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

$$\mathbf{B} = \mathbf{U}\mathbf{U}^{\mathrm{T}}$$

$$\delta \mathbf{x}_s = \mathbf{U}\mathbf{v}$$

Ensemble covariance part

Localization matrix $\mathbf{C} = \mathbf{C}'\mathbf{C}'^{\mathrm{T}}$

Alpha control variable $\boldsymbol{\alpha} = \begin{bmatrix} \mathbf{C}' & & 0 \\ & \ddots & \\ 0 & & \mathbf{C}' \end{bmatrix} \tilde{\boldsymbol{\alpha}}$

$$\delta \mathbf{x}_e = \sum_{i=1}^{N} (\mathbf{x}'_{bi} \circ \mathbf{C}'\tilde{\boldsymbol{\alpha}}_i)$$

Analysis increment and cost function

$$\beta_1^{\,2} + \beta_2^{\,2} = 1$$

$$\delta \mathbf{x}_0 = \beta_1 \delta \mathbf{x}_s + \beta_2 \delta \mathbf{x}_e = \beta_1 \mathbf{U}\mathbf{v} + \beta_2 \sum_{i=1}^{N} (\mathbf{x}'_{bi} \circ \mathbf{C}'\tilde{\boldsymbol{\alpha}}_i)$$

$$J(\mathbf{v}, \tilde{\boldsymbol{\alpha}}) = \frac{1}{2}\mathbf{v}^T\mathbf{v} + \frac{1}{2}\tilde{\boldsymbol{\alpha}}^{\mathrm{T}}\tilde{\boldsymbol{\alpha}} + \frac{1}{2}\sum_{t=1}^{I} [\mathbf{H}_t \mathbf{L}_t \delta \mathbf{x}_0 + \mathbf{d}_t]^T \mathbf{R}^{-1} [\mathbf{H}_t \mathbf{L}_t \delta \mathbf{x}_0 + \mathbf{d}_t]$$

innovation

4DEnVar-**NPC** (**N**on-**P**ropagation of alpha **C**ontrol variable)

Two approximations

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

$$\mathbf{L}_t \delta \mathbf{x}_e = \mathbf{L}_t \left(\sum_{i=1}^{N} \mathbf{x}'_{bi} \circ \mathbf{C}' \tilde{\boldsymbol{\alpha}}_i \right) = \sum_{i=1}^{N} \mathbf{L}_t (\mathbf{x}'_{bi} \circ \mathbf{C}' \tilde{\boldsymbol{\alpha}}_i) \approx \sum_{i=1}^{N} (\mathbf{L}_t \mathbf{x}'_{bi}) \circ \mathbf{C}' \tilde{\boldsymbol{\alpha}}_i$$

$$\nabla_{\tilde{\boldsymbol{\alpha}}_i} J(\tilde{\boldsymbol{\alpha}}_i) = \tilde{\boldsymbol{\alpha}}_i + \sum_{t=1}^{m} \mathbf{C}'^{\mathrm{T}} \underline{(\mathbf{L}_{\mathrm{t}} \mathbf{x}'_{bi})} \circ \mathbf{H}_t^{\mathrm{T}} \mathrm{R}^{-1} [\mathbf{H}_t \sum_{i=1}^{N} \{ (\mathbf{L}_{\mathrm{t}} \mathbf{x}'_{bi}) \circ \mathbf{C}' \tilde{\boldsymbol{\alpha}}_i \} + \mathbf{d}_t]$$

AJM avoided !!

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

$$\mathbf{L}_t \mathbf{x}'_{bi} \approx M_t(\mathbf{x}_{bi}) - \overline{M_t(\mathbf{x}_b)}$$

TLM avoided !!

Hybrid 4DEnVar

Static **B** part

$$\mathbf{B} = \mathbf{U}\mathbf{U}^{\mathrm{T}}$$

$$\delta \mathbf{x}_s = \mathbf{U}\mathbf{v}$$

Ensemble covariance part
(Liu et al 2008, 2009)

Localization matrix $\mathbf{C} = \mathbf{C}'\mathbf{C}'^{\mathrm{T}}$

perturbation matrix $\mathbf{S}'_i = \left(\mathbf{x}'_{bi} \quad \ldots \quad \mathbf{x}'_{bi}\right)$

Localized perturbation matrix

$$\mathbf{Z}'_b = [\mathbf{S}'_{b1} \circ \mathbf{C}' \quad \mathbf{S}'_{b2} \circ \mathbf{C}' \quad \ldots. \quad \mathbf{S}'_{bN} \circ \mathbf{C}']$$

Project $\mathbf{Z}'_b$ to obs space $(\mathbf{H}\mathbf{L}\mathbf{Z}'_b)$

$$J(\mathbf{v},\mathbf{w}) = \frac{1}{2}\mathbf{v}^T\mathbf{v} + \frac{1}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w} +$$

