Polynomial Regression and Serial Ensemble Kalman Filtering

Dan Hodyss[1], Bill Campbell[1], Jeff Anderson[2], Nancy Collins[2], Alex Reinecke[1]

1. *Naval Research Laboratory, Monterey, California*
2. *National Center for Atmospheric Research, Boulder, Colorado*

Posterior Mean

Our goal in this talk is to discuss the estimation of the posterior mean

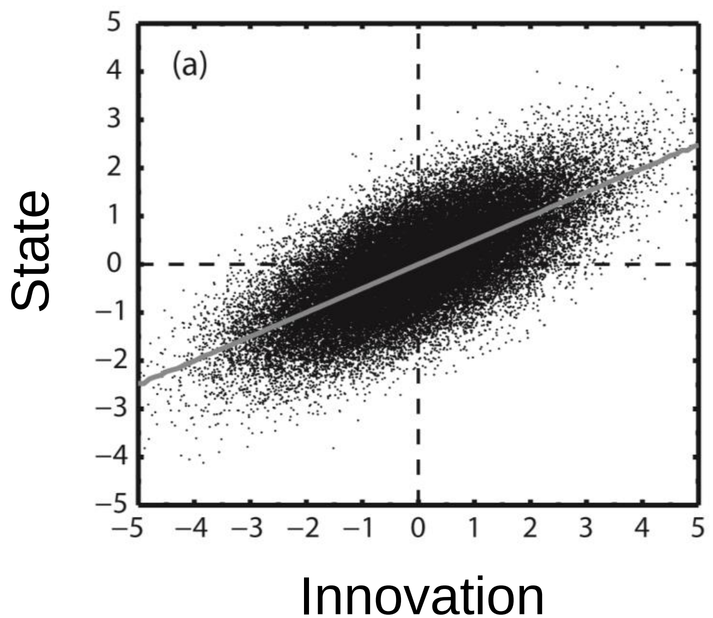
$$\bar{x}(y) = \int_{-\infty}^{\infty} xp(x \mid y)\,dx$$

where $p(x \mid y)$ is the posterior density.

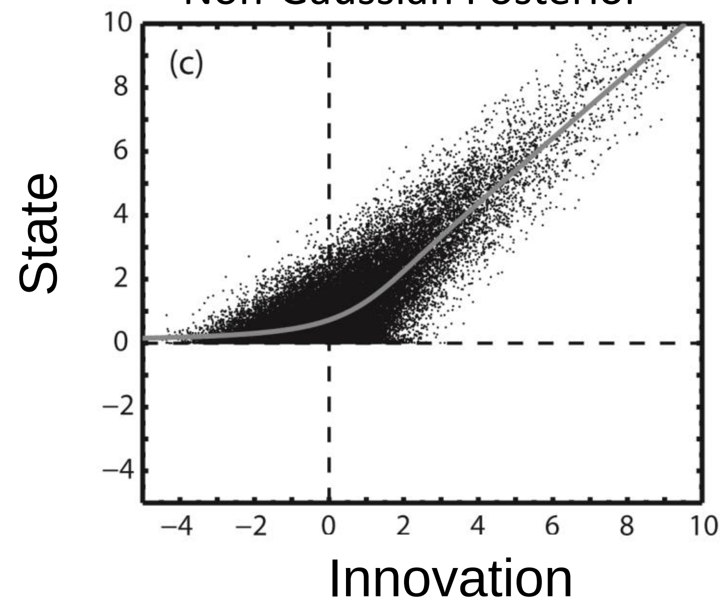
The Posterior Mean is a function of y

$$\bar{x}(y) = \int_{-\infty}^{\infty} x p(x \mid y)\, dx$$

Gaussian Posterior



Non-Gaussian Posterior



The Posterior Mean is Curved

Hodyss (2011; MWR) proved that the curvature was determined by the posterior third moment:

$$\frac{d^2\bar{x}}{dy^2} = \frac{\mathrm{T}(y)}{\mathrm{R}^2}$$

$\mathbf{T}(y)$ = Posterior Third Moment
$\mathbf{R}$ = Observation error variance

Multivariate generalization can be found in Hodyss (2011)

Polynomial Filtering

Hodyss (2011; MWR) shows how to expand the posterior mean as

$$\bar{x}(y) = \bar{x}_f + \overset{\boxed{\text{Kalman Gain}}}{\overset{\downarrow}{G_1}}\left[y - \bar{x}_f\right] + G_2\left[\left(y - \bar{x}_f\right)^2 - \alpha_2\right] + G_3\left[\left(y - \bar{x}_f\right)^3 - \alpha_3\right] + \ldots$$

The Kalman filter is “optimal” when:

1. The innovation is small
2. Prior is not too skewed

Polynomial Filtering

What's remarkable about this is it looks just like a Kalman filter!

$$\overline{x}\left(y\right)=\overline{x}_f+\mathbf{Gd}$$

Gain: $\mathbf{G}=\left[G_1 \quad G_2 \quad G_3 \quad \ldots\right]$

Innovation vector with new pseudo-obs!

$$\mathbf{d}=\begin{bmatrix} y-\overline{x}_f \\ \left(y-\overline{x}_f\right)^2-\alpha_2 \\ \left(y-\overline{x}_f\right)^3-\alpha_3 \\ \vdots \end{bmatrix}$$

