Model space localization in serial ensemble filters

Anna Shlyaeva[1,2], Jeff Whitaker[1] and Chris Snyder[3]

[1]NOAA/Earth System Research Lab, Boulder, Colorado, USA

[2]Colorado University/CIRES, Boulder, Colorado, USA

[3]NCAR, Boulder, Colorado, USA

Localization: model-space vs obs-space

Model-space localization

based on distances between state variables

- EnVar: using square root of $L \circ B$
- Modulated ensemble approach in ensemble filters:

Background ensemble is expanded to approximate square root of $L \circ B$

Obs-space localization

based on distances between state variables and observations and/or distances between observations

- Serial ensemble filters: localization of BH^T:

$C_1 \circ (BH^T)$ and $C_2 \circ (HBH^T)$

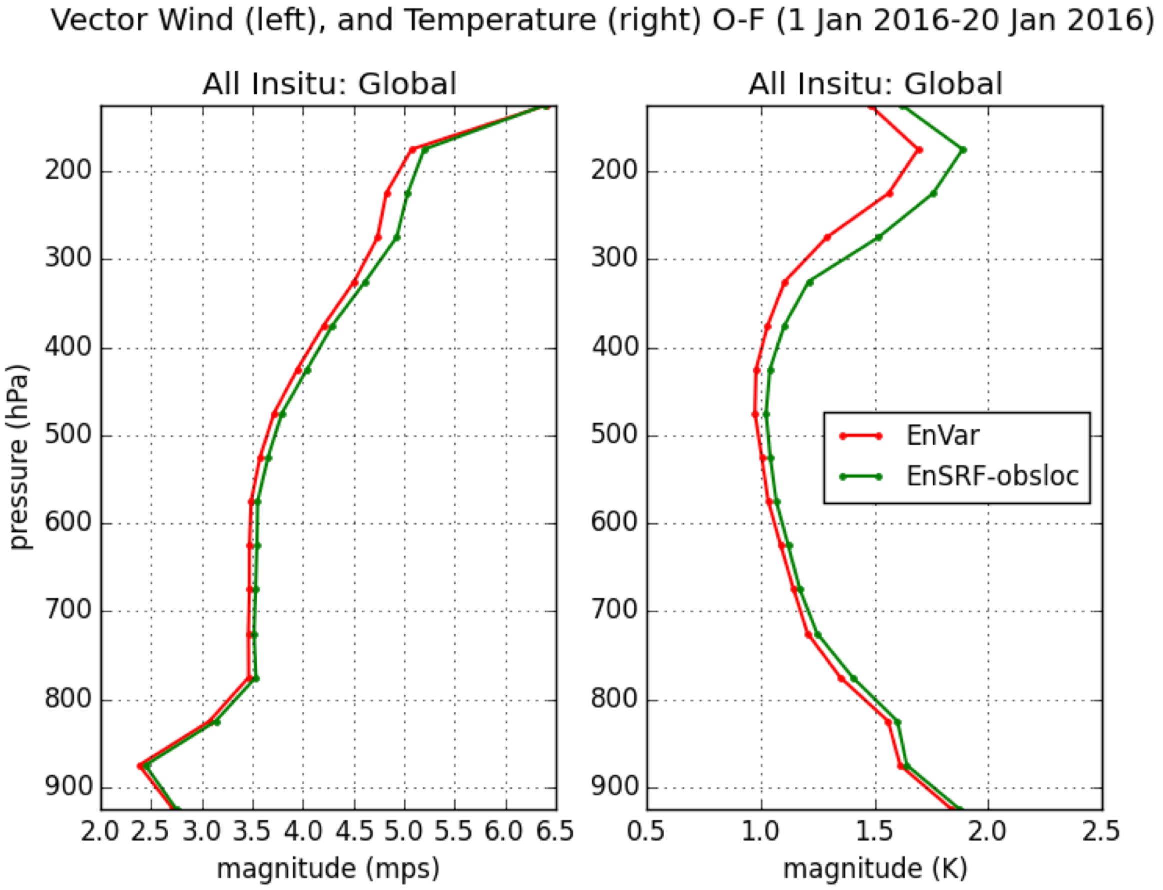
- LETKF-type filters:

Localization through increasing obs error variances based on the distance between the obs and the state variable being updated

Why model-space localization?