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

$$\frac{1}{2}\sum_{t=1}^{I}\left[\beta_1\mathbf{H}_t\mathbf{L}_t\delta\mathbf{x}_s + \beta_2(\mathbf{H}_t\mathbf{L}_t\mathbf{Z}'_b)\mathbf{w} + \mathbf{d}_t\right]^{\mathrm{T}}\mathbf{R}^{-1}[(\beta_1\mathbf{H}_t\mathbf{L}_t\delta\mathbf{x}_s + \beta_2(\mathbf{H}_t\mathbf{L}_t\mathbf{Z}'_b)\mathbf{w} + \mathbf{d}_t]$$

$$\nabla_{\mathbf{w}}J(\mathbf{w}) = \mathbf{w} + \sum_{t=1}^{I}\beta_1(\mathbf{H}_t\mathbf{L}_t\mathbf{Z}'_b)^{\mathrm{T}}\mathbf{R}^{-1}[\beta_1\mathbf{H}_t\mathbf{L}_t\mathbf{U}\mathbf{v} + \beta_2\mathbf{H}_t\mathbf{L}_t\mathbf{Z}'_b\mathbf{w} + \mathbf{d}_t]$$

Avoid AJM

4DEnVar-**NPL** (**N**o **P**ropagation of **L**ocalization)

Two approximations

(1) The localization matrix propagation by TLM is neglected

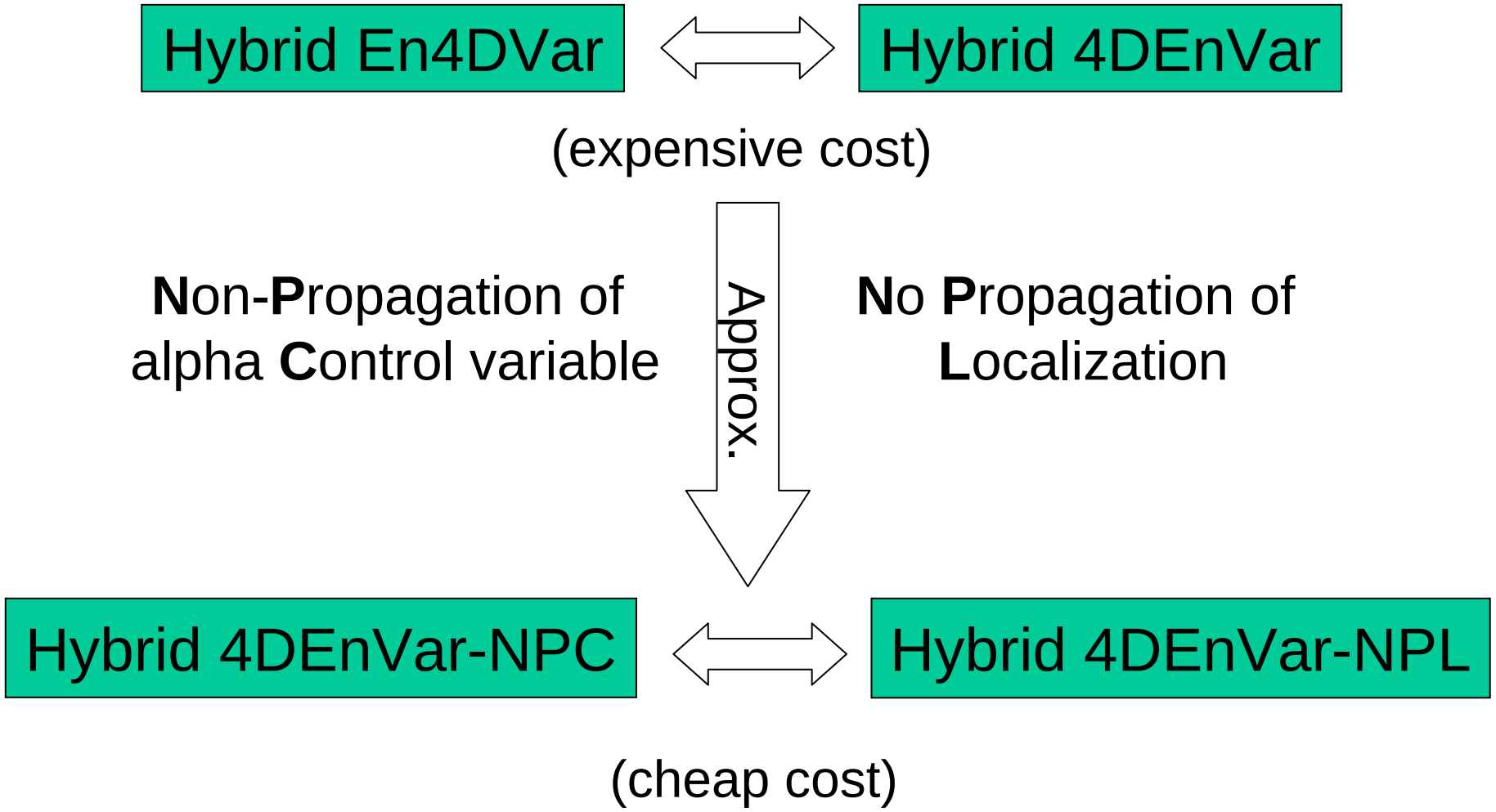
$$\mathbf{L}_t[\mathbf{S}'_{b1} \circ \mathbf{C}' \quad \mathbf{S}'_{b2} \circ \mathbf{C}' \quad .... \quad \mathbf{S}'_{bN} \circ \mathbf{C}']$$
$$\approx [(\mathbf{L}_t \mathbf{S}'_{b1}) \circ \mathbf{C}' \quad (\mathbf{L}_t \mathbf{S}'_{b2}) \circ \mathbf{C}' \quad .... \quad (\mathbf{L}_t \mathbf{S}'_{bN}) \circ \mathbf{C}']$$

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

$$\mathbf{L}_t \mathbf{x}'_{bi} \approx M_t(\mathbf{x}_{bi}) - \overline{M_t(\mathbf{x}_b)}$$

TLM avoided !!