All-at-Once vs. Serial-Solve

- Hodyss (2012; MWR) showed how to do the all-at-once solve.
  - Uses a minimization-based technique for the matrix inverse and a post-multiply step
- This talk will introduce the serial formulation.
  - There is a version of DART that now includes quadratic polynomial regression:
    - Ensemble Adjustment Quadratic Filter (EAQF)
    - Ensemble Kalman Quadratic Filter (EnQF)

Serial Quadratic Nonlinear Regression

Define the "squared pseudo-ob" as: $y_i^s = \left(y_i - \mathbf{h}_i \overline{\mathbf{x}}^f\right)^2 - r_i$

Define the "squared state" as: $s_j = \left(x_j - \overline{x}_j^f\right)^2$

Step 1: Assimilate regular ob

$$\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^f + \mathbf{P}_2 \mathbf{h}_i^T \frac{y_i - \mathbf{h}_i \overline{\mathbf{x}}^f}{P_{2ii} + r_i} \qquad \mathbf{s}^a = \mathbf{s}^f + \mathbf{P}_3^T \mathbf{h}_i^T \frac{y_i - \mathbf{h}_i \overline{\mathbf{x}}^f}{P_{2ii} + r_i}$$

Step 2: Assimilate the "pseudo-ob"

$$\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^f + \mathbf{P}_3 \mathbf{h}_i^T \frac{y_i^s - s_i^f}{P_{4ii} + r_{2i}} \qquad \mathbf{s}^a = \mathbf{s}^f + \mathbf{P}_4 \mathbf{h}_i^T \frac{y_i^s - s_i^f}{P_{4ii} + r_{2i}}$$

Serial Cubic Nonlinear Regression

Step 1: Assimilate regular ob

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^f + \mathbf{P}_2\mathbf{h}_i^T \frac{y_i - \mathbf{h}_i\bar{\mathbf{x}}^f}{P_{2ii} + r_i} \qquad \mathbf{s}^a = \mathbf{s}^f + \mathbf{P}_3^T\mathbf{h}_i^T \frac{y_i - \mathbf{h}_i\bar{\mathbf{x}}^f}{P_{2ii} + r_i} \qquad \mathbf{c}^a = \mathbf{c}^f + \mathbf{P}_4\mathbf{h}_i^T \frac{y_i - \mathbf{h}_i\bar{\mathbf{x}}^f}{P_{2ii} + r_i}$$

Step 2: Assimilate the “squared” ob

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^f + \mathbf{P}_3\mathbf{h}_i^T \frac{y_i^s - s_i^f}{P_{4ii} + r_{2i}} \qquad \mathbf{s}^a = \mathbf{s}^f + \mathbf{P}_4\mathbf{h}_i^T \frac{y_i^s - s_i^f}{P_{4ii} + r_{2i}} \qquad \mathbf{c}^a = \mathbf{c}^f + \mathbf{P}_5^T\mathbf{h}_i^T \frac{y_i^s - s_i^f}{P_{4ii} + r_{2i}}$$

Step 3: Assimilate the “cubed” ob

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^f + \mathbf{P}_4\mathbf{h}_i^T \frac{y_i^c - c_i^f}{P_{6ii} + r_{3i}} \qquad \mathbf{s}^a = \mathbf{s}^f + \mathbf{P}_5\mathbf{h}_i^T \frac{y_i^c - c_i^f}{P_{6ii} + r_{3i}} \qquad \mathbf{c}^a = \mathbf{c}^f + \mathbf{P}_6\mathbf{h}_i^T \frac{y_i^c - c_i^f}{P_{6ii} + r_{3i}}$$

Scalar Test Problem: Large Ensemble

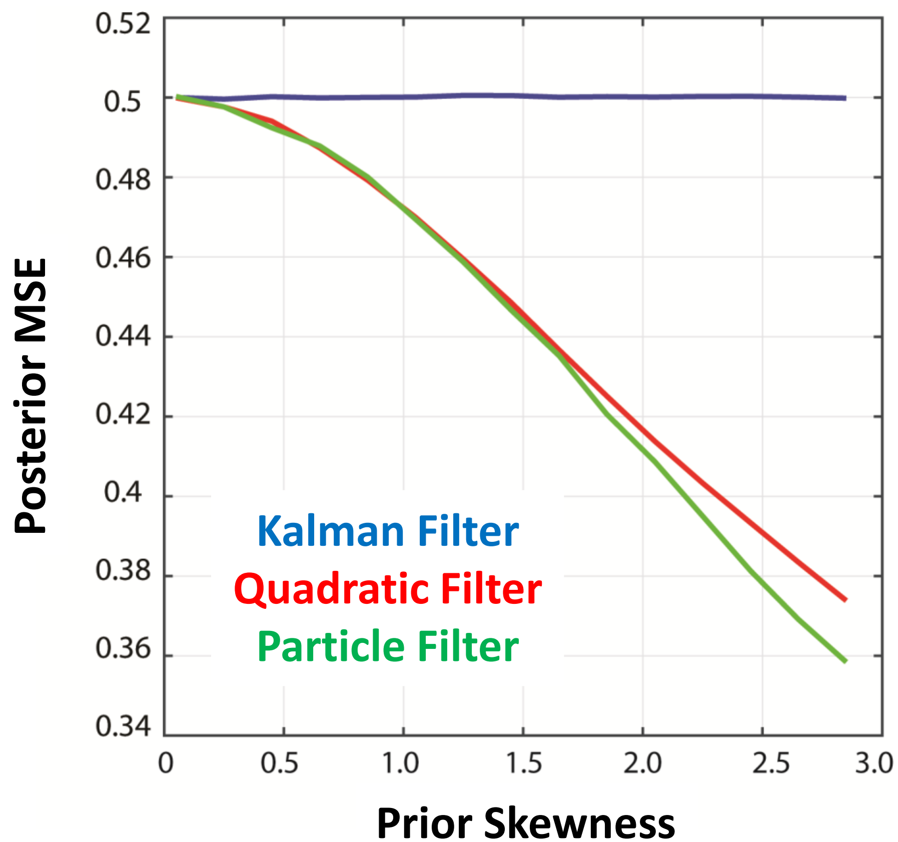
<u>DA Experiment</u>:
Gaussian Ob Likelihood (R=1)
Gamma Prior with (k,θ) chosen such that P = 1 but varying skewness

