DEPARTMENT OF ATMOSPHERIC & OCEANIC SCIENCE

Adaptive observation experiments with 3D-Var and Ensemble Kalman Filtering: Implications for intermittent lidar wind observations

Junjie Liu and Eugenia Kalnay

Feb. 7, 2007

UNIVERSITY OF MARYLAND

Thanks!

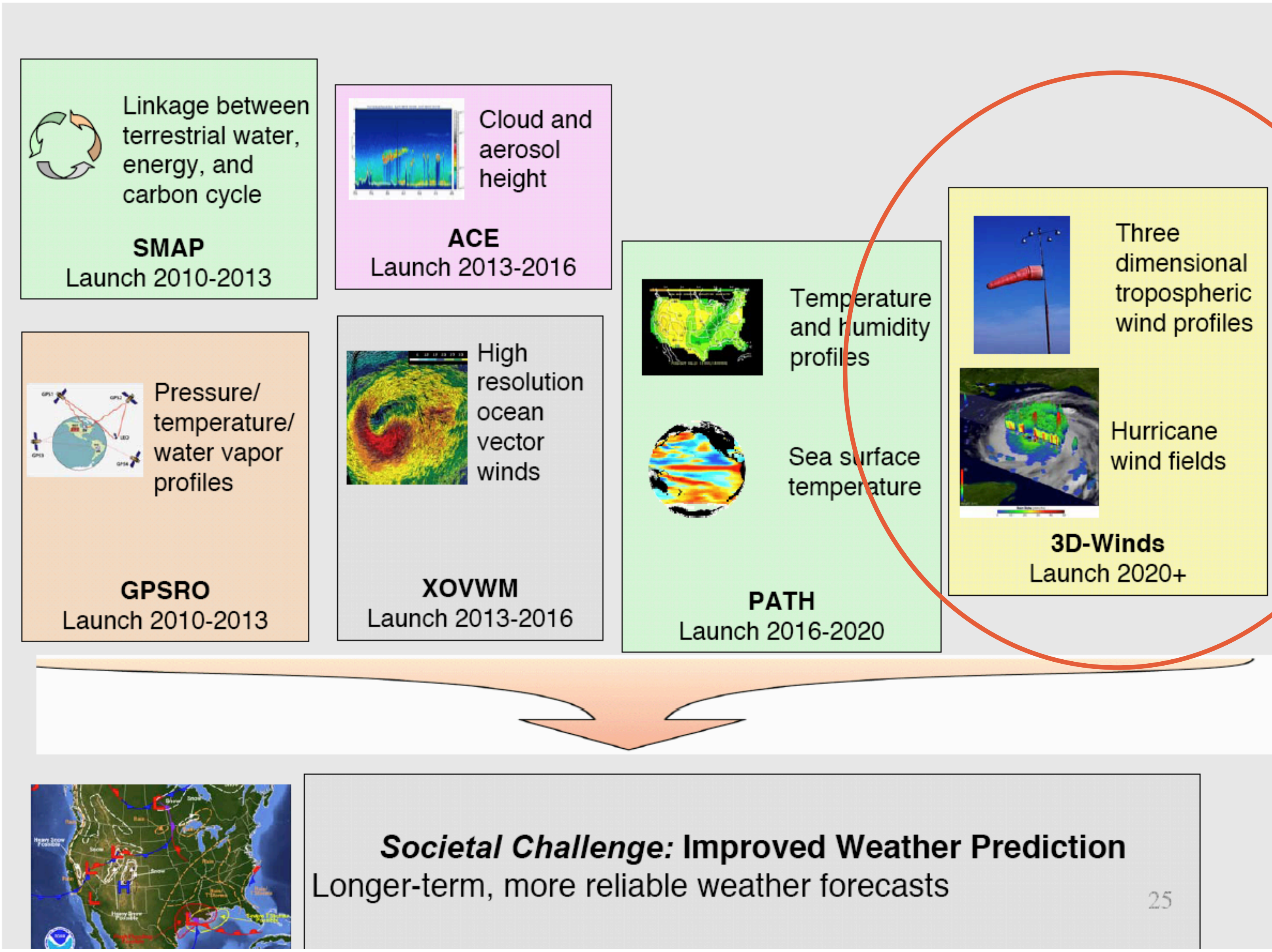
Background

- Lack of wind observations is a major factor for NWP uncertainty, especially over tropics and ocean.

- Future Doppler Wind Lidar (DWL, Stoffelen et al., 2005; "Earth Science and Applications from Space National Imperatives for the Next Decade and Beyond" recommendations) will provide much denser wind profile observations. To optimize the investment output, DWL could operate in adaptive targeting mode.

- How to allocate these observation resources (e.g., getting 90% of the impact observing 2-10% of the time) could maximize effectiveness of DWL observations.

Decadal Study Recommendations (AMS 2007)



Why are winds so important? Two reasons:

1. Geostrophic adjustment (i.e., remembering potential vorticity)

The impact of a mass observation $\delta\phi$ after adjustment is

$$\delta\phi_g = \frac{1}{1+n^2R^2}\delta\phi$$

The impact of a wind observation $\delta\psi$ after adjustment is

$$\delta\psi_g = \frac{n^2R^2}{1+n^2R^2}\delta\psi$$

$$n = \frac{2\pi}{L}$$

n is wavenumber, *L* is wavelength

$$R = \sqrt{gD / f^2}$$

Rossby radius of deformation

$n^2R^2 >> 1$ means $L << R$, i.e., short waves

Why are winds so important? Two reasons:

1. Geostrophic adjustment (i.e., remembering potential vorticity)

The impact of a mass observation $\delta\phi$ after adjustment is

$$\delta\phi_g = \frac{1}{1+n^2R^2}\delta\phi$$

The impact of a wind observation $\delta\psi$ after adjustment is

$$\delta\psi_g = \frac{n^2R^2}{1+n^2R^2}\delta\psi$$

For short waves $n^2R^2 >> 1$ winds dominate

For long waves $n^2R^2 << 1$ mass dominates

For most atmospheric flow waves are short compared to R

(For most ocean flow waves are long compared to R)

Why are winds so important? Second reason:

2. Because of the information (differential measurement)

The information (inverse of error variance) of mass and wind observations is added in the analysis:

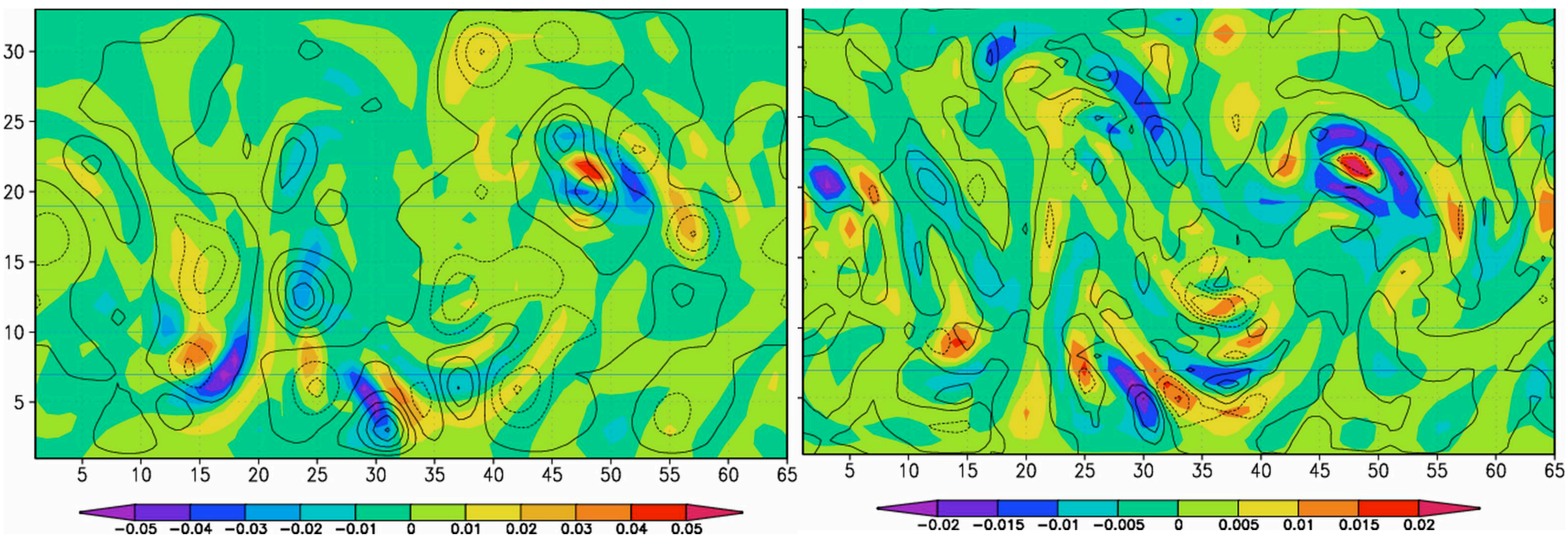
$$\frac{1}{|\delta\psi|^2} = \frac{f^2}{g^2|\delta z_{ob}|^2} + \frac{n^2}{|\delta \mathbf{v}_{ob}|^2}$$

This indicates that winds contribute more information in the tropics and for short waves.

