Earth System Research Laboratory
SCIENCE, SERVICE, & STEWARDSHIP

Ensemble-based data assimilation research in NOAA

Jeff Whitaker, Xuguang Wang, Tom Hamill, and Xue Wei
(*NOAA / ESRL / PSD*)

Zoltan Toth, Yucheng Song, and Dick Wobus
(*NOAA / NCEP*)

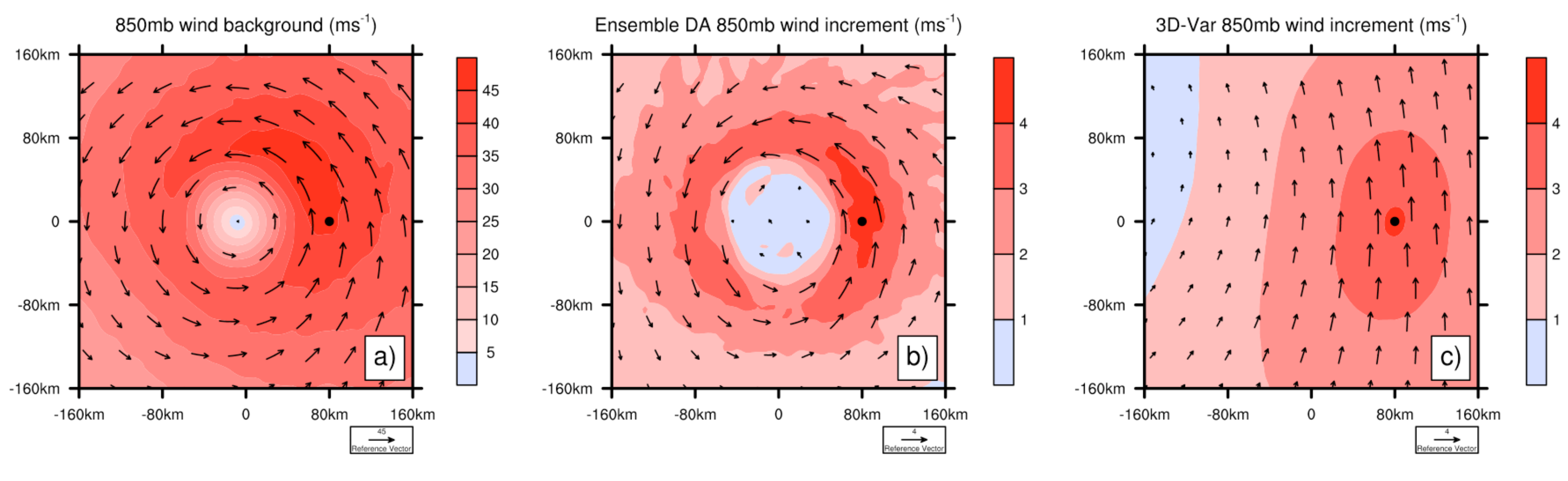
funding: NOAA THORPEX & NSF grants

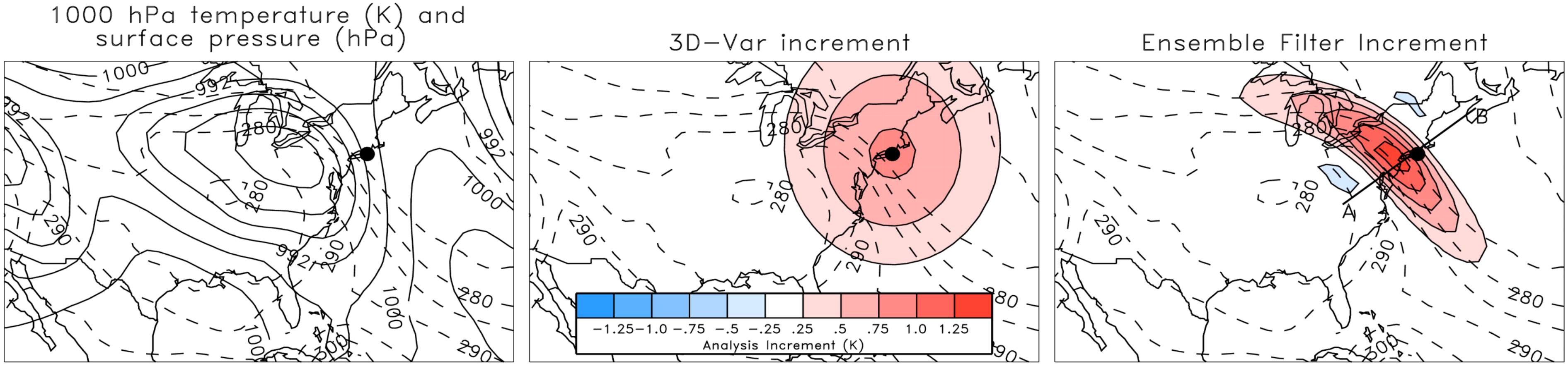
Outline

- Review of ensemble-based data assimilation
- Conventional-data experiments in T62-version of NCEP Global Forecast System (GFS); Jeff Whitaker, lead.
- Perfect-model test of a hybrid Ensemble Transform Kalman Filter / 3D-Var; Xuguang Wang, lead.

Ensemble-based data assimilation

- Parallel forecast and analysis cycles
- Ensemble of forecasts is used to estimate forecast-error statistics during the data assimilation





Advantages of ensemble-based data assimilation

- Potentially very accurate: equivalent to optimal Kalman-filter solution under special assumptions (infinite ensemble, Gaussian, perfect model, known **R**, linear growth of errors).
- Automatic initialization of ensemble forecasts; provides a distribution of analyses.
- Easy to code: algorithmically simple compared to *4D-Var.*

Disadvantages of ensemble-based data assimilation

- Computationally expensive, probably on par with 4D-Var. In conventional filters, costs scale with:
  - Number of observations
  - Dimension of model state
  - Size of ensemble
- Use of "covariance localization" (usually necessary to avoid filter divergence) may introduce imbalances to initial conditions.
- Relative improvements over 3D-Var largest when data sparse (not situation in modern NWP).

(note: will discuss possible algorithmic variants to address cost, imbalance)

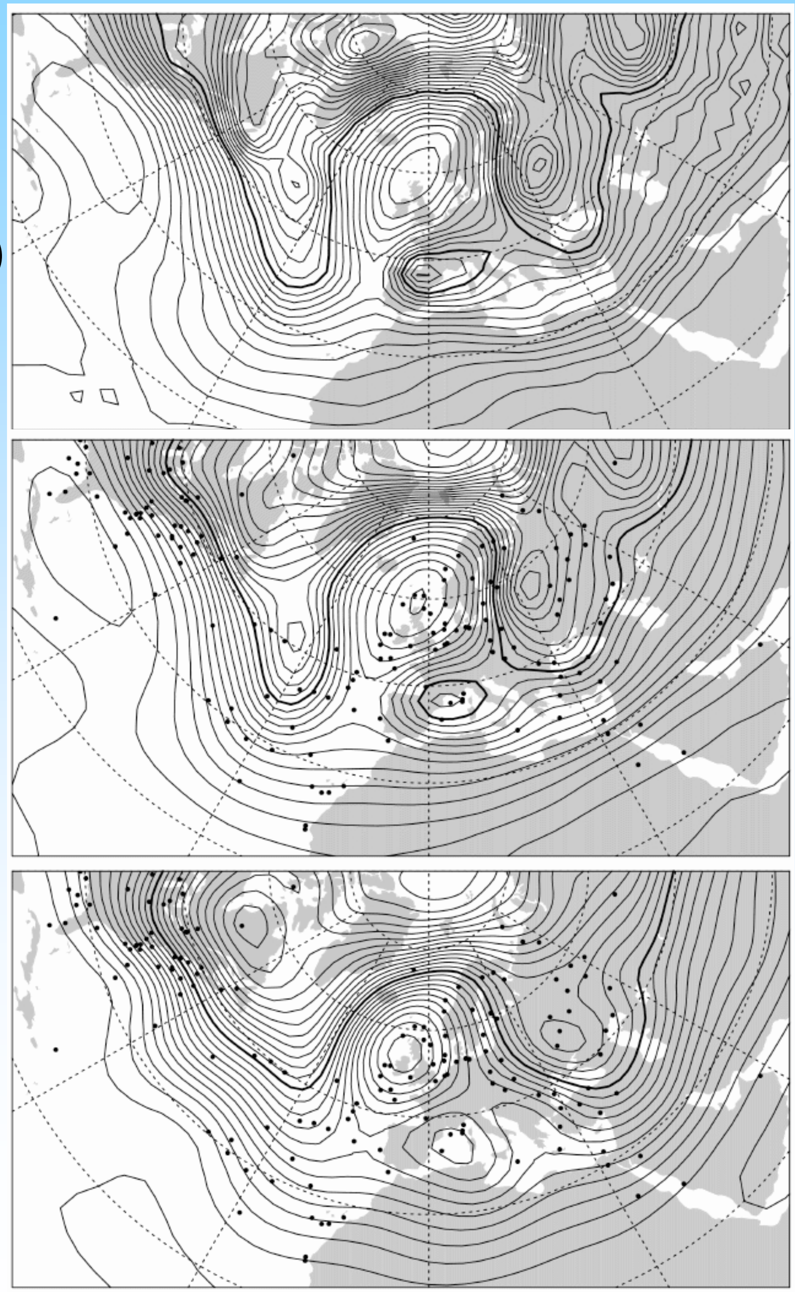
Example: Sparse Network (P_s obs only)

Whitaker et al. 2004, MWR, p.1190

Full NCEP-NCAR Reanalysis (3D-Var) **(200,000+ obs)**

Ensemble Filter **(214 surface pressure obs)**

Climatological covariances **(214 surface pressure obs)**

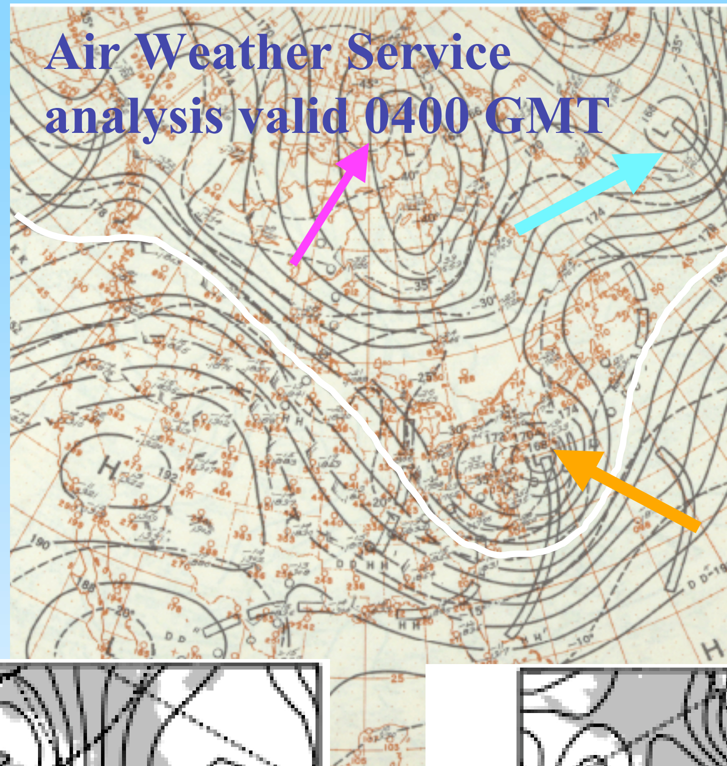


Black dots show pressure ob locations

RMS = 39.8 m

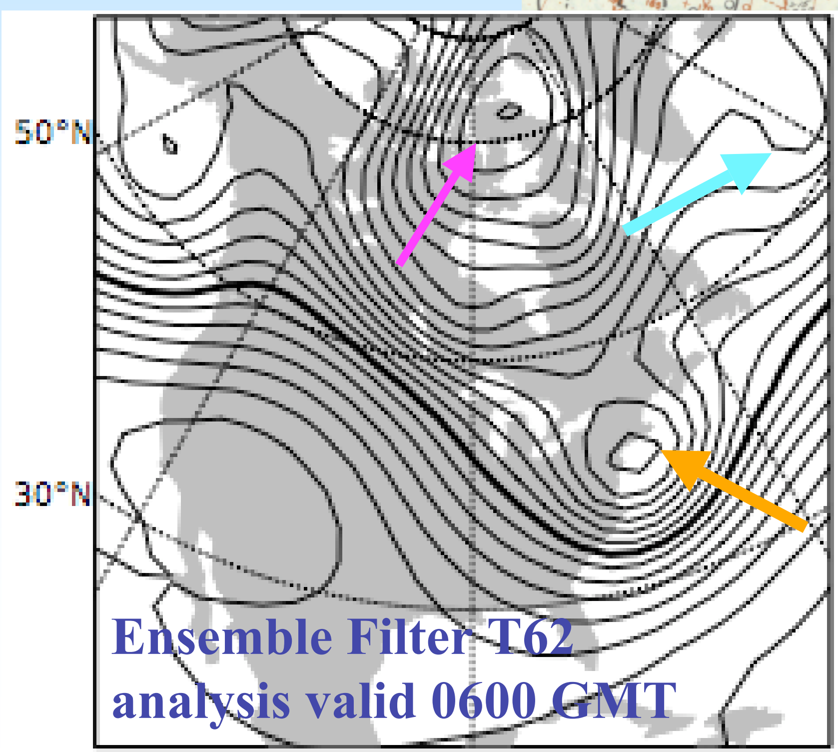
RMS = 82.4 m (3D-Var is worse!)