The relationship of hybrid 4D ensemble variational algorithms



Equivalence of hybrid En4DVar and hybrid 4DEnVar

$$\underline{\mathbf{Z}_b^{'}\mathbf{w}} = (\mathbf{S}'_{b1} \circ \mathbf{C}', \mathbf{S}'_{b2} \circ \mathbf{C}', \ldots, \mathbf{S}'_{bN} \circ \mathbf{C}') \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \vdots \\ \mathbf{w}_N \end{bmatrix} = \sum_{i=1}^{N} (\mathbf{S}'_{bi} \circ \mathbf{C}')\mathbf{w}_i = \sum_{i=1}^{N} \mathbf{x}'_{bi} \circ (\mathbf{C}'\mathbf{w}_i)$$

From Hybrid 4DEnVar

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

Hybrid En4DVar Cost function ⟷ **Hybrid 4DEnVar Cost function**

Equivalence of 4DEnVar-NPC and 4DEnVar-NPL

$$\underline{\mathbf{T}_t'\mathbf{w}} = [(\mathbf{L}_t\,\mathbf{S}_{b1}')\circ\mathbf{C}' \quad (\mathbf{L}_t\mathbf{S}_{b2}')\circ\mathbf{C}' \quad \ldots. \quad (\mathbf{L}_t\mathbf{S}_{bN}')\circ\mathbf{C}']\begin{bmatrix}\mathbf{w}_1\\ \mathbf{w}_2\\ \vdots\\ \mathbf{w}_N\end{bmatrix} = \sum_{i=1}^{N}\mathbf{L}_t\mathbf{x}_{bi}'\circ(\mathbf{C}'\mathbf{w}_i).$$

From Hybrid 4DEnVar-NPL

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

Hybrid 4DEnVar-NPC Cost function ⟷ **Hybrid 4DEnVar-NPL Cost function**

Single observation tests with a 1D linear advection model

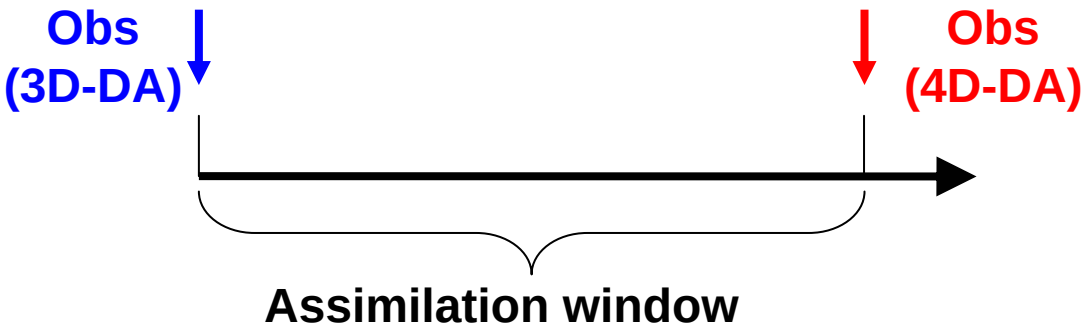
$$\frac{\partial u}{\partial t} + \mathrm{U}\frac{\partial u}{\partial x} = 0$$

The numerical result also proven:

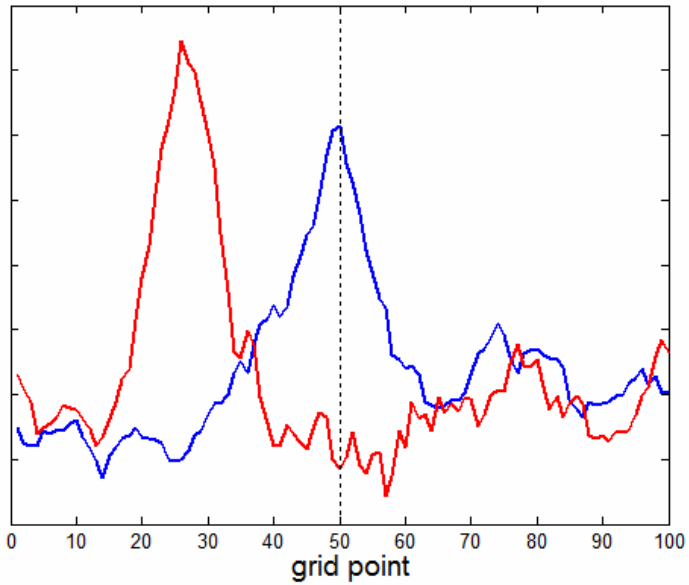
En4DVar and 4DEnVar (with no approximation to the flow-following-localization) are the same