- Location of an observation may not be well defined (e.g. satellite radiances).
- Correlations between model variables are often higher than correlations between model and observed variables.
- Multiscale localization methods are more natural to implement in model space.
- Cross-channel observation error correlations are easier to implement with model-space localization.

Background errors when assimilating radiances only in EnVar and EnSRF:



Serial ensemble filter

In a serial filter observations are processed one at a time. Update from the first $k-1$ observations is used as the background for assimilating the k-th observation.

The i-th state variable update with k-th observation:

$$x_i^{(k)} = x_i^{(k-1)} + K_i^{(k)}\left(y_k - h_k(x^{(k-1)})\right)$$

$x^{(k-1)}$ is the state after update with $k-1$ observations,

y_k is the k-th observation with observation error variance σ_k^2, and $K_i^{(k)}$ is a scalar Kalman gain.

Localization in serial filters

Observation-space localization

$$K_i^{(k)} = \frac{C_{ik}\left(XX^T\mathbf{h}_\mathbf{k}^\mathbf{T}\right)_i}{\mathbf{h}_\mathbf{k}XX^T\mathbf{h}_\mathbf{k}^\mathrm{T} + \sigma_k^2}$$

C is the $nx \times ny$ background-obs error covariance localization matrix

$X = (x - \bar{x})/\sqrt{nens - 1}$ are the normalized ensemble perturbations,
$\mathbf{h}_\mathbf{k}$ is the k-th row of the forward operator Jacobian $\mathbf{H}$.

Localization in serial filters

Model-space localization

$$K_i^{(k)} = \frac{(L \circ XX^T)_i \mathbf{h}_\mathbf{k}^\mathbf{T}}{\mathbf{h}_\mathbf{k}(L \circ XX^T)\mathbf{h}_\mathbf{k}^\mathbf{T} + \sigma_k^2}$$

L is the $nx \times nx$ background error covariance localization matrix

Observation-space localization

$$K_i^{(k)} = \frac{C_{ik}\left(XX^T \mathbf{h}_\mathbf{k}^\mathbf{T}\right)_i}{\mathbf{h}_\mathbf{k} XX^T \mathbf{h}_\mathbf{k}^\mathbf{T} + \sigma_k^2}$$

C is the $nx \times ny$ background-obs error covariance localization matrix

$X = (x - \bar{x})/\sqrt{nens - 1}$ are the normalized ensemble perturbations,
$\mathbf{h}_\mathbf{k}$ is the k-th row of the forward operator Jacobian $\mathbf{H}$.

Localization in serial filters

Model-space localization

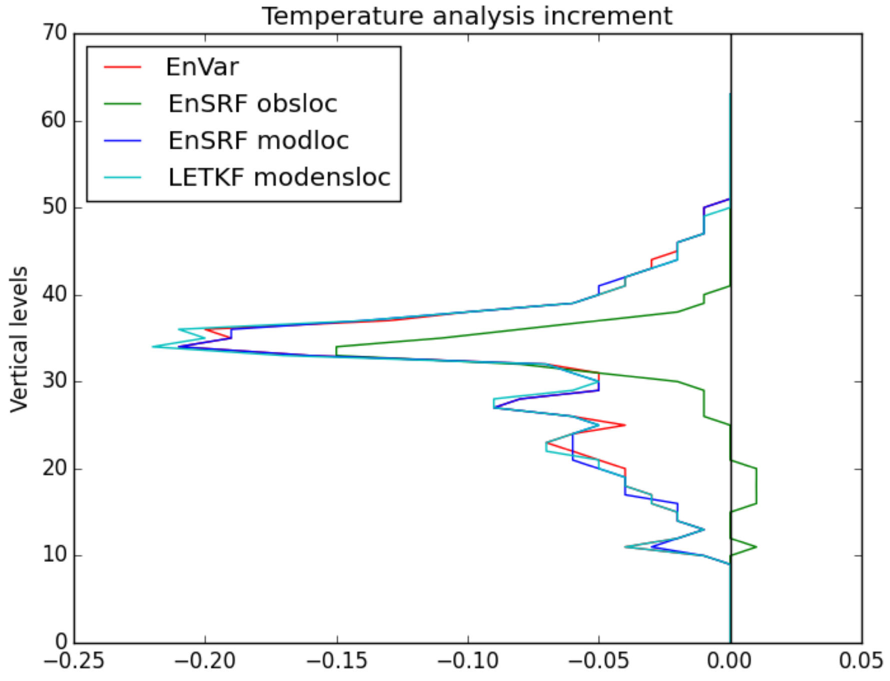
$$K_i^{(k)} = \frac{(L \circ XX^T)_i \mathbf{h}_\mathbf{k}^\mathbf{T}}{\mathbf{h}_\mathbf{k}(L \circ XX^T)\mathbf{h}_\mathbf{k}^\mathbf{T} + \sigma_k^2}$$