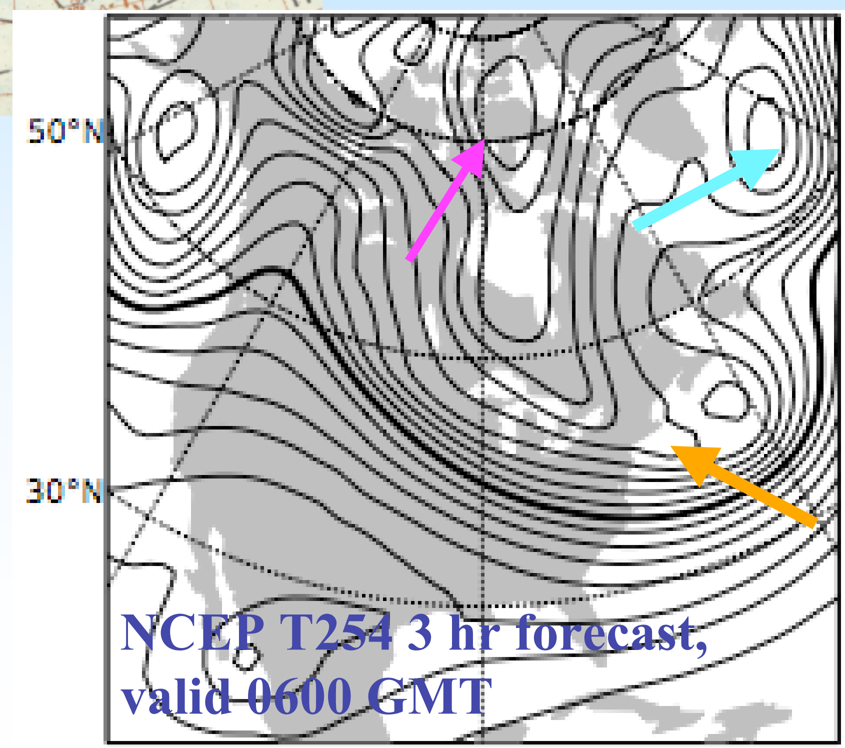
500 hPa geopotential height, 27 December 1947, record New York City snowfall



Ensemble Filter analysis better than NCEP 3D-VAR

5500 m (18000 ft) contour is thickened





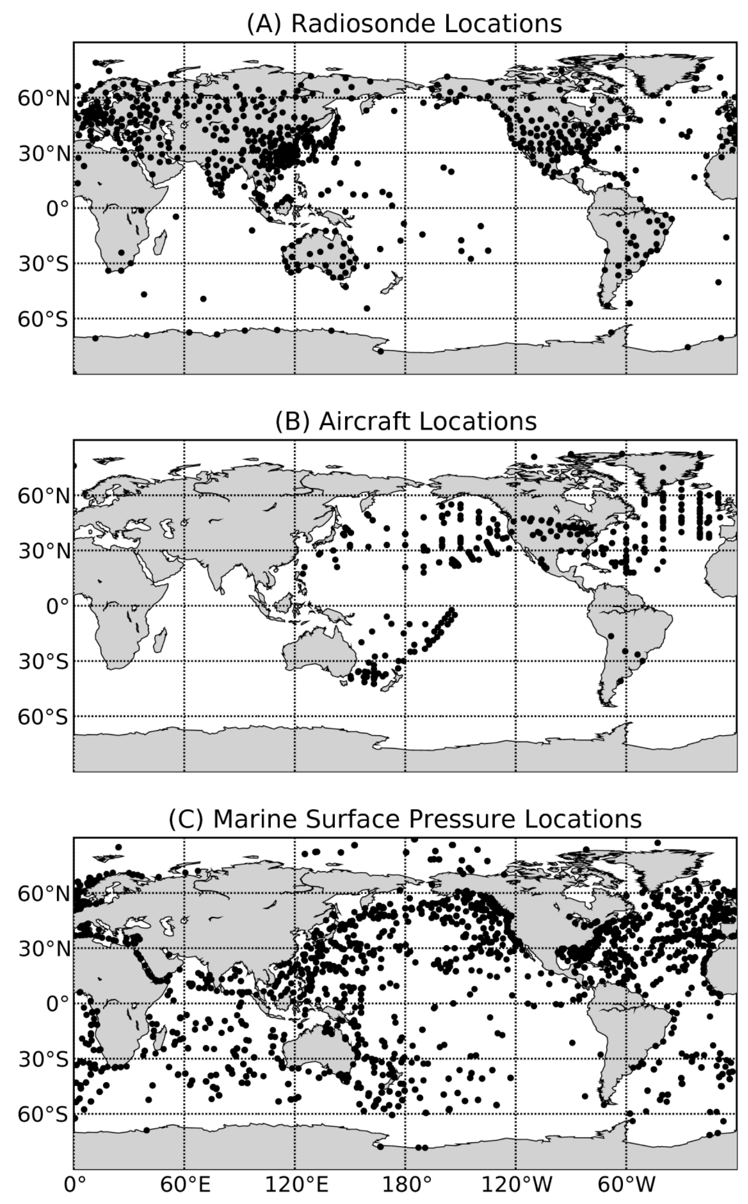
Motivation for GFS real-data experiment

- How do ensemble-based data assimilation algorithms compare with existing NCEP 3D-Var with *full current observational data set?*
- **Problem**: At NCEP's current operational (T254) resolution, too expensive for us to assimilate radiances while running on research computers.
- **Compromise**: compare against 3D-Var in reduced-resolution (T62) model with all observations *except* satellite radiances.

Experiment design

- **Model**: NCEP GFS, T62 L28, March 2004 physics. 100 members.
- **Observations**: *Almost all non-radiance data*; raobs, ACARS, profilers, cloud-drift winds, surface observations.
  - 200K observations @ 1200 UTC, 100K@ 1800 UTC
  - Surface pressure observations adjusted to model's orography
  - No non-surface pressure observations below $\sigma = 0.9$
  - Same observation error statistics as NCEP 3D-Var
  - Assimilate every 6 h, time-interpolate background to obs time if asynoptic
- **Period of test**: January 2004; throw out the first week as spin-up.
- **Compare against**:
  - T62 3D-Var with March 2004 GFS code, data specified above.
  - Operational T254 3D-Var analysis with all data

Observation locations



Ensemble Square-Root Filter (EnSRF)

$$\mathbf{X} = (\mathbf{x}_1^b - \overline{\mathbf{x}^b}, \ldots, \mathbf{x}_n^b - \overline{\mathbf{x}^b})$$

$$\mathbf{P}^b = \rho \circ \frac{1}{n-1}\mathbf{X}\mathbf{X}^{\mathrm{T}}$$

← background-error covariances estimated from ensemble, with "localization"

$$\mathbf{K} = \mathbf{P}^b\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}^b\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}$$

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{K}\left(\mathbf{y} - H(\bar{\mathbf{x}}^b)\right).$$

← Mean state updated, correcting background to new observations, weighted by **K**, the Kalman gain

$$\widetilde{\mathbf{K}} = \left(1 + \sqrt{\frac{\mathbf{R}}{\mathbf{H}\ \mathbf{P}^b\mathbf{H}^{\mathrm{T}} + \mathbf{R}}}\right)^{-1}\mathbf{K}.$$

$$\mathbf{x}_i'^a = \mathbf{x}_i'^b - \widetilde{\mathbf{K}}\left(H(\mathbf{x}_\mathbf{i}'^\mathbf{b})\right).$$

← "reduced" Kalman gain calculated to update perturbations around mean

$$\mathbf{x}_i^b(t+1) = M(\mathbf{x}_i^b(t) + e), e \sim N(0, \mathbf{Q})$$

↑ Forecast forward to the next time when data is available. Add noise in some fashion to simulate model error.

EnSRF details

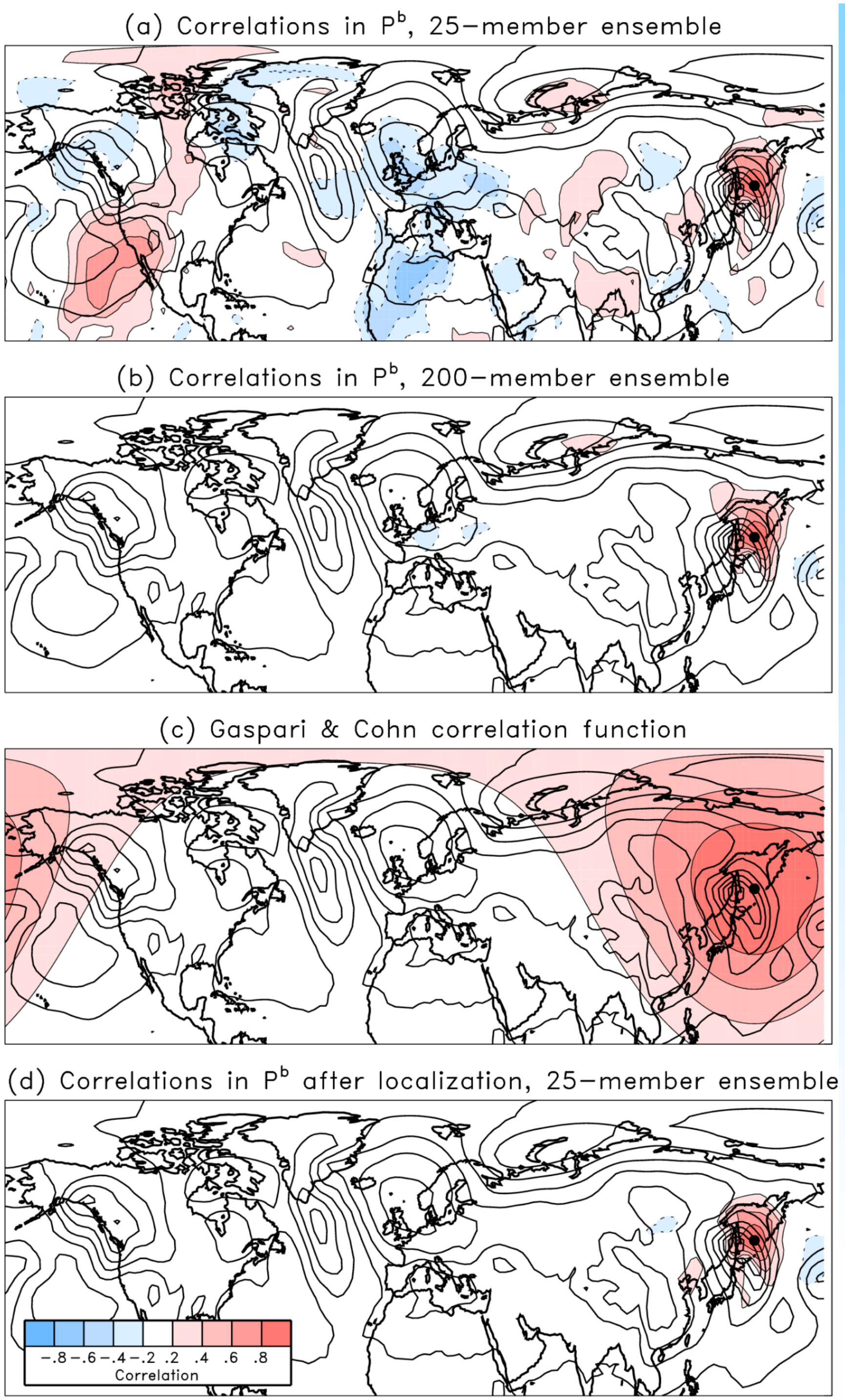
- **Covariance localization**
  - ***Horizontal***: Blackman window function, tapers to zero at 2800 km
  - ***Vertical***: Tapers to zero at 3 scale heights for surface pressure, 2 scale heights otherwise.
  - **Lynch filter** to control gravity-wave noise (3h forecast Gaussian-weighted average of 0-6 h forecast)

- **Influence check:** assimilate observation only if F-test shows that it will significantly reduce variance (>1 percent reduction from prior)

- **Model Error**:
  - ***Covariance inflation***, 30% NH, 24% SH, taper in between. Inflation amount tapers in vertical to 0.0 at 6 scale heights (problem with top boundary).
  - ***Relaxation to prior***: Snyder and Zhang (MWR, 2003), relax analysis ensemble back toward prior (15% analysis, 85% prior). $x'^a = cx'^a + (1-c)x'^b$
  - ***Additive errors***, random 6-h model tendencies scaled by 33 %. Samples from NCEP-NCAR reanalysis, '71-'00, for similar time of the year.

