Research on Atmospheric Data Assimilation Techniques for Antarctic Applications: Schemes and Preliminary Results

Qingnong Xiao[1], Chengsi Liu[2], and Kekuan Chu[1]

1. College of Marine Science, University of South Florida
2. Institute of Atmospheric Physics/LASG, Beijing, China

Outline

- Motivations
- Technical schemes
- Preliminary results
- Summary

Advanced data assimilation

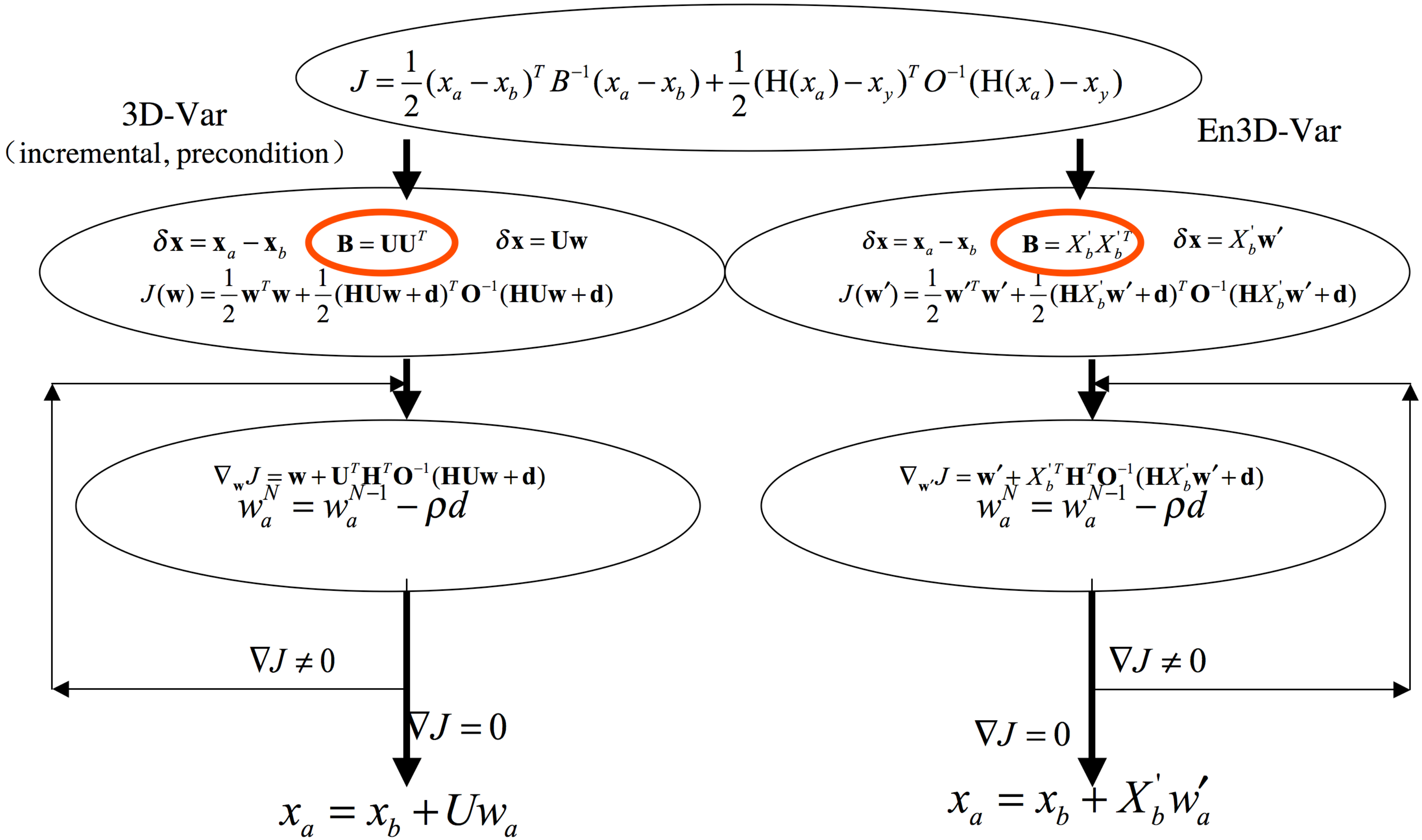
- 3D-Var
  - ✓ All data can be assimilated simultaneously, with a pre-defined background error covariance matrix.
  - ✓ Equivalent to OI, but it can ingest non-conventional data.
  - ✓ It is possible to add constraints to the cost function to control spurious noise.
- 4D-Var
  - ✓ It is a non-sequential data assimilation technique, fitting observations in the whole assimilation window (optimal trajectory).
  - ✓ It is applied in many operational centers.
  - ✓ However, there are disadvantages compared with EnKF technique (TL and AD are difficult to code; background error covariance is evolved only within assimilation window and it is usually static at analysis time).
- Ensemble Kalman filter
  - ✓ It is a hot topic in recent years, and research shows promising results.
  - ✓ It is easy to design and code, and can include any physical process as needed.
  - ✓ One of the prominent advantages is its flow-dependent background error covariance.

Advanced data assimilation

- Ensemble-based variational data assimilation, En3/4D-Var
  - ✓ It is proposed by Liu et al. (2008; 2009):

    Liu, C., Q. Xiao, and B. Wang, 2008: An ensemble-based four-dimensional variational data assimilation scheme: Part I: Technical formulation and preliminary test. *Mon. Wea. Rev.*, **136**, 3363-3373.
    Liu, C., Q. Xiao, and B. Wang, 2009: An ensemble-based four-dimensional variational data assimilation scheme: Part II: Observing system simulation experiments with Advanced Research WRF (ARW). *Mon. Wea. Rev.*, **137**, 1687-1704.
  - ✓ It is variational approach, minimizing a cost function to find the optimal analysis state.
  - ✓ It adopts the technique of EnKF to include the flow-dependent background error covariance from ensemble forecast.
  - ✓ It can be implemented in the existing variational data assimilation system without significant changes of the system setup.
  - ✓ Preliminary results from WRF En3/4D-Var are satisfactory.

En3D-Var（Lorenc 2003）



En4D-Var

$$J = \frac{1}{2}(x_a - x_b)^T B^{-1}(x_a - x_b) + \sum_{i=0}^{N} \frac{1}{2}(HM_{0\sim i}(x_a) - y_i)^T O^{-1}(HM_{0\sim i}(x_a) - y_i)$$

En4D-Var
(generalized En3D-Var)

$$\mathbf{M}X_b^{'} \approx \frac{1}{\sqrt{N-1}}(MX_{bi} - \overline{MX_b})$$

$$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}\sum_{i=0}^{I}(\mathbf{HM}_{0\sim i}X_b^{'}\mathbf{w} + \mathbf{d}_i)^T\mathbf{O}^{-1}(\mathbf{HM}_{0\sim i}X_b^{'}\mathbf{w} + \mathbf{d}_i)$$

$$\nabla_{\mathbf{w}} J = \mathbf{w} + \sum_{i=0}^{I} X_b^{'T}\mathbf{M}^T{}_{i\sim 0}\mathbf{H}^T\mathbf{O}^{-1}(\mathbf{HM}_{0\sim i}X_b^{'}\mathbf{w} + \mathbf{d}_i)$$

$$w_a^N = w_a^{N-1} - \rho d$$

$\nabla J \neq 0$

$\nabla J = 0$

$$x_a = x_b + X_b^{'} w_a^{\prime}$$

En4D-Var

$$J = \frac{1}{2}(x_a - x_b)^T B^{-1}(x_a - x_b) + \sum_{i=0}^{N} \frac{1}{2}(HM_{0\sim i}(x_a) - y_i)^T O^{-1}(HM_{0\sim i}(x_a) - y_i)$$

En4D-Var
(generalized En3D-Var)

$$\mathbf{M}X_b' \approx \frac{1}{\sqrt{N-1}}(MX_{bi} - \overline{MX_b})$$

$$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}\sum_{i=0}^{I}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w} + \mathbf{d}_i)^T \mathbf{O}^{-1}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w} + \mathbf{d}_i)$$

$$\nabla_{\mathbf{w}} J = \mathbf{w} + \sum_{i=0}^{I} X_b'^T \mathbf{M}^T_{i\sim 0}\mathbf{H}^T \mathbf{O}^{-1}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w} + \mathbf{d}_i)$$

