Implementation and Testing of the Four-Dimensional Ensemble-Variational (4DEnVAR) Hybrid Algorithms within the ARPS Data Assimilation Framework

Chengsi Liu, Ming Xue, and Rong Kong
Center for Analysis and Prediction of Storms
University of Oklahoma
Outline

• A unified framework for four 4D ensemble-variational algorithms (En4DVar and 4DEnVar)

• The relationship among the EnVar algorithms (Approximations and equivalence )

• ARPS 4DEnVar framework design

• OSSE for a storm case
Hybrid En-4DVar (Lorenc 2003, Clayton et al 2012)

Static $B$ part

$$B = UU^T$$

Ensemble covariance part

Localization matrix

$$C = C'C'^T$$

Alpha control variable

$$\alpha = \begin{bmatrix} C' & 0 \\ \vdots & \ddots \\ 0 & C' \end{bmatrix} \tilde{\alpha}$$

$$\delta x_s = Uv$$

$$\delta x_e = \sum_{i=1}^{N} (x_{bi} \circ C' \tilde{\alpha}_i)$$

Analysis increment and cost function

$$\beta_1^2 + \beta_2^2 = 1$$

$$\delta x_0 = \beta_1 \delta x_s + \beta_2 \delta x_e = \beta_1 Uv + \beta_2 \sum_{i=1}^{N} (x_{bi} \circ C' \tilde{\alpha}_i)$$

$$J(v, \tilde{\alpha}) = \frac{1}{2} v^T v + \frac{1}{2} \tilde{\alpha}^T \tilde{\alpha} + \frac{1}{2} \sum_{t=1}^{T} [H_t L_t \delta x_0 + d_t]^T R^{-1} [H_t L_t \delta x_0 + d_t]$$
4DEnVar-NPC (Non-Propagation of alpha Control variable)

Two approximations

(1) Neglecting temporal propagation of alpha control variable by TLM

\[ L_t \delta x_e = L_t \left( \sum_{i=1}^{N} x'_{bi} \circ C' \tilde{\alpha}_i \right) = \sum_{i=1}^{N} L_t(x'_{bi} \circ C' \tilde{\alpha}_i) \approx \sum_{i=1}^{N} (L_t x'_{bi}) \circ C' \tilde{\alpha}_i \]

\[ \nabla \tilde{\alpha}_i J(\tilde{\alpha}_i) = \tilde{\alpha}_i + \sum_{i=1}^{m} C'^T (L_t x'_{bi}) \circ H_t^T R^{-1} \left[ H_t \sum_{i=1}^{N} \{(L_t x'_{bi}) \circ C' \tilde{\alpha}_i \} + d_t \right] \]

AJM avoided!!

(2) Use nonlinear model ensemble forecasts to replace the temporal propagation of perturbations by the TLM

\[ L_t x'_{bi} \approx M_t(x_{bi}) - M_t(x_b) \]

TLM avoided!!
**Hybrid 4DEnVar**

**Static \( B \) part**

\[ B = UU^T \]

\[ \delta x_s = Uv \]

**Ensemble covariance part**

(Liu et al 2008, 2009)

Localization matrix \( C = C'C'T \)

Perturbation matrix \( S'_i = \begin{pmatrix} x'_{bi} & \ldots & x'_{bi} \end{pmatrix} \)

Localized perturbation matrix \( Z'_b = [S'_{b1} \circ C' \quad S'_{b2} \circ C' \quad \ldots \quad S'_{bN} \circ C'] \)

Project \( Z'_b \) to obs space \( (HLZ'_b) \)

(\( w \) is the same as \( \tilde{\alpha} \) of En4DVar)

\[
J(v, w) = \frac{1}{2} v^T v + \frac{1}{2} w^T w + \sum_{t=1}^{I} \beta_1 H_t L_t \delta x_s + \beta_2 (H_t L_t Z'_b) w + d_t \]

\[
\nabla_w J(w) = w + \sum_{t=1}^{I} \beta_1 (H_t L_t Z'_b)^T R^{-1} [\beta_1 H_t L_t Uv + \beta_2 H_t L_t Z'_b w + d_t] \]

**Avoid AJM**
4DEnVar-NPL (No Propagation of Localization)

Two approximations

(1) The localization matrix propagation by TLM is neglected

\[
L_t[S'_{b1} \circ C' \quad S'_{b2} \circ C' \quad \ldots \quad S'_{bN} \circ C'] \\
\approx [(L_t S'_{b1}) \circ C' \quad (L_t S'_{b2}) \circ C' \quad \ldots \quad (L_t S'_{bN}) \circ C']
\]

(2) Ensemble forecasts are used to replace the temporal propagation of perturbations by the TLM (as in En4DVar-NPC)

\[
L_t x'_{bi} \approx M_t(x_{bi}) - \overline{M_t(x_b)}
\]

TLM avoided !!
The relationship of hybrid 4D ensemble variational algorithms

Hybrid En4DVar ⇔ Hybrid 4DEnVar
(expensive cost)

Non-Propagation of alpha Control variable

No Propagation of Localization

Hybrid 4DEnVar-NPC ⇔ Hybrid 4DEnVar-NPL
(cheap cost)
Equivalence of hybrid En4DVar and hybrid 4DEnVar

\[ Z'_b w = (S'_{b1} \circ C', S'_{b2} \circ C', \ldots, S'_{bN} \circ C') \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} = \sum_{i=1}^{N} (S'_{bi} \circ C')w_i = \sum_{i=1}^{N} x'_{bi} \circ (C'w_i) \]