Kalman and Quadratic formulas evaluated with known analytic moments (i.e. infinite ensemble)

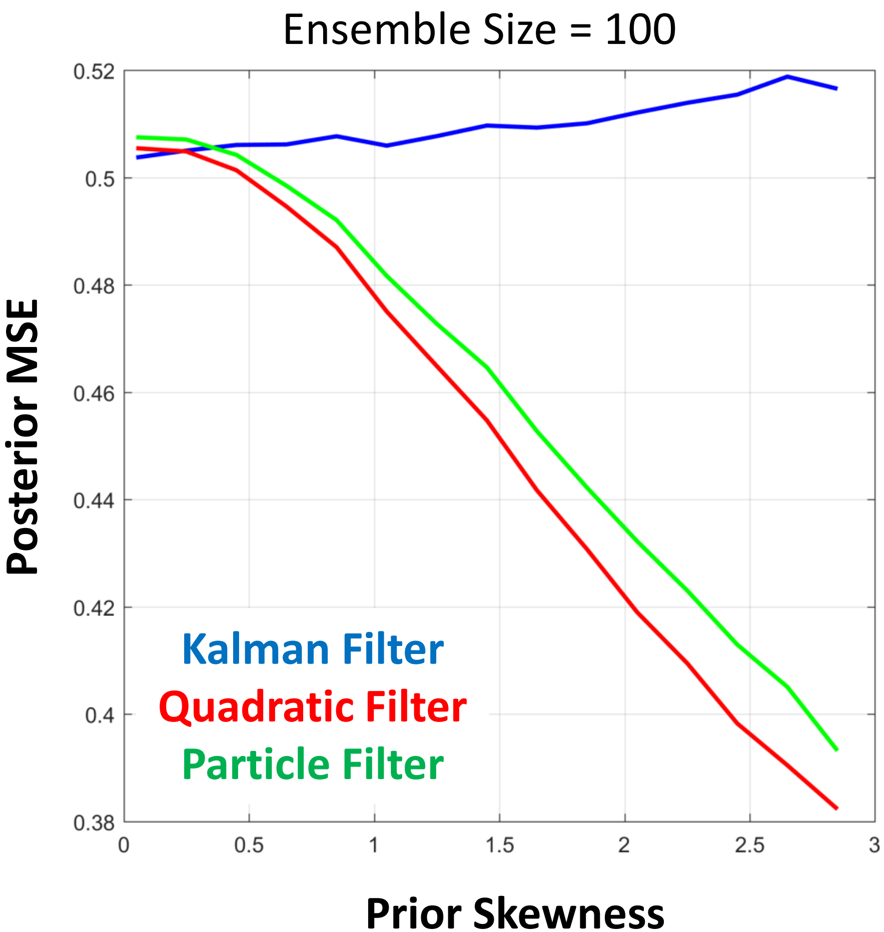
Particle filter uses 10^6 particles

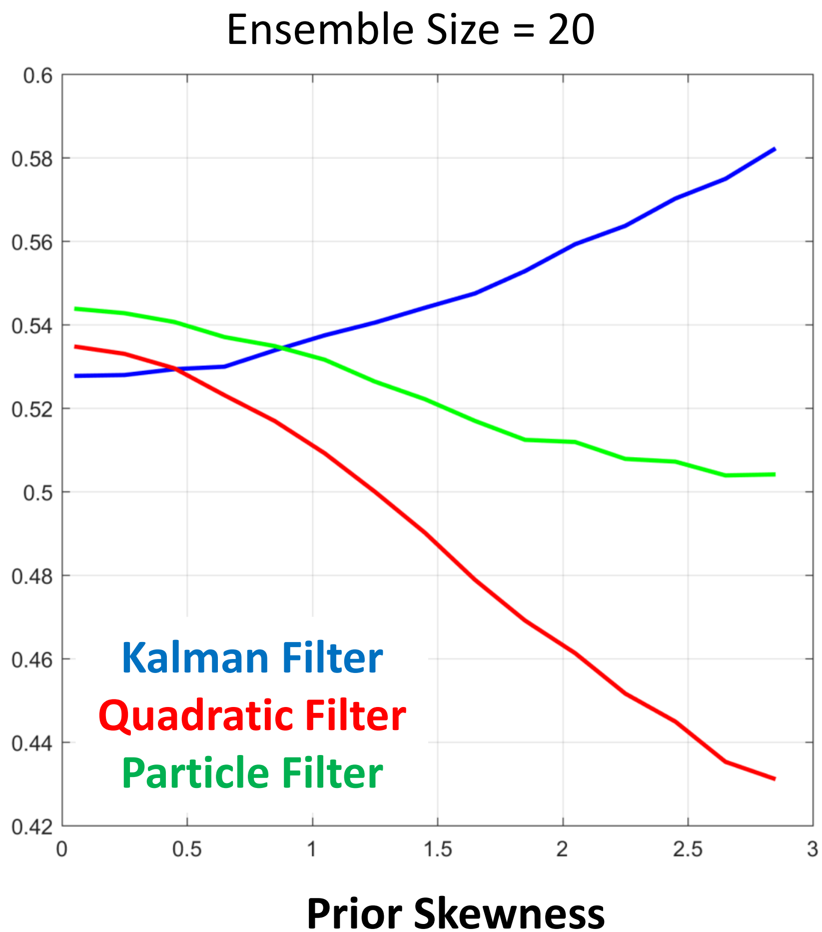
Experiment is repeated 10^6 times for different truths and different obs