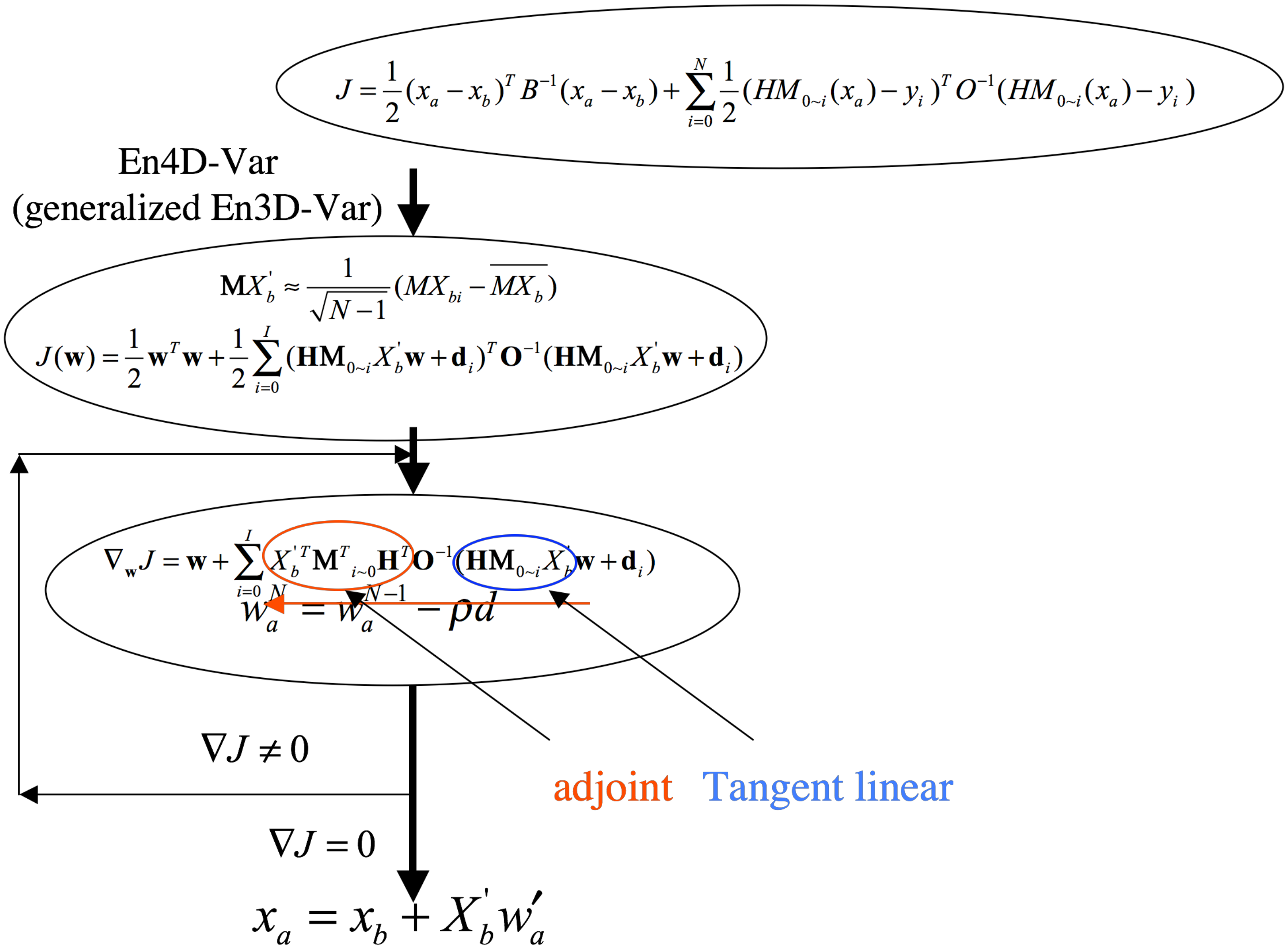
$$w_a^N = w_a^{N-1} - \rho d$$

$$\nabla J \neq 0$$

$$\nabla J = 0$$

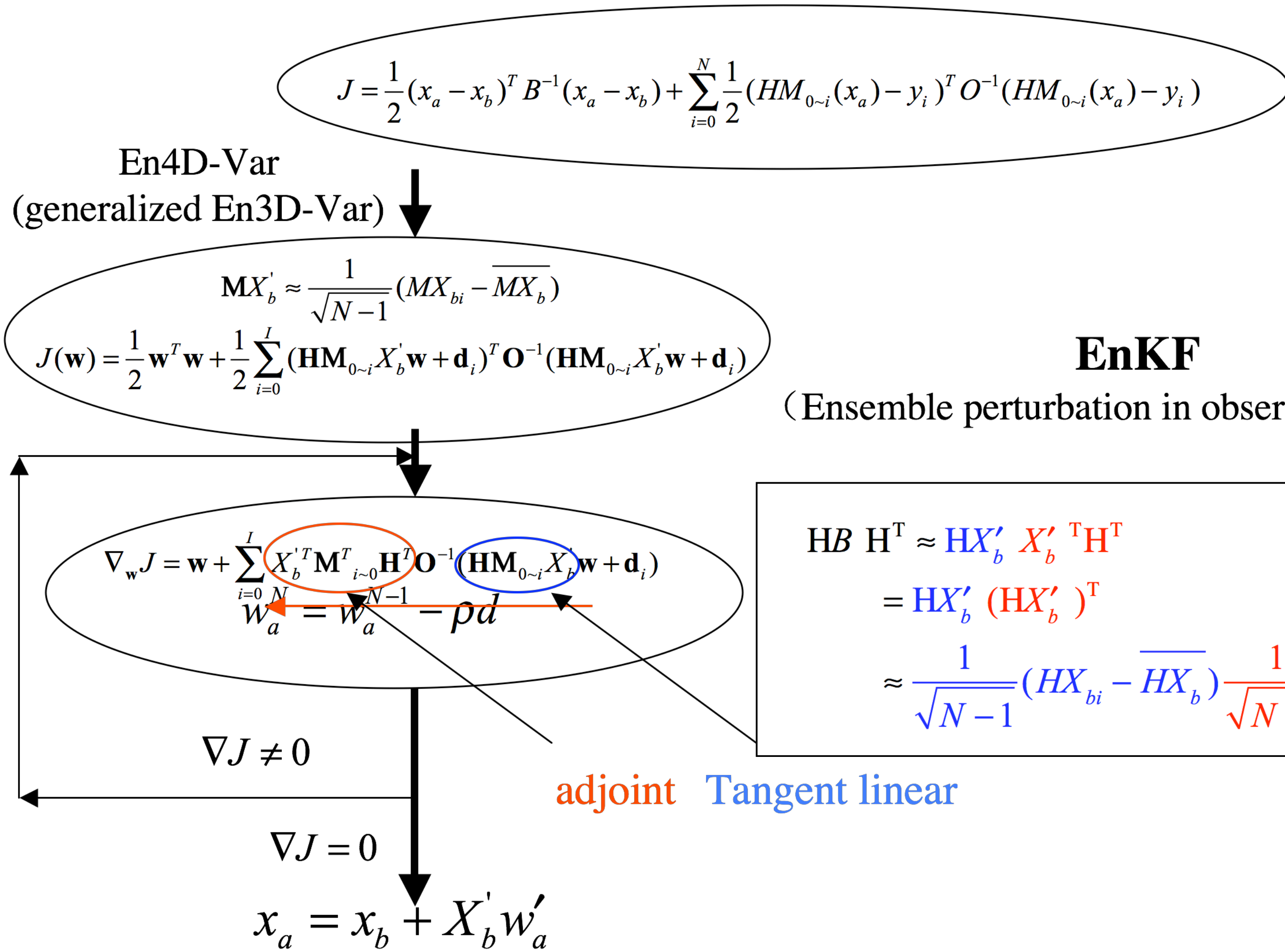
$$x_a = x_b + X_b' w_a'$$

En4D-Var



$J = \frac{1}{2}(x_a - x_b)^T B^{-1}(x_a - x_b) + \sum_{i=0}^{N}\frac{1}{2}(HM_{0\sim i}(x_a) - y_i)^T O^{-1}(HM_{0\sim i}(x_a) - y_i)$
En4D-Var
(generalized En3D-Var)
$\mathbf{M}X_b' \approx \frac{1}{\sqrt{N-1}}(MX_{bi} - \overline{MX_b})$
$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}\sum_{i=0}^{I}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w} + \mathbf{d}_i)^T\mathbf{O}^{-1}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w} + \mathbf{d}_i)$
$\nabla_{\mathbf{w}}J = \mathbf{w} + \sum_{i=0}^{I} X_b'^T \mathbf{M}^T_{i\sim 0}\mathbf{H}^T\mathbf{O}^{-1}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w} + \mathbf{d}_i)$
$w_a^N = w_a^{N-1} - \rho d$
$\nabla J \neq 0$
$\nabla J = 0$
$x_a = x_b + X_b' w_a'$
adjoint
Tangent linear