Covariance Localization

A way of dealing with inappropriate covariance estimates due to small ensemble size. Increases dimensionality of background-error covariance.



Ordinary EnSRF cycle (serial processing)

Loop over analysis times:

- run 6-h forecast for each ensemble member from the previous analysis

Loop over observations:

- Do we need this ob? (will it significantly reduce ensemble uncertainty estimate?) If not continue to Next observation
- Update the ensemble mean and the ensemble perturbations using the KF update equations.

End loop over observations

- Add variance to account for errors outside span of ensemble

End loop over analysis times

Revised "Local" LEnSRF cycle (more parallelized processing)

Loop over analysis times (every 6 h):

- Run 9-h forecast for each ensemble member from the previous analysis
- Compute $H\overline{\mathbf{x}}^b,\ H(\mathbf{x}^b)'$ at every observation location between 3 and 9 h (linear interpolation of background in time).
- Divide up state vector elements, randomly shuffled among processors

Loop over each state vector element on each processor

Loop over observations within localization radius of this grid point.

- Do we need this ob? (will it significantly reduce ensemble uncertainty estimate?) If not continue to Next observation
- Update the ensemble mean, perturbations, as well as $H\overline{\mathbf{x}}^b,\ H(\mathbf{x}^b)'$ using the KF update equations.

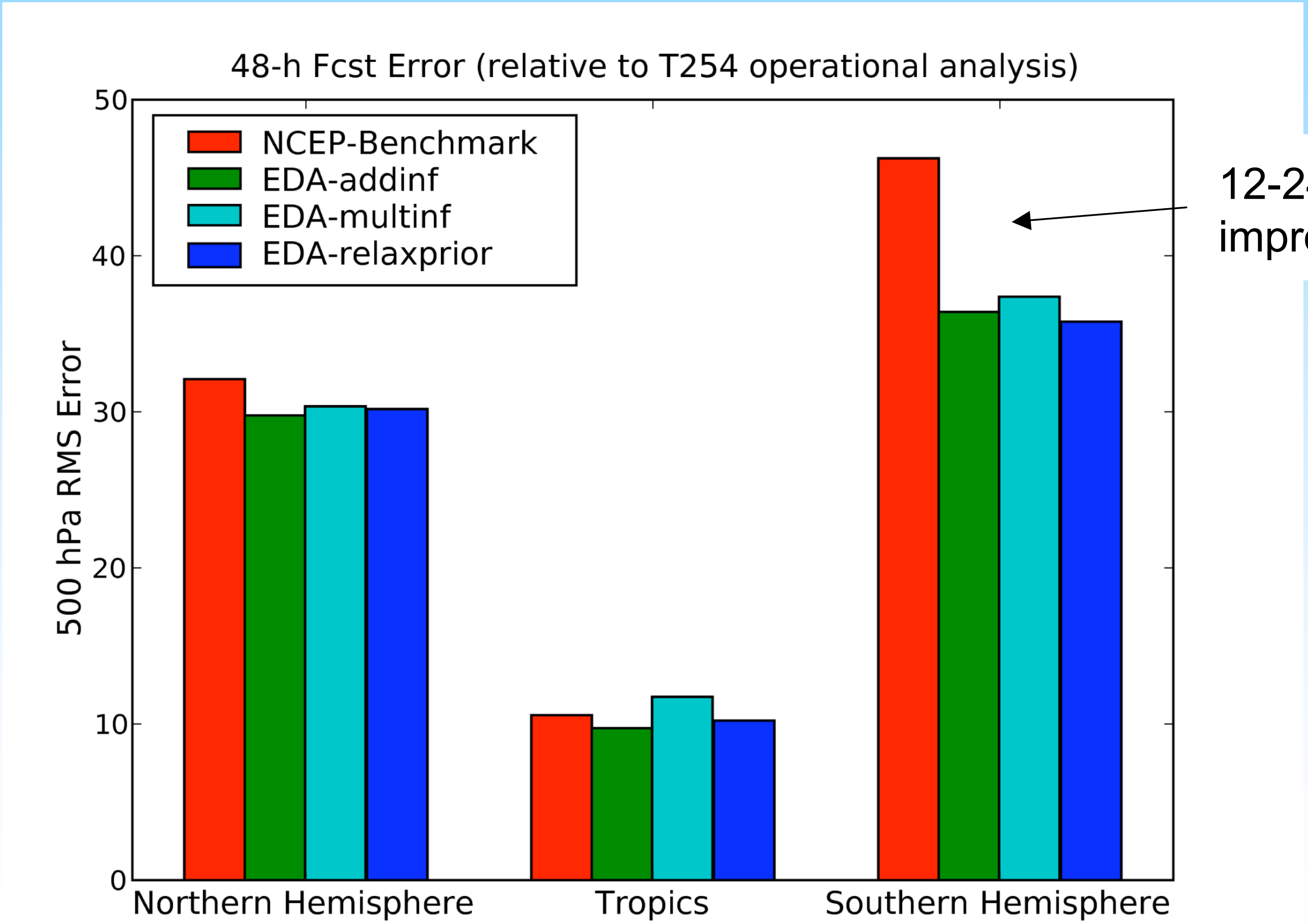
End loop over observations

End loop over state vector elements

- Add variance to account for errors outside span of ensemble.

End loop over analysis times

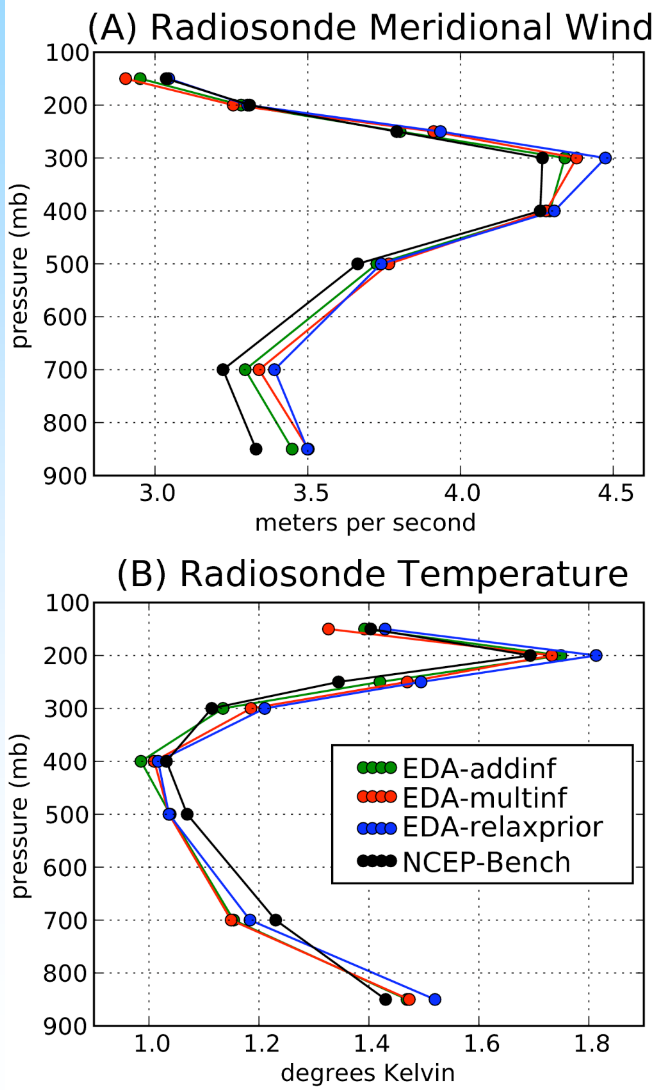
Comparison of model-error parameterizations, T62 GFS (500 hPa Z)



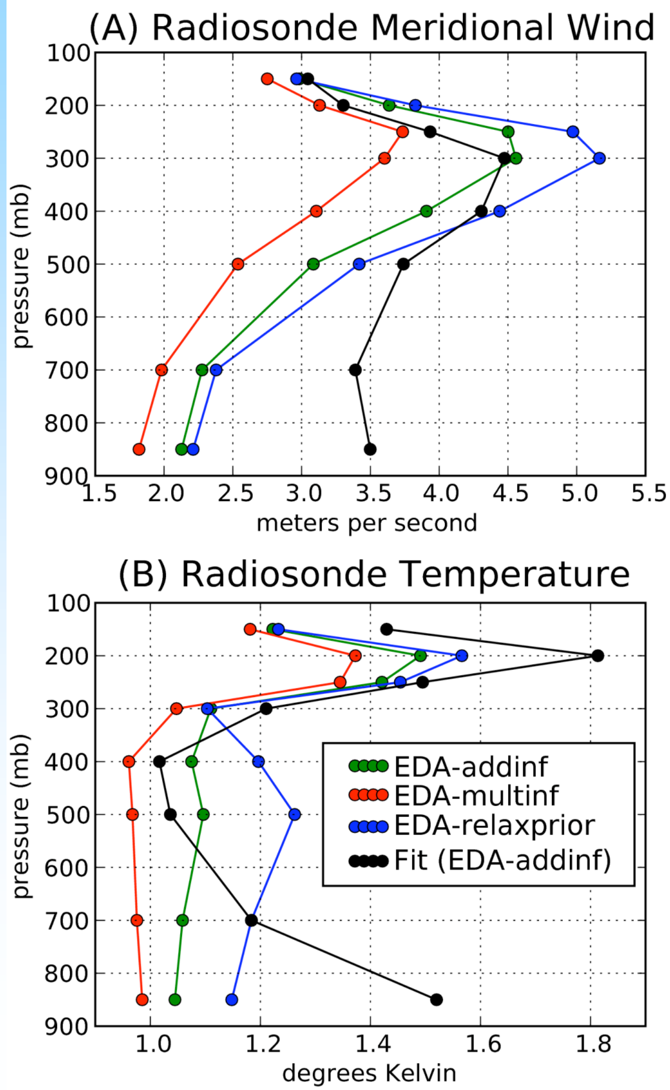
12-24 h improvement

(Note: T255 analysis truncated to T62 for verification)