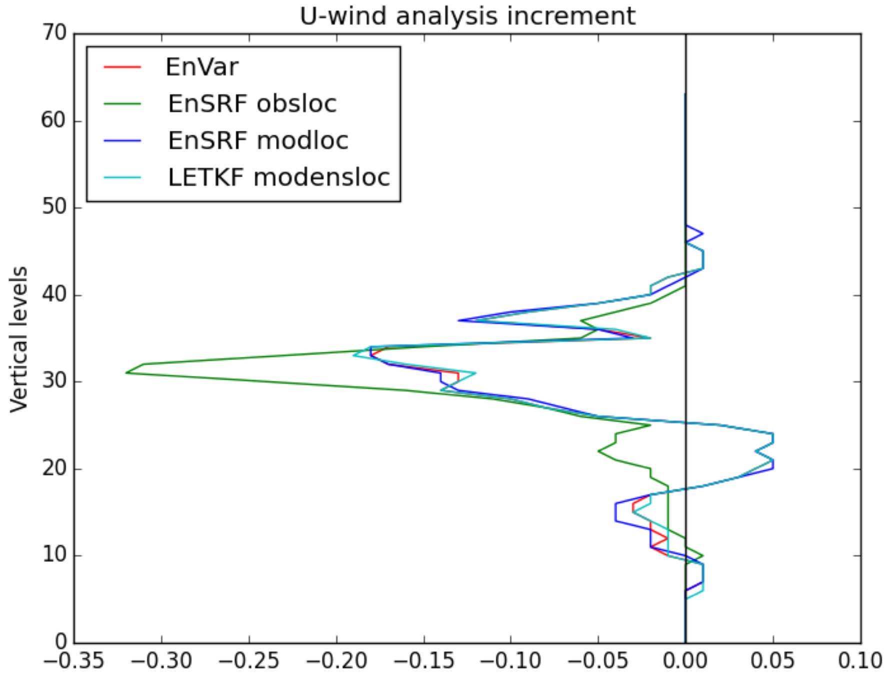
L is the $nx \times nx$ background error covariance localization matrix

- Need to have linearization of the observation operator.
- Additional cost (vs obs-space loc):
  - Factor of nnz for each updated y-x pair (nnz = number of non-zero elements in linearized obs operator);
  - in practice less, depending on how tight localization is.

Implementation

- Implemented for NOAA ensemble square-root filter
- Only vertical model-space localization (important for radiances)
- Single obs experiments (AMSUA channel 7):