Winds are derivatives of the mass field, so for short waves winds are more accurate than the mass field.

From a QG model OSSE (Corazza et al, 2003):
Forecast errors (colors) and analysis increments contours)

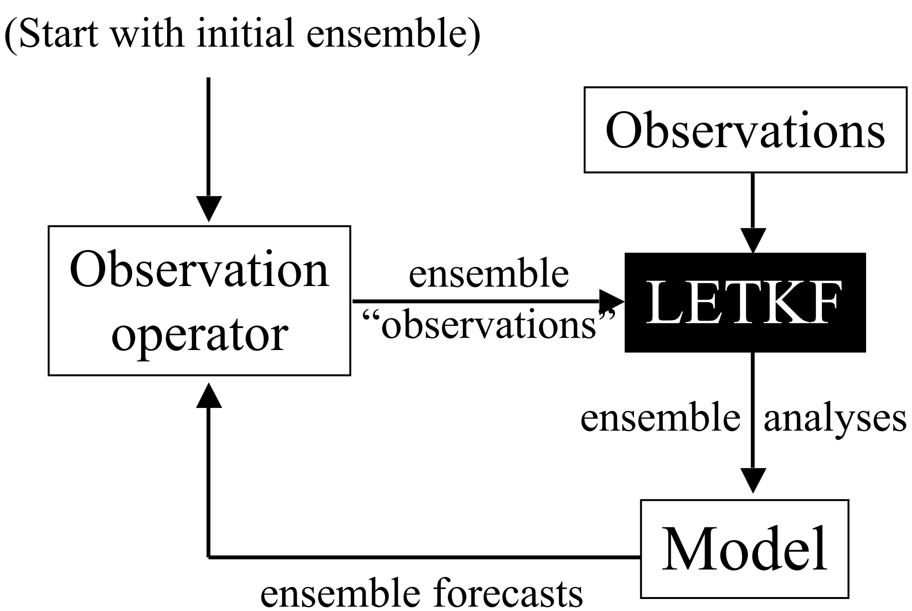
3D-Var EnKF



3D-Var does not capture the "errors of the day"

The EnKF ensemble **B** knows about the errors of the day, and uses the observations more efficiently

At the University of Maryland we developed the Local Ensemble Transform Kalman Filter (LETKF) (Ott et al, 2004, Hunt et al, 2004, 2007)



- Model independent (black box)
- Obs. assimilated simultaneously at each grid point
- 100% parallel: very fast
- 4D LETKF extension

Faster, cheaper and better than 4D-Var

Local Ensemble Transform Kalman Filter (LETKF)

Forecast step: $\mathbf{x}^{b}_{n,k} = M_n\left(\mathbf{x}^{a}_{n-1,k}\right)$

Analysis step: construct

$$\mathbf{X}^b = \left[\mathbf{x}_1^b - \overline{\mathbf{x}}^b, \ldots, \mathbf{x}_K^b - \overline{\mathbf{x}}^b\right];$$

$$\mathbf{y}_i^b = H(\mathbf{x}_i^b);\ \mathbf{Y}_n^b = \left[\mathbf{y}_1^b - \overline{\mathbf{y}}^b, \ldots, \mathbf{y}_K^b - \overline{\mathbf{y}}^b\right]$$

Choose for each grid point the observations to be used, and compute the local analysis error covariance and perturbations in ensemble space:

$$\tilde{\mathbf{P}}^a = \left[\left(K-1\right)\mathbf{I} + \mathbf{Y}^{bT}\mathbf{R}^{-1}\mathbf{Y}^b\right]^{-1};\ \mathbf{W}^a = [(K-1)\tilde{\mathbf{P}}^a]^{1/2}$$

Analysis in ensemble space: $\overline{\mathbf{w}}^a = \tilde{\mathbf{P}}^a\mathbf{Y}^{bT}\mathbf{R}^{-1}(\mathbf{y}^o - \overline{\mathbf{y}}^b)$

and add to $\mathbf{W}^a$ to get the analysis ensemble in ensemble space

The new ensemble analyses are the columns of

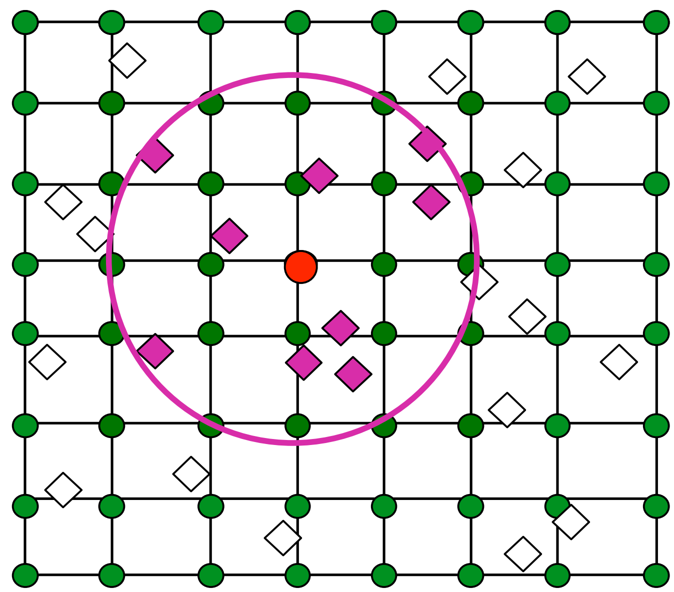
$$\mathbf{X}_n^a = \mathbf{X}_n^b\mathbf{W}^a + \overline{\mathbf{x}}^b$$

Localization based on observations

Perform data assimilation in a local volume, choosing observations

The state estimate is updated at the central grid red dot

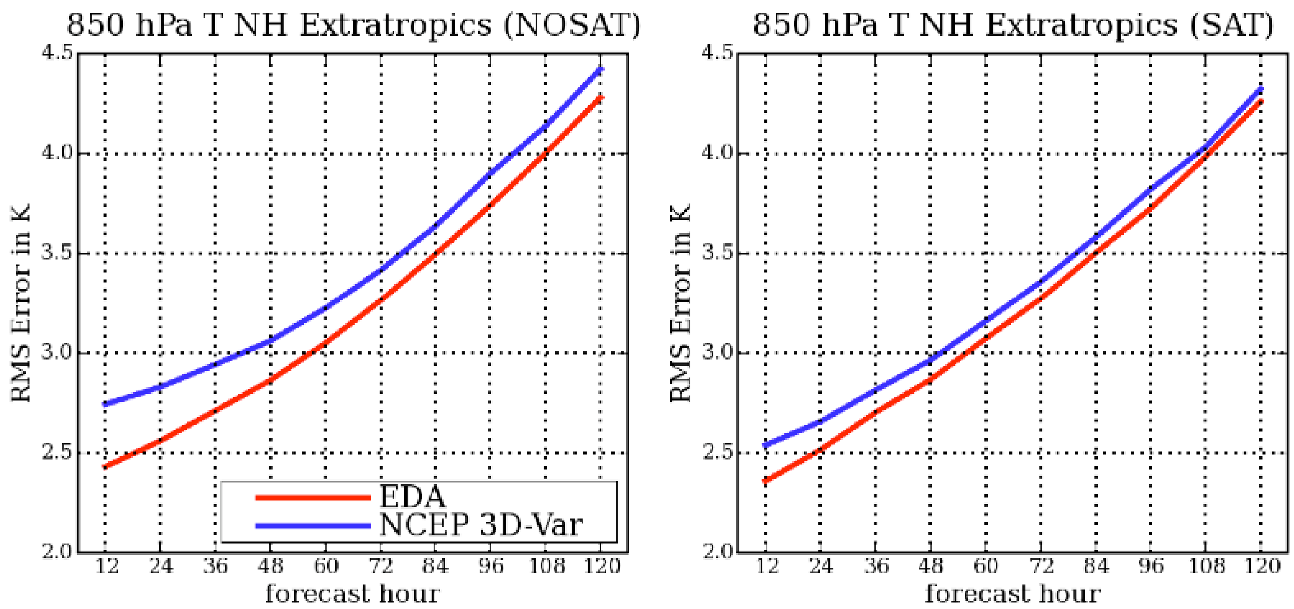
All observations (purple diamonds) within the local region are assimilated



