Advanced data assimilation methods-
EKF and EnKF

Hong Li and Eugenia Kalnay
University of Maryland

17-22 July 2006
Recall the basic formulation in OI

- **OI for a Scalar:**
  \[ T_a = T_b + w(T_0 - T_b) \]
  \[ 0 \leq w \leq 1 \]
  Optimal weight to minimize the analysis error is:
  \[ w = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_o^2} \]

- **OI for a Vector:**
  \[ x_i^f = Mx_i^{a_{i-1}} \]
  \[ x_i^a = x_i^b + W(y_i^o - Hx_i^b) \]
  \[ W = BH^T[HBH^T + R]^{-1} \]
  \[ P^a = [I - WH]B \]
  \[ B = \delta x_b \cdot \delta x_b^T \]
  \[ \delta x_b = x_b - x_t \]
  \[ \delta x_b = x_b - x_t \]

- B is statistically pre-estimated, and constant with time in its practical implementations. Is this a good approximation?
OI and Kalman Filter for a scalar

- **OI for a scalar:**

  \[ T_b(t_{i+1}) = M(T_a(t_i)); \quad \sigma_b^2 = (1 + a)\sigma_a^2 = \frac{1}{1-w} \sigma_a^2 = \text{const.} \]

  \[ w = \frac{\sigma_b^2}{\sigma_a^2 + \sigma_o^2}; \quad \sigma_a^2 = (1-w)\sigma_b^2 \]

  \[ T_a = T_b + w(T_0 - T_b) \]

- **Kalman Filter for a scalar:**

  \[ T_b(t_{i+1}) = M(T_a(t_i)); \quad \sigma_b^2(t) = (L\sigma_a)(L\sigma_a)^T; \quad L = dM / dT \]

  \[ w = \frac{\sigma_b^2}{\sigma_a^2 + \sigma_o^2}; \quad \sigma_a^2 = (1-w)\sigma_b^2 \]

  \[ T_a = T_b + w(T_0 - T_b) \]

- Now the background error variance is **forecasted** using the linear tangent model L and its adjoint \( L^T \)
“Errors of the day” computed with the Lorenz 3-variable model: compare with rms (constant) error

- Not only the amplitude, but also the structure of B is constant in 3D-Var/OI
- This is important because analysis increments occur only within the subspace spanned by B

$\sigma_z^2 = (z_b - z_t)^2$

$= 0.08$
“Errors of the day” obtained in the reanalysis (figs 5.6.1 and 5.6.2 in the book)

- Note that the mean error went down from 10m to 8m because of improved observing system, but the “errors of the day” (on a synoptic time scale) are still large.

- In 3D-Var/OI not only the amplitude, but also the structure of B is fixed with time.
Flow independent error covariance

- In OI (or 3D-Var), the scalar error correlation between two points in the same horizontal surface is assumed homogeneous and isotropic. (p162 in the book)

- If we observe only Washington, D.C, we can get estimate for Richmond and Philadelphia corrections through the error correlation (off-diagonal term in B).

\[
B = \begin{pmatrix}
\sigma_1^2 & \text{cov}_{1,2} & \cdots & \text{cov}_{1,n} \\
\text{cov}_{1,2} & \sigma_2^2 & \cdots & \text{cov}_{2,n} \\
\vdots & \vdots & \ddots & \vdots \\
\text{cov}_{1,n} & \text{cov}_{2,n} & \cdots & \sigma_n^2
\end{pmatrix}
\]

- In OI(or 3D-Var), the scalar error correlation between two points in the same horizontal surface is assumed homogeneous and isotropic. (p162 in the book)
Typical 3D-Var error covariance

In OI(or 3D-Var), the error correlation between two mass points in the same horizontal surface is assumed homogeneous and isotropic. (p162 in the book)

$$B = \begin{bmatrix} \sigma_1^2 & \text{cov}_{1,2} & \ldots & \text{cov}_{1,n} \\ \text{cov}_{1,2} & \sigma_2^2 & \ldots & \text{cov}_{2,n} \\ \ldots & \ldots & \ldots & \ldots \\ \text{cov}_{1,n} & \text{cov}_{2,n} & \ldots & \sigma_n^2 \end{bmatrix}$$
Flow-dependence – a simple example (Miyoshi, 2004)

There is a cold front in our area…
What happens in this case?