10% reduction in MSE when prior skewness is about 1.5



Scalar Test Problem: Small Ensemble



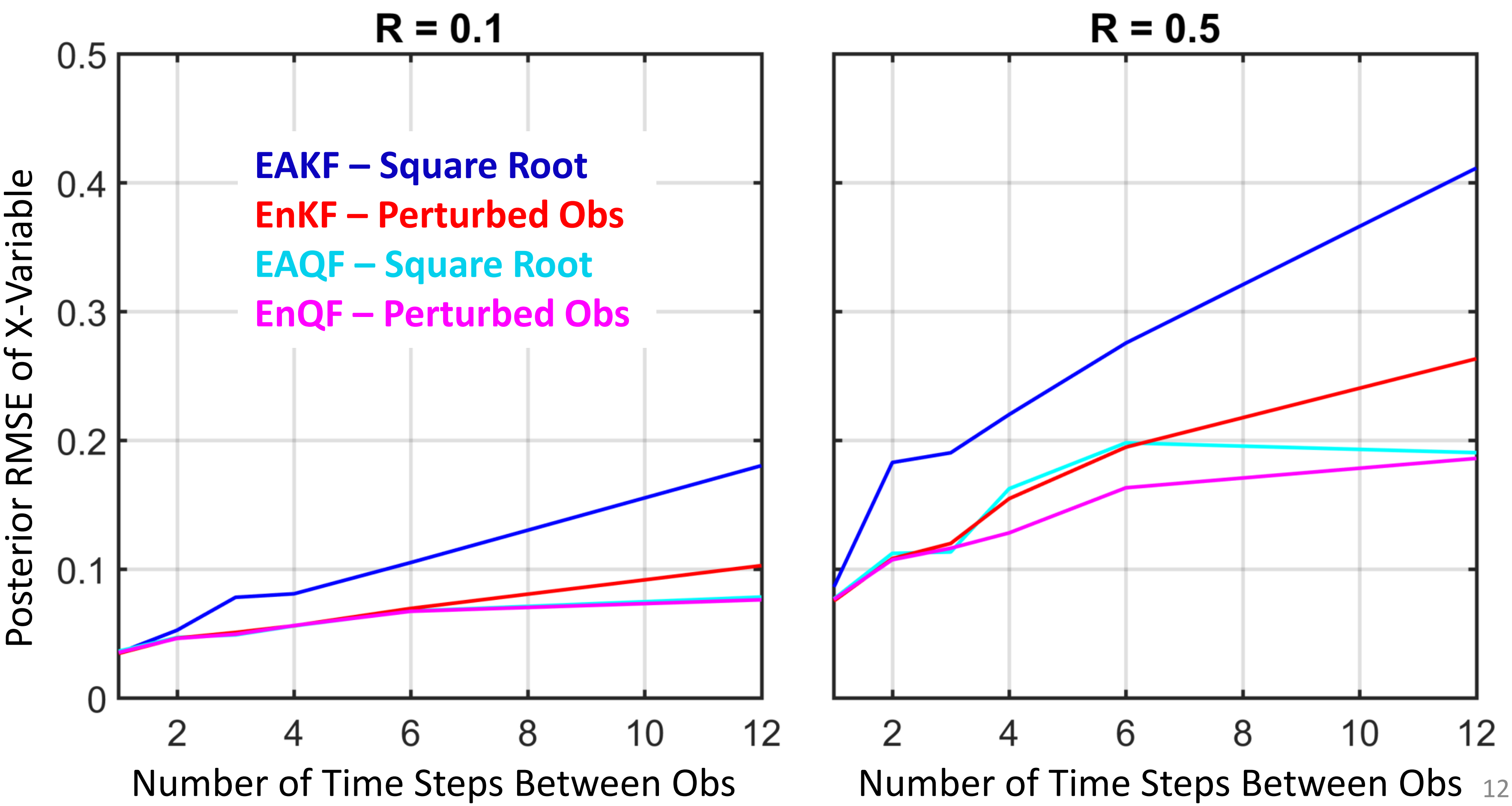


Is truncating the expansion a useful method for reducing sampling error?