DA cycling experiments

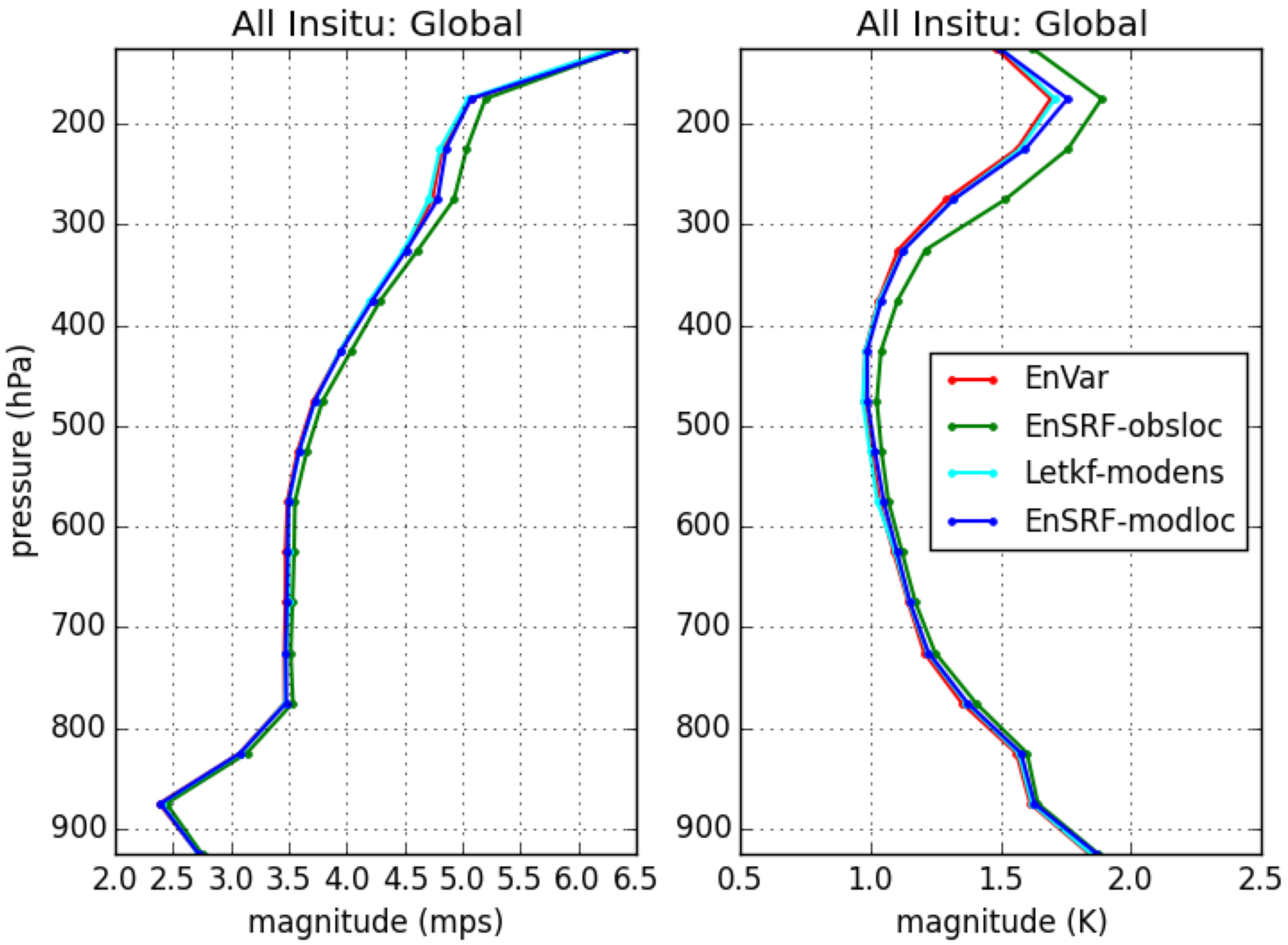
- NOAA Global Forecast System (beta-version), analysis at ~0.7^o resolution
- Compare EnSRF-obsloc, EnSRF-modloc, LETKF-modens (Jeff Whitaker and Lili Lei) & EnVar

Experimental design:

- Radiance-only assimilation (to amplify effect of vertical localization)
- EnVar (for the purpose of fair comparison of localizations):
  - 100% ensemble covariances
  - No outer loops (no relinearization of H)
  - Additional balance constraints off
- All experiments used operational bias correction coefficients