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

Flow-following & non-flow-following localization



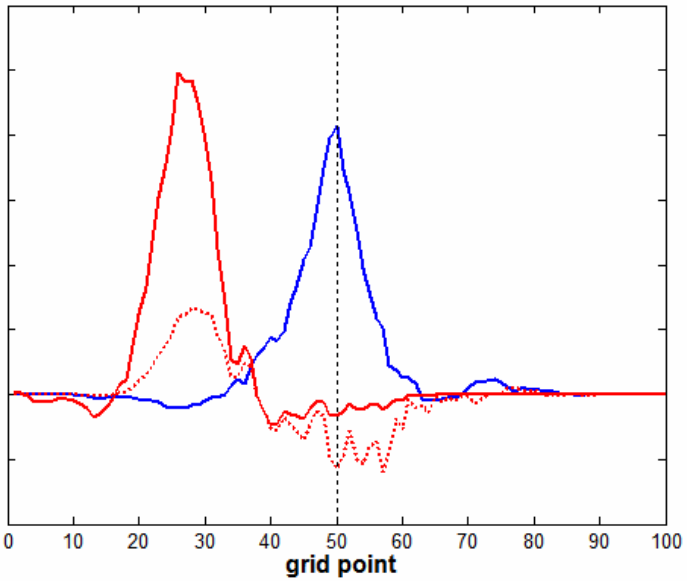
Ensemble DA **W/O localization**



(four algorithms are the same)

advection

Ensemble DA **W/ localization**



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

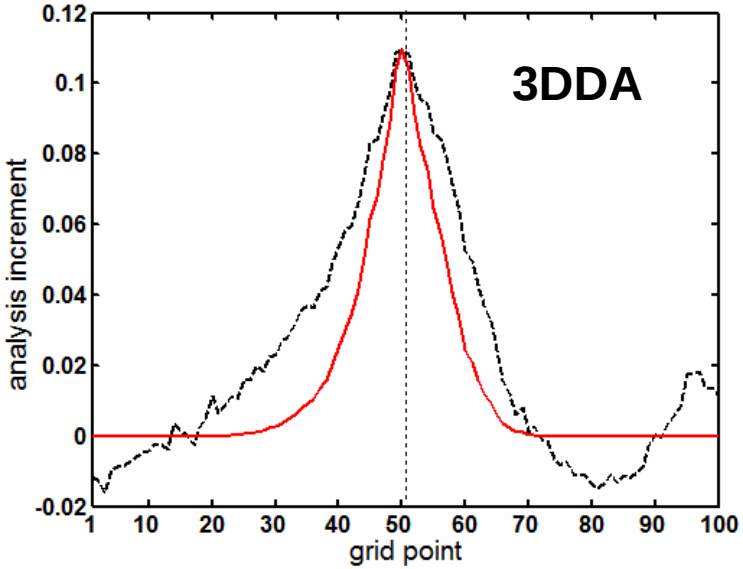
sensitivity experiments

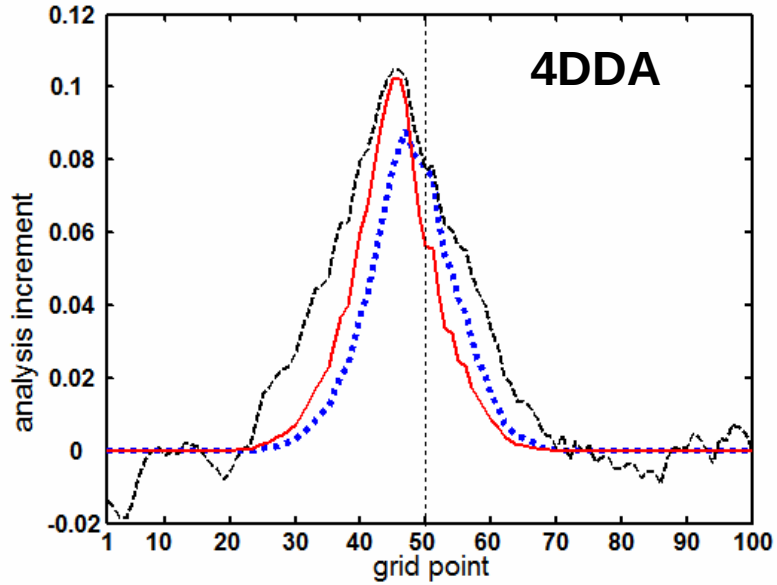
Exp1 (3DDA): obs at beginning of DA window