Whitaker et al. (presented at AMS)

Used LETKF because it is fast with many observations (10^6 radiances). Verified against independent AIRS retrievals.

Forecast error (vs. AIRS T) - NOSAT and SAT



In the NH the advantage of EnKF is smaller than in the SH but still significant.

Ensemble spread adaptive observation strategy

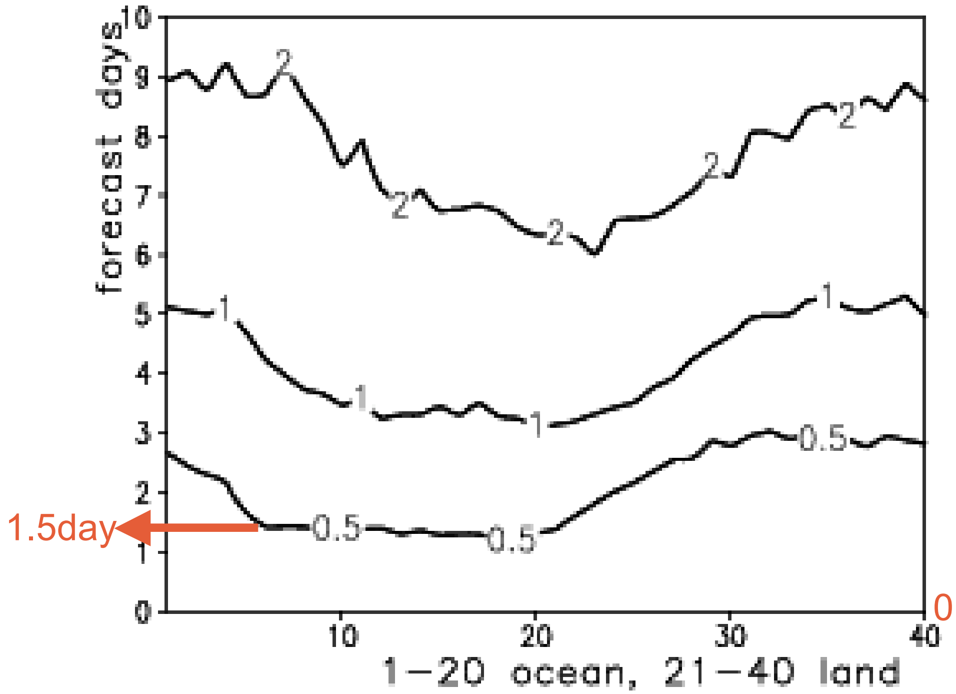
- The natural choice for an adaptive observation strategy in an ensemble Kalman Filter is the ensemble spread among the forecasts.

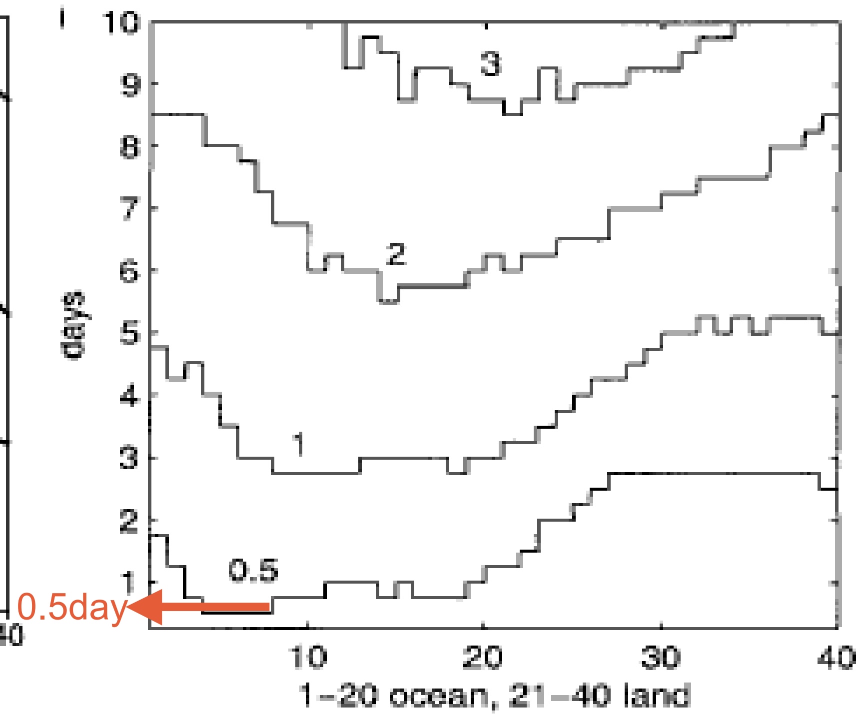
- Ensemble spread estimated from ensemble Kalman filter (EnKF) reflects the forecast (dynamical) uncertainties related with the flow of the day.

- In EnKF the ensemble spread strategy is very simple: we add the adaptive observations where the ensemble spread is large.

Background: ensemble spread adaptive strategy in Lorenz-40 variable model

10-day forecast RMS error of Lorenz-40 variable model





Ensemble spread method with 15 ensemble-member LETKF (Liu et al., 2006)

Singular vector method with 1024 ensemble-member EnKF (Hansen and Smith, 2000)

Ensemble spread sampling strategy with 15-member LETKF gives better results than singular vector method with a 1024 ensemble-member EnKF scheme

Possible strategies (10% obs)

- Uniform observations
- Random observations
- Climatological observations (e.g., storm tracks)
- “Ideal” observations where error is largest (cannot be done)
- Observe everywhere (100% coverage)
- **Ensemble spread**

Questions

- How effective is the ensemble spread in a global model with a simple experimental setup?
- How much impact can we get if we only observe 10% wind observations instead of 100% (full coverage)?
- If we observe only 2%?
- How different is the impact from adaptive observations with different data assimilation schemes? (Compare 3D-Var and LETKF)