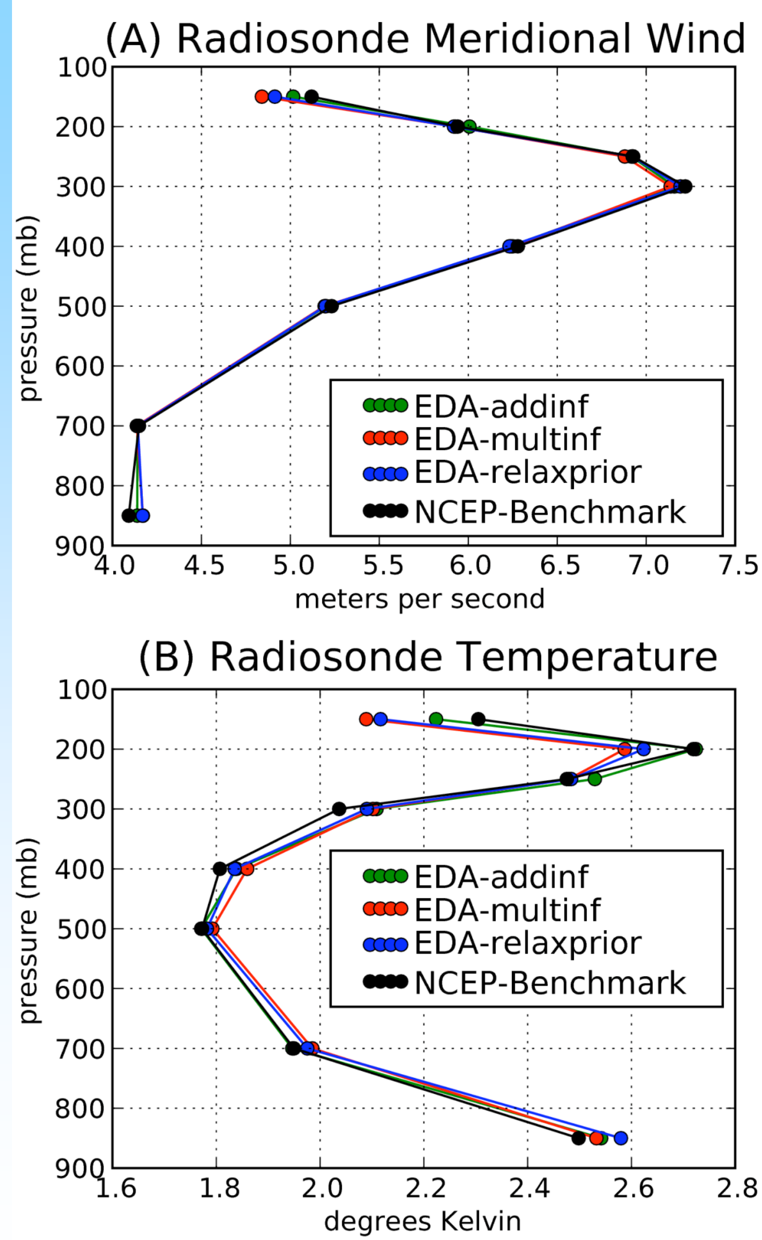
6-h forecast fit to observations



6-h forecast spread



Fit to observations, 48-h forecast



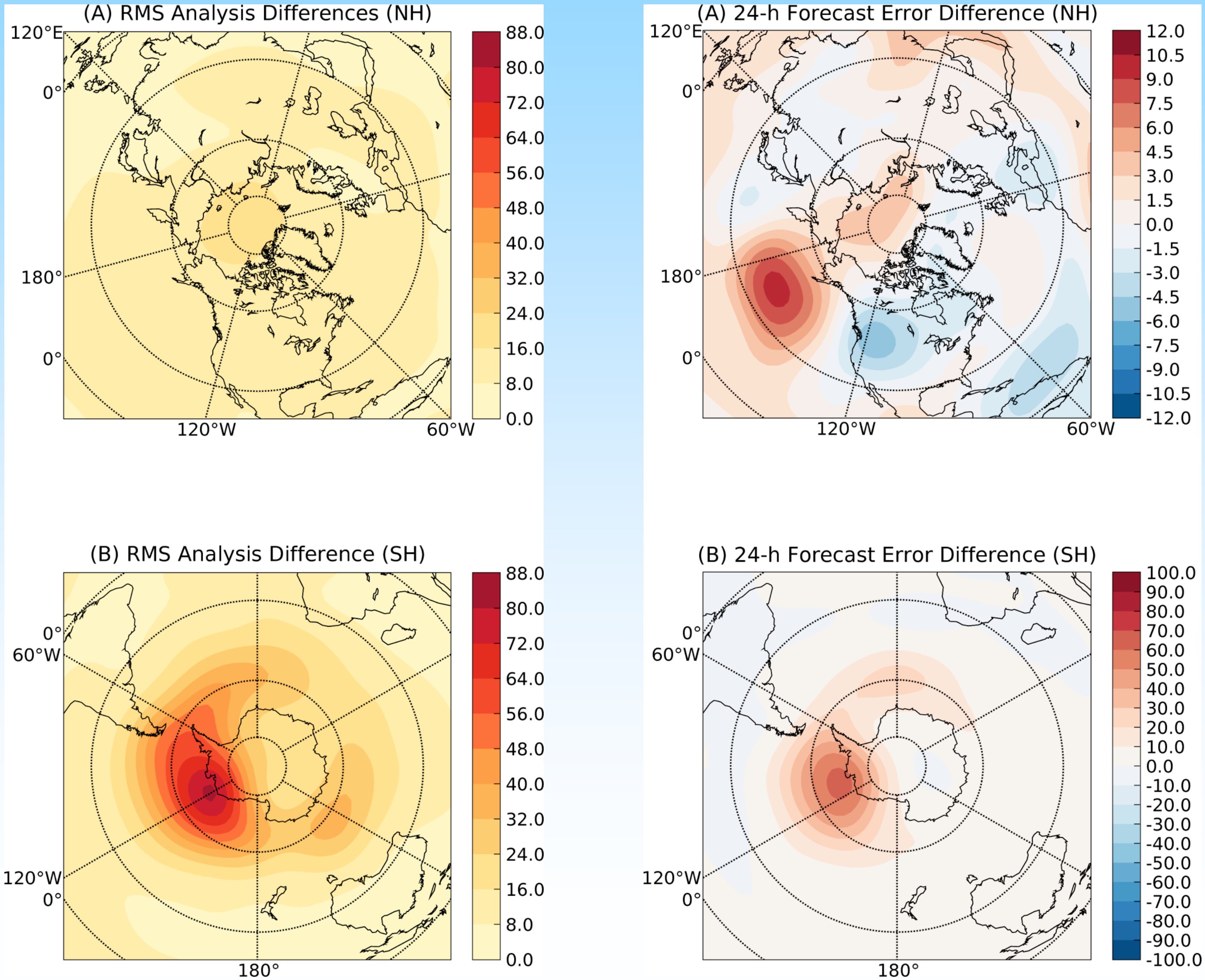
Fits to aircraft data and marine surface-pressure observations

	EDA-addinf	EDA-multinf	EDA-relaxprior	NCEP-Benchmark
aircraft meridional wind (m/s)	7.47 (88%)	7.44 (93%)	7.38 (74%)	7.62
surface marine pressure (hPa)	3.18 (99%)	3.24 (99%)	3.21 (99%)	3.45

Table 1: Fits of 48-h forecasts initialized from 00 UTC and 12 UTC EDA ensemble mean and NCEP-Benchmark analyses to aircraft (airep and pirep) observations between 300 and 150 hPa and surface marine (ship and buoy) pressure observations for 2004010800 to 2004012900. Only those observations between one hour 00 UTC and 12 UTC are used. Three different EDA analyses are shown, each employing a different method for parameterizing system error (see text for details). The confidence level (in percent) for the difference between each of the EDA results and the NCEP-Benchmark result is also given. This confidence level is computed using a two-tailed t-test for the difference between means of paired samples, taking into account serial correlations in the data.

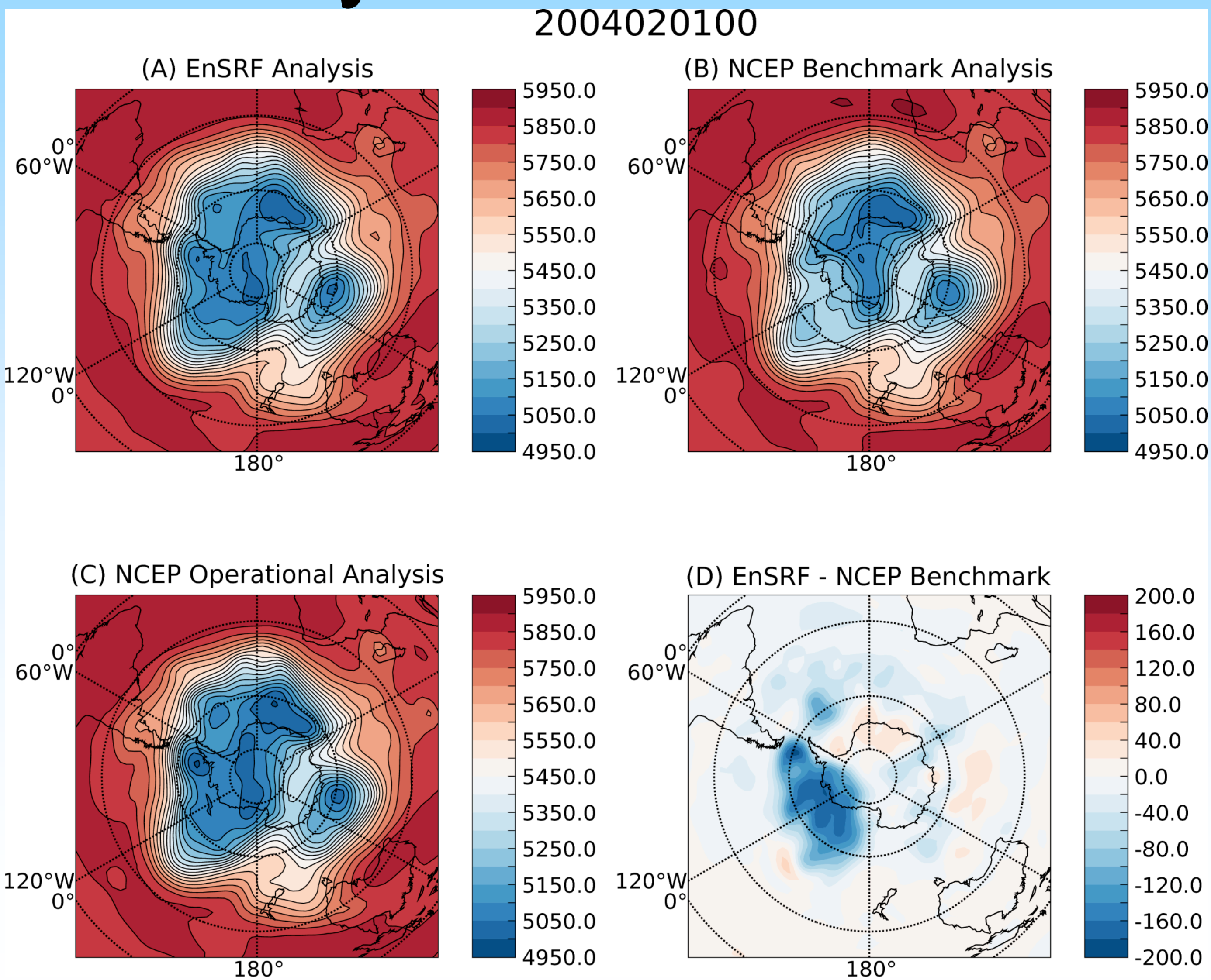
More noticeable difference in fit between EDA and benchmark here; more data-sparse area.

Where are the differences largest? (benchmark - EDA/addinf)