DA in Lorenz-63

Obs of x and z
Ensemble size: 1000 members
Ran for 1000 cycles

Adaptive prior inflation (inf_flavor = 2)



Prior Skewness in Lorenz-63

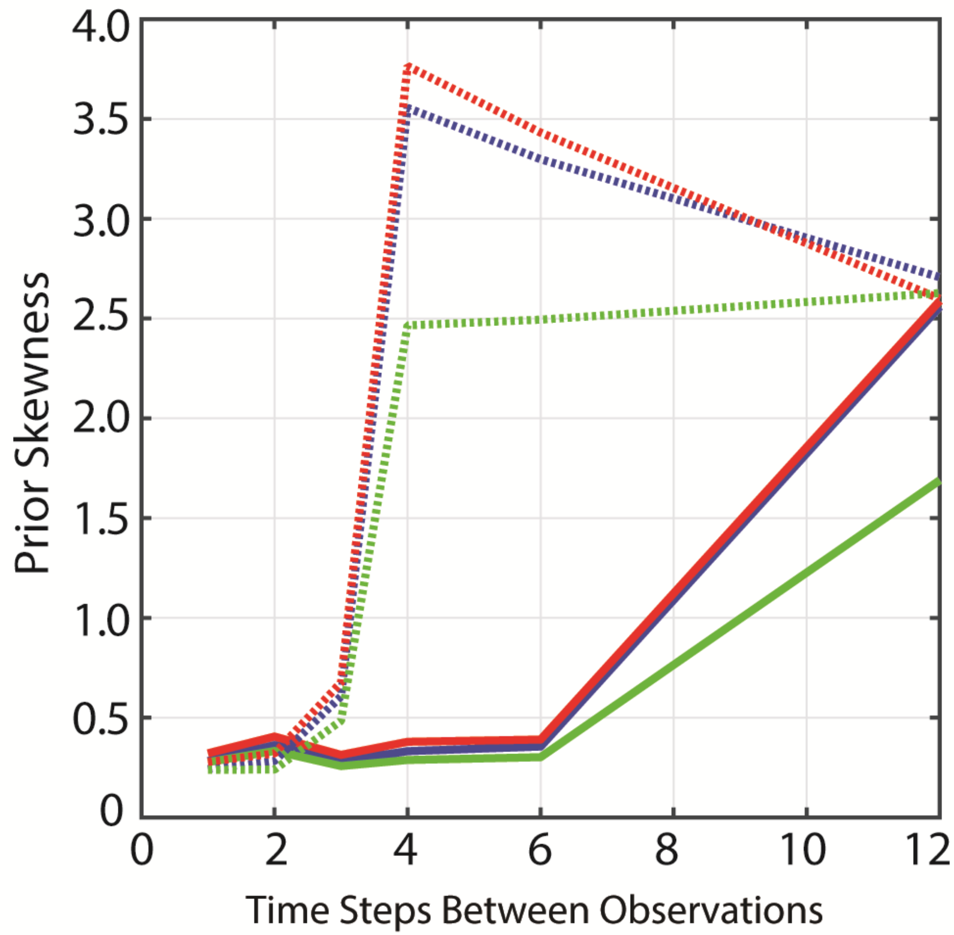
Average absolute value of skewness over last 800 cycles.

Solid: R = 0.1
Dashed: R = 0.5

For R = 0.1 skewness is substantial after 6 time steps between observations

For R = 0.5 skewness is substantial after 3 time steps between observations

This appears to explain the posterior separation between KPO and QPO on previous slide.