This is not appropriate
This does reflects the flow-dependence.
Extended Kalman Filter (EKF)

- **Forecast step**
  \[ \mathbf{x}_i^b = M \mathbf{x}_{i-1}^a \]
  \[ \mathbf{P}_i^b = \mathbf{L}_{i-1} \mathbf{P}_{i-1}^a \mathbf{L}^T_{i-1} + \mathbf{Q} \]

- **Analysis step**
  \[ \mathbf{x}_i^a = \mathbf{x}_i^b + \mathbf{K}_i (\mathbf{y}_i^o - \mathbf{Hx}_i^f) \]
  \[ \mathbf{K}_i = \mathbf{P}_i^b \mathbf{H}^T [\mathbf{HP}_i^b \mathbf{H}^T + \mathbf{R}]^{-1} \]
  \[ \mathbf{P}_i^a = \mathbf{I} - \mathbf{K}_i \mathbf{H} \]

- Using the flow-dependent \( \mathbf{P}_i^b \), analysis is expected to be improved significantly

However, it is computationally expensive. \( \mathbf{P}_i^b, \mathbf{L}_i \) \( n \times n \) matrix, \( n \sim 10^7 \)
computing equation directly is impossible.
Ensemble Kalman Filter (EnKF)

\[
P_i^b = L_{i-1} P_i^a L_{i-1}^T + Q
\]

- Although the dimension of \( P_i^f \) is huge, the rank (\( P_i^f \)) << \( n \) (dominated by the errors of the day)
  \[
P_i^b \approx \frac{1}{m} \sum_{k=1}^{m} (x_k^f - x^f)(x_k^f - x^f)^T
  
  \text{Ideally} \quad m \to \infty
  
\]

- Using ensemble method to estimate
  \[
P_i^b \approx \frac{1}{K - 1} \sum_{k=1}^{K} (x_k^f - \overline{x^f})(x_k^f - \overline{x^f})^T
  
  = \frac{1}{K - 1} \mathbf{X}^b \cdot \mathbf{X}^{bT}
  
  K \text{ ensemble members, } K<<(\text{n})
  
\]

- Problem left: How to update ensemble?
  i.e.: How to get \( \mathbf{x}_i^d \) for each ensemble member? Using ** K times?

Physically,
- "errors of day" are the instabilities of the background flow. Strong instabilities have a few dominant shapes (perturbations lie in a low-dimensional subspace).
- It makes sense to assume that large errors are in similarly low-dimensional spaces that can be represented by a low order EnKF.
Ensemble Update: two approaches

1. Perturbed Observations method:
An “ensemble of data assimilations”

- It has been proven that an observational ensemble is required (e.g., Burgers et al. 1998). Otherwise \( P^a_{i n \times n} = [I - K_i H] P^b_i \) is not satisfied.

- Random perturbations are added to the observations to obtain observations for each independent cycle

\[
y^{o}_i(k) = y^{o}_i + \text{noise}
\]

- However, it introduces a source of sampling errors when perturbing observations.
Ensemble Update: two approaches

2. Ensemble square root filter (EnSRF)

- Observations are assimilated to update only the ensemble mean.
  \[
  \overline{x}_i^a = \overline{x}_i^b + K_i (y_i^o - H \overline{x}_i^b)
  \]

- Assume analysis ensemble perturbations can be formed by transforming the forecast ensemble perturbations through a transform matrix.

\[
\frac{1}{k-1} X^a X^{aT} = P_{i,n \times n}^a = [I - K_i H] P_i^b = [I - K_i H] \frac{1}{k-1} X^b X^{bT}
\]

\[
x_i^b = M x_{i-1}^a
\]

\[
P_i^b \approx \frac{1}{K-1} \sum_{k=1}^K (x_k^b - \overline{x}_i^b)(x_k^b - \overline{x}_i^b)^T
\]

\[
K_i = P_i^b H^T [H P_i^b H^T + R]^{-1}
\]

\[
\overline{x}_i^b = \overline{x}_i^b + K_i (y_i^o - H \overline{x}_i^b)
\]

\[
X_i^a = T_i X_i^b
\]

\[
x_i^a = x_i^a + X^a
\]
Several choices of the transform matrix

- EnSRF, Andrews 1968, Whitaker and Hamill, 2002)
  \[ X^a = (I - \tilde{K}H)X^b, \tilde{K} = \alpha K \]