En4D-Var

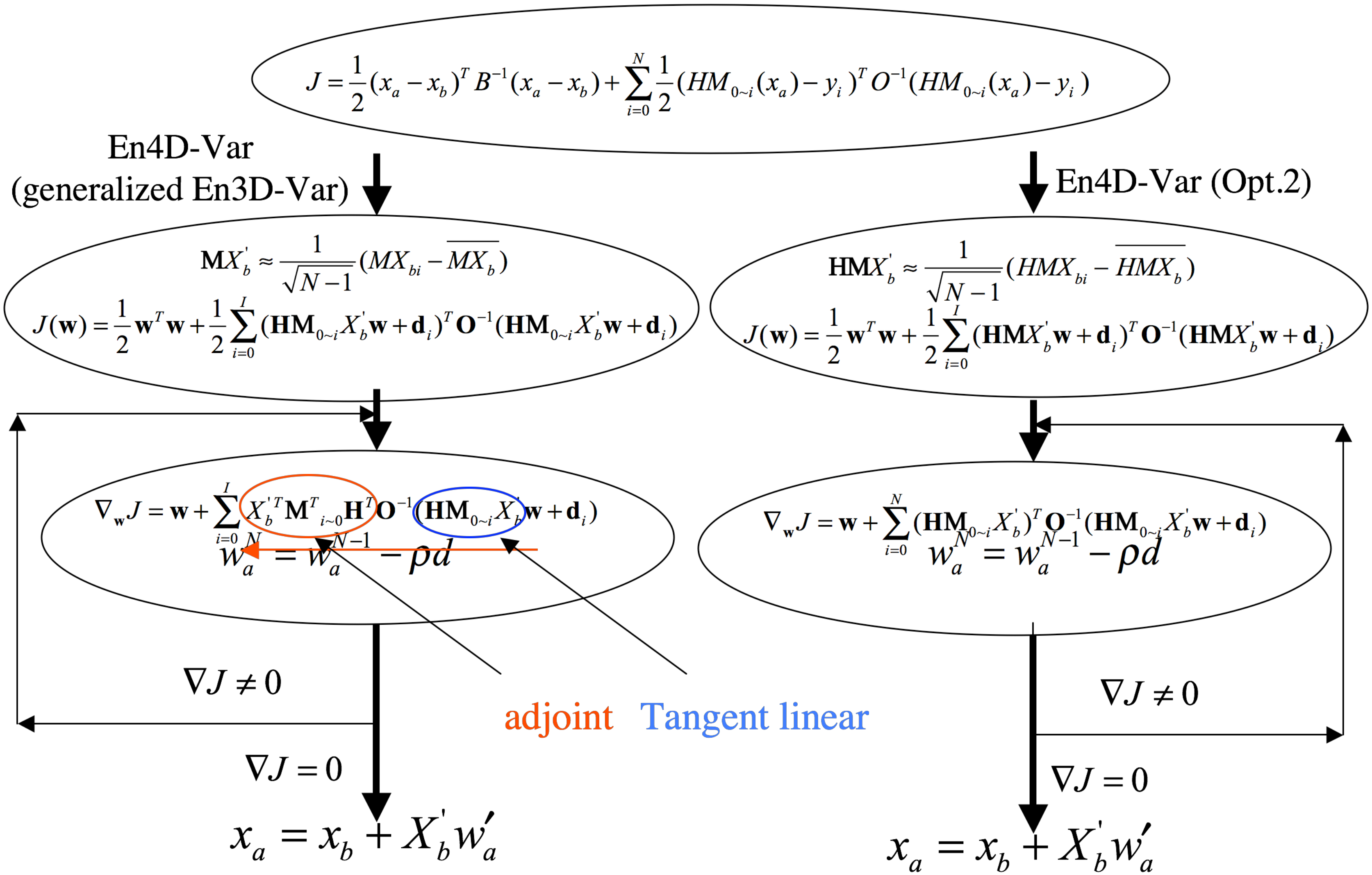


EnKF

（Ensemble perturbation in observation space）

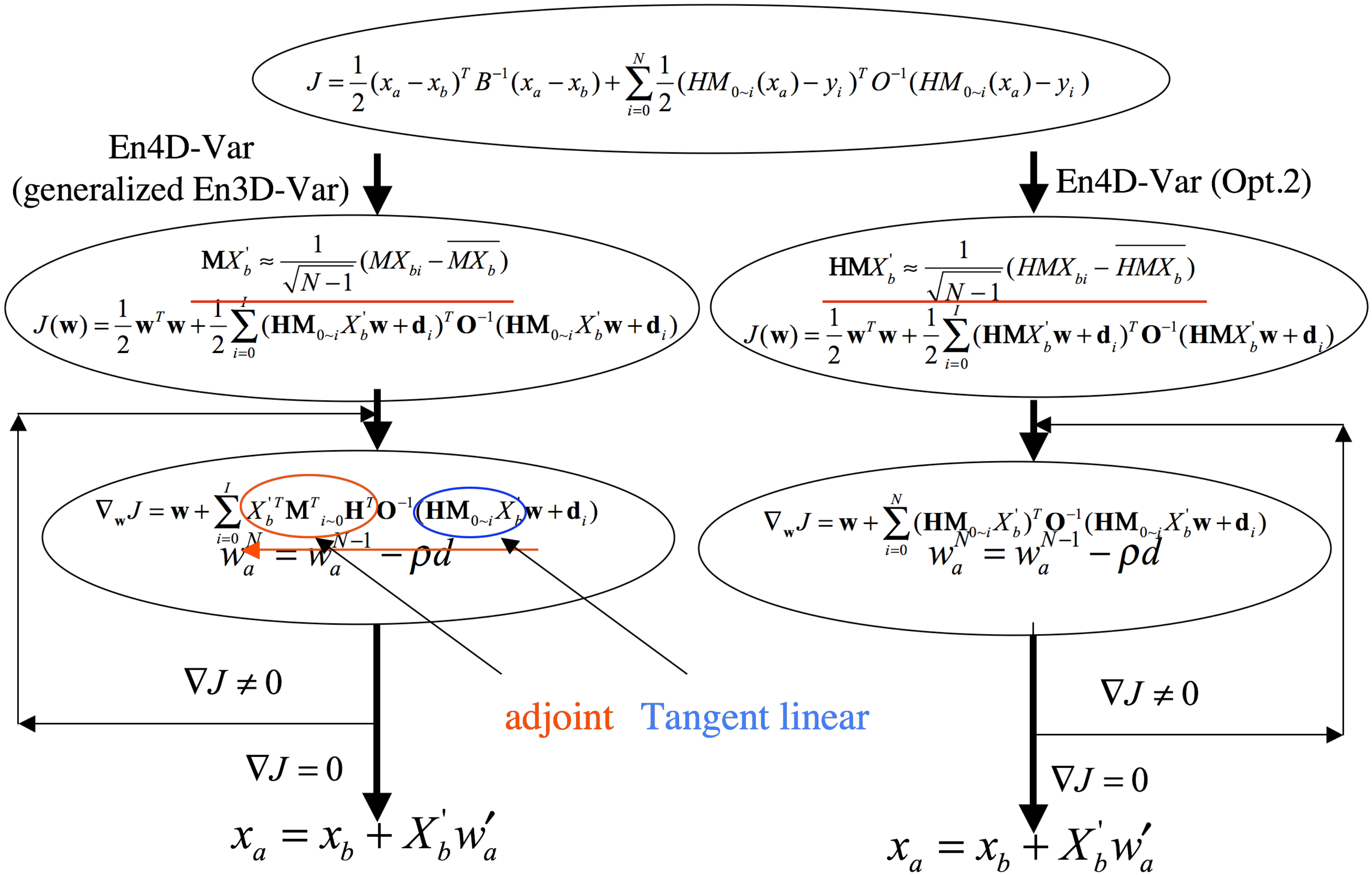
$$\mathrm{H}B\,\mathrm{H}^{\mathrm{T}}\approx \mathrm{H}X_b^{\prime}\,X_b^{\prime\,\mathrm{T}}\mathrm{H}^{\mathrm{T}}$$
$$=\mathrm{H}X_b^{\prime}\,(\mathrm{H}X_b^{\prime})^{\mathrm{T}}$$
$$\approx\frac{1}{\sqrt{N-1}}(HX_{bi}-\overline{HX_b})\frac{1}{\sqrt{N-1}}(HX_{bi}-\overline{HX_b})^{\mathrm{T}}$$

En4D-Var



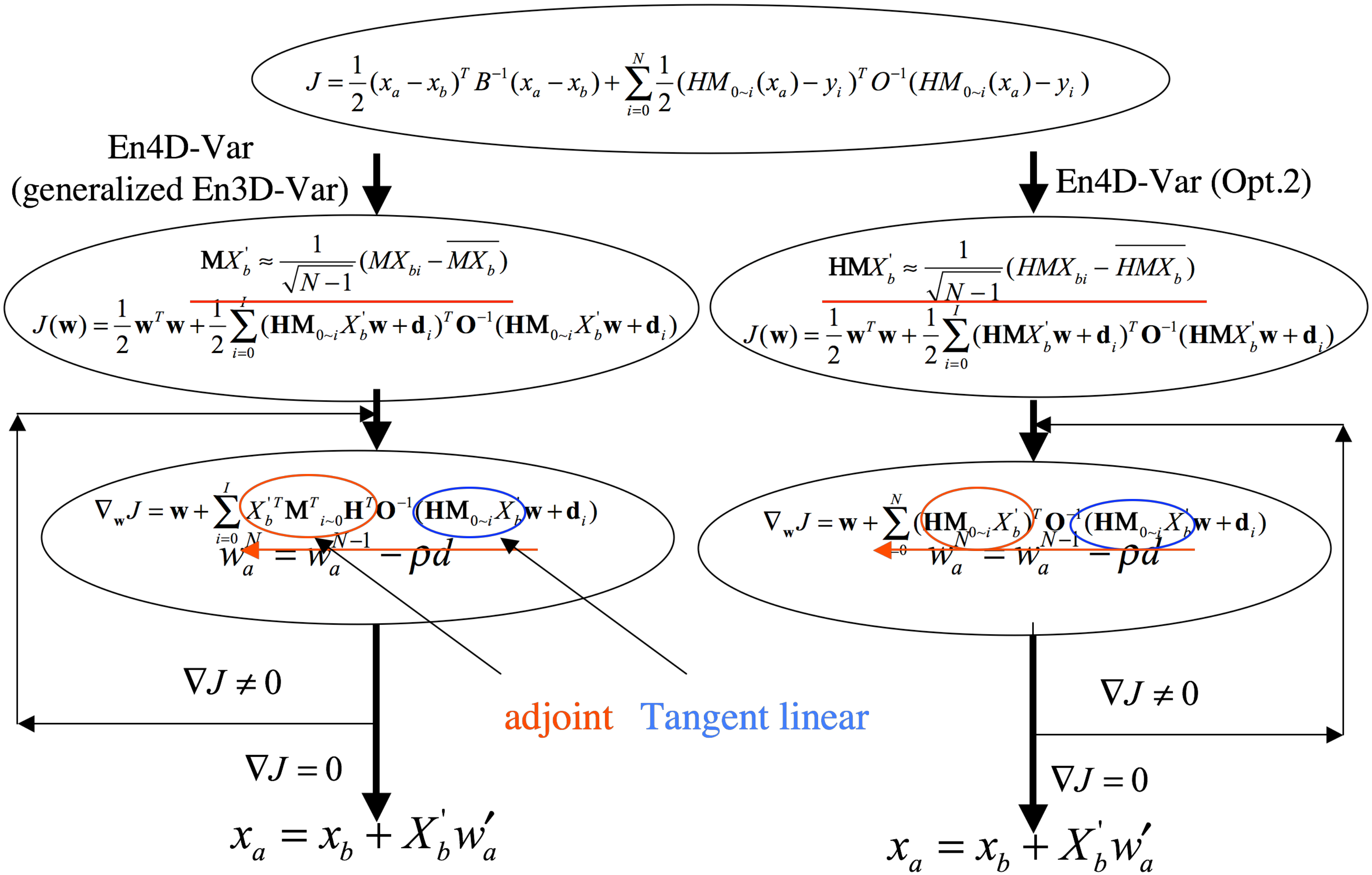
$J = \frac{1}{2}(x_a - x_b)^T B^{-1}(x_a - x_b) + \sum_{i=0}^{N} \frac{1}{2}(HM_{0\sim i}(x_a) - y_i)^T O^{-1}(HM_{0\sim i}(x_a) - y_i)$
En4D-Var
(generalized En3D-Var)
En4D-Var (Opt.2)
$\mathbf{M}X_b' \approx \frac{1}{\sqrt{N-1}}(MX_{bi} - \overline{MX_b})$
$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}\sum_{i=0}^{I}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w} + \mathbf{d}_i)^T\mathbf{O}^{-1}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w} + \mathbf{d}_i)$
$\mathbf{HM}X_b' \approx \frac{1}{\sqrt{N-1}}(HMX_{bi} - \overline{HMX_b})$
$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}\sum_{i=0}^{I}(\mathbf{HM}X_b'\mathbf{w} + \mathbf{d}_i)^T\mathbf{O}^{-1}(\mathbf{HM}X_b'\mathbf{w} + \mathbf{d}_i)$
$\nabla_{\mathbf{w}}J = \mathbf{w} + \sum_{i=0}^{I} X_b'^T\mathbf{M}^T_{i\sim 0}\mathbf{H}^T\mathbf{O}^{-1}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w} + \mathbf{d}_i)$
$w_a^N = w_a^{N-1} - \rho d$
$\nabla_{\mathbf{w}}J = \mathbf{w} + \sum_{i=0}^{N}(\mathbf{HM}_{0\sim i}X_b')^T\mathbf{O}^{-1}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w} + \mathbf{d}_i)$
$w_a^N = w_a^{N-1} - \rho d$
$\nabla J \neq 0$
$\nabla J = 0$
adjoint
Tangent linear
$x_a = x_b + X_b' w_a'$
$\nabla J \neq 0$
$\nabla J = 0$
$x_a = x_b + X_b' w_a'$