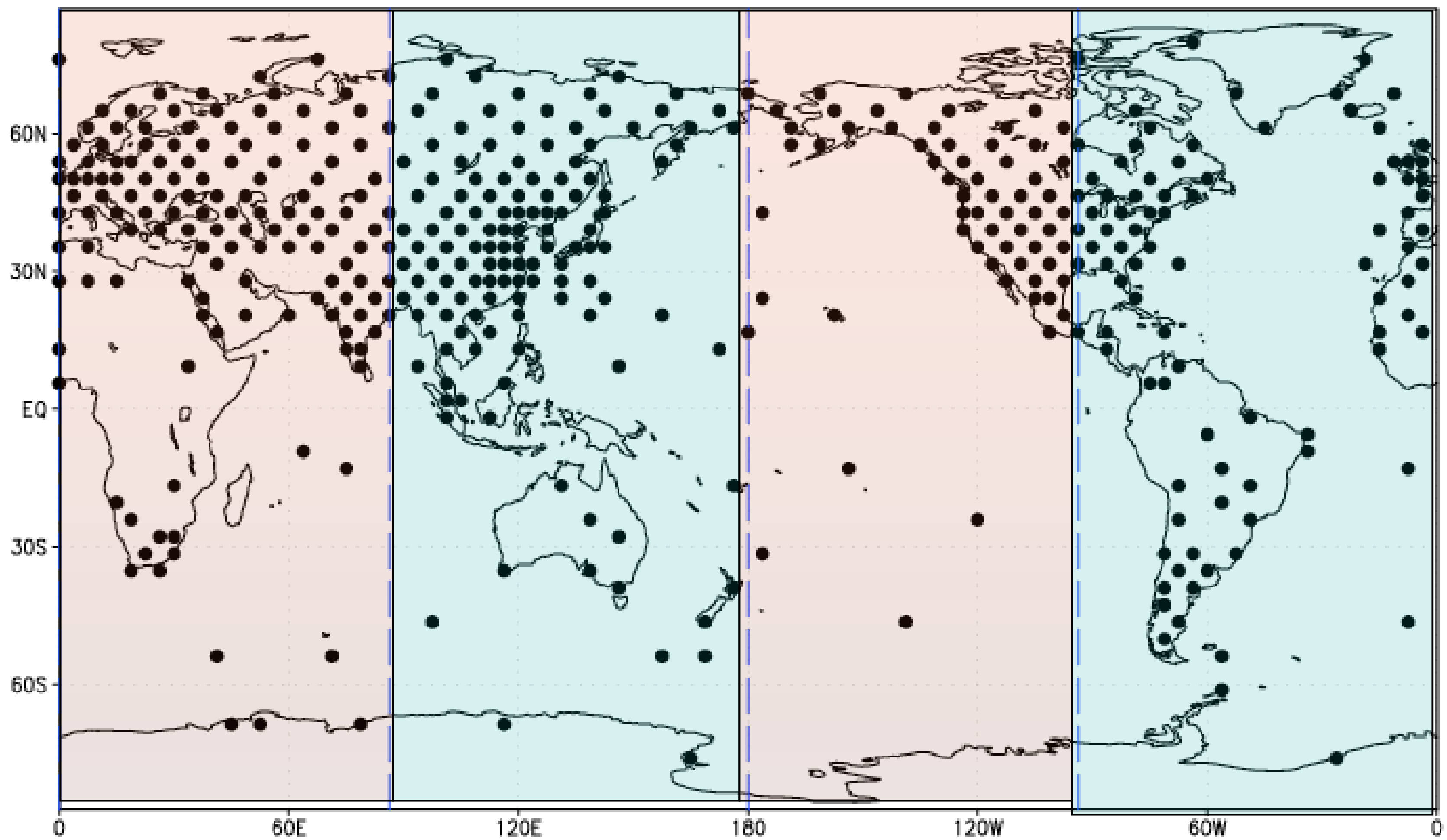
Experimental Design (very simple)

- SPEEDY model (Molteni, 2003, adapted by Miyoshi, 2005)
  - ➢ A global model with fast computation speed.
  - ➢ 96 grid points horizontally, and 48 grid points meridionally, and 7 vertical levels

- Data assimilation schemes
  - ➢ Local Ensemble Transform Kalman Filter (LETKF, Hunt et al., 2007)
  - ➢ 3D-Var (Miyoshi, 2005)

- Simulated observations
  - ➢ Obtained from “truth” (a long time integration) plus random perturbations.
  - ➢ Basic observation locations are rawinsonde locations, which observe all the dynamical variables.
  - ➢ We observe both zonal and meridional wind at adaptive observation points.

Rawinsonde observation locations (closed circles) and simulated satellite winds "scanning range"



00z and 12z   06z and 18z   00z and 12z   06z and 18z

→ The "orbit" allows simulated DWL observations potentially scanning each location twice a day.

→ 10% adaptive observations

Sampling strategies

- **Ensemble spread strategy** (from Local Ensemble Transform Kalman Filter)
  - → Adaptive observations are at locations with large ensemble wind spread at 500hPa.
  - → 3D-Var and LETKF have the same adaptive observation points
- **Random location**
  - → Randomly pick locations from potential locations.
- **Uniform distribution**
  - → Uniformly distributed.
- **Climatological large ensemble spread**
  - → Adaptive observations are at locations with large climatological average ensemble wind spread from rawinsonde assimilation.
  - → Constant with time, and same for 3D-Var and LETKF.
- **“Ideal” sampling**

  Adaptive observations are at locations with large background error obtained from the “truth”.

Sampling strategies

Changing

- **Ensemble spread strategy** (from Local Ensemble Transform Kalman Filter)
  → Adaptive observations are at locations with large ensemble wind spread at 500hPa.
  → 3D-Var and LETKF have the same adaptive observation points
- **Random location**
  → Randomly pick locations from potential locations.

Constant

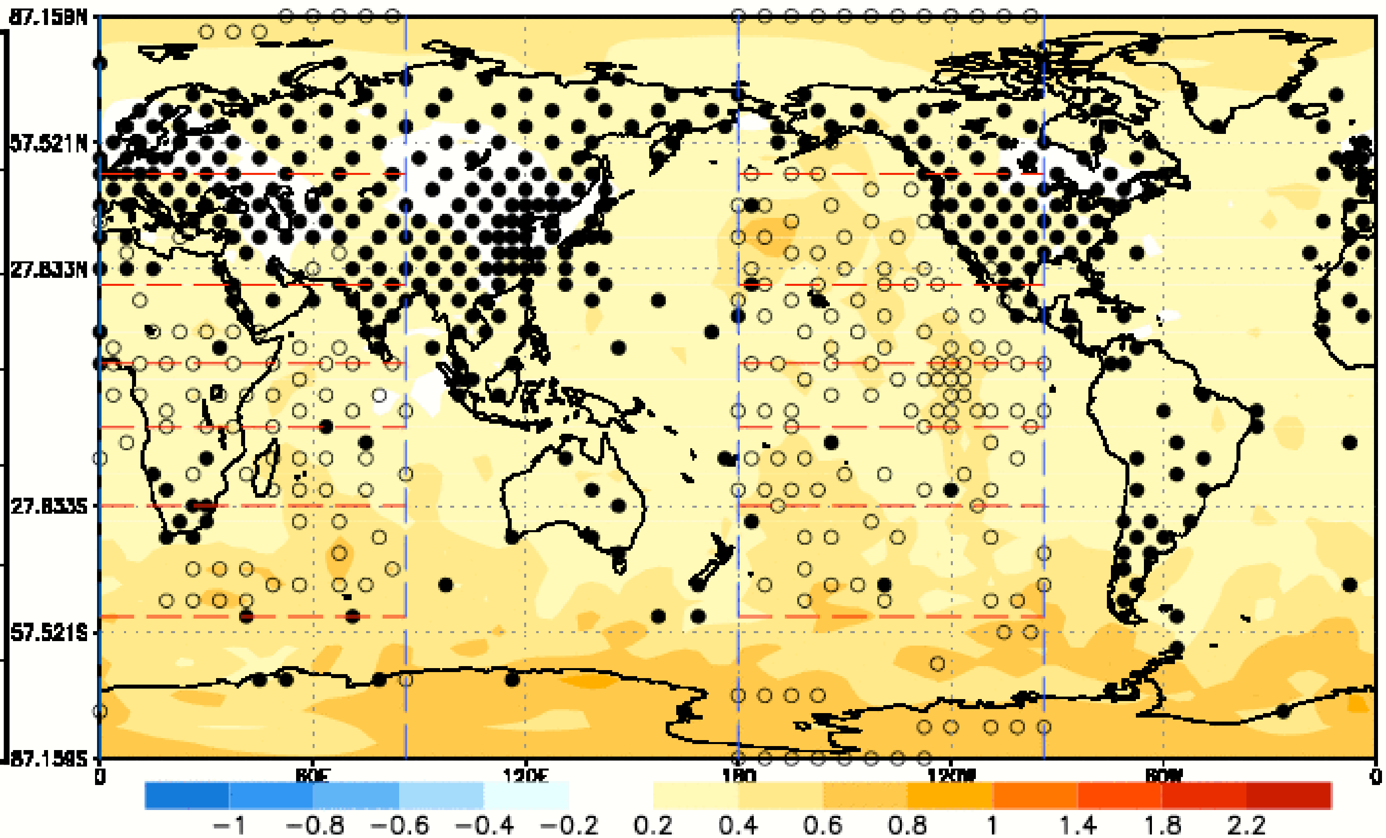
- **Uniform distribution**
  → Uniformly distributed.
- **Climatological large ensemble spread**
  → Adaptive observations are at locations with large climatological average ensemble wind spread from rawinsonde assimilation.
  → Constant with time, and same for 3D-Var and LETKF.