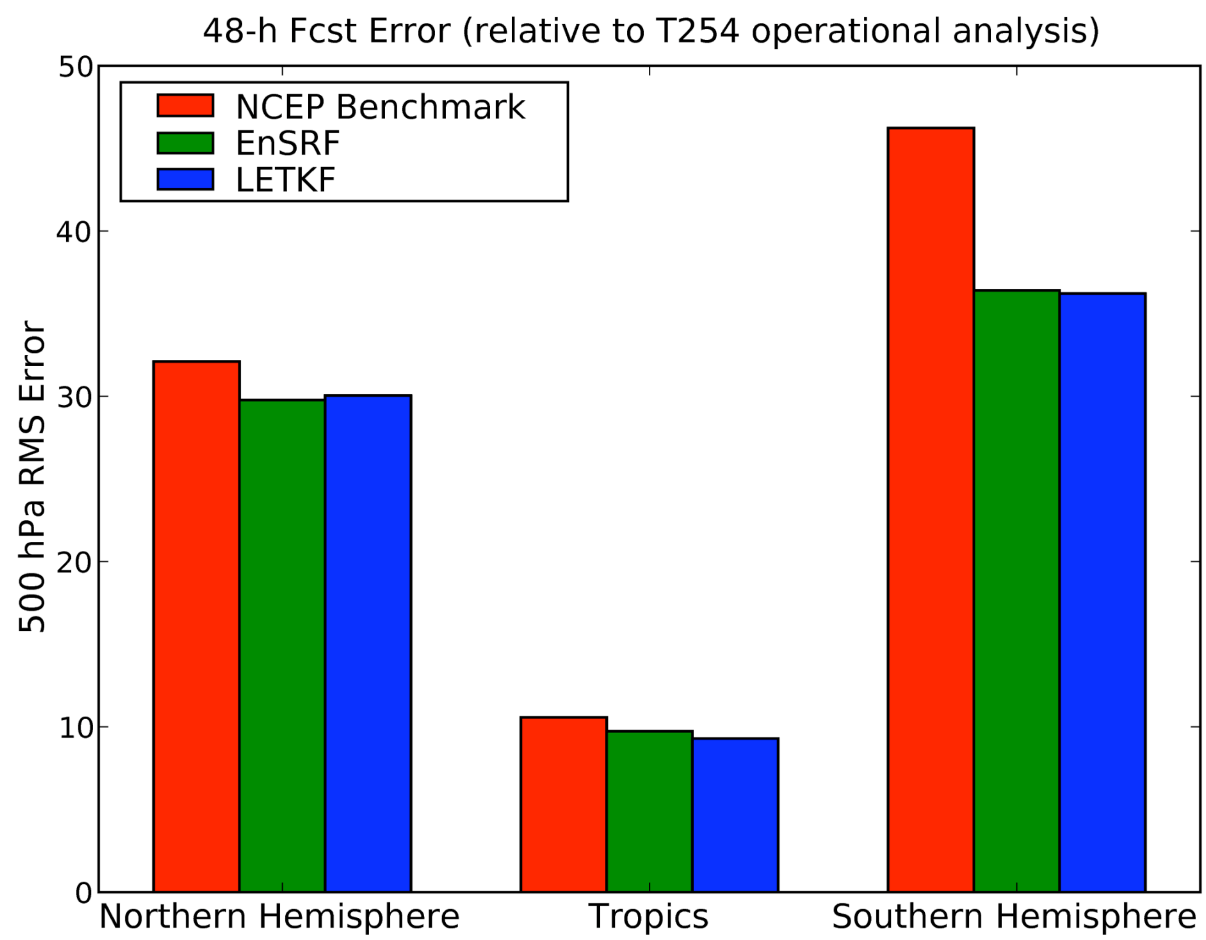


Note: T12 Gaussian smoother used

Example of Southern Hemisphere analysis differences



Comparison to University of MD's Local Ensemble Transform Kalman Filter



(very preliminary, very optimistic since their algorithm faster)

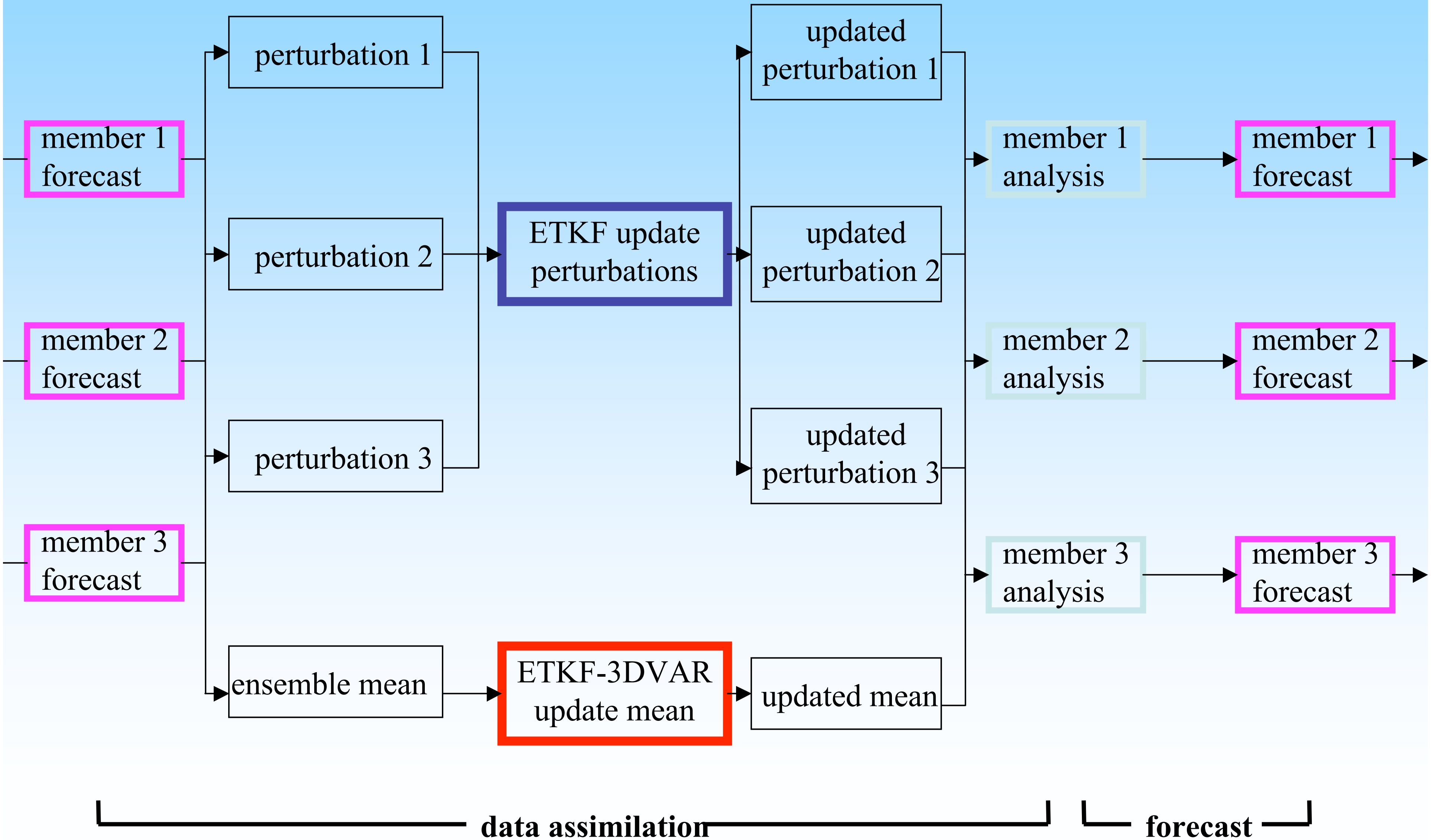
Conclusions on GFS real-data experiments

- Experimental EnSRF similar to (data dense areas) or outperforms (data-sparse areas) operational 3D-Var run at same resolution with same subset of observations. The sparser the network, the bigger the advantage for EnSRF (SH, historical reanalysis).

- Additive model error parameterization works slightly better than alternatives.

- Next:
  - More exploration of U. Maryland's Local ETKF
  - Assimilate radiances
  - Techniques for 'super-obbing'
  - Model error : include bias correction (never done!)
  - Parallel testing on NCEP machine?

Hybrid ETKF-3DVAR (Xuguang Wang)



Hybrid ETKF / 3D-Var : Why?

- Hybrid method is less expensive than EnSRF, while still benefitting from ensemble-estimated error statistics.
  - Costs don't necessarily scale linearly with number of observations
  - No parallel data assimilation cycles, just update of mean and computationally efficient rotation/rescaling of perturbations with ETKF.
- The hybrid may be more robust for small ensemble size, since can adjust the amount of 3D-Var vs. ensemble covariance used.
- The hybrid can be conveniently adapted to the existing variational framework.

Hybrid ETKF/ 3D-Var update of mean

$$J(\mathbf{x}) = \frac{1}{2}\left(\mathbf{x}-\mathbf{x}^b\right)^T\left(\mathbf{P}^b\right)^{-1}\left(\mathbf{x}-\mathbf{x}^b\right) + \frac{1}{2}\left(H(\mathbf{x})-\mathbf{y}\right)^T\mathbf{R}^{-1}\left(H(\mathbf{x})-\mathbf{y}\right)$$

- Background error covariance is approximated by a linear combination of the sample covariance matrix of the ETKF forecast ensemble and the static covariance matrix.

$$\mathbf{P}^b = (1-\alpha)\mathbf{P}^e + \alpha\mathbf{B}, \qquad 0 \le \alpha \le 1$$

$$\mathbf{P}^e = \rho \circ \left(\frac{1}{n-1}\mathbf{X}^b\left(\mathbf{X}^b\right)^T\right)$$

$$\mathbf{X}^b = \left(\mathbf{x}_1^b - \overline{\mathbf{x}}^b, \ldots, \mathbf{x}_n^{'} - \overline{\mathbf{x}}^b\right)$$

- Can be conveniently adapted into the operational 3D-Var through augmentation of control variables.

ETKF update of perturbations

- ETKF transforms forecast perturbations $\mathbf{X}^b$ into analysis perturbations $\mathbf{X}^a$ by

$$\mathbf{X}^a = \mathbf{X}^a\mathbf{T}$$

where $\mathbf{T}$ is chosen by trying to solve the Kalman filter error covariance update equation, with forecast error covariance approximated by ensemble covariance.

- Latest formula for $\mathbf{T}$ (Wang et al. 2004;2005, MWR)

$$\mathbf{T} = \mathbf{C}\left(\rho\Gamma + \mathbf{I}\right)^{-1/2}\mathbf{C}^T$$

$\mathbf{C}$ and Γ contain eigenvectors and eigenvalues of the nxn (ens. size) matrix

$$\frac{1}{n-1}\left(\mathbf{X}^b\right)^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{X}^b$$

ρ is the estimated fraction of forecast error variance projected onto ensemble subspace

- Computationally inexpensive for ensemble size of o(100), because tranformation fully in subspace of perturbations.

Update of mean (OI) and formation of $\mathbf{B}$

- In this experiment, the hybrid updates the mean like OI, using

$$\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^b + \mathbf{P}^b\mathbf{H}^T\left(\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R}\right)^{-1}\left(\mathbf{y} - \mathbf{H}\overline{\mathbf{x}}^b\right)$$

$$\mathbf{P}^b\mathbf{H}^T = (1-\alpha)\mathbf{P}^e\mathbf{H}^T + \alpha\mathbf{B}\mathbf{H}^T$$

$$\mathbf{H}\mathbf{P}^b\mathbf{H}^T = (1-\alpha)\mathbf{H}\mathbf{P}^e\mathbf{H}^T + \alpha\mathbf{H}\mathbf{B}\mathbf{H}^T$$

This is equivalent to the variational solution under our experiment design,which will assume normality of errors.

- The static error covariance model is constructed iteratively from a large sample of 24h fcst. errors.

First estimate $\mathbf{BH}^T$ and $\mathbf{HBH}^T$ is constructed from 250 24h fcst. errors with covariance localization (iteratively)

→ Run a huge number of DA cycles (7000 > number of model dimension) with $\mathbf{BH}^T$ and $\mathbf{HBH}^T$

→ Construct $\mathbf{BH}^T$ and $\mathbf{HBH}^T$ with this huge sample of 24h forecast errors (→ back to Run a huge number of DA cycles)

Experiment design

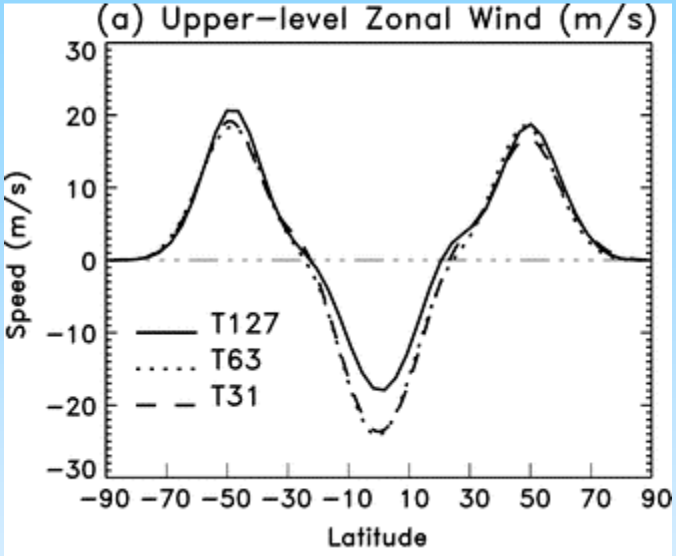
- Numerical model

➢Dry 2-layer spectral PE model run at T31;

➢Model state consists of vorticity, divergence and layer thickness of Exner function

➢Error doubling time is 3.78 days at T31

➢Perfect model assumption



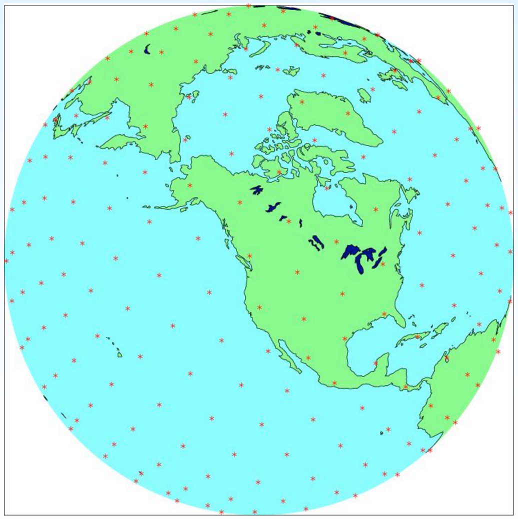
(Hamill and Whitaker MWR, 2005)