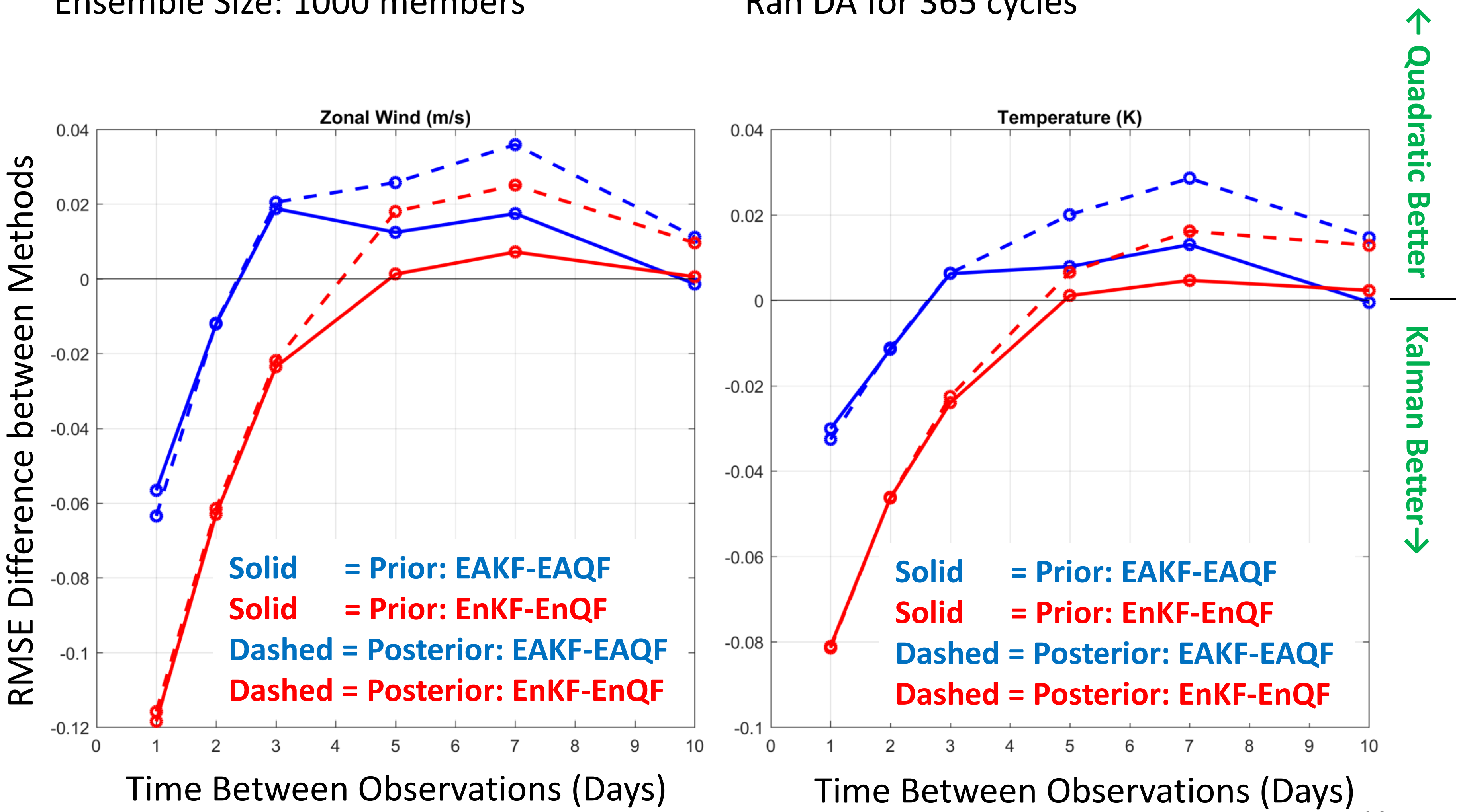


X-variable, **Y-variable**, **Z-variable**

DA in the Bgrid GCM

Resolution: 60x30x5 (approximately 6 degrees)
Obs: 10 soundings of U, V, T,Ps with R = 1
Ensemble Size: 1000 members

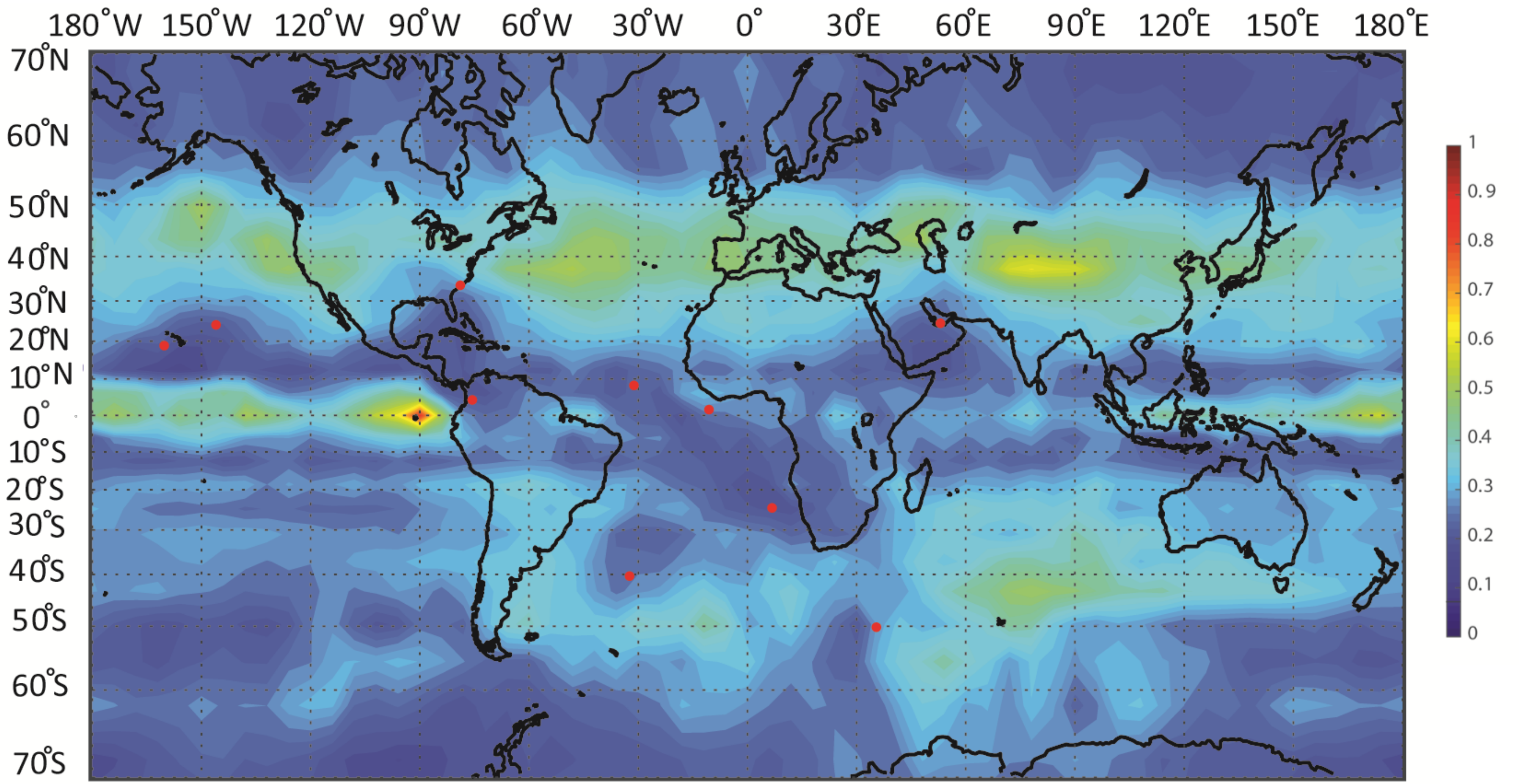
Adaptive prior inflation (inf_flavor = 2)
Ran DA for 365 cycles