From Hybrid 4DEnVar

Substituting the above equation to the **Hybrid 4DEnVar** cost function

Hybrid En4DVar Cost function \rightleftharpoons \text{Hybrid 4DEnVar Cost function}
Equivalence of 4DEnVar-NPC and 4DEnVar-NPL

\[ T_i'w = \left[ (L_i S'_{b1}) \circ C' \ (L_i S'_{b2}) \circ C' \ \ldots \ (L_i S'_{bN}) \circ C' \right] \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} = \sum_{i=1}^{N} L_i x'_b \circ (C'w_i). \]

From Hybrid 4DEnVar-NPL

Substituting the above equation to the **Hybrid 4DEnVar-NPL** cost function

Hybrid 4DEnVar-NPC Cost function \leftrightarrow Hybrid 4DEnVar-NPL Cost function
Single observation tests with a 1D linear advection model

\[
\frac{\partial u}{\partial t} + U \frac{\partial u}{\partial x} = 0
\]

The numerical result also proven:

En4DVar and 4DEnVar (with no approximation to the flow-following-localization) are the same

En4DVar-NPC and 4DEnVar-NPL (with non-flow-following localization approximation) are the same
Flow-following & non-flow-following localization

Assimilation window

Ensemble DA W/O localization

Ensemble DA W/ localization

Obs (3D-DA)

Obs (4D-DA)

(four algorithms are the same)

solid: flow-following localization (En4DVar/4DEnVar)
dot: non-flow-following localization (En4DVar-NPC/4DEnVar-NPL)
Two factors affecting approximations:
Advection speed & spatial de-correlation scale

sensitivity experiments

Exp1 (3DDA): obs at beginning of DA window

Exp2 (4DDA): obs at end of DA window
(by slow speed and large scale)

Exp3 (4DDA_2S): same as Exp2,
but doubling the advection speed

Exp4 (4DDA_L/2): same as Exp2,
but half of the error spatial de-correlation scale
black no-localization
red flow-following localization (En4DVar/4DEnVar)
blue non-flow-following localization (4DEnVar-NPC/4DEnVar-NPL)
The hybrid 4DEnVAR DA system based on the ARPS variational DA framework

- 4DEnVar-NPC is adopted as ARPS hybrid variational algorithm because of some reasons
  1. Adjoint and tangent linear are not good approximation in convective scale DA.
  2. High resolution observations, like Radar, are main observation source for convective DA. Observation-space-based 4DEnVar algorithm will get expansive computational cost.
The ARPS 4DEnVar characteristics

• 3D-recursive filter is used for ARPS-4DEnVar localization.

• The capabilities for convective-scale radar DA (Vr and reflectivity)

• Physical constraint terms, like the divergence constraint and Gradient Wind Balance Constrain, can be considered in the cost function.
OSSE for a storm case

- Tested with simulated data from a classic supercell storm of 20 May 1977 near Del City, Oklahoma
- Domain: 67 x 67 x 35 grids, vertical stretching. 2km horizontal resolution
- 70-min length of simulation
- the de-correlation scale is 9.6km for horizontal and 3km for vertical.
Single observation test

- Single observation analysis is on 25th minute after forecast
- Two kinds of perturbations,
  1. 25 minutes forecast from background
  2. 4 cycles (5min cycle interval) by EnKF
RMSE of EnKF cycling

Single obs test time
wind vector near observation (Vr)

True, Background, 3DVar, En3DVar
(EnKF is very closed to En3DVar)
horizontal V Wind (k=12)

T-B

A-B (EnKF)

A-B (3DVar)

A-B (En3DVar)
v wind vertical cross section (y=31)

T-B

A-B (3DVar)

A-B (EnKF)

A-B (En3DVar)
wind vector near observation (Vr)
Non-cycle-perturbation .VS. cycle-perturbation

Background, true, En3DVar_NCP , En3DVar_CP
W on vertical cross-section

- T-B
- A-B (3DVar)
- A-B (En3DVar Non-cycle-pert)
- A-B (En3DVar cycle-pert)
PT on vertical cross-section

T-B

A-B (Non-cycle-pert)

A-B (cycle-pert)
Reflectivity DA

True, background and 3DEnVar
(there are little affect to wind by reflectivity)
Qr (k=12)
Qr on vertical cross-section

A-B (3DVar)

A-B (En3DVar)
Summary

• **En4DVar** based on the alpha control variable and **4DEnVar** based on Liu et al. algorithm are mathematically equivalent. They use flow-following ensemble covariance localization implicitly or explicitly.

• **4DEnVar-NPC** and **4DEnVar-NPL** are approximate algorithms and both algorithm are equivalent. They use non-flow-following localization implicitly or explicitly.

• For relatively slow signal propagation and/or relatively larger error correlation sales, **En4DVar-NPC** and **4DEnVar-NPL** are a good approximation to **En4DVar** and **4DEnVar**

• **4DEnVar-NPC** is adopted as ARPS hybrid DA scheme.

• ARPS **En3DVar** and **EnKF OSSE** result is very similar. Both have flow-dependent analysis

• **Cycling updated perturbations** give a better analysis for **En3DVar**
On-going research

• **Optimal horizontal and vertical localization decorrelation scales** will be found by sensitivity experiments.

• **The optimal relative weighing** between the ensemble-based and static background error covariance will be obtained by sensitivity experiments.

• **Physical constraints** (e.g. 2D and 3D mass continuity constraint (Hu 2006)) will also be tested in 3/4D-EnVar to help recover the cross-beam velocity.

• **Time localization** will be tested on 4DEnVar.

• **The Cycle and forecast** will be compared among ARPS-3DVar, EnKF, En3DVar and 4DEnVar.