- EAKF (Anderson 2001)
  \[ X^a = AX^b \]

- ETKF (Bishop et al. 2001)
  \[ X^a = X^b T \]

- LETKF (Hunt, 2005)
  Based on ETKF but perform analysis simultaneously in a local volume surrounding each grid point
An Example of the analysis corrections from 3D-Var (Kalnay, 2004)

An example with the QG system (Corazza et al, 2003)

Background error (color) and 3D-Var analysis correction (contours)

The analysis corrections due to the observations are isotropic because they don’t know about the errors of the day
An Example of the analysis corrections from EnKF (Kalnay, 2004)

QG model example of Local Ensemble KF (Corazza et al)

Background error (color) and LEKF analysis correction

The LEKF does better because it captures the errors of the day
Summary of LETKF (Hunt, 2005)

Forecast step (done globally): advance the ensemble 6 hours
\[ x_n^{b(i)} = M(x_{n-1}^{a(i)}) \quad i = 1, \ldots, k \]

Analysis step (done locally). Local model dimension m, locally used s obs
\[ X = \begin{bmatrix} x^{b(1)} & \ldots & x^{b(k)} \end{bmatrix} \quad \text{Matrix of ensembles (mxk)} \]
\[ X^b = X - \bar{x}^b \quad \text{Matrix of ensemble perturbations (mxk)} \]
\[ Y^b = H(X) - H(\bar{x}^b) \approx HX^b \quad \text{Matrix of ens. obs. perturbations (sxk)} \]
\[ P^b = (k - 1)^{-1}X^bT X^b \quad \text{in model space, so } \tilde{P}^b = (k - 1)^{-1}I \quad \text{in ensemble space (kxk)} \]
\[ (P^a)^{-1} = (P^b)^{-1} + H^T R^{-1} H \quad \text{in model space, so that} \]
\[ (\tilde{P}^a)^{-1} = (k - 1)I + (HX^b)^T R^{-1} (HX^b) \quad \text{in ensemble space (kxk)} \]
Summary of LETKF (cont.)

Forecast step (done globally): advance the ensemble 6 hours
\[ x_n^{b(i)} = M(x_{n-1}^{a(i)}) \quad i = 1, \ldots, k \]

Analysis step (done locally). Local model dimension m, locally used obs s
\[ X^b = X - \bar{x}^b \quad Y^b = H(X) - H(\bar{x}^b) \approx HX^b \]
\[ (\tilde{P}^a) = \left[(k-1)I + (HX^b)^T R^{-1}(HX^b)\right]^{-1} \quad \text{in ensemble space (kxk)} \]
\[ P^a = X^{aT}X^a = X^{bT}\tilde{P}^aX^b \quad \text{in model space (mxm)} \]
\[ X^a = X^b (\tilde{P}^a)^{1/2} \quad \text{Ensemble analysis perturbations in model space (mxk)} \]
\[ w_n^a = \tilde{P}^aY^bT R^{-1}(y_n^o - \bar{y}_n^b) \quad \text{Analysis increments in ensemble space (kx1)} \]
\[ \bar{x}_n^a = \bar{x}_n^a + X^b w_n^a \quad \text{Analysis in model space (mx1)} \]

We finally gather all the analysis and analysis perturbations from each grid point and construct the new global analysis ensemble (nxk)
Summary steps of LETKF

1) Global 6 hr ensemble forecast starting from the analysis ensemble
2) Choose the observations used for each grid point
3) Compute the matrices of forecast perturbations in ensemble space $X^b$
4) Compute the matrices of forecast perturbations in observation space $Y^b$
5) Compute $P^b$ in ensemble space space and its symmetric square root
6) Compute $w^a$, the k vector of perturbation weights
7) Compute the local grid point analysis and analysis perturbations.
8) Gather the new global analysis ensemble.
EnKF vs 4D-Var

- EnKF is simple and model independent, while 4D-Var requires the development and maintenance of the adjoint model (model dependent).
- 4D-Var can assimilate asynchronous observations, while EnKF assimilate observations at the synoptic time.
- Using the weights $w_a$ at any time 4D-LETKF can assimilate asynchronous observations and move them forward or backward to the analysis time.

Disadvantage of EnKF:
- Low dimensionality of the ensemble in EnKF introduces sampling errors in the estimation. ‘Covariance localization’ can solve this problem.