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

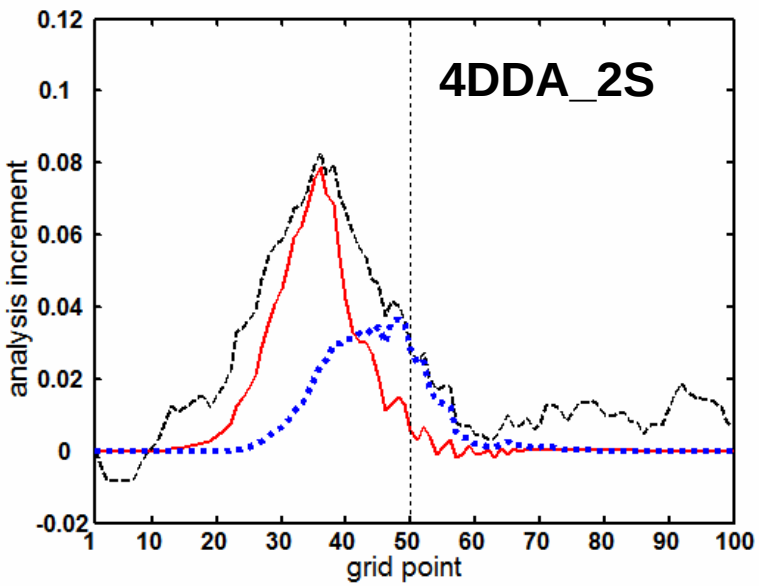


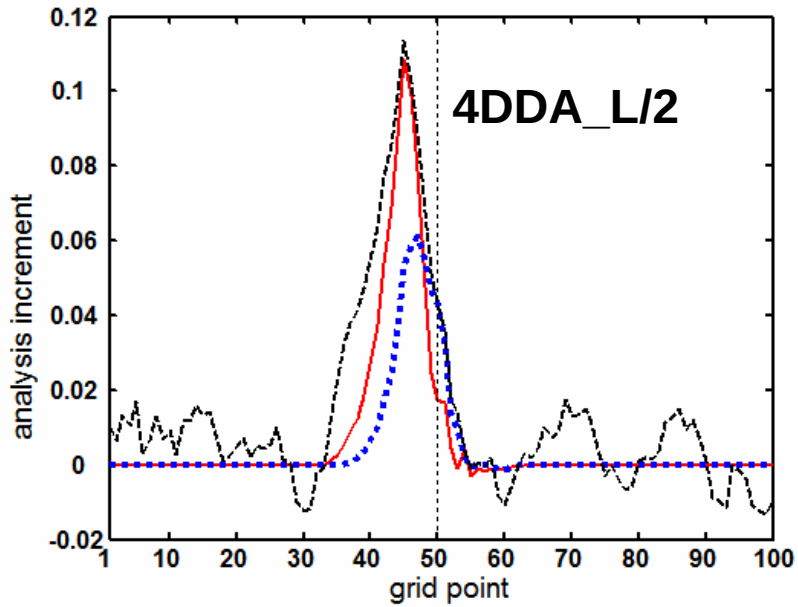


black no-localization

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

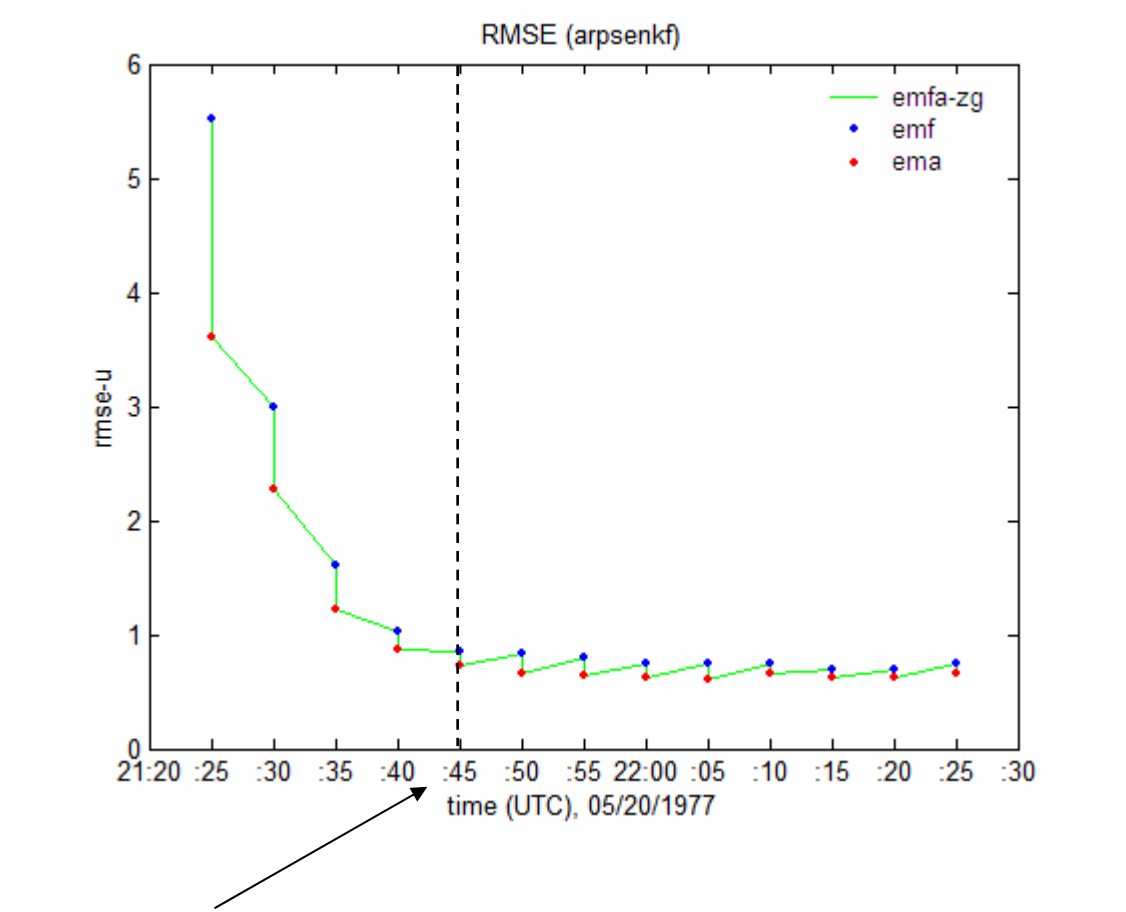
- Tested with simulated data from a classic supercell storm of 20 May 1977 near Del City, Oklahoma
- Domain : 67 x 67 x 35 grids, vertical stretching. 2km horizontal resolution
- 70-min length of simulation
- the de-correlation scale is 9.6km for horizontal and 3km for vertical.