Bgrid Prior Skewness: Obs 1 Day Apart

Average absolute value of skewness over last 200 cycles.

Zonal Wind: Lowest Layer

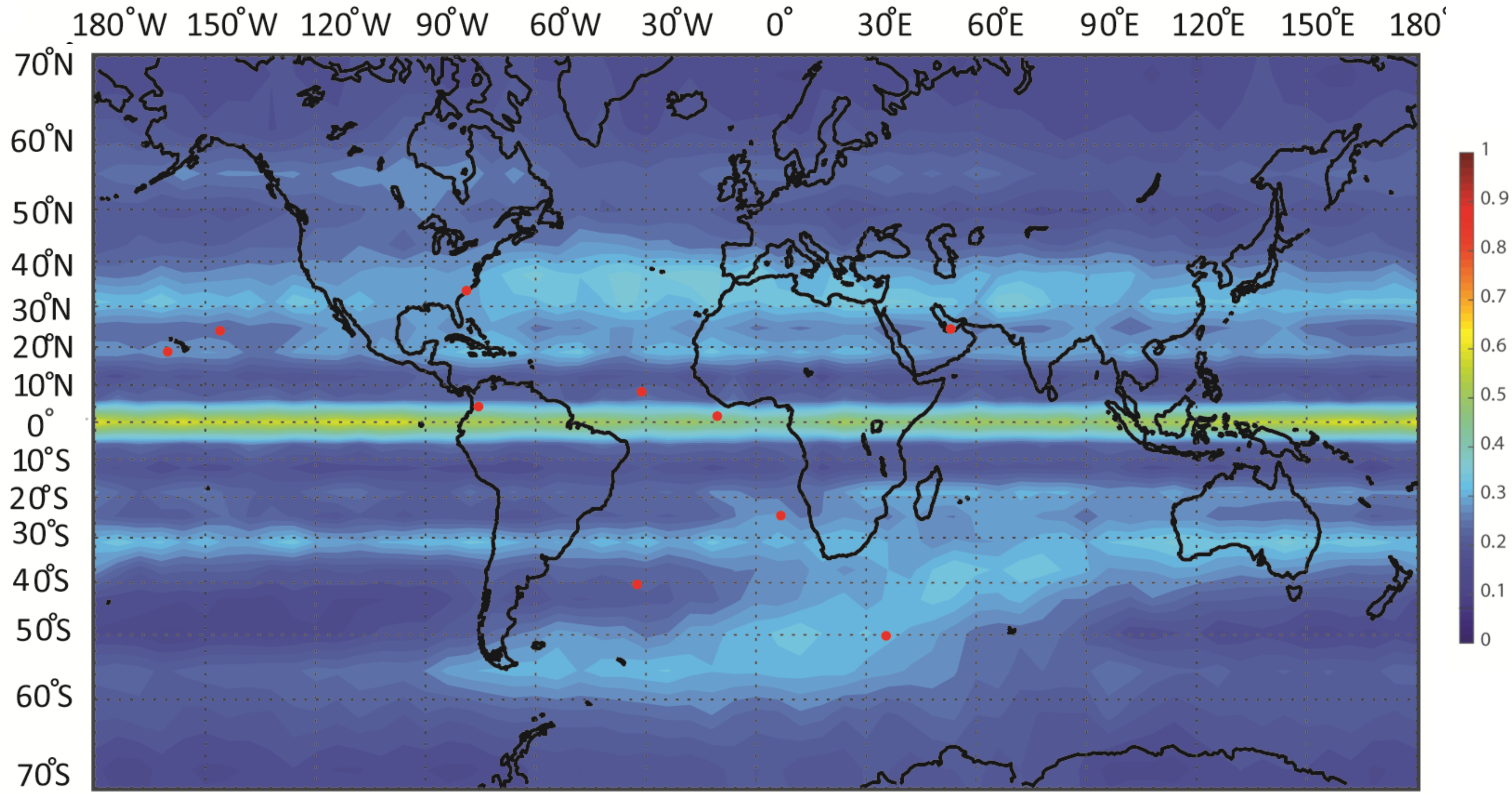


Red Circles = Ob locations

Bgrid Prior Skewness: Obs 10 Days Apart

Average absolute value of skewness over last 200 cycles.