- Observations

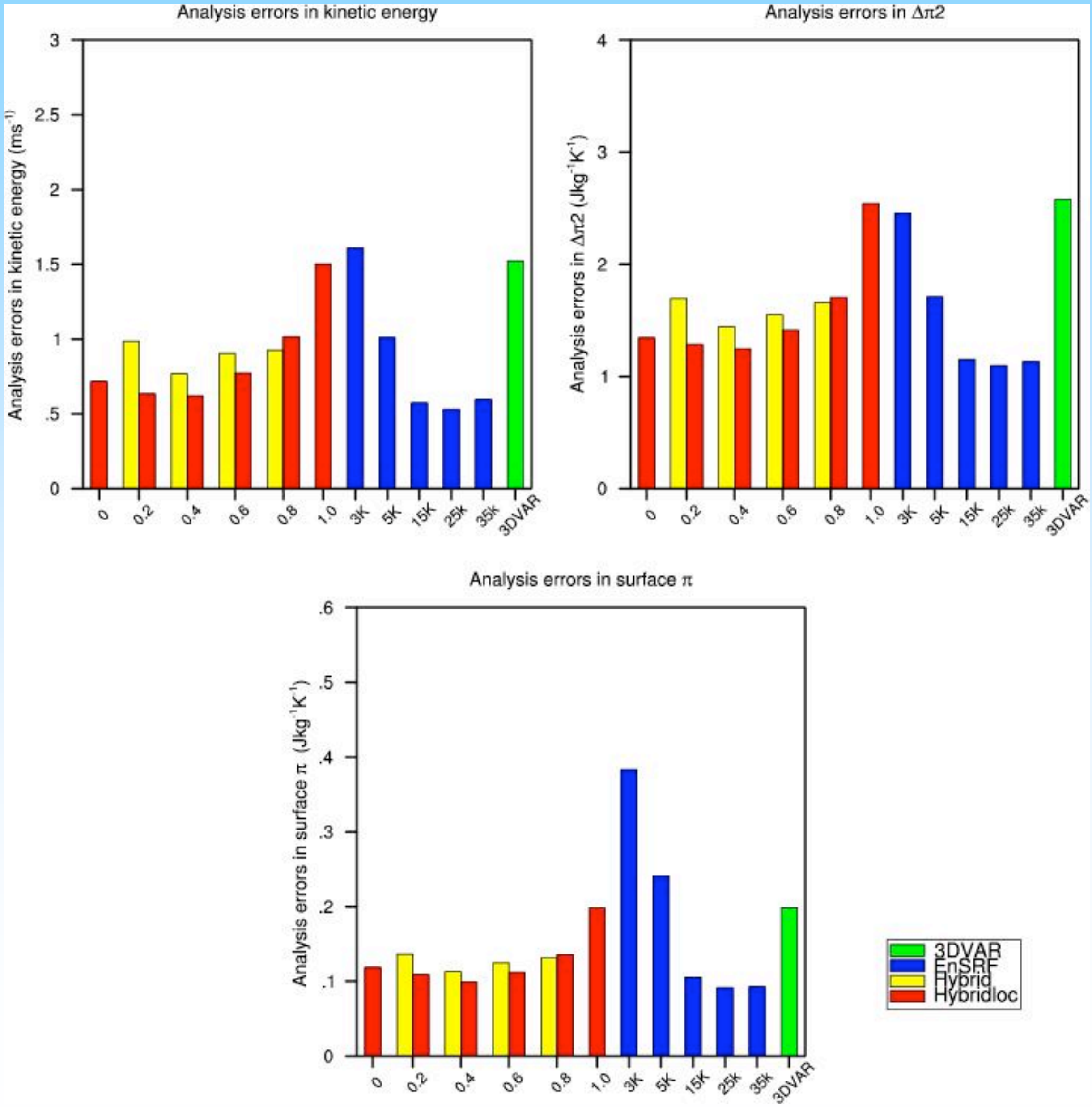
➢362 Interface and surface Exner functions taken at equally spaced locations

➢Observation values are T31 truth plus random noise drawn from normal distribution; RMS error equivalent to approximately 1K, 1 hPa.

➢Assimilated every 24 h



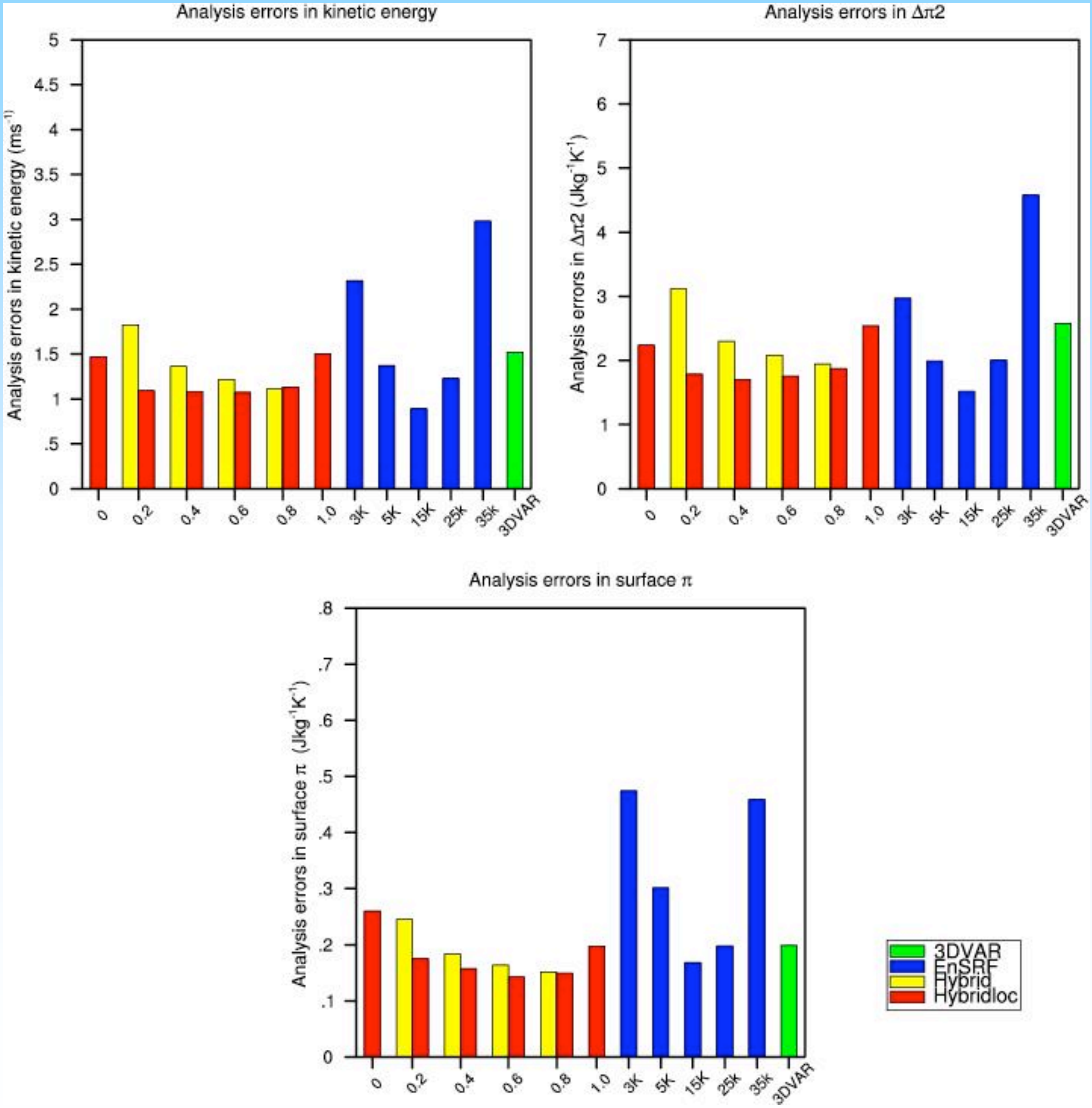
RMS analysis errors, 50 members



- Improved accuracy of EnSRF over 3D-Var can be mostly achieved by the hybrid.

- Covariance localization applied on the ETKF ensemble when updating the mean (but *not* applied when updating the perturbations) improved the analyses of the hybrid.

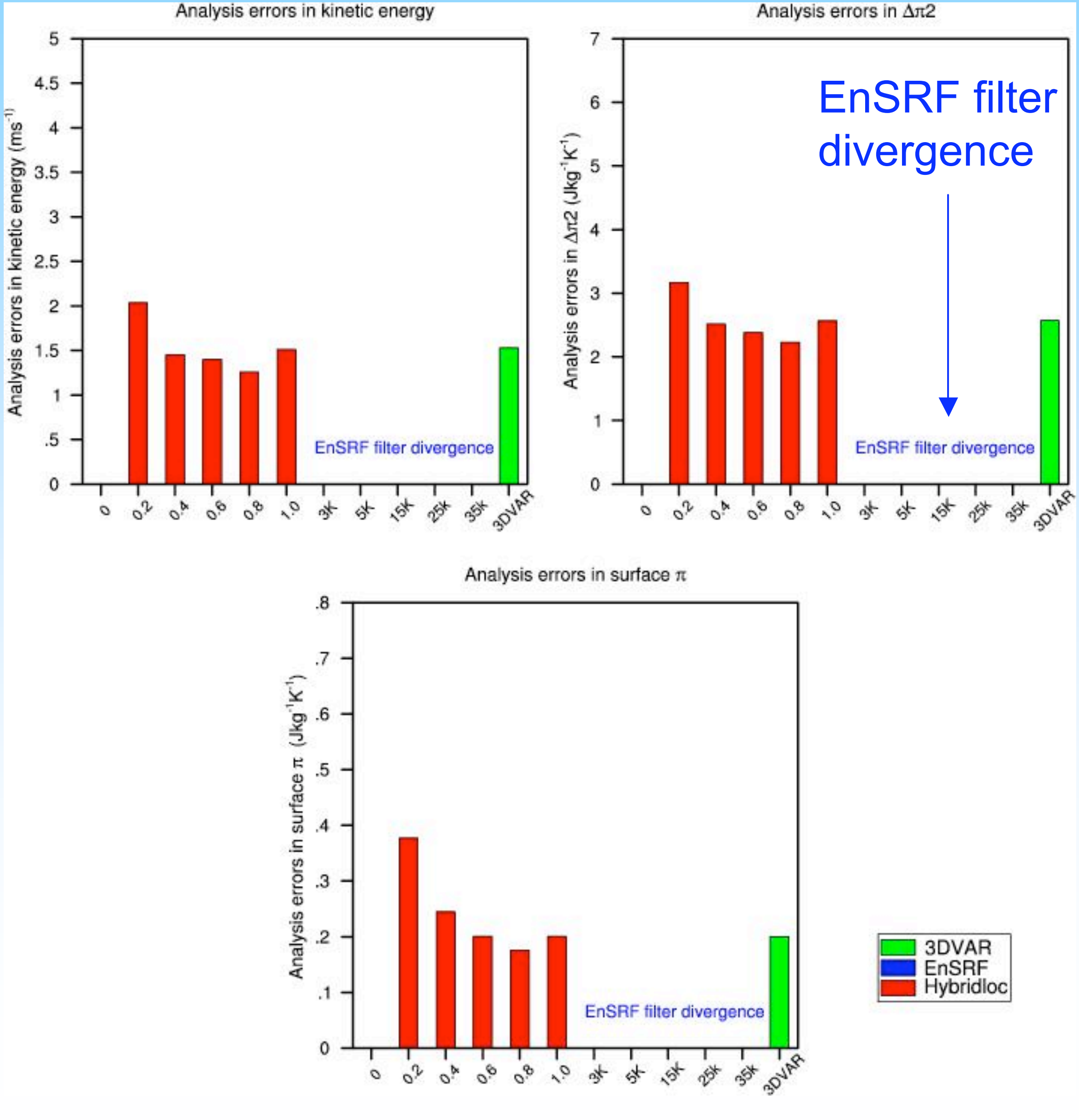
RMS analysis errors, 20 members



- Both 20-member hybrid and EnSRF worse than 50-member, but still better than 3D-Var