Single observation test

- Single observation analysis is on 25th minute after forecast
- Two kinds of perturbations,

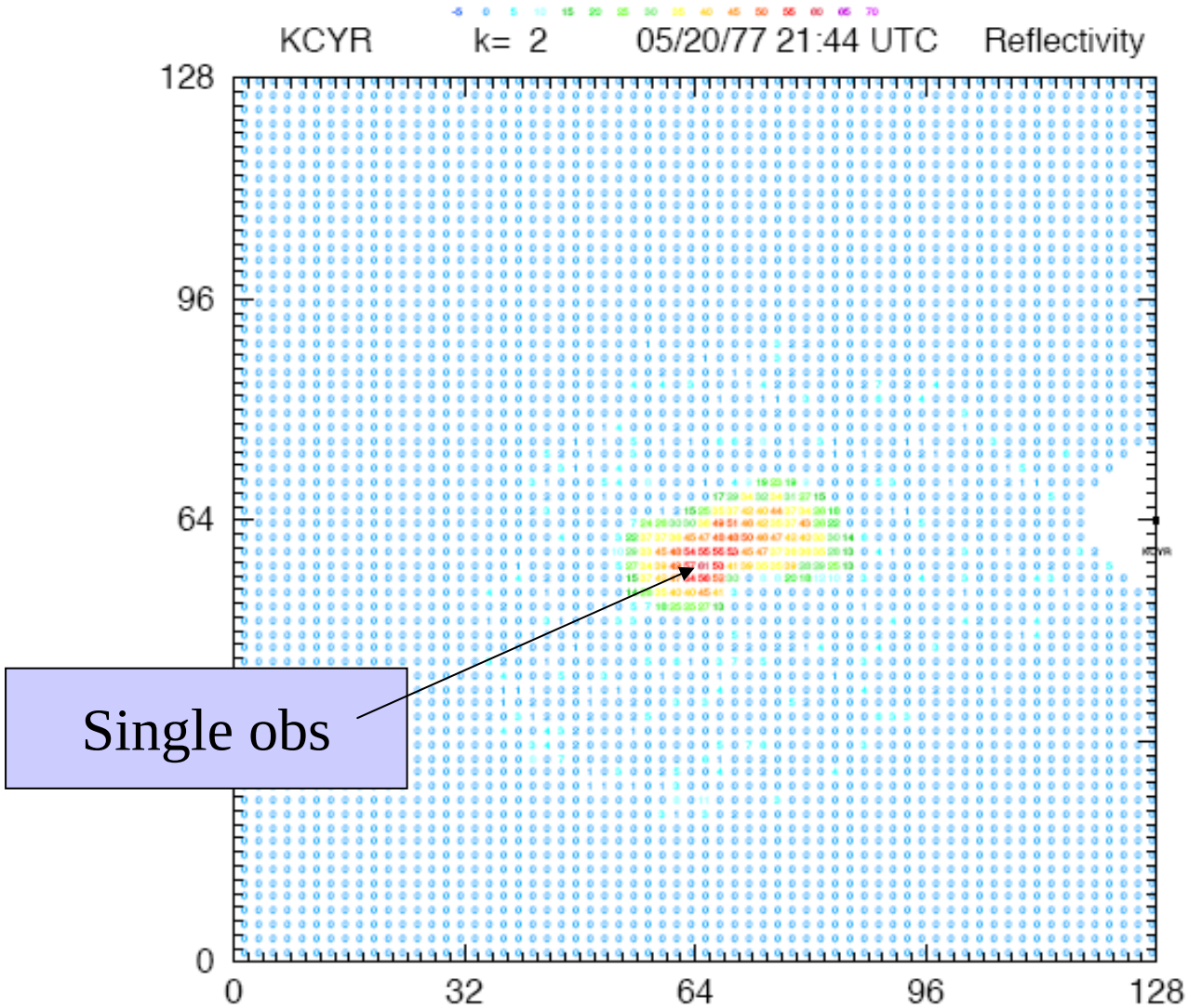
1. 25 minutes forecast from background
2. 4 cycles (5min cycle interval) by EnKF

