The Mollified Ensemble Kalman Filter (MEnKF)...
and other ideas

Javier Amezcua
(discussing work by Bergemann and Reich)

Chaos-Weather meeting 06/01/10
Works studied

- **2010b: MEnKF**
  - Impact of discontinuous analysis increments.
  - Technique to minimize impacts tested in a novel modified Lorenz 40var model.

- **2010a: Localization.**
  - Localization based on a ODE formulation of the analysis step.
  - ODE formulation.

- **2009: Propagation and continuous factorization.**
  - SVD decomposition of ensembles (continuous).
  - ODE formulation.
Problem

\[ \frac{d}{dt} x(t) = f(x,t) \]

Discontinuous analysis increments

Spurious high frequency oscillations.

E.g. generation of artificial gravity waves and loss of [geostrophic] balance.
Alternatives

- Nudging or Newtonian Relaxation

\[
\frac{d}{dt} x^b(t) = f(x^b(t), t) + \alpha(x^o - x^b)
\]

- Incremental Analysis Update (IAU)
Mollified Ensemble Kalman Filter

$n$ state variables, $m$ ensemble members, $k$ observations.

- We have a system described by:
  \[
  \frac{d}{dt} x(t) = f(x, t) \quad x(t_0) \sim N(x_0, B) \quad y(t_j) - H \bar{x}(t_j) \sim N(0, R)
  \]

- We have an ensemble of $m$ solutions. As always we define:
  \[
  \bar{x}(t) = \frac{1}{m} \sum_{i=1}^{m} x_i(t) \quad P(t) = \frac{1}{m-1} \sum_{i=1}^{m} (x_i(t) - \bar{x}(t))(x_i(t) - \bar{x}(t))^T
  \]
Formulation of the MEnKF

- The evolution of the $i^{th}$ member:

$$\frac{d}{dt}x_i(t) = f(x_i, t) - \sum_{j=1}^{M} \delta(t - t_j) P \nabla_{x_i} V_j(X)$$

- With an [observation-related] ‘potential’ defined as:

$$V_j(X) = \frac{m}{2} \left[ S_j(\bar{x}) + \frac{1}{m} \sum_{i=1}^{m} S_j(x_i) \right]$$

$$S_j(x) = \frac{1}{2} (H x - \underline{y}(t_j))^T R^{-1} (H x - \underline{y}(t_j))$$

- The observations enter the system as impulses.
The Dirac Delta Function

- Generalized function, distribution or measure with the following properties:

\[ \delta(x - a) = \begin{cases} 0, & x \neq a \\ \infty, & x = a \end{cases} \]

\[ \int_{-\infty}^{+\infty} \delta(x - a) \, dx = 1 \]

- By consequence:

\[ \int_{-\infty}^{+\infty} f(x) \delta(x - a) \, dx = f(a) \]
The Dirac Delta Function

- As a limit:
  - Of a rectangular function
    \[
    \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} [H(x + \varepsilon) - H(x - \varepsilon)]
    \]
  - Of a Gaussian
    \[
    \lim_{\sigma \to 0} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2} \left( \frac{x}{\sigma} \right)^2}
    \]
The Dirac Delta Function

- Of a standard hat function

\[ \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \psi \left( \frac{s}{\varepsilon} \right) \]

\[ \psi(s) = \begin{cases} 
1 - |s| & |s| \leq 1 \\
0 & \text{else}
\end{cases} \]
Mollify the discontinuous adjustments in our ODE:

\[
\frac{dx_i(t)}{dt} = f(x_i, t) - \sum_{j=1}^{M} \delta(t - t_j) \text{PV}_{\bar{x}_i} V_j(X)
\]

Mollify the Dirac Delta [with a hat function]
**Modified Lorenz 40var model**

- A model with the usual [slow] variables and coupled fast variables.

\[
\frac{d}{dt} x_i = (x_{i+1} - x_{i-2})x_{i-1} - x_i + 8
\]

\[
\frac{d^2}{dt^2} h_i = \frac{1}{\varepsilon^2} \left[ -h_i + \alpha^2 (h_{i+1} - 2h_i + h_{i-1}) \right]
\]