- Hybrid nearly as accurate (KE, $\Delta\pi_2$ norms) or even better (surface π norm) compared to EnSRF

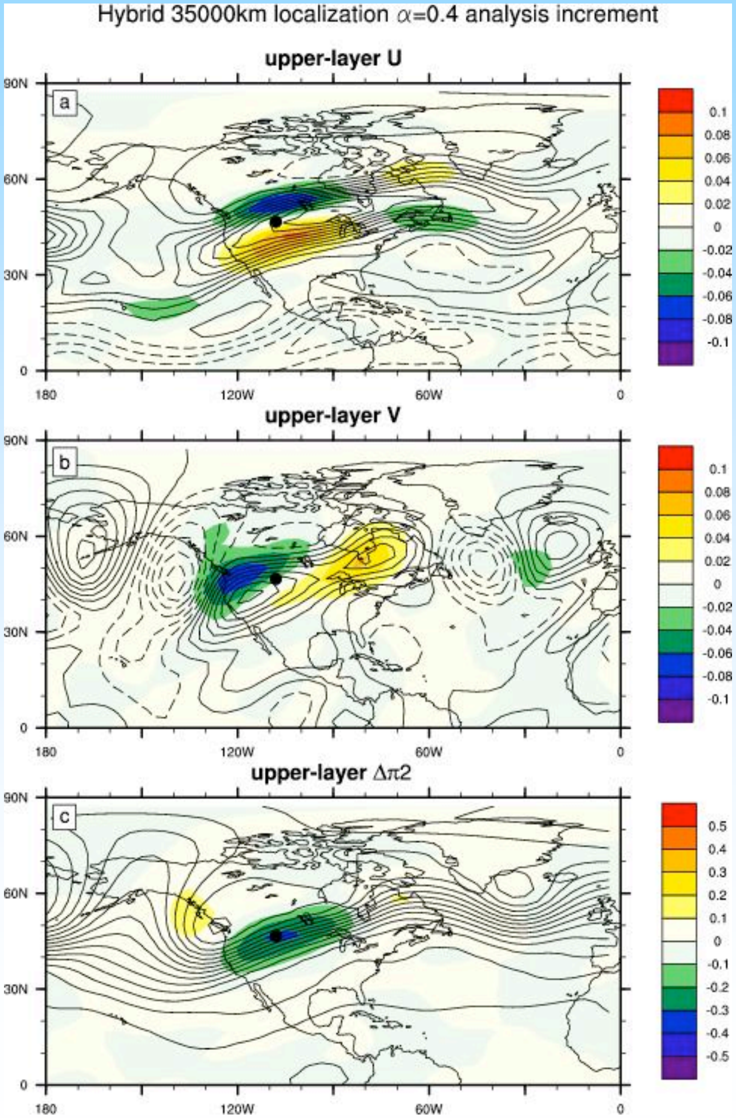
RMS analysis errors, 5 members



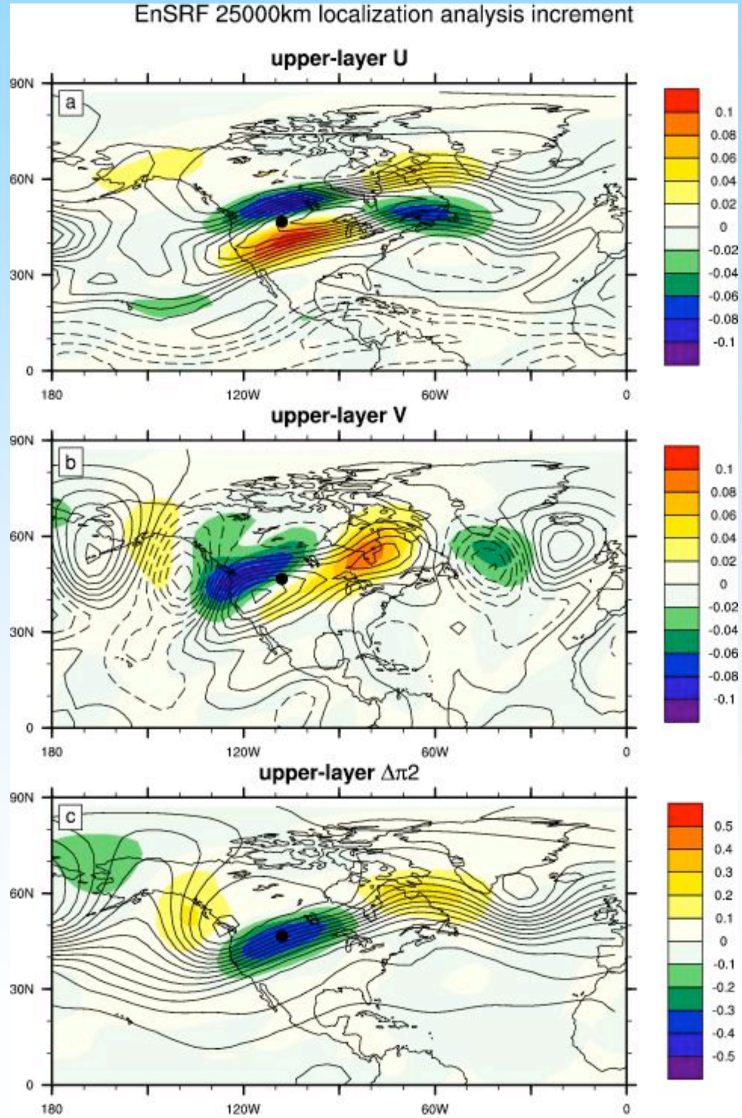
- EnSRF experienced filter divergence for all localization scales tried.
- Hybrid was still more accurate than the 3D-Var.
- Hybrid is more robust in the presence of small ensemble size.

Comparison of flow-dependent background-error covariance models

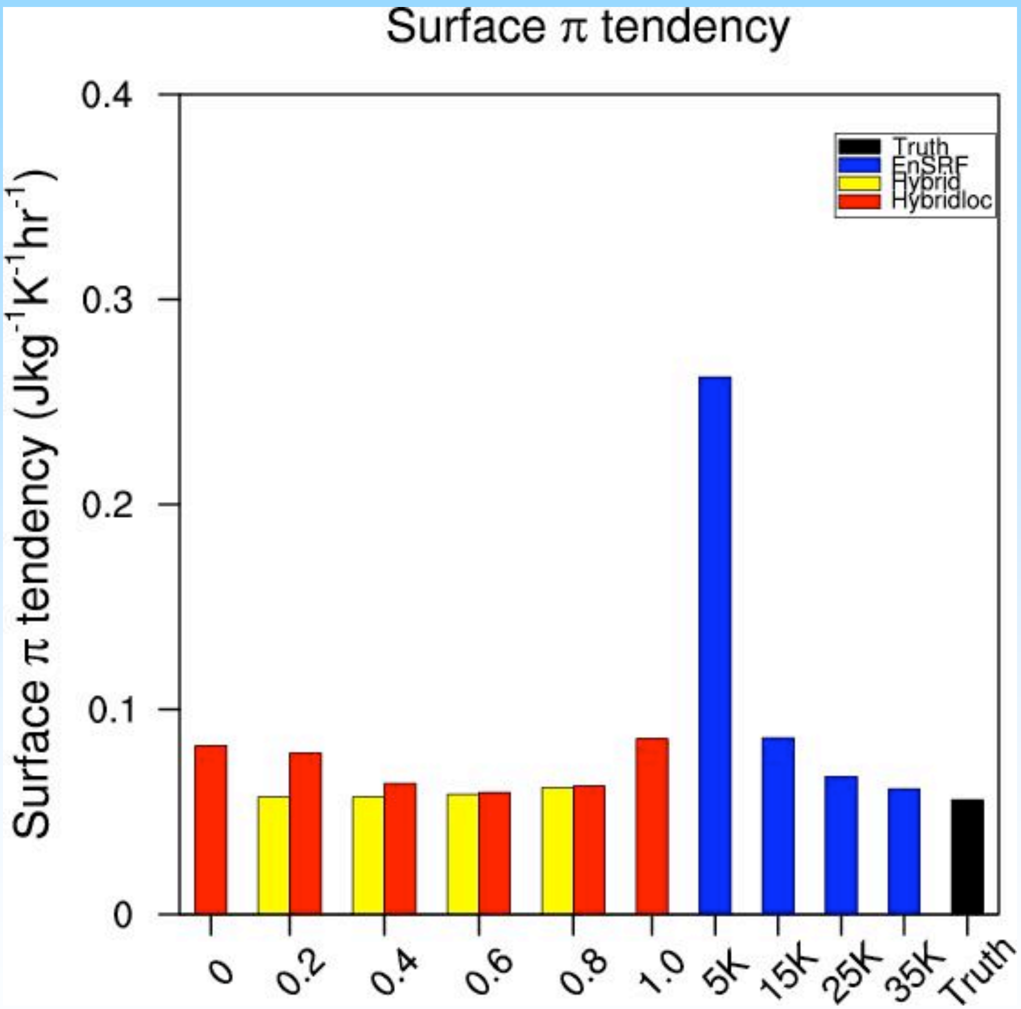
(Hybrid 50 mem. result)



(EnSRF 50 mem. result)



Initial-condition balance

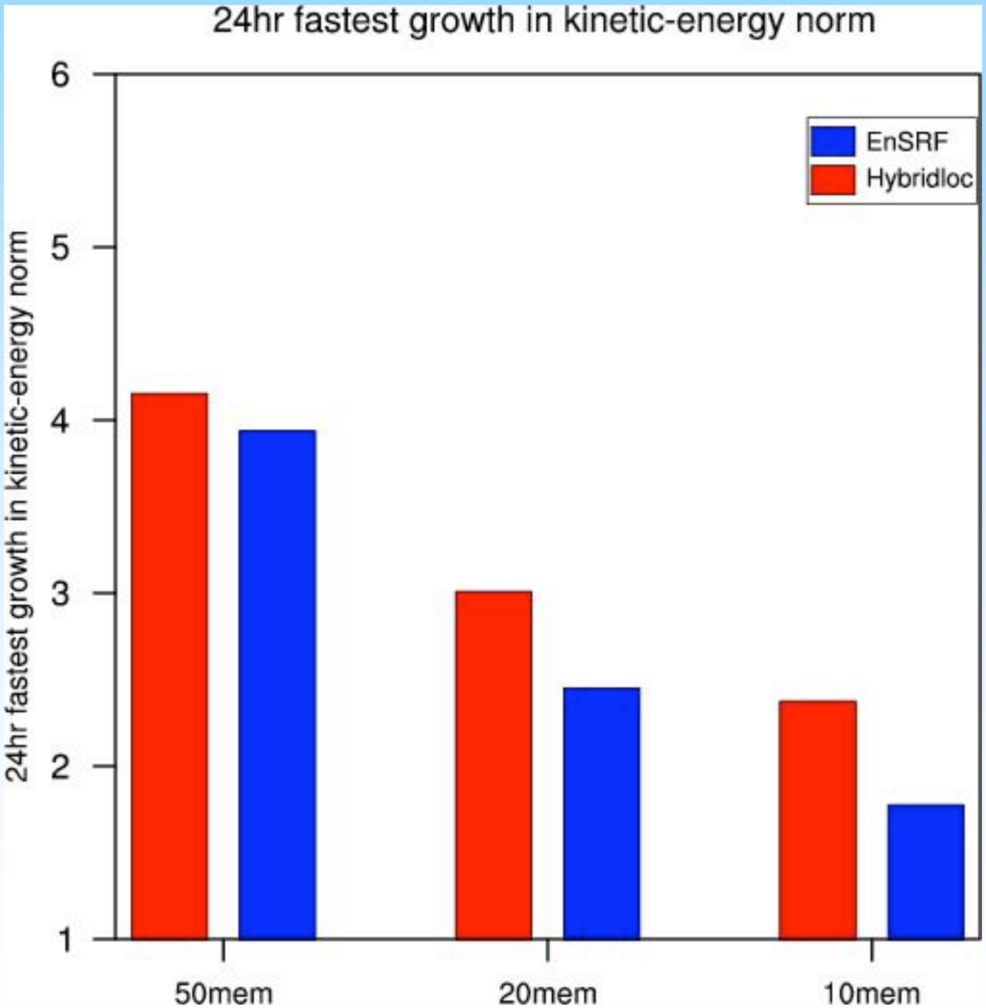


(50-member result)

- Analysis is more imbalanced with more severe localization.