RMSE of EnKF cycling

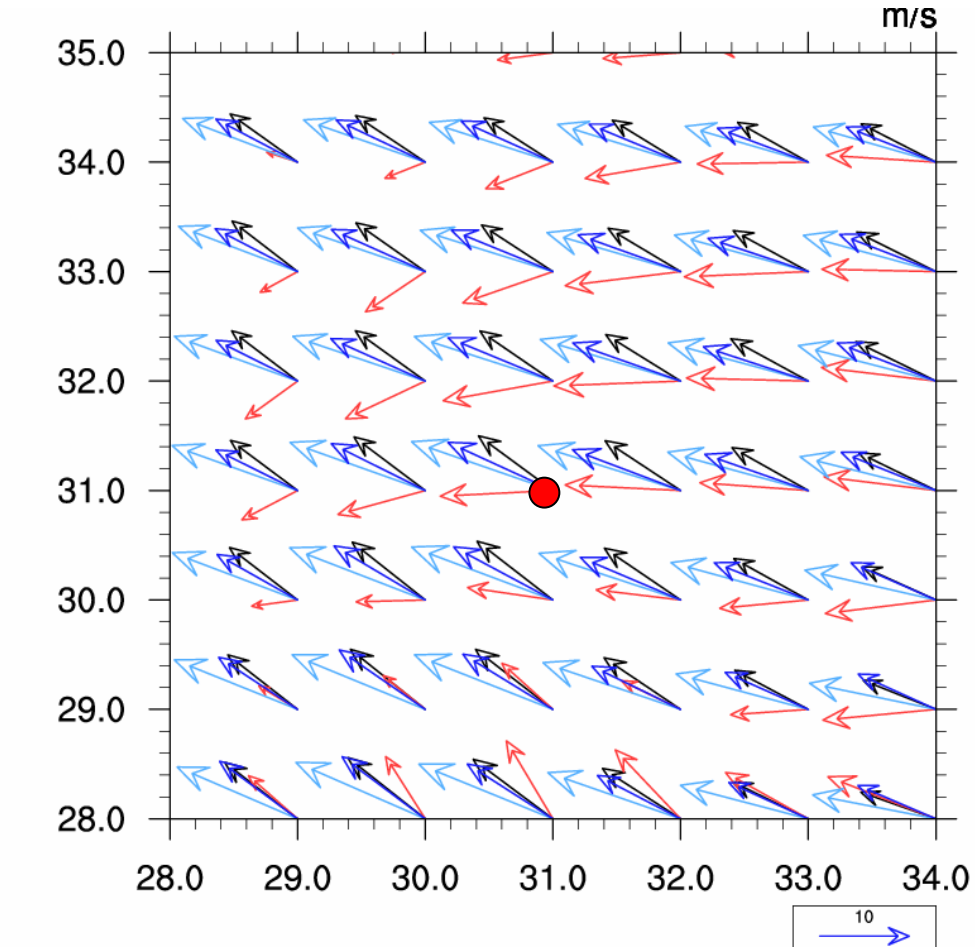


Single obs test time

Observations



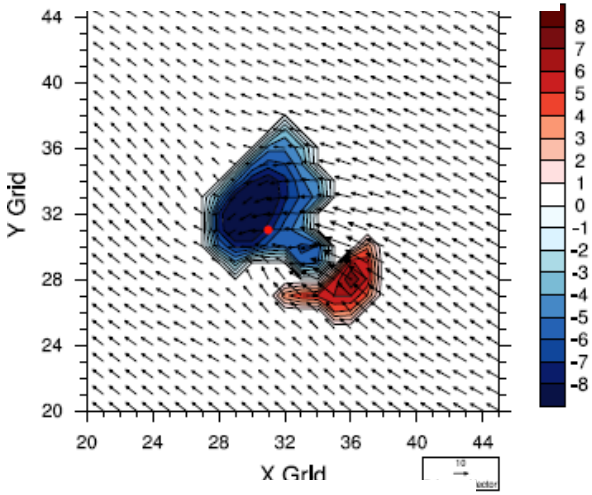
wind vector near observation (Vr)



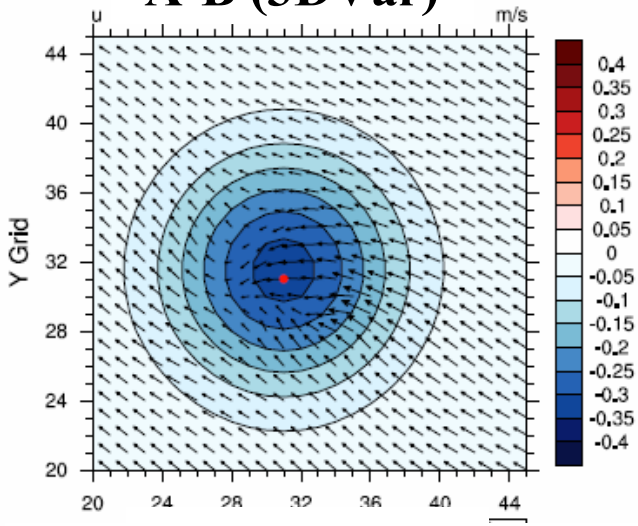
True, Background, 3DVar, En3DVar
(EnKF is very closed to En3DVar)

horizontal V Wind (k=12)