En4D-Var



$J=\frac{1}{2}(x_a-x_b)^T B^{-1}(x_a-x_b)+\sum_{i=0}^{N}\frac{1}{2}(HM_{0\sim i}(x_a)-y_i)^T O^{-1}(HM_{0\sim i}(x_a)-y_i)$
En4D-Var
(generalized En3D-Var)
En4D-Var (Opt.2)
$\mathbf{M}X_b' \approx \frac{1}{\sqrt{N-1}}(MX_{bi}-\overline{MX_b})$
$J(\mathbf{w})=\frac{1}{2}\mathbf{w}^T\mathbf{w}+\frac{1}{2}\sum_{i=0}^{I}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w}+\mathbf{d}_i)^T\mathbf{O}^{-1}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w}+\mathbf{d}_i)$
$\mathbf{HM}X_b' \approx \frac{1}{\sqrt{N-1}}(HMX_{bi}-\overline{HMX_b})$
$J(\mathbf{w})=\frac{1}{2}\mathbf{w}^T\mathbf{w}+\frac{1}{2}\sum_{i=0}^{I}(\mathbf{HM}X_b'\mathbf{w}+\mathbf{d}_i)^T\mathbf{O}^{-1}(\mathbf{HM}X_b'\mathbf{w}+\mathbf{d}_i)$
$\nabla_{\mathbf{w}}J=\mathbf{w}+\sum_{i=0}^{I}X_b'^T\mathbf{M}^T_{i\sim 0}\mathbf{H}^T\mathbf{O}^{-1}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w}+\mathbf{d}_i)$
$w_a^N = w_a^{N-1}-\rho d$
$\nabla_{\mathbf{w}}J=\mathbf{w}+\sum_{i=0}^{N}(\mathbf{HM}_{0\sim i}X_b')^T\mathbf{O}^{-1}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w}+\mathbf{d}_i)$
$w_a^N = w_a^{N-1}-\rho d$
$\nabla J \neq 0$
$\nabla J = 0$
adjoint
Tangent linear
$\nabla J \neq 0$
$\nabla J = 0$
$x_a = x_b + X_b' w_a'$
$x_a = x_b + X_b' w_a'$

En4D-Var

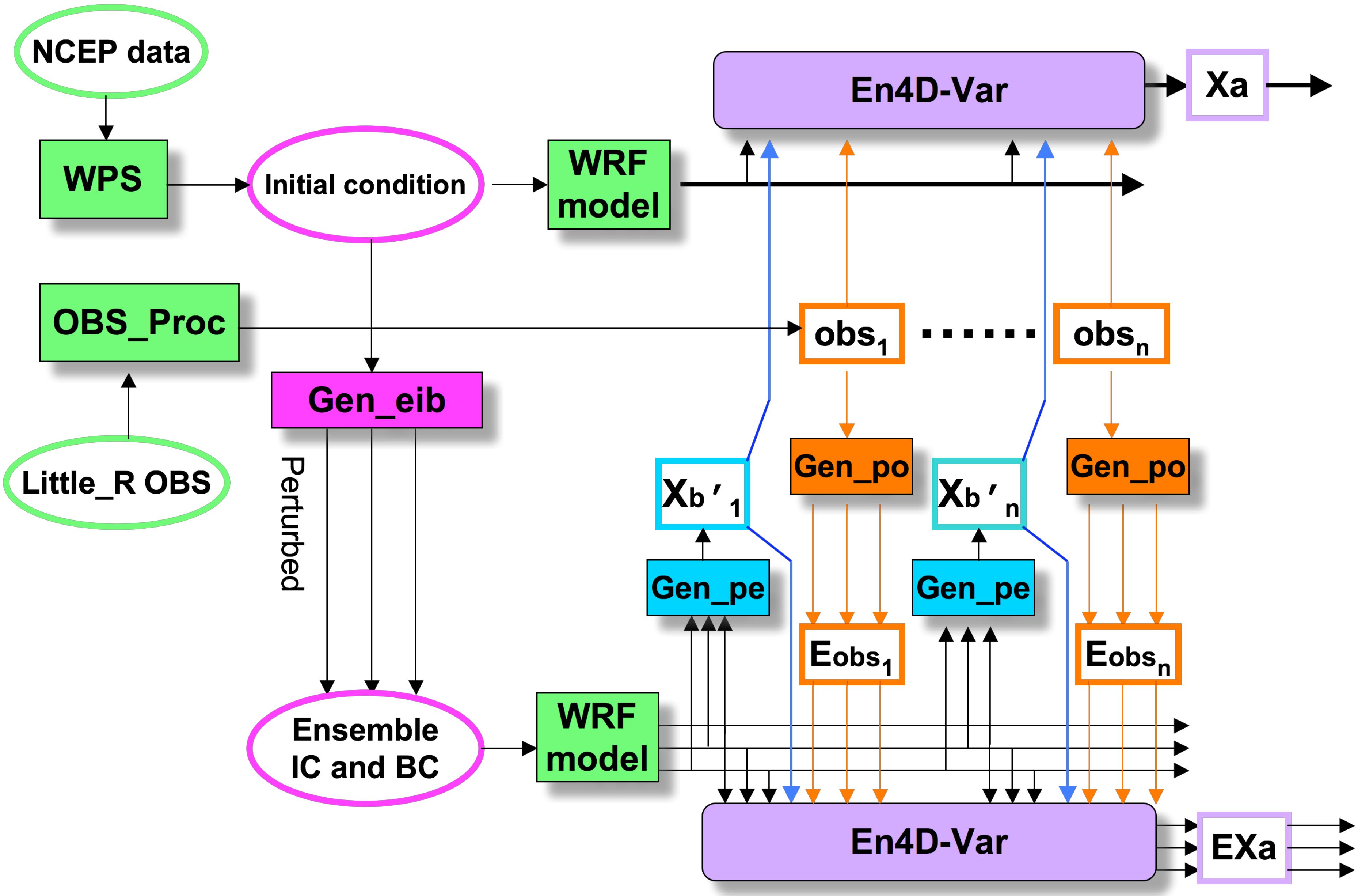


$J = \frac{1}{2}(x_a - x_b)^T B^{-1}(x_a - x_b) + \sum_{i=0}^{N} \frac{1}{2}(HM_{0\sim i}(x_a) - y_i)^T O^{-1}(HM_{0\sim i}(x_a) - y_i)$
En4D-Var
(generalized En3D-Var)
En4D-Var (Opt.2)
$\mathbf{M}X_b' \approx \frac{1}{\sqrt{N-1}}(MX_{bi} - \overline{MX_b})$
$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}\sum_{i=0}^{I}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w} + \mathbf{d}_i)^T\mathbf{O}^{-1}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w} + \mathbf{d}_i)$
$\mathbf{HM}X_b' \approx \frac{1}{\sqrt{N-1}}(HMX_{bi} - \overline{HMX_b})$
$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}\sum_{i=0}^{I}(\mathbf{HM}X_b'\mathbf{w} + \mathbf{d}_i)^T\mathbf{O}^{-1}(\mathbf{HM}X_b'\mathbf{w} + \mathbf{d}_i)$
$\nabla_{\mathbf{w}}J = \mathbf{w} + \sum_{i=0}^{I} X_b'^T\mathbf{M}^T_{i\sim 0}\mathbf{H}^T\mathbf{O}^{-1}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w} + \mathbf{d}_i)$
$w_a^N = w_a^{N-1} - \rho d$
$\nabla_{\mathbf{w}}J = \mathbf{w} + \sum_{i=0}^{N}(\mathbf{HM}_{0\sim i}X_b')^T\mathbf{O}^{-1}(\mathbf{HM}_{0\sim i}X_b'\mathbf{w} + \mathbf{d}_i)$
$w_a^N = w_a^{N-1} - \rho d$
adjoint
Tangent linear
$\nabla J \neq 0$
$\nabla J = 0$
$x_a = x_b + X_b' w_a'$
$\nabla J \neq 0$
$\nabla J = 0$
$x_a = x_b + X_b' w_a'$