Impossible

- **"Ideal" sampling**
  Adaptive observations are at locations with large background error obtained from the "truth".

10% adaptive observation (open circles) distribution from ensemble spread (shaded area; Unit: m/s) strategy of LETKF

22	(53.81N-87.26N)
33	(31.55N-50.10N)
35	(9.28N-27.83N)
52	(5.57S-5.57N)
35	(24.12S-9.27S)
33	(50.10S-27.83S)
22	(87.26S-53.81S)

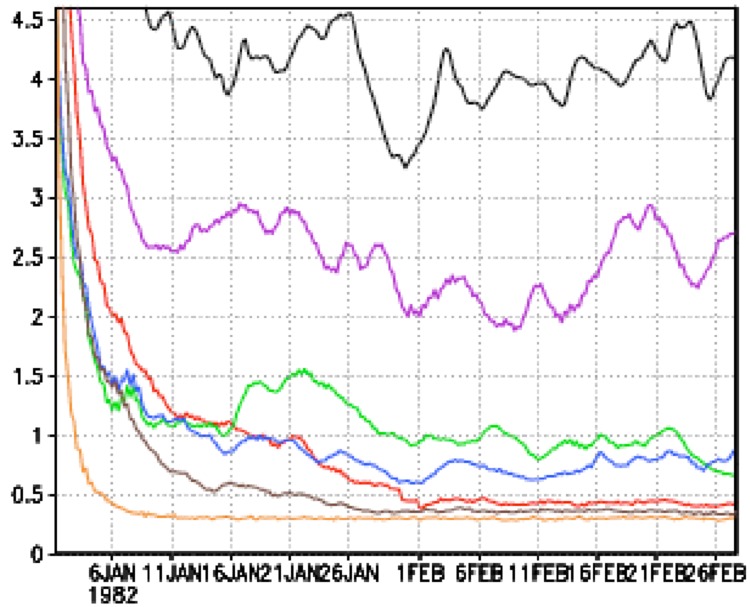


In red: number of adaptive observations in each band separated by red line, proportional to the area of each band.

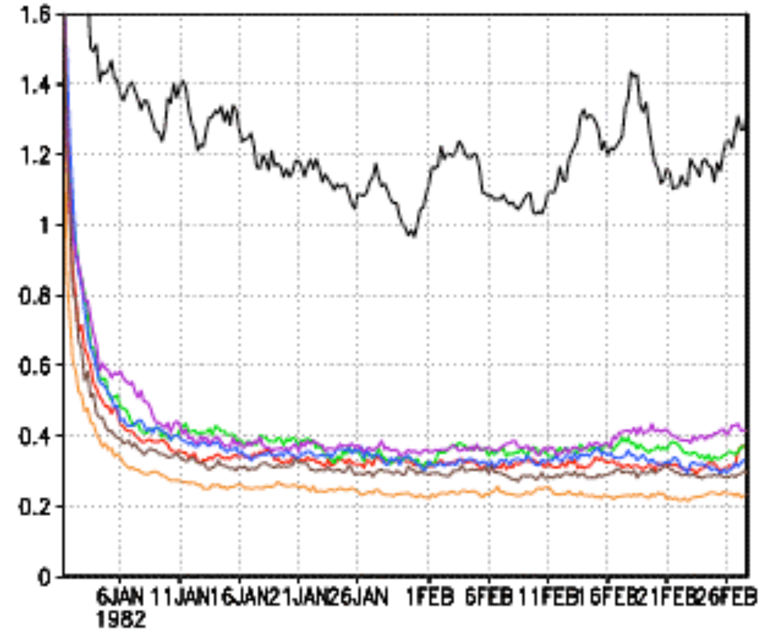
500hPa zonal wind RMS error

Rawinsonde; climatology; uniform; random; ensemble spread; “ideal”; 100%

3D-Var



LETKF



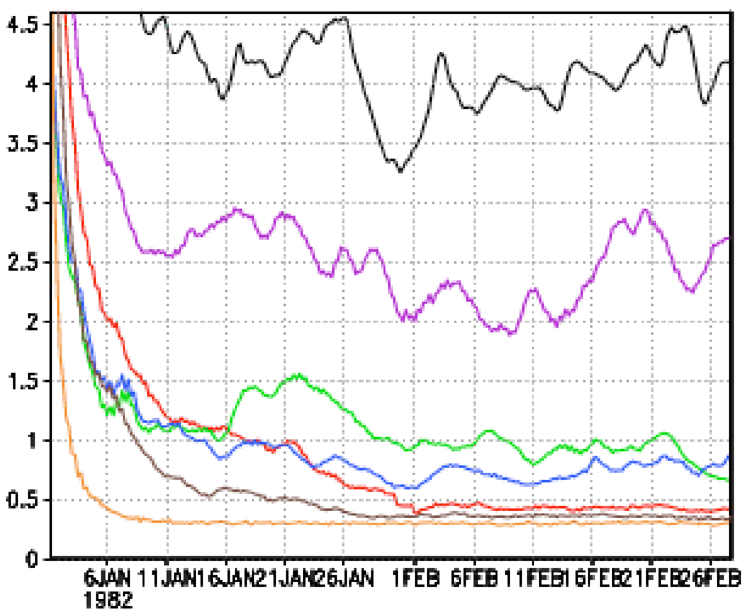
RMSE	4.04	2.36	0.92	0.74	0.43	0.36	0.30	1.18	0.38	0.36	0.33	0.32	0.29	0.23

- With 10% adaptive observations, the analysis accuracy is significantly improved for both 3D-Var and LETKF.
- 3D-Var is more sensitive to adaptive strategies than LETKF. Ensemble spread strategy gets best result among operational possible strategies