- Analyses of the hybrid with the smallest rms error are as balanced or more balanced than those of the EnSRF, especially for small ensemble size (not shown).

Maximal perturbation growth in subspace of ensemble

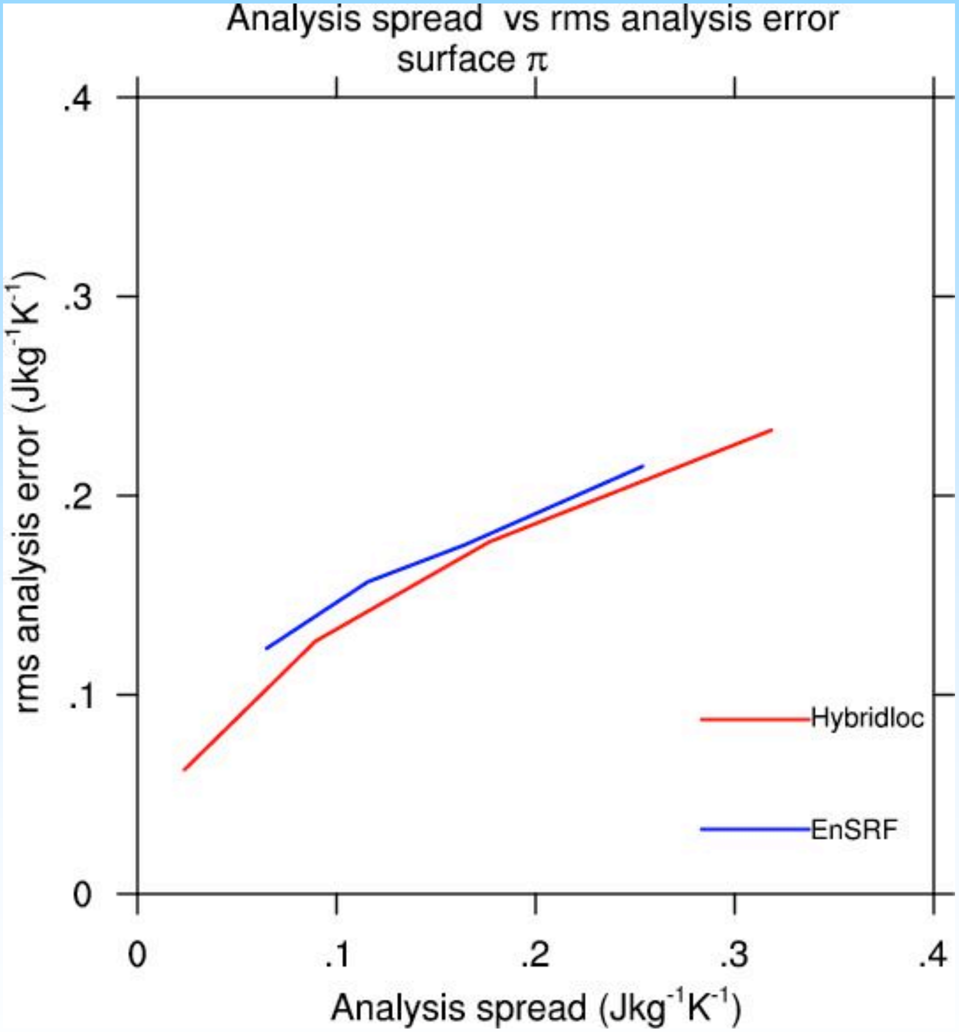


- Find linear combination coefficients $\mathbf{b}$ to maximize

$$\frac{\mathbf{b}^T (\mathbf{X}^b)^T \mathbf{S} \mathbf{X}^b \mathbf{b}}{\mathbf{b}^T (\mathbf{X}^a)^T \mathbf{S} \mathbf{X}^a \mathbf{b}}$$

- Maximal growth in the ETKF ensemble perturbation subspace is faster than that in the EnSRF ensemble perturbation subspace; initially unbalanced EnSRF perturbations grow slower?

Spread-skill relationships



(20-member result)

- Overall average of the spread is approximately equal to overall average of rms error for both EnSRF and hybrid.

- Abilities to distinguish analyses of different error variances are similar for EnSRF and hybrid.

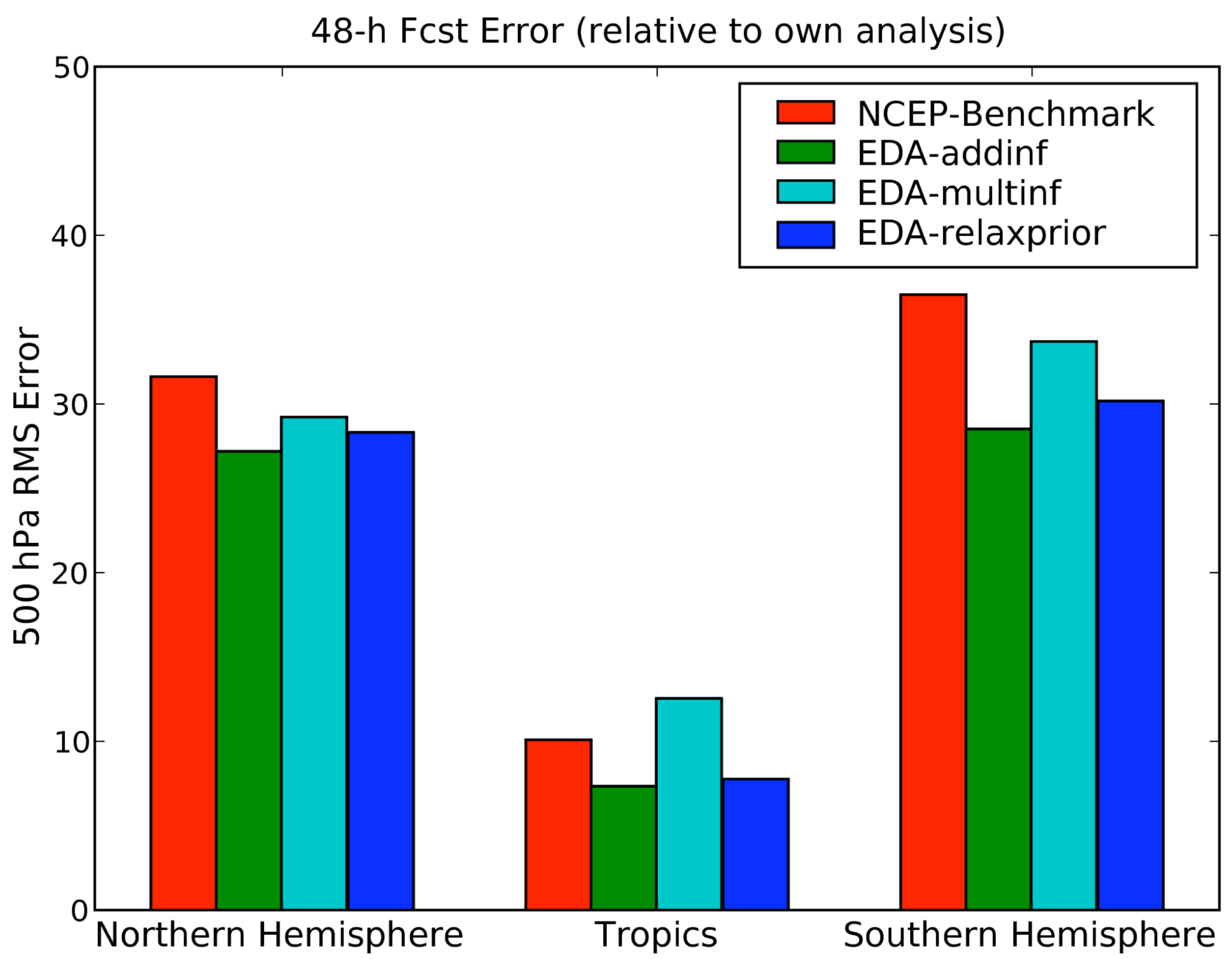
Summary of ETKF/3D-Var hybrid

- The hybrid analyses achieved similar improved accuracy of the EnSRF over 3D-Var.
- The hybrid was more robust when ensemble size was small.
- The hybrid analyses were more balanced than the EnSRF analyses.
- The ETKF ensemble maximal growth was faster than the EnSRF.
- The ETKF ensemble variance was as skillful as the EnSRF.
- The hybrid can be conveniently adapted into the existing operational 3D-Var framework.
- The hybrid is expected to be less expensive than the EnSRF.

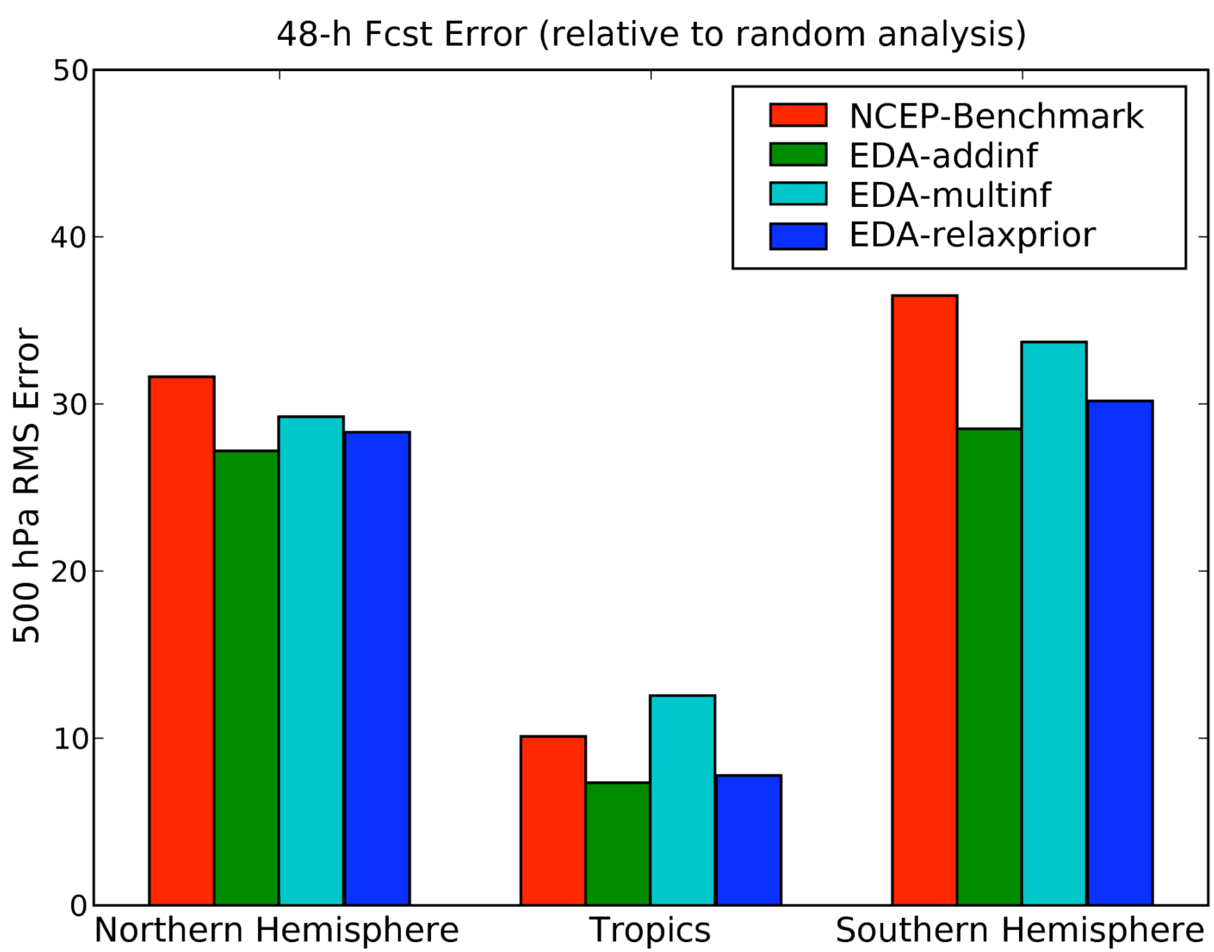
Upcoming work (we hope)

- Build WRF and GFS hybrids
  - Test WRF in tropical cyclones (data sparse, unusual background-error covariances)
- Extend the idea of the hybrid ETKF-3DVAR to the 4D-Var framework.

48-h error relative to own analysis



48-h error verifying against random analysis



Bottom Line - *T126 EnSRF without radiances close to T254 Operational 3D-Var with radiances in Northern Hemisphere*