Some characteristics of En4D-Var

- En4D-Var uses the flow-dependent B matrix from ensemble forecast.
- It avoids tangent linear and adjoint models in its formulation (in Opt.2).
- It couples incremental approach with preconditioning using ensemble perturbation matrix.
- But sampling errors are introduced to En4D-Var (in Opt.2).

Flow Chart for WRF-En4DVar



Experimental designs for OPP project

- Experiments with various data assimilation techniques for the Antarctic weather predictions through case studies and a month long verification.

- The data assimilation techniques to be tested are WRF 3D-Var, 4D-Var, En3D-Var and En4D-Var.

- The case selected is a cyclone penetrating the Western Antarctic Ice Sheet (WAIS) from 1200 UTC 3 through 1200 UTC 6 October 2007.

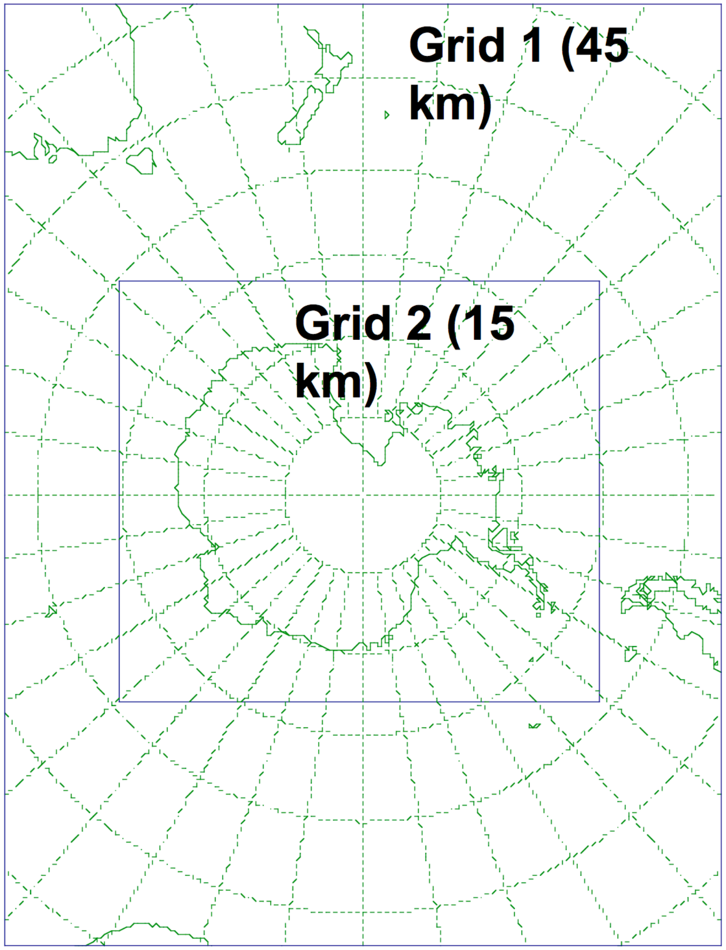
- Verification will be performed for the whole month of October 2007.

- Some preliminary results from case study using WRF 3D-Var has been finished.

Model Domains and Physics

Model Grids:
Two-way nesting
Outer grid: 220*290 (45 km)
Inner grid: 442*418 (15 km)
43 vertical levels

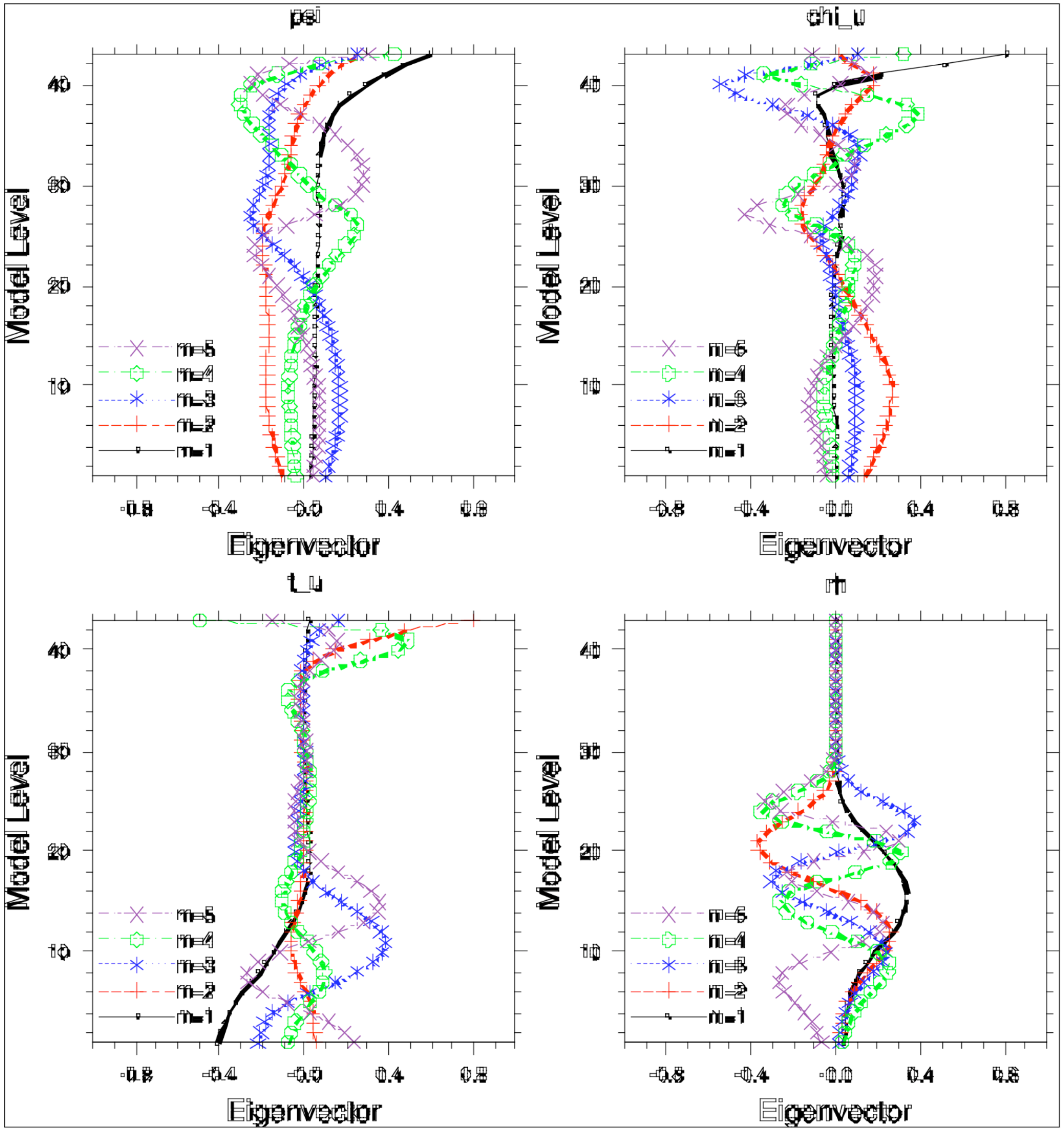
Model Physics:
WSM 5-class scheme
RRTM long wave radiation scheme
Goddard short wave radiation scheme
Mellor-Yamada-Janjic TKE PBL scheme
Kain-Fritsch (new Eta) cumulus scheme



Background Error Covariance

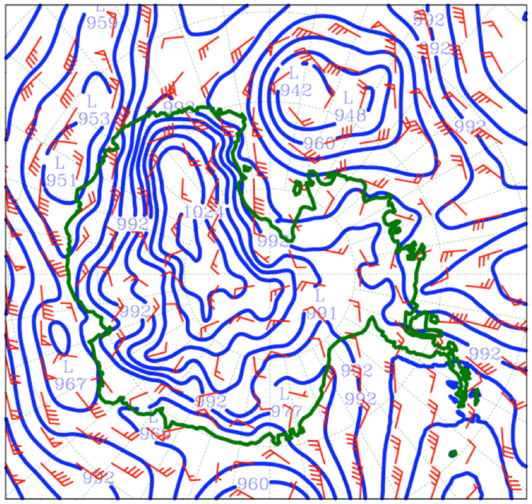
National Meteorological Center method (Parrish and Derber 1992)

The differences between 24- and 12-h forecasts in October 2007 were taken as background errors to calculate the background error covariance.