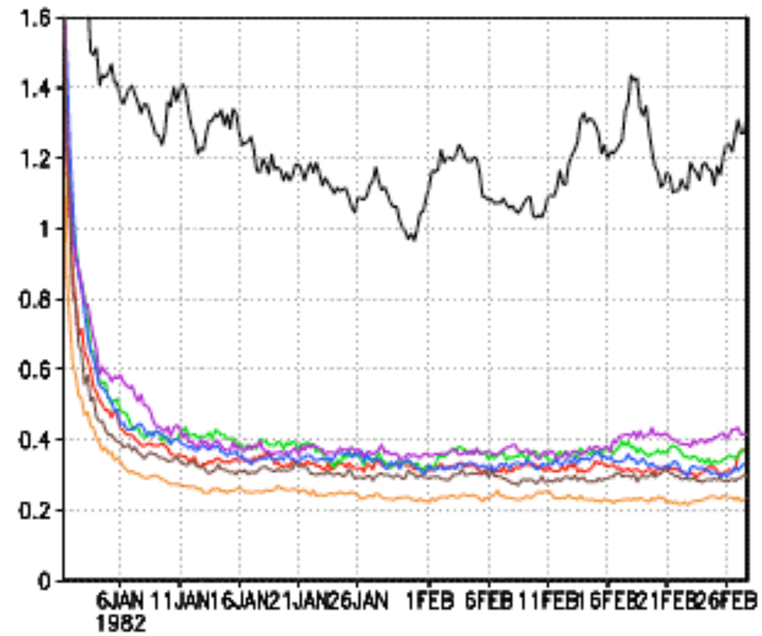
500hPa zonal wind RMS error

Rawinsonde; climatology; uniform; random; ensemble spread; "ideal"; 100%

3D-Var



LETKF



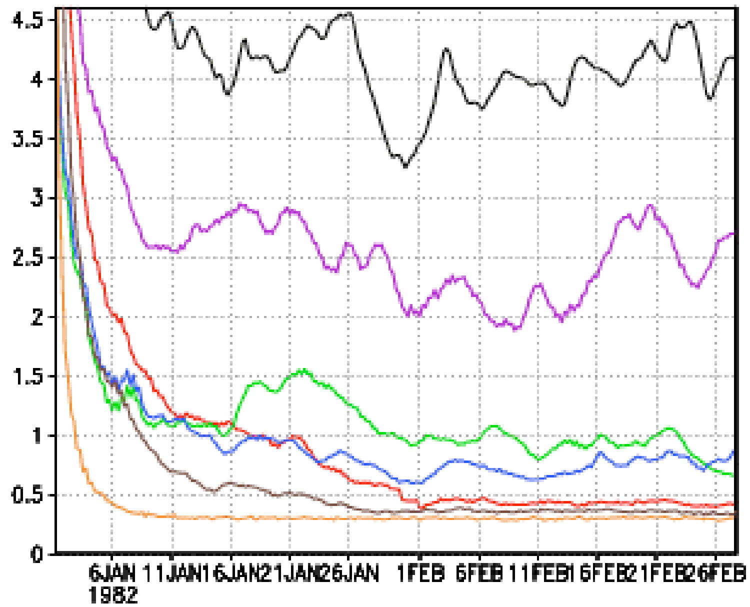
RMSE	4.04	2.36	0.92	0.74	0.43	0.36	0.30	1.18	0.38	0.36	0.33	0.32	0.29	0.23

- ➢With 10% adaptive observations, the analysis accuracy is significantly improved for both 3D-Var and LETKF.
- ➢ 3D-Var is more sensitive to adaptive strategies than LETKF. Ensemble spread strategy gets best result among operational possible strategies

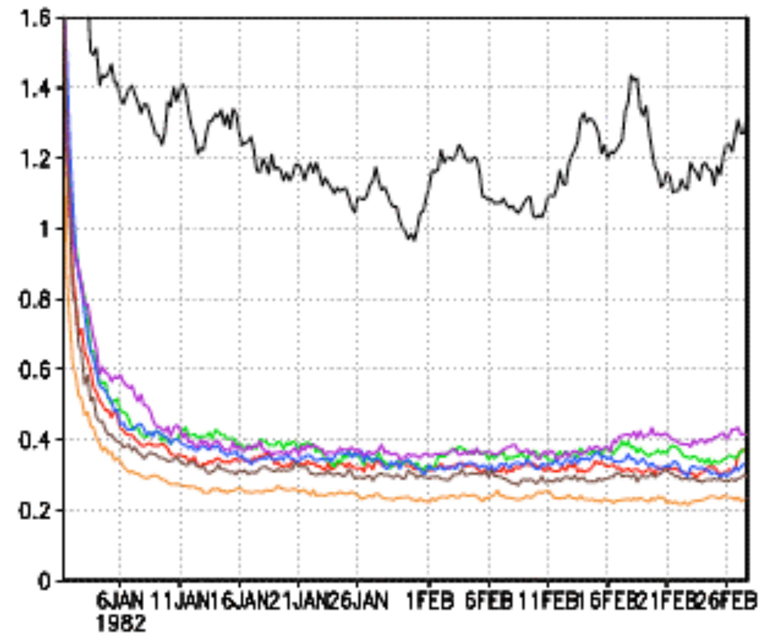
500hPa zonal wind RMS error

Rawinsonde; climatology; uniform; random; ensemble spread; “ideal”; 100%

3D-Var



LETKF

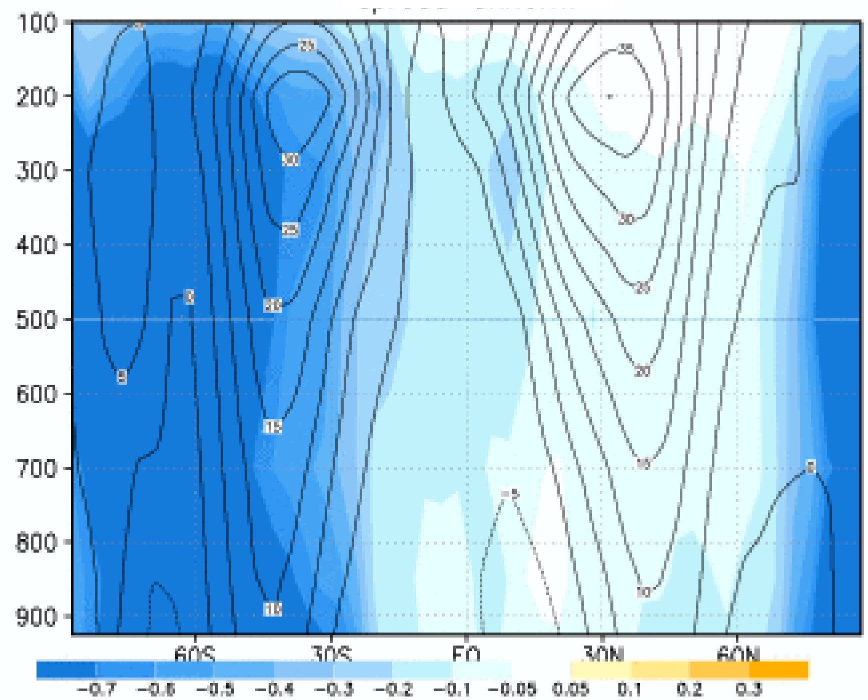


RMSE	4.04	2.36	0.92	0.74	0.43	0.36	0.30	1.18	0.38	0.36	0.33	0.32	0.29	0.23

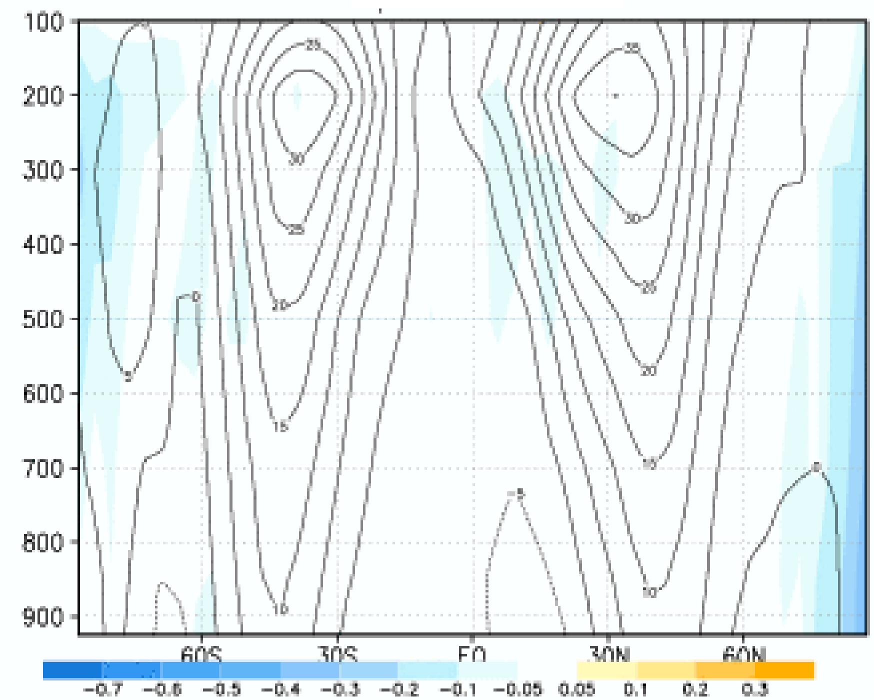
- With 10% adaptive observations, the analysis accuracy is significantly improved for both 3D-Var and LETKF.
- 3D-Var is more sensitive to adaptive strategies than LETKF. Ensemble spread strategy gets best result among operational possible strategies

The zonal mean RMS error difference between ensemble spread strategy and uniform distribution for zonal wind

