The Sixth WMO Symposium on Data Assimilation
College Park, MD, 7-11 October 2013

A Hybrid Variational-Ensemble Data Assimilation Method
with an Implicit Optimal Hessian Preconditioning

Milija Zupanski

Cooperative Institute for Research in the Atmosphere
Colorado State University
Fort Collins, Colorado, U. S. A.

Acknowledgements:
- National Science Foundation (NSF) - Collaboration in Mathematical Geosciences (CMG)
- NASA Global Precipitation Mission (GPM) Program
- NCAR CISL high-performance computing support (“Yellowstone”)
Hybrid variational-ensemble methods

Hessian preconditioning and static error covariance

Hybrid data assimilation with WRF model and real observations

Future development
Take advantage of both variational and ensemble DA methodologies

Hybrid methods generally address two major aspects:

1) Error covariance
   - flow-dependence
   - rank
   - uncertainty feedback
   **Combine flow-dependent and static error covariance**

2) Nonlinearity
   - iterative minimization
   - Hessian preconditioning
   **Use iterative minimization to obtain optimal analysis solution**
Practical issues of hybrid data assimilation

Combine flow-dependent and static error covariance

1- Linear combination of full matrices or square-root matrices

\[ P_{f}^{1/2} = P_{ENS}^{1/2} + (1 - a)P_{VAR}^{1/2} \]

2- What is the optimal way of combining static and flow dependent matrices?

Use iterative minimization to obtain optimal analysis solution

1- Iterative minimization from variational methods

2- Can this be improved by using an independent iterative minimization with optimal Hessian preconditioning?

[Note: Optimal Hessian preconditioning is defined here as an inverse square-root of the Hessian matrix (e.g., Axelsson and Barker 1984)]

\[ G = EE^T \quad G^{1/2} = E^T \]
Limitations of optimal Hessian preconditioning in hybrid data assimilation

- Assume standard cost function
  \[ J(x) = \frac{1}{2} [x \ x^f]^T P_f^{-1} [x \ x^f] + \frac{1}{2} [y \ h(x)]^T R^{-1} [y \ h(x)] \]
  \[ x^a = x^f + P_f^{1/2} w \]
- Apply common change of variable
- Optimal Hessian preconditioning is
  \[ G^{1/2} = \left( I + P_f^{T/2} H^T R^{-1} H P_f^{1/2} \right)^{1/2} \]

In \textit{variational} data assimilation the inversion is practically impossible due to high dimension of state \((N_s \sim 10^7)\) and static error covariance matrix \((N_s \times N_s)\)

In \textit{ensemble} data assimilation the inversion is possible due to reduced rank ensemble error covariance, implying the preconditioning matrix of smaller size \((N_{\text{ens}} \times N_{\text{ens}})\)

In \textit{hybrid} data assimilation the inversion is limited by requirements of the (full-rank) static error covariance

- **Option #1**: variational framework (use preconditioning from variational methods)
- **Option #2**: ensemble framework (\textbf{define reduced-rank static error covariance} first, then use preconditioning from ensemble methods)

If feasible, the option #2 allows optimal Hessian preconditioning in hybrid data assimilation methods
Reduced-rank static error covariance

1. Assume that a full rank static error covariance square root has been defined

\[ P^{1/2} \]

2. Construct an orthonormal reduced rank matrix \( Q \), and

3. Define a reduced-rank static covariance \( P_{RR} \) as

\[ P_{RR}^{1/2} = P^{1/2} Q \]

**How to define \( Q \)?**

1. Use SVD of local matrix and truncate (preserve similarity with global matrix)
2. Build *global* block-circulant matrix from *local* singular vectors (preserve orthogonality)
3. Scale by diagonal matrix to account for SVD truncation
Processing reduced rank matrix:

Global horizontal response to a single observation

Horizontal response (truth)

<table>
<thead>
<tr>
<th>Horizontal response level = 0.16</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
</tr>
</tbody>
</table>

Horizontal response (RR)

<table>
<thead>
<tr>
<th>Horizontal response level = 0.16</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.18</td>
</tr>
</tbody>
</table>

Horizontal response
( RR + localization + scaling)

<table>
<thead>
<tr>
<th>Horizontal response level = 0.16</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
</tr>
</tbody>
</table>

Sufficient rank covariance becomes acceptable after post-processing
Preliminary assessment the proposed hybrid methodology: Experimental design

- **Model:** WRF-ARW mesoscale model at 27 km / 28 layer resolution
  - 80 x 75 x 28 grid points
- **Control variables:** wind, perturbation potential temperature, specific humidity
- **DA algorithm:** Maximum Likelihood Ensemble Filter (MLEF)
  1. **static:** Reduced rank static forecast error covariance with 40 columns/ensembles
  2. **dynamic:** Standard ensemble algorithm with 32 ensembles
  3. **hybrid:** Combined static and dynamic forecast error covariance with 72 columns/ensembles
- **Observation operator:** Forward component of Gridpoint Statistical Interpolation (GSI)
  - NCEP operational observations and quality control
- **Experimental setup:**
  - May 20, 2013, central United States
  - 6-hour assimilation window $a = 0.7$
  - Linear combination coefficient
Full rank static error covariance

- Toeplitz matrix as a covariance for stationary process
- Simplified cross-correlations between variables

Variable 1: Auto-correlation

Variables 2, 3, 4: Cross-correlation
Synoptic situation

- **Severe weather with tornadoes over Oklahoma**
- **Front associated with a low in upper midwest**

Surface weather map valid 1200 UTC on May 20, 2013

Specific humidity (700 hPa)  
Temperature (700 hPa)

Analysis increments ($x^a-x^i$) of standard MLEF (32 ensembles) show dominant analysis adjustments along the front
Experiment 2: Analysis increments \((x^a-x^b)\) at 700 hPa
(valid 00 UTC 20 May 2013)

Hybrid produces a mixture of dynamic and static information: either one can prevail locally
Summary and future work

- Proof of concept that the presented hybrid system can work with cross-covariances
- Reduced rank static error covariance approach may be feasible for realistic applications – allows optimal Hessian preconditioning
- Preliminary experiments with new hybrid system encouraging
  - realistic model
  - real data
- The anticipated performance has been achieved

- Future improvements of reduced rank static error covariance
  - high-dimensional state and realistic variational covariance
  - examine alternative bases: Fourier, wavelet
- Future improvements of mixing static and dynamic information
  - diagonal matrix instead of alpha (e.g., augmented control variable)
  - define orthogonally complement subspaces
- Tests new hybrid method in realistic weather systems
  - all-sky satellite radiance assimilation
  - coupled land-atmosphere-chemistry models