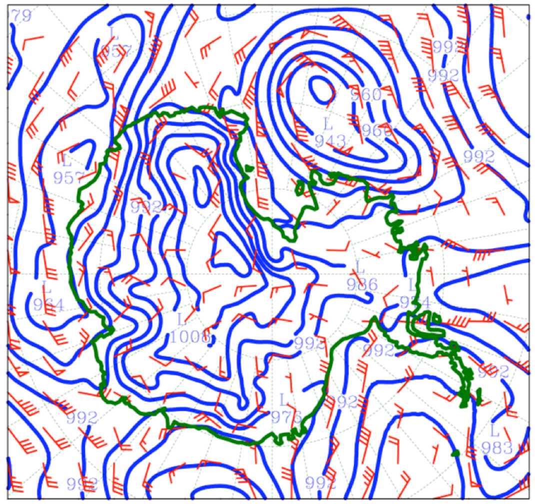


The Antarctic cyclone analysis (FNL)

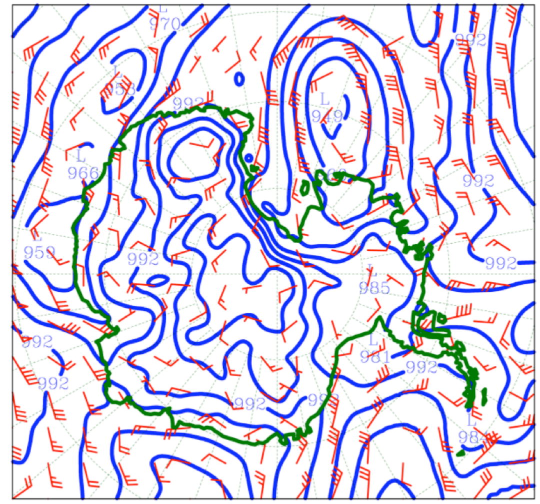
0312



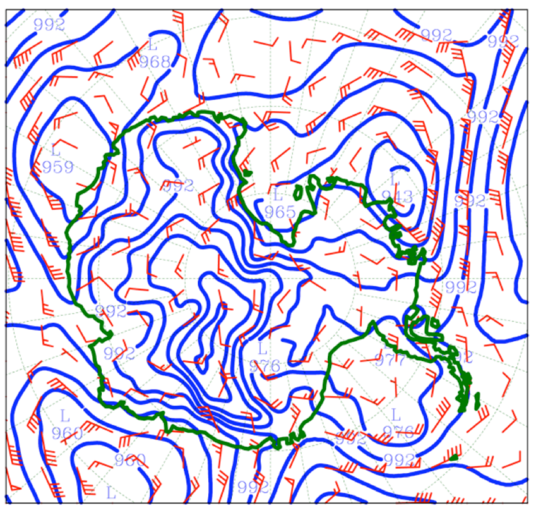
0400



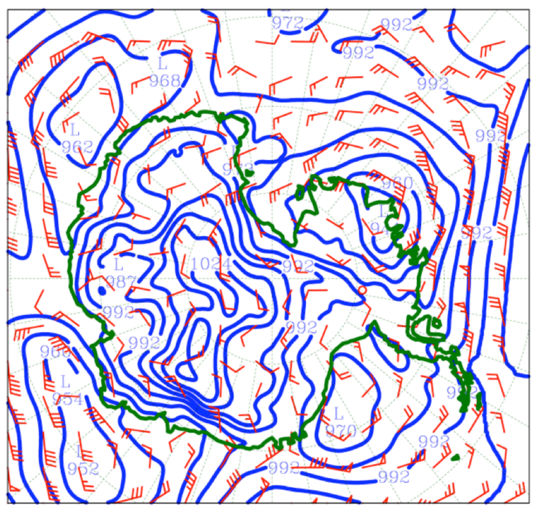
0412



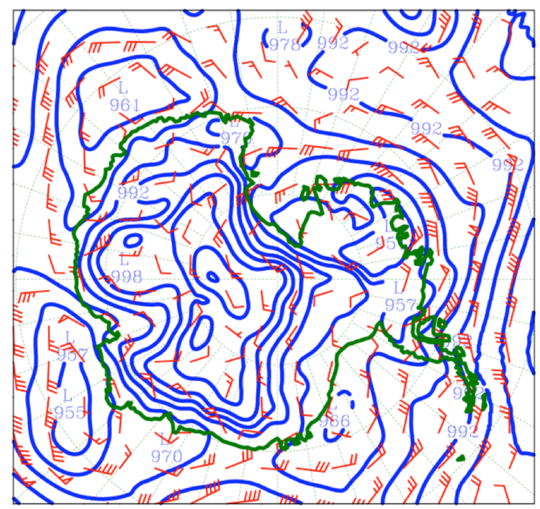
0512



0600



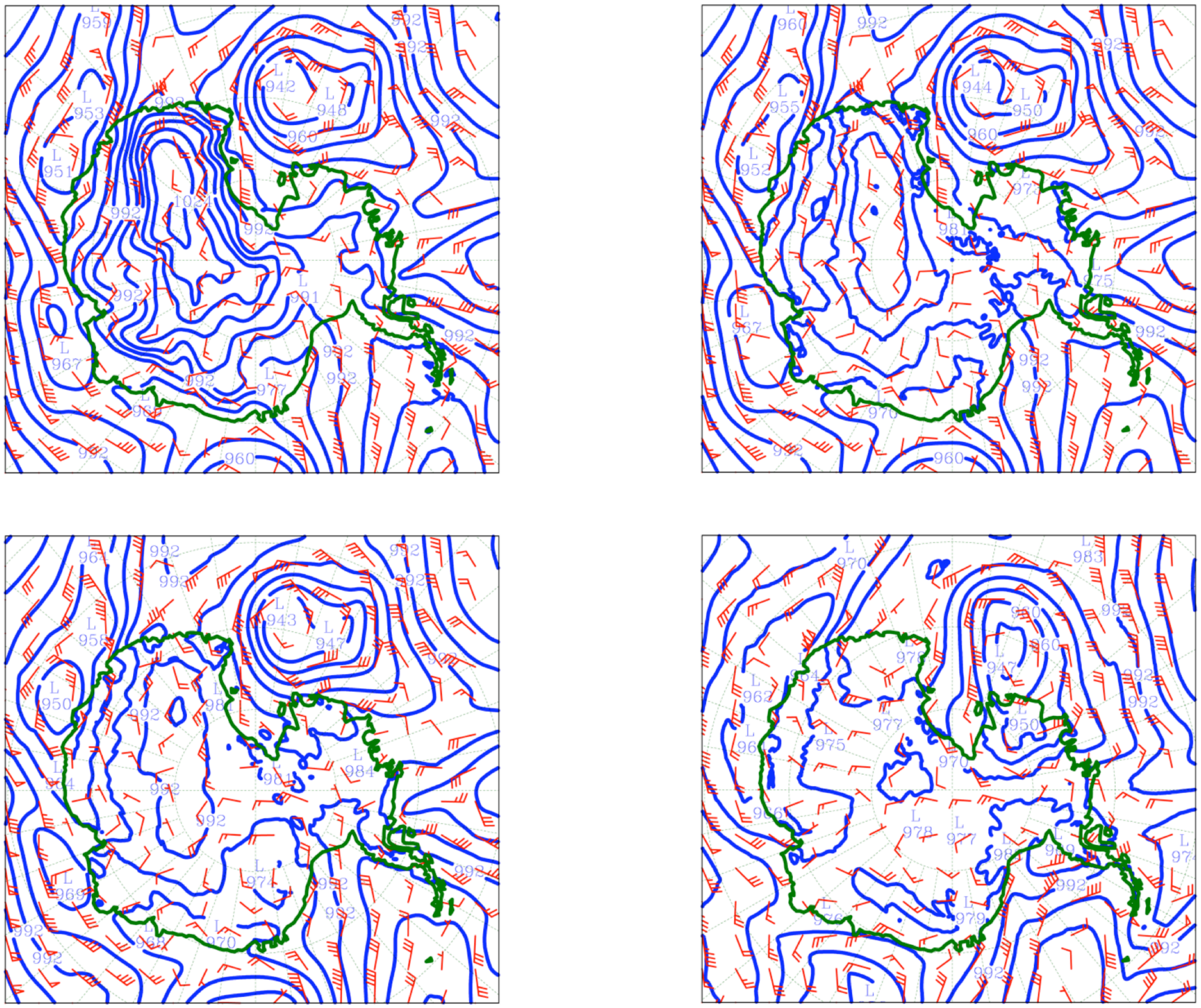
0612



Initial state at 1200 UTC 3 Oct.

