Spectral Ensemble Kalman Filters

Jan Mandel$^{12}$, Ivan Kasanický$^2$, Martin Vejmelka$^2$, Kryštof Eben$^2$, Viktor Fuglík$^2$, Marie Turčičová$^2$, Jaroslav Resler$^2$, and Pavel Juruš$^2$

$^1$University of Colorado Denver  $^2$Academy of Sciences of the Czech Republic

European Meteorological Society Annual Meeting
Prague, October 6, 2014
Supported by the Czech Science Foundation grant GA13–34856S and the U.S. National Science Foundation grant DMS–1216481.
Ensemble Kalman Filter

- Incorporate the observation $HX \approx Y$ in the probability distribution of the state $X$, represented by the forecast ensemble $X^{f,1}, \ldots, X^{f,N}$.

- Analysis step:

$$X^{a,i} = X^{f,i} - P_N H^* (HP_N H^* + R)^{-1} (HX^{f,i} - Y^i)$$

where

- the superscript * denotes the transpose
- $P_N = \frac{1}{N-1} \sum_{i=1}^{N} (X^{f,i} - \bar{X}^f)(X^{f,i} - \bar{X}^f)^*$ is the ensemble covariance
- $\bar{X}^f = \frac{1}{N} \sum_{i=1}^{N} X^{f,i}$ is the ensemble mean
- $Y^i = Y + \epsilon^i$, $\epsilon^i \sim N(0, R)$ are the perturbed data vectors
- $H$ is the observation operator
The need for localization

- Ensemble covariance is low rank \( N - 1 \Rightarrow \) has large numbers far away from the diagonal

- \( \mathbf{X} \) is a random field and \( \text{Cov}(\mathbf{X}(x), \mathbf{X}(y)) \approx 0 \) for large distance \( y - x \)

- Sampling error: \( \mathbf{P}_N \rightarrow \mathbf{P} \) only asymptotically for \( N \rightarrow \infty \), but real ensembles are small: \( N \approx 20-\text{few} \ 100 \ \text{max} \)

- Tapering fix: \( \mathbf{P}_N \rightarrow \mathbf{P}_N \circ \mathbf{T} \), multiply term-by-term by a fixed tapering matrix to force small entries when \( x - y \) large. But how far away exactly? Depends on \( N \), for large \( N \) the convergence \( \mathbf{P}_N \rightarrow \mathbf{P} \) should take over

- Expensive/hard to parallelize implementation, banded/sparse matrix operations.
Covariance of random fields

- Covariance between the values at two points $x, y$ is the covariance function $f_x(x - y) = \text{Cov}(X(x), X(y))$

- If the covariance function does not change with location, the covariance matrix is diagonal in the Fourier basis $u_1, \ldots, u_n$ (sines, cosines or complex exponential)

\[
\text{Cov}(X) = \mathbf{F}^* \begin{bmatrix}
\lambda_1 \\
\vdots \\
\lambda_n
\end{bmatrix} \mathbf{F}, \quad D = \mathbf{F} \text{Cov}(X) \mathbf{F}^*
\]

- Multiplication by $F = [u_1, \ldots, u_n]^*$ is a discrete Fourier transform

- Other orthogonal bases (e.g., wavelets) and frames allow variability with location.
Analysis of spectral diagonal covariance

**Theorem.** Suppose Cov \((X)\) after the transformation is diagonal:

\[
F \text{Cov} (X) F^* = \begin{bmatrix}
\lambda_1 \\
\vdots \\
\lambda_n
\end{bmatrix}
\]

- Transform the **sample covariance** \(P_N\) in the same way, keep only the diagonal part \(D\).
- \(D\) transformed back is better than \(P_N\) in the Frobenius norm:

\[
\mathbb{E} \left\| F^* DF - \text{Cov} (X) \right\|_F^2 = \frac{2}{N} \sum_{i=1}^{n} \lambda_i^2 < \mathbb{E} \left\| P_N - \text{Cov} (X) \right\|_F^2 = \frac{2}{N} \sum_{i=1}^{n} \lambda_i^2 + \frac{1}{N} \sum_{i \neq j} \lambda_i \lambda_j
\]
Spectral EnKF - Simple case: single variable, all observed