3D-Var



LETKF



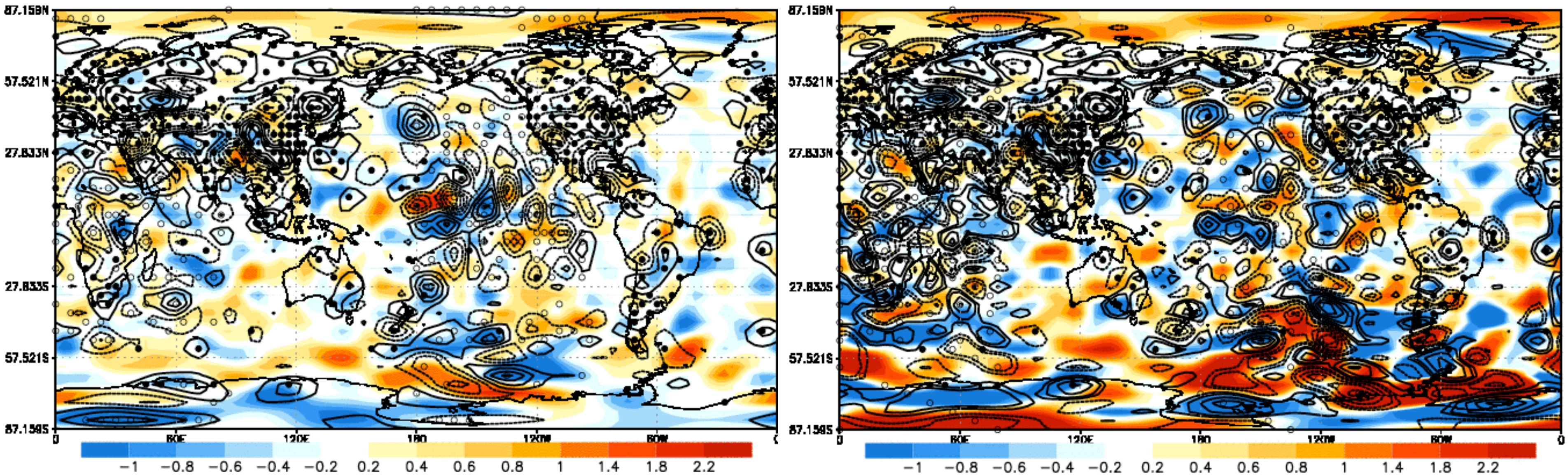
- 3D-Var shows much more significant difference between different strategies than that of LETKF.
- The largest differences between different adaptive strategies with 3D-Var is over data sparse regions since there the adaptive observations have most impact.

Analysis increment (contour, Unit: m/s, interval: 0.2m/s), background error (shaded, Unit: m/s), and adaptive observation locations (open circles)

3D-Var

Ensemble spread sampling strategy

Uniform distribution



➢Analysis increment of 3D-Var is centered around observation locations.

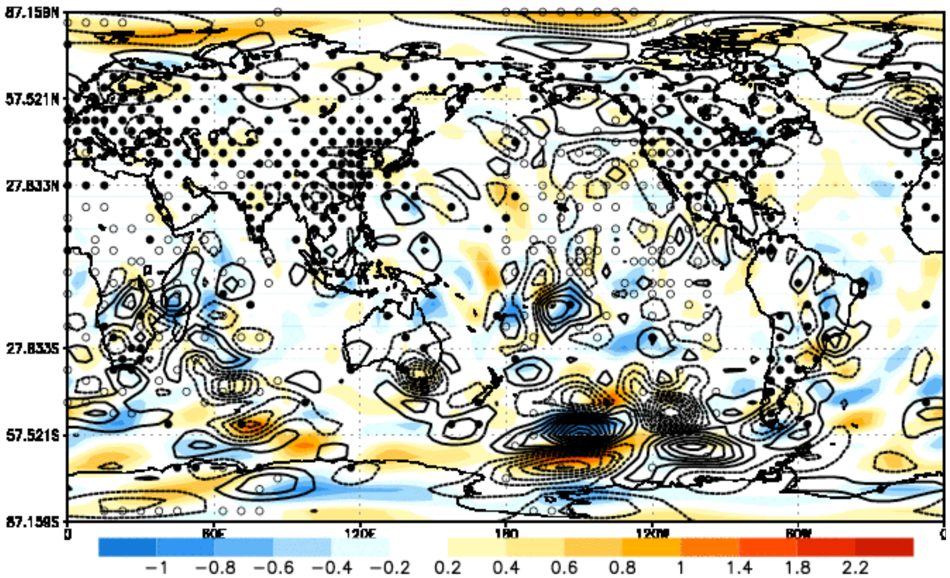
➢With adaptive observations from ensemble spread sampling strategy, analysis increment is pretty consistent with background error, like introducing the time-changing background error.

➢With adaptive observation with uniform distribution, analysis increment is not consistent with background error.

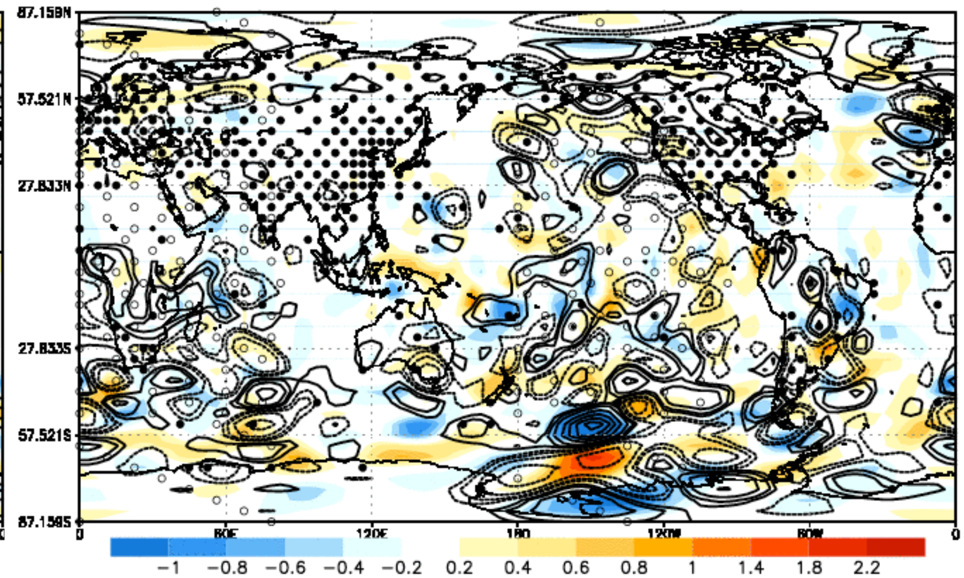
Analysis increment (contour, Unit: m/s, interval: 0.1m/s), background error (shaded, Unit: m/s), and adaptive observation locations (open circles)

LETKF

Ensemble spread sampling strategy



Uniform distribution



- Analysis increment of LETKF is not centered around observation locations, but line along background error

