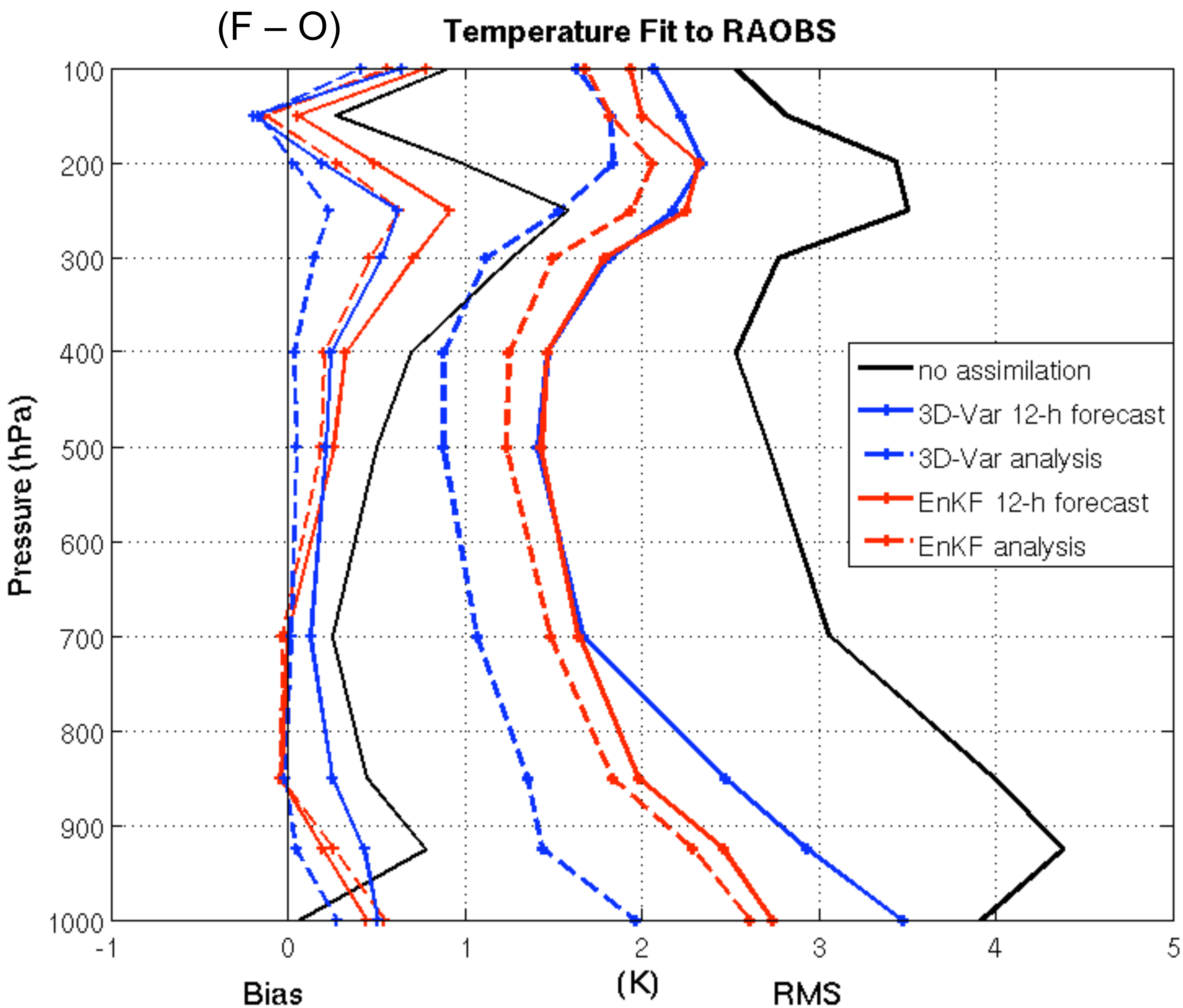
EnKF vs 3D-Var



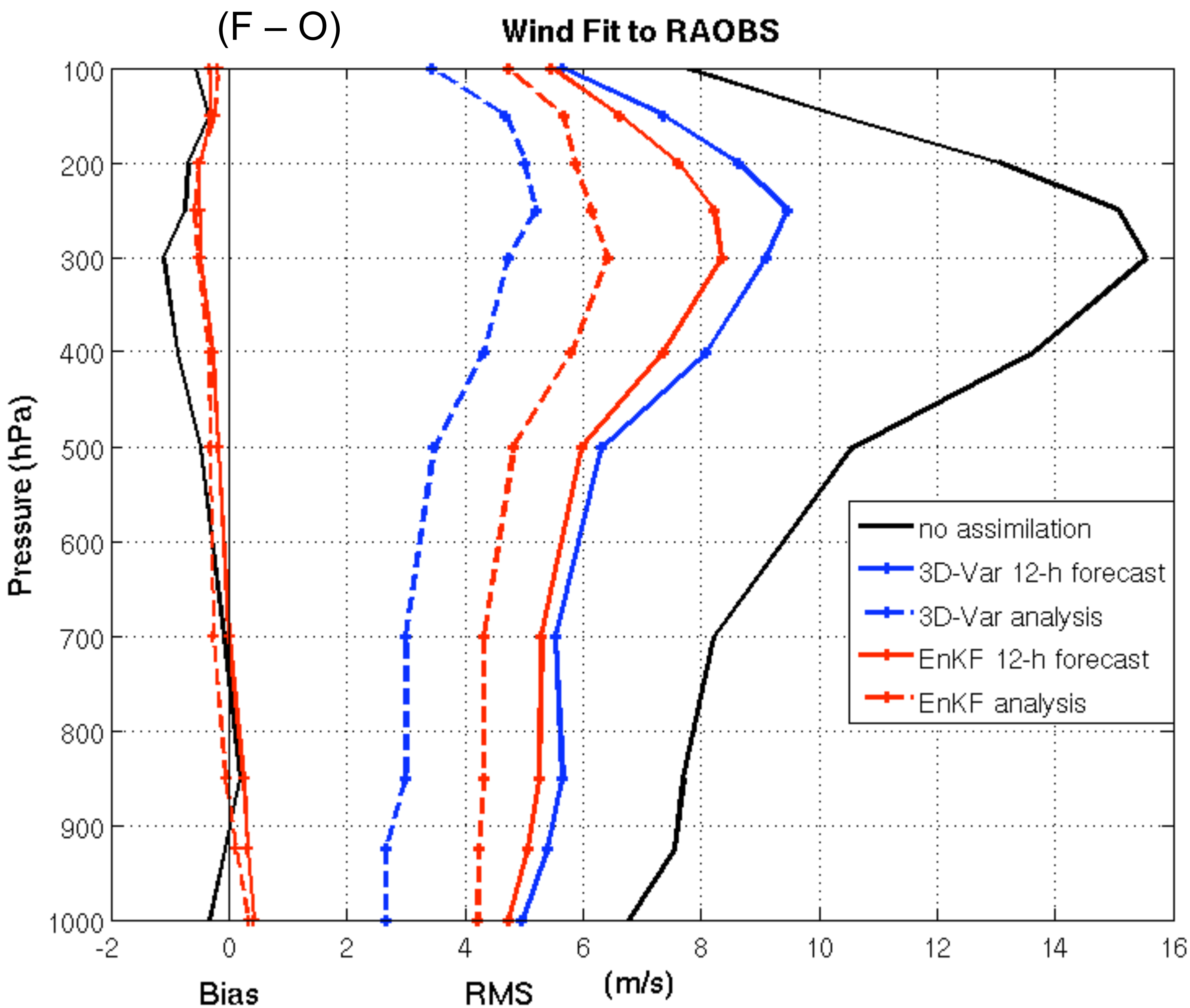
EnKF vs 3D-Var



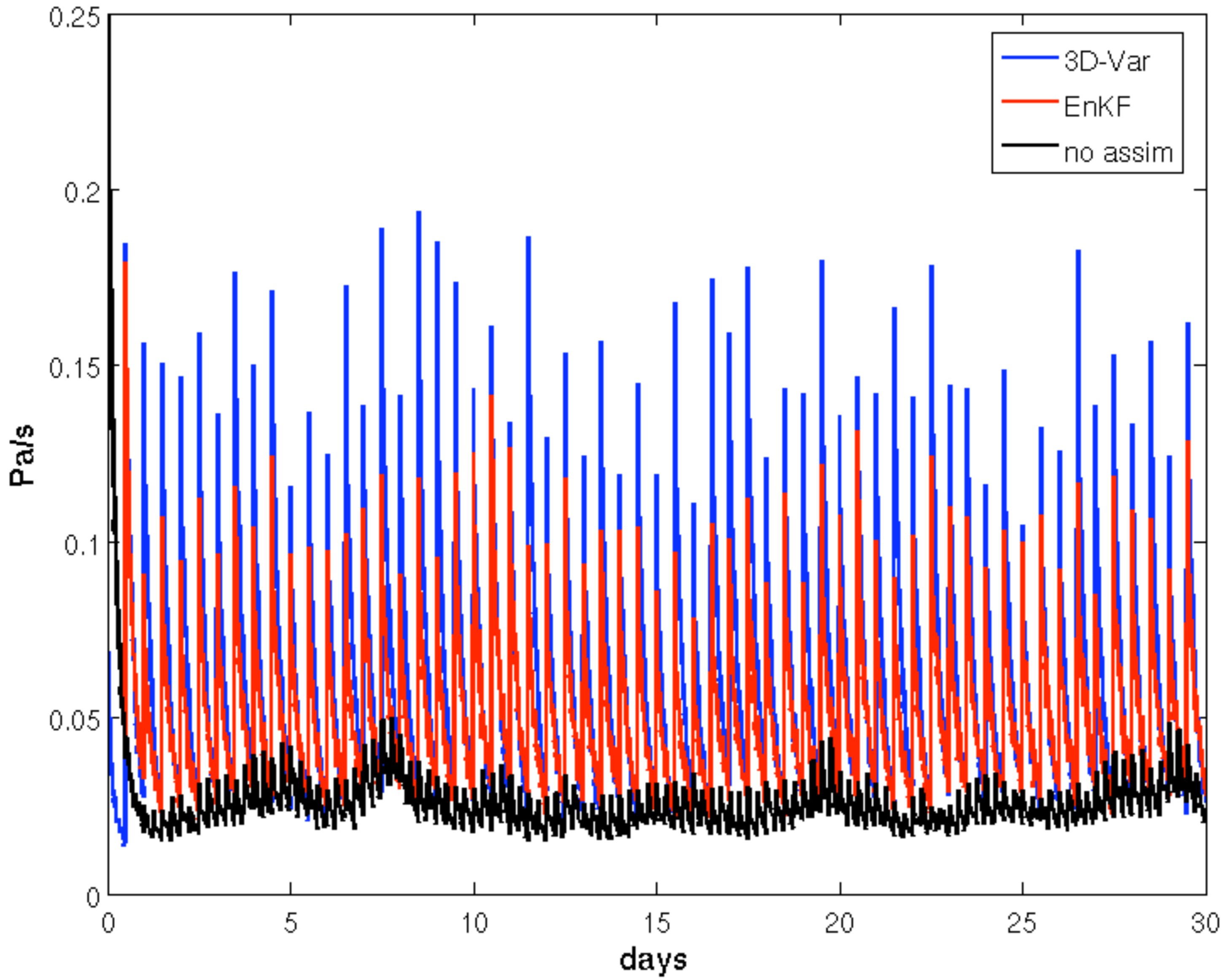
EnKF vs 3D-Var



EnKF vs 3D-Var



Imbalance



3D-Var: ~0.057

EnKF: ~0.048

Free: ~0.026

Conclusion 1

- Results indicate good performance of the EnKF at synoptic scale with WRF.
- The 12-hour forecasts from EnKF are closer to the observations than 3D-Var 12-hour forecasts:
  - Consistent with earlier OSSE experiments.
  - Non-stationary forecast-error statistics in the EnKF are beneficial.
  - Both schemes may improve with further tuning.

Conclusion 2

- The EnKF introduces somewhat less imbalance than 3D-Var:
  - OSSE experiments reported earlier indicated that more imbalance is introduced by the EnKF
  - Initialization? (digital filter, build balance constraints in the EnKF scheme?)
  - Need larger ensemble?
  - More sophisticated localization? (group filter)

Ongoing work with WRF EnKF

- Humidity observations assimilation
- Satellite data assimilation (retrievals)
- Surface observations
- Radar data assimilation:
  - Radial velocities
  - Reflectivities