- \(\alpha > 0\) Controls de dispersion in the grid.
- \(0 < \varepsilon << 1\) Controls the faster evolution (wrt the original variables).
Modified Lorenz 40var model

- Variables coupled through an energy exchange term.
  \[ E_{\text{coupling}} = -\delta \sum_{i=1}^{40} h_i x_i \]
  \( \delta > 0 \) Coupling strength

- The new system

\[
\frac{d}{dt} x_i = (1 - \delta)(x_{i+1} - x_{i-2})x_{i-1} + \delta(x_{i-1}h_{i+1} - x_{i-2}h_{i-1}) - x_i + 8
\]

\[
\varepsilon^2 \frac{d^2}{dt^2} h_i = -h_i + \alpha^2 \left[ h_{i+1} - 2h_i + h_{i-1} \right] + x_i
\]
Modified Lorenz 40var model

- The fast dynamics of the model is entirely conservative.
- We can notice a balance relationship:

\[ x_i = h_i - \alpha^2 [h_{i+1} - 2h_i + h_{i-1}] \]
Results of the MEnKF in the modified Lorenz 40var model
Considerations on the MEnKF

- Spreading the analysis.

- Can be viewed as nudging with the nudging coefficients determined from ensemble prediction.

- IAU can become instable faster than MEnKF.
Finding the analysis ensemble

- The ensembles
  \[ X = [x_1 - \bar{x} \mid \cdots \mid x_m - \bar{x}] \]

- Usual transformations
  \[ X^a = A X^b \]
  \[ X^a = X^b W \]
  \[ W = \left[ (k-1) \tilde{P}^a \right] \]
Finding the analysis ensemble $X^b \rightarrow X^a$

- The ODE formulation (Simons, 2006)
  \[
  \frac{d}{ds} X = -\frac{1}{2(m-1)} XX^T H^T R^{-1} HX
  \]
  Pseudotime $0 \leq s \leq 1$
  \[X(0) = X^b \quad X(1) = X^a\]

- Also
  \[
  \frac{d}{ds} \bar{x} = -\frac{1}{m-1} XX^T H^T R^{-1} (H \bar{x} - \underline{y})
  \]
  \[\bar{x}(0) = \bar{x}^b \quad \bar{x}(1) = \bar{x}^a\]
The ODE formulation

- Putting together

\[
\begin{align*}
\frac{d}{ds} X &= -\frac{1}{2(m-1)} XX^T H^T R^{-1} H X \\
\frac{d}{ds} \bar{x} &= -\frac{1}{m-1} XX^T H^T R^{-1} (H \bar{x} - \bar{y})
\end{align*}
\]

- And defining

\[
S(x) = \frac{1}{2} (H \bar{x} - \bar{y}) R^{-1} (H \bar{x} - \bar{y})^T
\]

- Then

\[
\frac{d}{dt} x_i = -\frac{1}{2} P \left[ \nabla_{x_i} S(x_i) + \nabla_{\bar{x}} S(\bar{x}) \right]
\]
The ODE formulation

- Moreover, \( \forall s, \quad 0 \leq s \leq 1 \) we have closure:

\[
\bar{x}(s) = \frac{1}{m} \sum_{i=1}^{m} x_i(s) \quad P(s) = \frac{1}{m-1} \sum_{i=1}^{m} (x_i(s) - \bar{x}(s))(x_i(s) - \bar{x}(s))^T
\]

- So that

\[
\frac{d}{ds} x_i = -P \nabla \bar{x_i} \left[ \frac{m}{2} \left( S(\bar{x}) + \sum_{i=1}^{m} S(x_i) \right) \right]
\]

\[
V(X)
\]

- Finally for the system

\[
\frac{d}{dt} x_i(t) = f(x_i, t) - \sum_{j=1}^{M} \delta(t - t_j) P \nabla \bar{x_i} V_j(X)
\]
Plans with the MEnKF

- Apply the formulation in the Lorenz 3D variable model.
  - No balance to be checked, but to see what happens!

- Apply the formulation to a 2D Shallow Water Equations system.
  - Creation of spurious gravity waves.
  - Geostrophic balance.