FNL

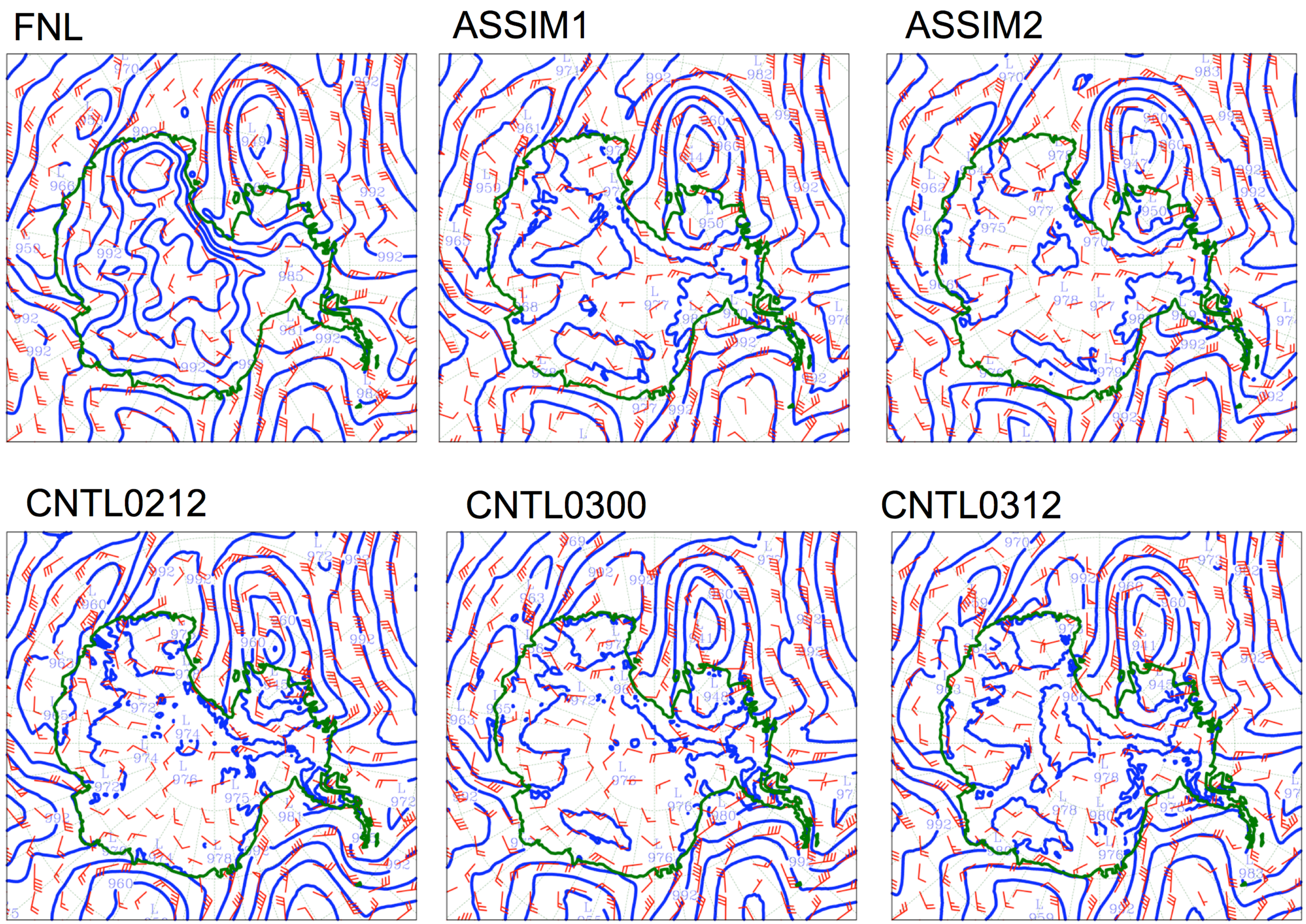
CNTL



ASSIM1 (12-h 3DVAR cycling)

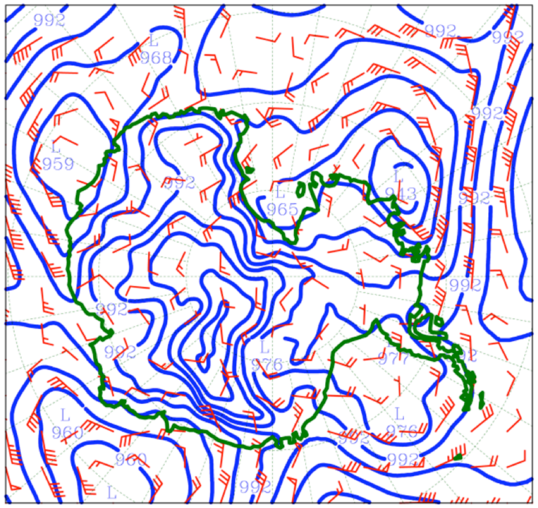
ASSIM2 (24-h 3DVAR cycling)

24-hr forecast at 1200 UTC 4 Oct.

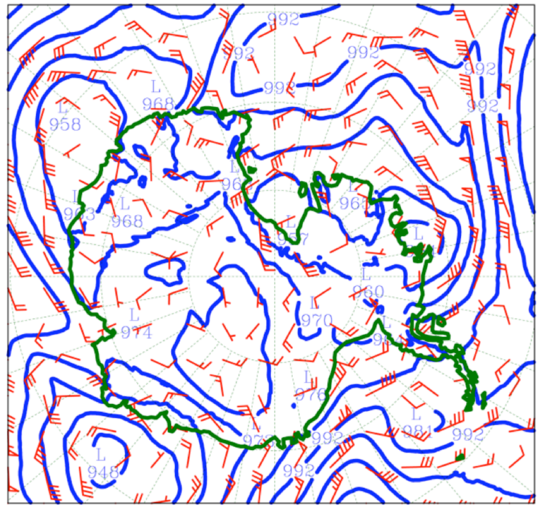


48-hr forecast at 1200 UTC 5 Oct.