Zonal Wind: Lowest Layer

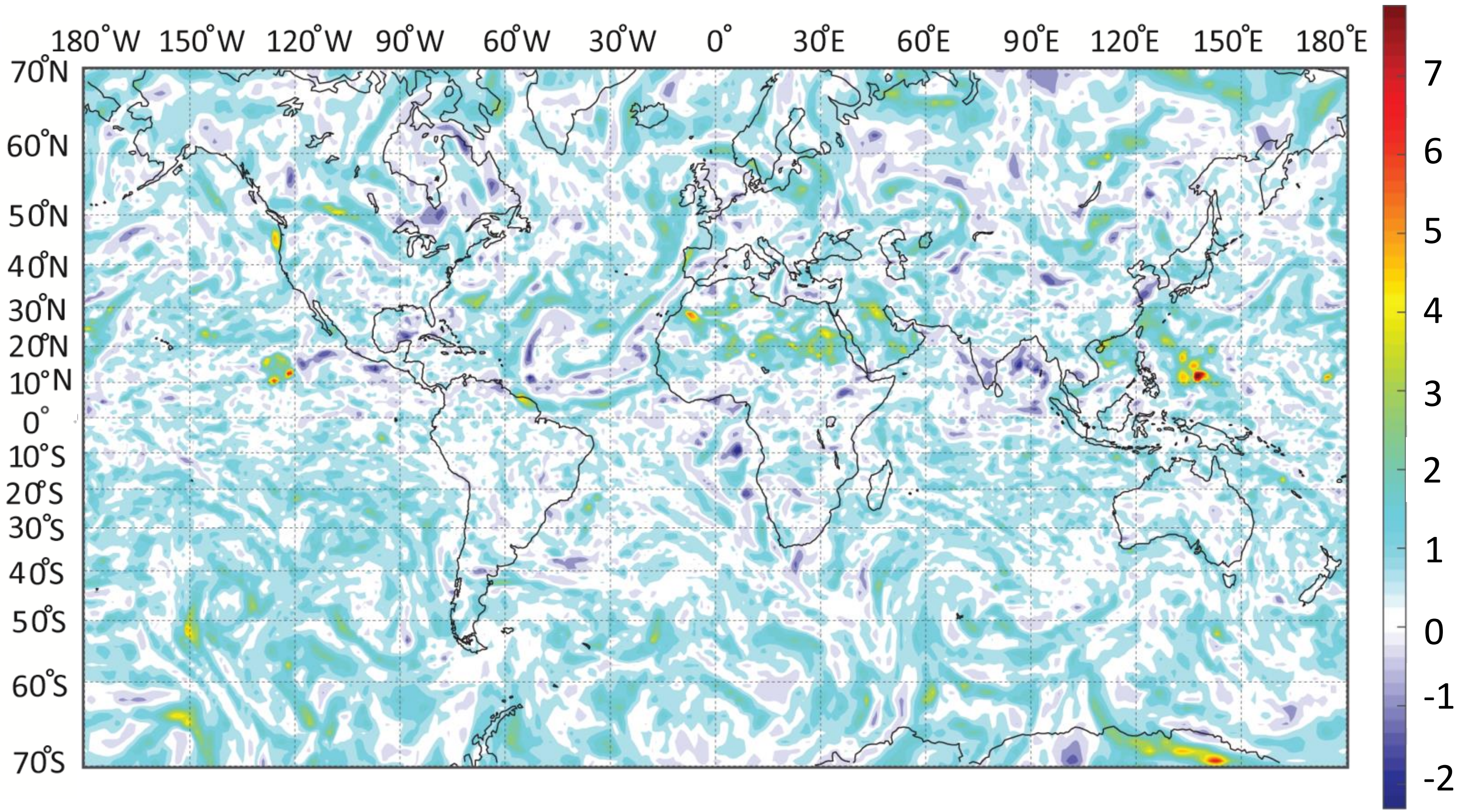


Red Circles = Ob locations

Summary and Future Work

- Polynomial filtering is easily implementable within an already constructed Ensemble-Based Kalman filter.
- There’s little point to non-Gaussian methods if the skewness is small.
  - QC: Only create “pseudo-obs” for obs in regions of high skewness … ?
- Future Work: Testing these results with nonlinear ob operators, higher resolution, and trying this out in the Navy’s aerosol model (NAAPS).

NAAPS Prior Skewness: Total Aerosol Optical Depth



Importance of the Third Moment

If you believe that using a nonlinear observation operator in the numerator of the Kalman gain is useful, viz.

$$\left\langle (x-\overline{x})\left(h(x)-\overline{h(x)}\right)\right\rangle$$

Then note the Taylor-series about $x=\overline{x}$:

$$h(x)=h(\overline{x})+\frac{dh}{dx}(x-\overline{x})+\frac{1}{2}\frac{d^2h}{dx^2}(x-\overline{x})^2+\ldots$$

Use this in the covariance to obtain

$$\left\langle (x-\overline{x})\left(h(x)-\overline{h(x)}\right)\right\rangle=\frac{dh}{dx}\left\langle (x-\overline{x})^2\right\rangle+\frac{1}{2}\frac{d^2h}{dx^2}\left\langle (x-\overline{x})^3\right\rangle+\ldots$$

The lowest-order impact is from the third moment!