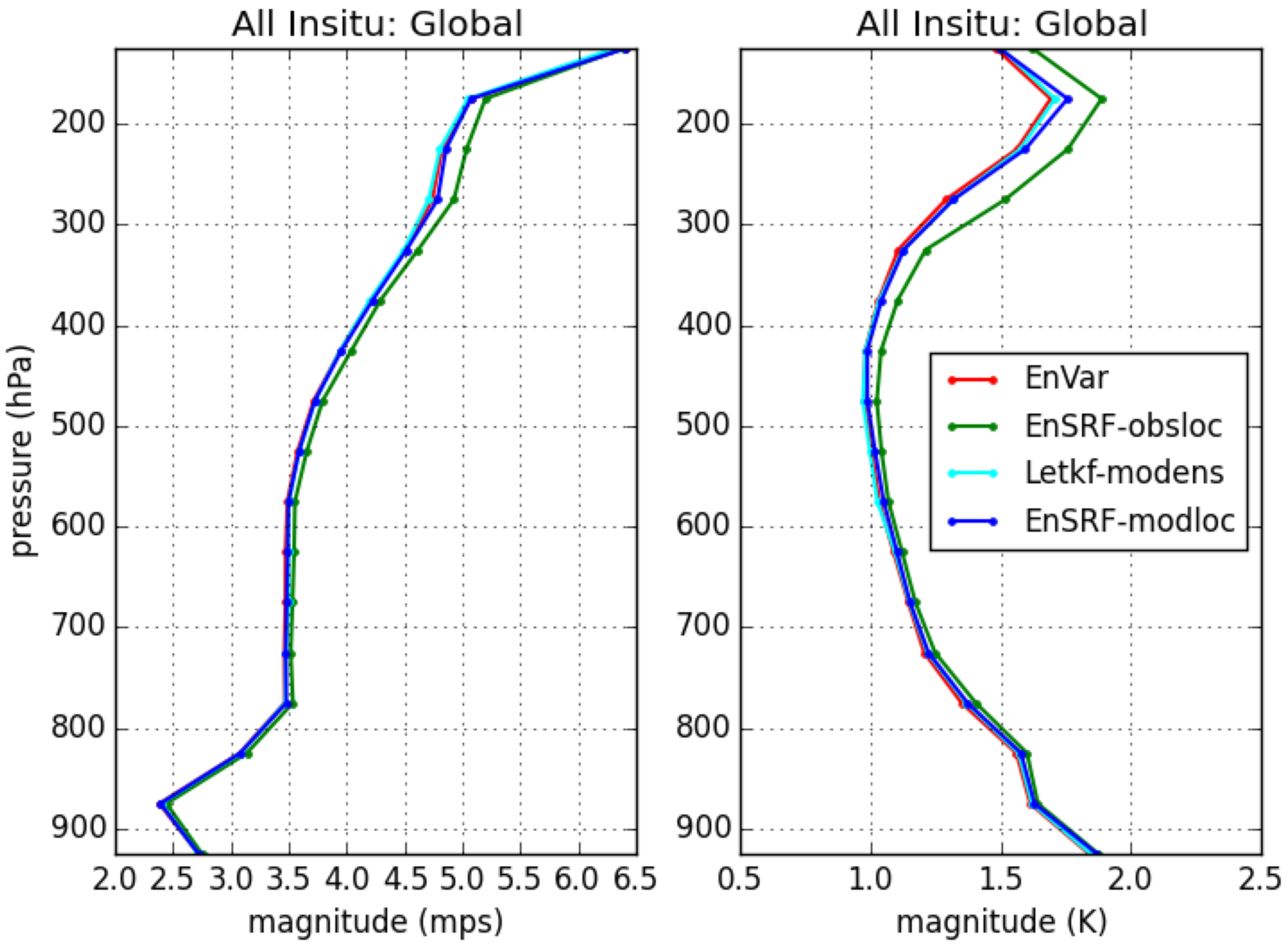
DA cycling results (6-h forecast vs in-situ obs)

Vector Wind (left), and Temperature (right) O-F (1 Jan 2016-20 Jan 2016)



DA cycling results (6-h forecast vs in-situ obs)

Vector Wind (left), and Temperature (right) O-F (1 Jan 2016-20 Jan 2016)



- The differences between obs-space and model-space localization are smaller if
  - localization lengthscale is increased
  - In-situ observations are also included
- EnVar results are better if outer loops are used or additional balance constraint is on

Computational cost of model-space localization

In the experiments shown, EnSRF with model-space localization was ~8 times more expensive than with obs-space localization (when using only radiances).

Possible optimizations:

- Additional computations are thread-parallelizable: optimize threads per node.
- We expect that for the system that distributes model points and observations in local area the data transfer overhead should be less than in our system (with random distribution).

Conclusions

- Model-space localization can be applied in serial ensemble filters directly, if the linearized observation operator is available.
- The additional computational cost depends on how sparse of the observation operator Jacobian is.

- The results of radiance-only assimilation with global NOAA atmospheric system show the benefit of model-space over observation-space localization.