- Assuming observation operator $H = I$, data covariance $R = I$
- Compute the diagonal $D$ of the covariance of the transformed forecast ensemble $\left[FX^f,1, \ldots, FX^f,N \right]$
- Analysis update of the transformed forecast ensemble becomes multiplication by a diagonal matrix:

$$FX^{a,i} = FX^{f,i} - D(D+I)^{-1}F \left(X^{f,i} - Y^i \right)$$

- Inverse transform the analysis ensemble: $X^{a,i} = F^* \left(FX^{a,i} \right)$, $i = 1, \ldots, N$
Spectral EnKF - More general state and observation

- **Low-dimensional and scalar observations:**
  - Use spectral diagonal covariance $F^TDF$ in place of the ensemble covariance
  - Few matrix-vector multiplications to set up
  - Only need to invert a small matrix or a scalar

- **Multiple variables on the same grid, one completely observed:**
  - Spectral diagonal crosscovariances between the variables

- **Multiple variables on different grids, same dimension:**
  - Interpolate all variables to the same grid
  - Extend the analysis back to the original grids

- **Both 2D and 3D variables**
  - Treat 2D layers as separate variables
Spectral EnKF - Part of a variable observed

- Extend the data to whole domain by zeros
- Augment the state by a copy of the variable with 0 outside of the data region
- Wavelets, not Fourier, for locality
Shallow water equations

- Variables: fluid depth, horizontal velocities
- Equations: conservation of mass and horizontal momenta
- 64 x 64 grid with step 150 km, depth 10 km
- Background $\sigma=100$ m, spin-up 15 m, time step 1 s, assimilation cycle 10 seconds
Shallow water equations - Assimilation cycles

Mean RMSE from 10 repetitions

DST=discrete sine transform, DCT=discrete cosine transform,
DWT=wavelet transform, Coiflet 5
DST = discrete sine transform, DCT = discrete cosine transform,
DWT = wavelet transform, Coiflet 5
WRF - assimilation setup

- WRF 3.6, one domain with resolution 27 x 27 km covering Middle Europe, 39 vertical levels
- a common WRF configuration (NOAH Land-surface model, Lin et al. microphysics, Dudhia shortwave radiation, Yonsei PBL)
- initial ensemble: perturbations of a deterministic GFS-initialized run
- 6 hours spin-up for downscaling
- each 2D layer of 3D WRF variables is a separate variable for assimilation
- wind interpolated to cell centers for the same dimension
- assimilation every hour (6 assimilation cycles)
WRF - assimilation setup

- background covariance constructed from a one month simulation of WRF on the same domain, (NMC method, WRFDA routine gen_be)
- corresponding perturbations by means of WRFDA da_wrfvar in randomcv regime

V10 y-wind component for two initial ensemble members
WRF - assimilation of 2D potential temperature layer

- 9 ensemble members, 10th member = the truth
- observation = the lowest 2D layer of the potential temperature $T$ in the 10th ensemble member, $\sigma = 0.316$ K
WRF - assimilation of a station temperature

Difference between analysis and forecast mean in the lowest 2D layer of the potential temperature $T$

- Standard EnKF
- Spectral diagonal, sine transform

- ensemble with 10 members, observation = $T$ in the middle of the grid, mean forecast $+ 1$ K, $\sigma = 0.2$ K
Conclusions

- Covariance is diagonal or close in the spectral domain
- Fast - FFT or wavelet transform, diagonal matrices
- FFT is better in the spatially homogeneous case
- Wavelets are better in the spatially nonhomogeneous cases
- Important kinds of observation equations supported
- Automatic, no tuning of covariance distance
- Needs only very small ensembles, about 10
- Preserves convergence in the large ensemble limit

Future work:

- Better understanding of cross-covariances
- Mathematical analysis of more general cases
- Frames, better 2D wavelets
References

Jonathan D. Beezley, Jan Mandel, and Loren Cobb.

Jan Mandel, Jonathan D. Beezley, and Volodymyr Y. Kondratenko.

Olivier Pannekoucke, Loïk Berre, and Gerald Desroziers.

Reinhard Furrer and Thomas Bengtsson.
Tom Auligné and Gael Descombes.
Background error covariance and GEN_BE. *2014 GSI Community Tutorial*.

Jan Mandel, Loren Cobb, and Jonathan D. Beezley.