Mesoscale Weather Prediction with a hybrid EnKF-4DVAR system

Kosuke Ito

JAMSTEC/MRI

Dr. Masaru Kunii kindly provided the results of EnKF
Concept of 4DVAR

• Cost function

\[ J = \frac{1}{2} (x_0 - x_b)^T B^{-1} (x_0 - x_b) \]
\[ + \frac{1}{2} \sum_t (y_t - Hx_t)^T R^{-1} (y_t - Hx_t) \]

Misfit b/w init. value of analysis and first guess
Misfit b/w analysis and observation

• Optimization problem

Find \( x_0 \) that gives the minimum value of \( J \)
(with the constraint of the model dynamics
\( \delta x_t = M \delta x_0 \) within data assimilation window)
Solution of BLUE (Johnson et al. 2005)

• The solution of 4DVAR and KF is the same at the end of DA (Linearity, Gaussianity, same B and R assumed)

$$x_{a,t} = x_{b,t} + MB^{1/2} \sum_{l=1}^{m} \frac{\lambda_l^2}{1 + \lambda_l^2} u_l^T R^{-1/2} (y - Hx_{b,t}) \frac{\lambda_l}{\lambda_l} v_l$$

• Consider the singular value decomposition below

$$L = R^{-1/2} HMB^{1/2} = UDV^T = \sum_{l=1}^{m} u_l \lambda_l v_l^T$$

observability matrix

• Then, the solution can be rewritten as
SVD decomposition of analysis increment
(Linearity and Gaussianity assumed; Johnson et al. 2005)

$$\mathbf{L} \equiv \mathbf{R}^{-1/2} \mathbf{H} \mathbf{M} \mathbf{B}^{1/2} = \mathbf{U} \mathbf{D} \mathbf{V}^T = \sum_{l=1}^{m} \mathbf{u}_l \lambda_l \mathbf{v}_l^T$$

observability matrix

$$\mathbf{x}_{a,t} = \mathbf{x}_{b,t} + \mathbf{MB}^{1/2} \sum_{l=1}^{m} \left( \frac{\lambda_l^2}{1 + \lambda_l^2} \right) \mathbf{u}_l^T \mathbf{R}^{-1/2} (\mathbf{y} - \mathbf{H} \mathbf{x}_{b,t}) \mathbf{v}_l$$

analysis increment

- $\mathbf{v}$ is the eigenvector of $\mathbf{LL}^T$
  - $\mathbf{v}$ is the eigenvector of $\mathbf{MBM}^T$.
- $\lambda^2 \gg 1 \rightarrow$ singular vector constitutes the solution
- $\lambda^2 \ll 1 \rightarrow$ singular vector does not constitute the solution
- If $\mathbf{R}_{ij}$ is small, $\lambda$ is large.
Hybrid EnKF-4DVAR (Lorenc 2003)

\[ J = \frac{1}{2} (x_0 - x_b)^T B^{-1} (x_0 - x_b) + \frac{1}{2} \sum_{t'} (x_{t'} - y_{t'})^T R^{-1} (x_{t'} - y_{t'}) \]

- Conventionally, NMC-based \( B \) (climatological) is used, but EnKF-based \( B \) is preferable to represent flow-dependency.
Hybrid EnKF-4DVAR system
Anomaly correlation (Buehner et al., 2010)
Why EnKF-based $B$ is needed for severe weather prediction using 4DVAR?

- NMC-based $B$ represents only the climatological correlation and is not suitable for severe weather.
- $B$ has large influence on the structure of analysis increment away from observational data.
  --> EnKF-based $B$ helps to cover sparseness of observations to capture the phenomena of interest.
Implementation (Lorenc 2003, Wang et al. 2007)

For simplification, I don’t consider the mixture of NMC-B here.

- Since $B^{-1}$ is not easy to get, we use transformed variables $v$ defined $\delta x = B^{1/2} v$. It simplifies $J_B = (1/2)v^T v$ and $\partial J_B / \partial v = v$.
- We can obtain $v$ with $v=0|_{iter=1}$, $dJ/dv$.

<table>
<thead>
<tr>
<th></th>
<th>w/o localization</th>
<th>w/ localization (α-vector method)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B$</td>
<td>$XX^T [n \times n]$</td>
<td>$XX^T \circ S [n \times n]$</td>
</tr>
<tr>
<td>$B^{1/2}$</td>
<td>$X [n \times m]$</td>
<td>$(\text{diag}(x_1)S^{1/2}, \ldots, \text{diag}(x_m)S^{1/2}) [n \times mn]$</td>
</tr>
<tr>
<td>$dJ/dv$</td>
<td>$B^{1/2}dJ/d\delta x [m]$</td>
<td>$B^{1/2}dJ/d\delta x [mn]$</td>
</tr>
<tr>
<td>$\delta x$</td>
<td>$B^{1/2}v [n]$</td>
<td>$B^{1/2}v [n]$</td>
</tr>
</tbody>
</table>

[ ]: size of vector or matrix, $\circ$: Schur product
$X$: perturbations relative to ensemble mean for all members
$n$: degree of freedom, $m$: num of members, $S$: Localization operator
Localization (Lorenz96): 1-obs. test

30 mem. No localization

30 mem. w/localization

1000 mem. No localization

Analysis increment

Analysis increment

Analysis increment

B

B

B
Implementation to JNoVA

• JNoVA is “JMA Nonhydrostatic model”-based Variational data assimilation system used for the regional forecast in JMA.
• Adjoint-based 4DVAR
• Flow-dependent $\mathbf{B}$: EnKF-based 51 members
• Inner model: $\Delta x=15\text{km}$, Outer model: $\Delta x=5\text{km}$
• Large scale condensation
• Assimilation window: 3 hours.
• For simplicity, we show the increment of horizontal wind only.
1-obs DA around Typhoon Talas

- Obs: U (t=1h, z=1120m) indicated by ★
- Contour: background zonal wind
- Vector: analysis increment of hor. wind (z=1120m)

NMC-based B (Conventional) EnKF-based B (Hybrid EnKF-4DVAR)

Irrelevant to the vortex structure
Consistent with vortex displacement
JNoVA: 1-obs. DA around typhoon Talas

- Analysis increment (U)
  - EnKF-B w/o localization

- Analysis increment (U)
  - EnKF-B w/ localization (neighboring 3x3)

- Analysis increment (U)
  - NMC-B (JNoVA original)
Analysis increment at initial time of DA window (For real data during 9/100Z-03Z)

- Analysis increment by using NMC-based B is located in the dense observation area.
- Analysis increment by using EnKF-based B is large around TC and frontal clouds.

Spread in EnKF members
Summary

• **Solution of BLUE:**
  Singular vector of $R^{-1/2}HMB^{1/2}$ with large singular value determines the structure of analysis increment.

• **Hybrid EnKF-4DVAR**
  EnKF-based flow-dependent $\mathbf{B}$ is used instead of NMC-based climatological $\mathbf{B}$.

• **Implementation to JNoVA**
  Currently ongoing. Preliminary experiments seem to work well. Further tests for localization and other tasks are needed.