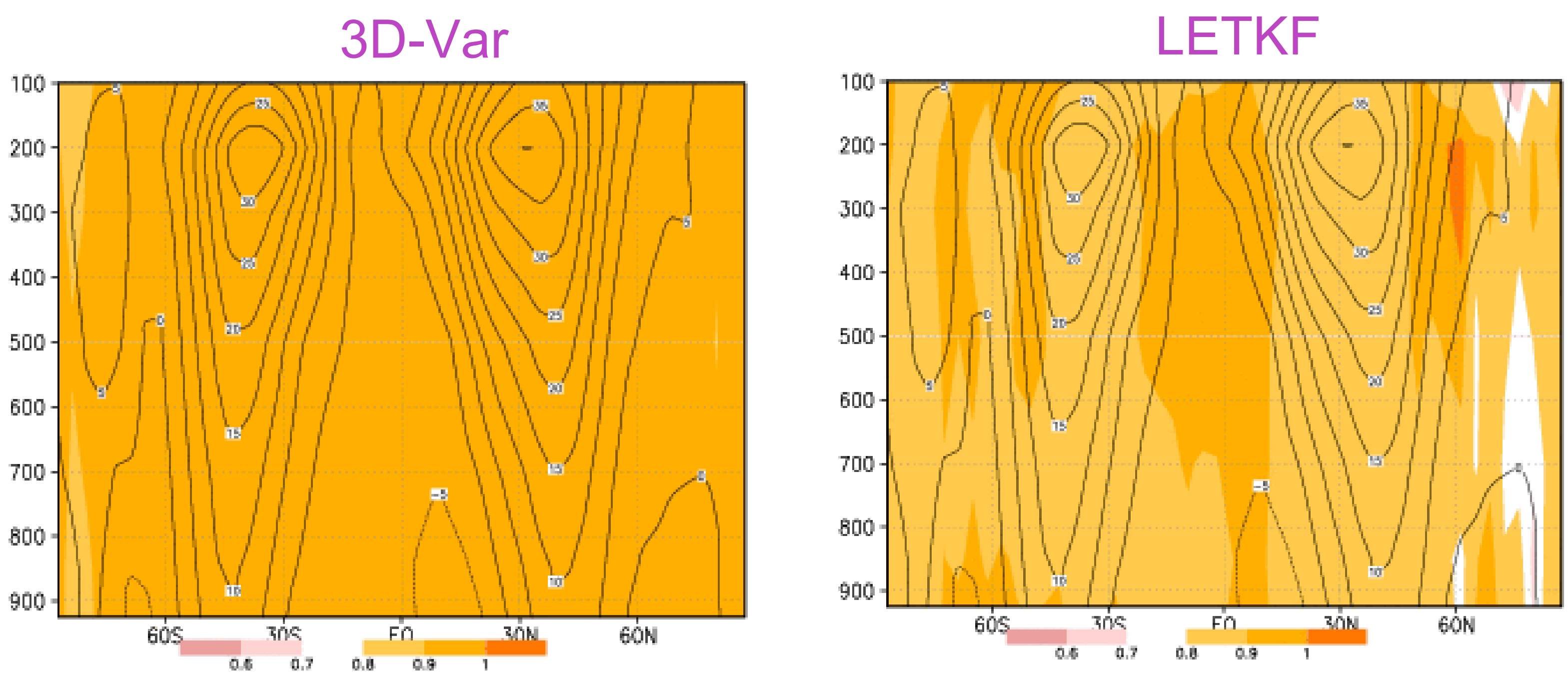
- Analysis increments from both ensemble spread sampling strategy and uniform distribution are consistent with background error

- What percentage improvement do we get from 10% adaptive observations compared with 100% adaptive observations?

$$B = \frac{RMS(10\%) - RMS(rawinsonde)}{RMS(100\%) - RMS(rawinsonde)}$$

Percentage effectiveness of 10% adaptive observation

Ensemble Spread

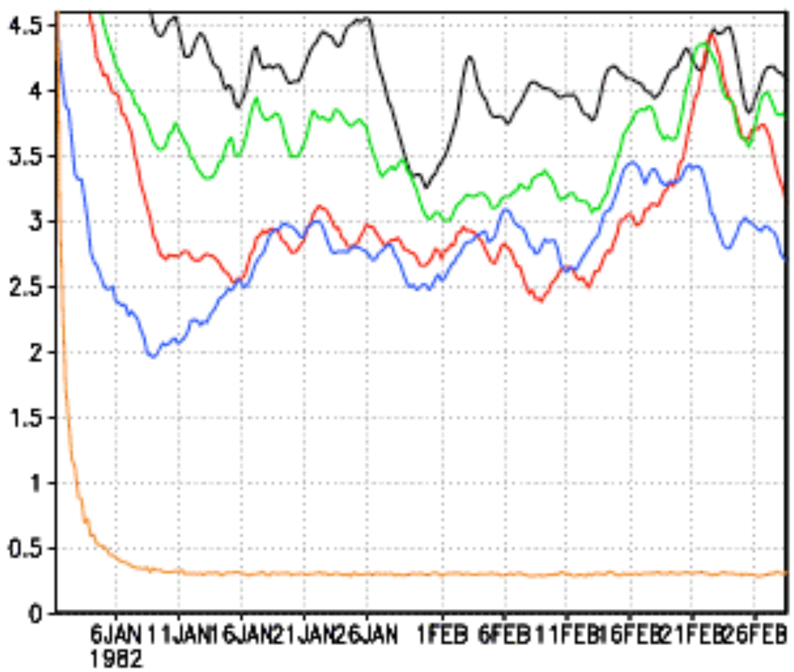


- ➢With ensemble spread, 10% adaptive observation can get more than 90% effect of 100% observation with 3D-Var.
- ➢The percentage effectiveness is more than 80% most of area with LETKF because it already accounts for “errors of the day”

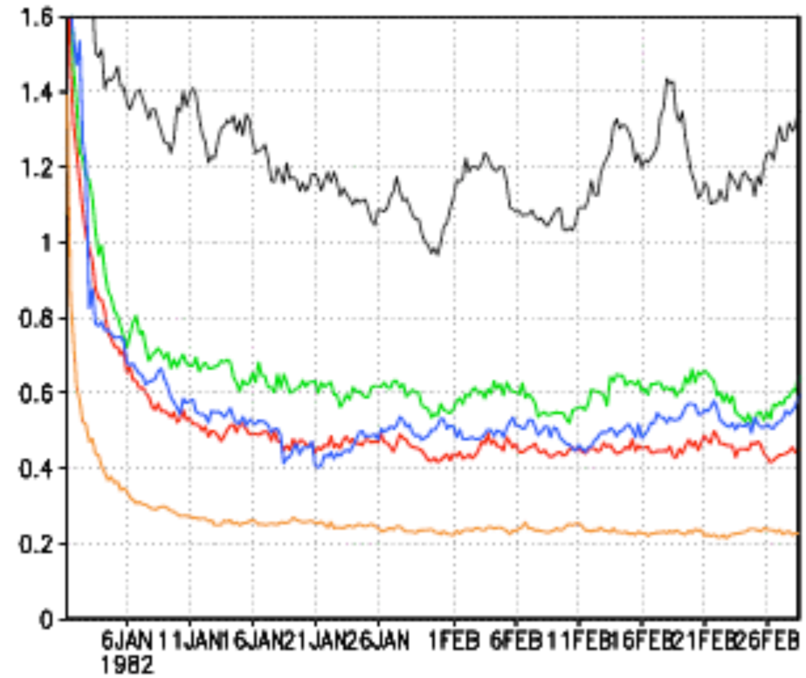
500hPa zonal wind RMS error (2% adaptive obs)

Rawinsonde; uniform; random; ensemble spread; 100%

3D-Var



LETKF



RMSE	4.04	3.53	3.00	3.11	0.30

1.18	0.59	0.51	0.45	0.23

- With fewer (2%) adaptive observations, ensemble spread sampling strategy outperforms the other methods in LETKF
- For 3D-Var 2% adaptive observations are clearly not enough

Conclusions and discussion

- With 10% adaptive wind observations, both 3D-Var and LETKF are much improved. Wind observations not only improve the wind analysis, but also the other variables (e.g., temperature).
- With 3D-Var, the largest improvement is from ensemble spread method derived from LETKF scheme.
- Changing observation locations with time is better than keeping constant observation locations with 3D-Var.
- With ensemble spread in 3D-Var, 10% adaptive observation can get over 90% effectiveness of full coverage over half the globe.
- 2% observations are not enough for 3D-Var, but ensemble spread still works well in LETKF, giving ~90% impact.
- Ensemble spread sampling strategy gives almost the same impact as the “ideal” (impossible in practice) sampling strategy.

Caveats and future

- Our experiments were done with a crude simulation and a low resolution global model. However, the main results should be valid for a much more realistic simulation:
  - The optimal adaptive strategy <u>even in 3D-Var</u> is the <u>ensemble spread of EnKF</u>. The reason is that ensemble spread reflects the dynamical uncertainty, and is telling 3D-Var where the “errors of the day” are.
  - With ensemble spread in 3D-Var, 10% adaptive observation can get over 90% effect of full coverage over half the globe. With every method (ensemble spread, uniform, random), LETKF can get more 80% improvement from 10% adaptive observations.
  - “Dwelling on areas of forecast uncertainty” based on ensemble spread should be considered in the design of any atmospheric instrument.

➔ We would like to examine the robustness of the results with more realistic NCEP OSSEs and the new nature run.