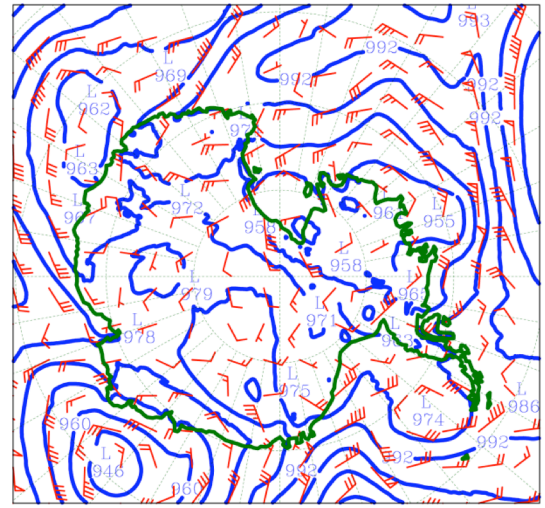
FNL



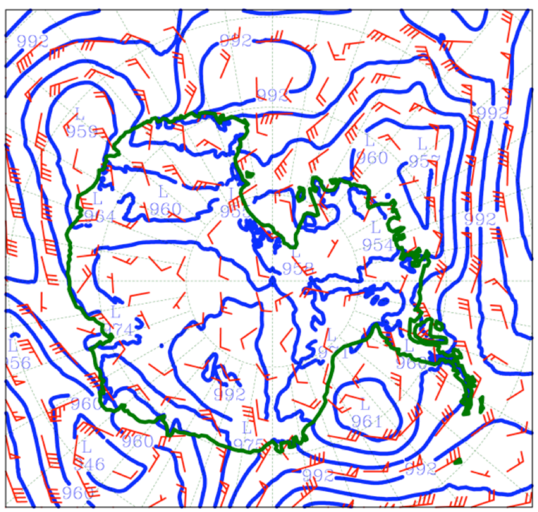
ASSIM1



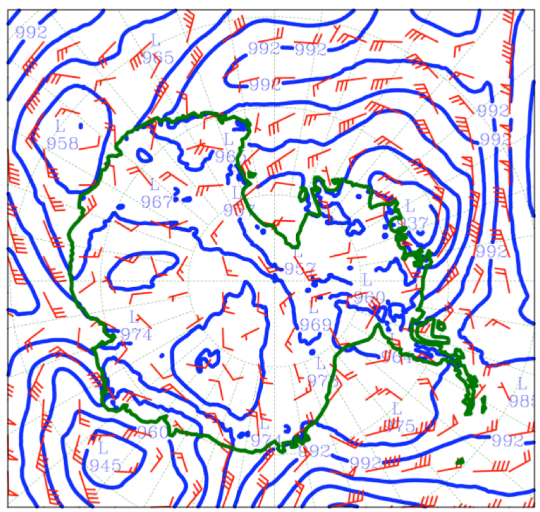
ASSIM2



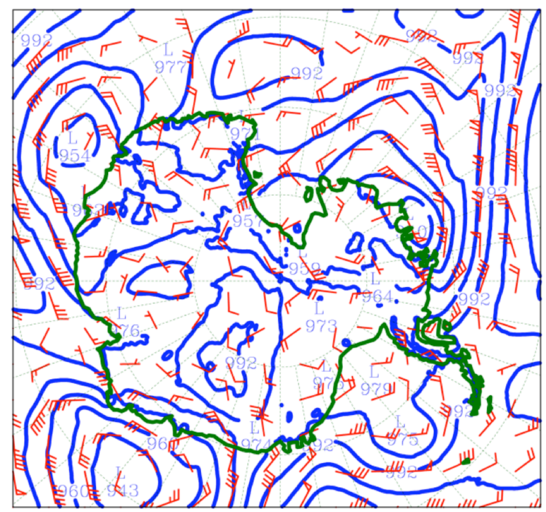
CNTL0212



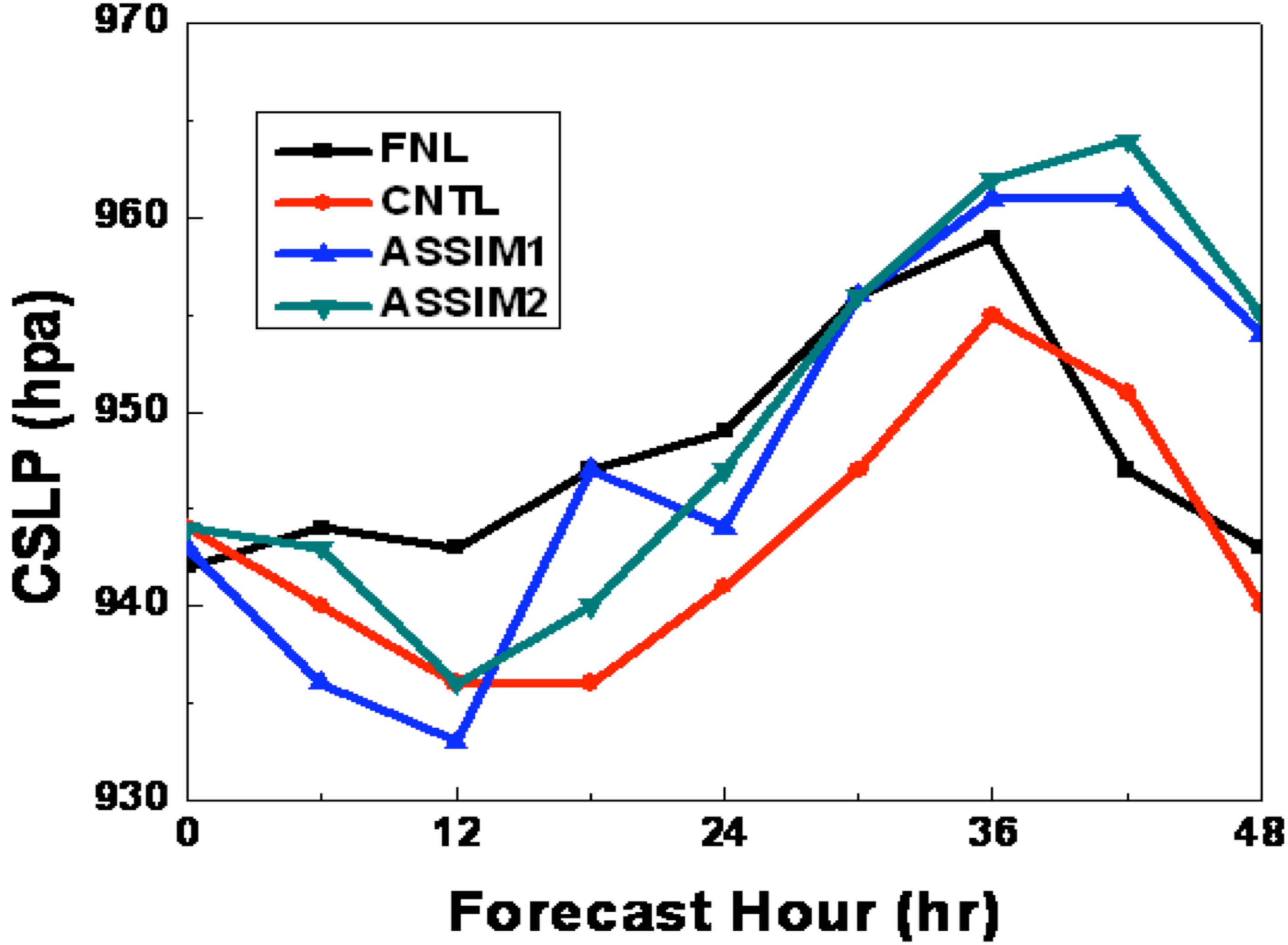
CNTL0230



CNTL0312



Intensity change



Summary and On-going Work

- Data assimilation schemes for the OPP project are designed. We will test 3DVAR, 4DVAR, En3DVAR and En4DVAR for the Antarctic applications.

- One-month runs from 0000 UTC 01 till 0000 UTC 31 October 2007, two times a day at 0000 UTC and 1200 UTC from the NCEP FNL analysis with AMPS domain configuration has been conducted. Background error covariance has been generated.

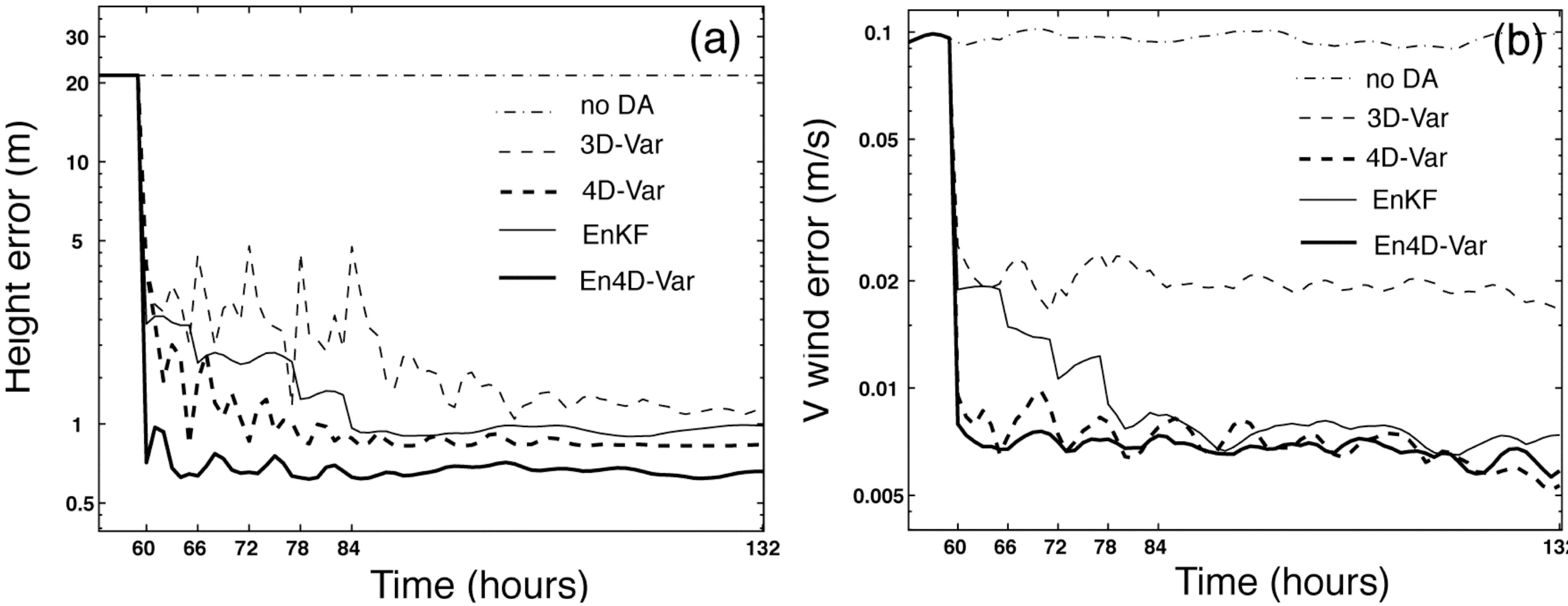
- WRF 3DVAR experiments for the Antarctic cyclone case on 3-6 October 2007 has been started. WRF 4DVAR and En3/4DVAR experiments will follow.

- We will perform one-month verifications for the designed data assimilation techniques in the future.

Thank you !

Questions and comments
are welcome.

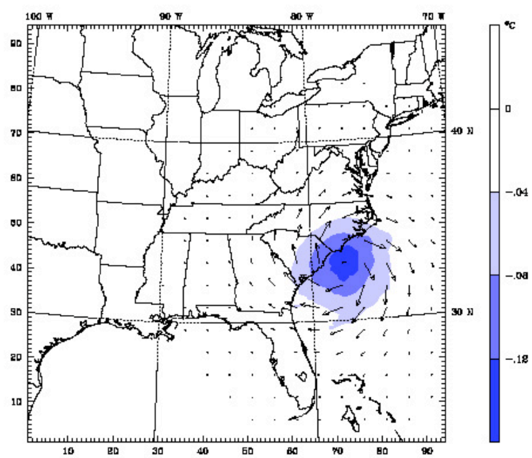
Proof-of-concept test with shallow water model



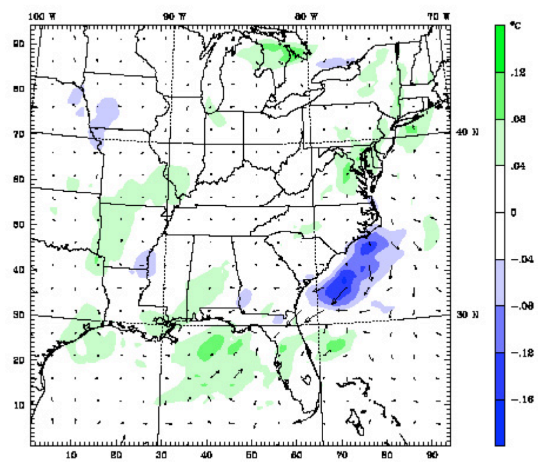
Evolution of domain-average RMSE

Single observation test

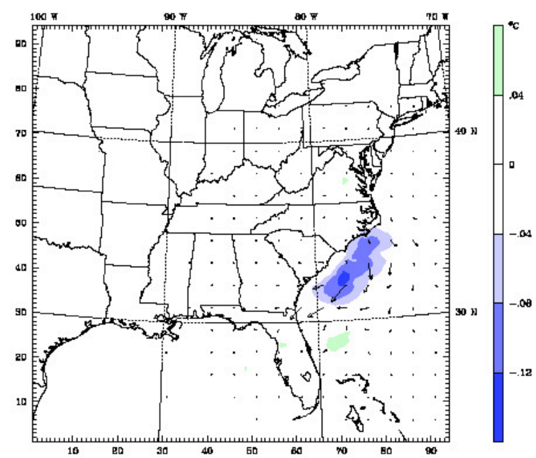
(single T observation at 850hpa at 24-12Z Jan.)



WRF-Var



En4D-Var without localization



En4D-Var with localization

Increments of wind vector and temperature at 1000hpa