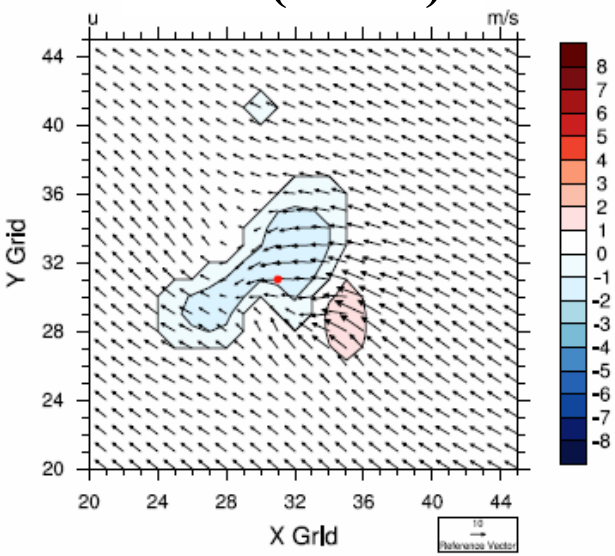
T-B



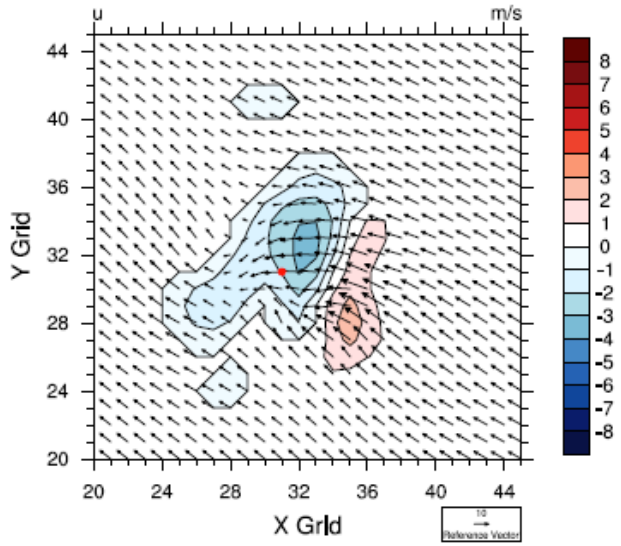
A-B (3DVar)



A-B (EnKF)

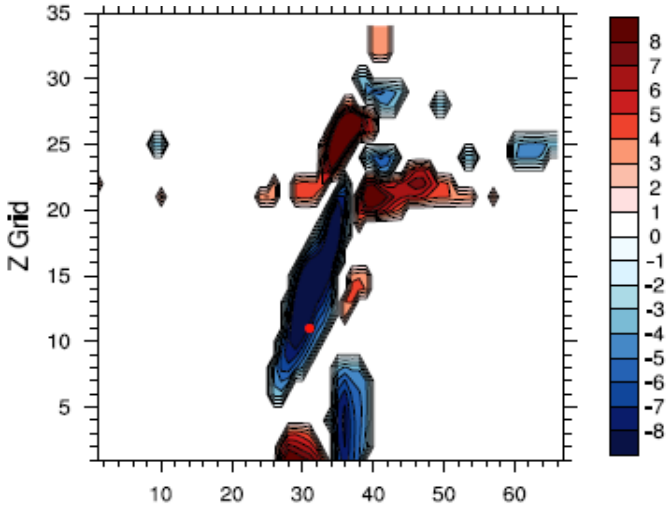


A-B (En3DVar)

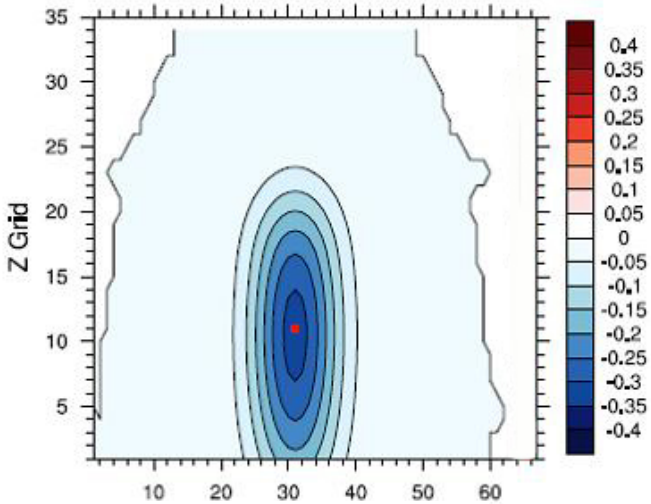


v wind vertical cross section (y=31)

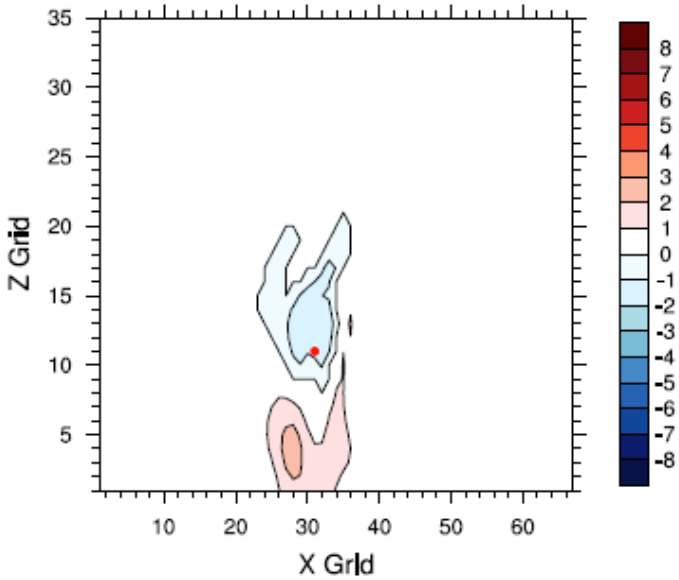
T-B



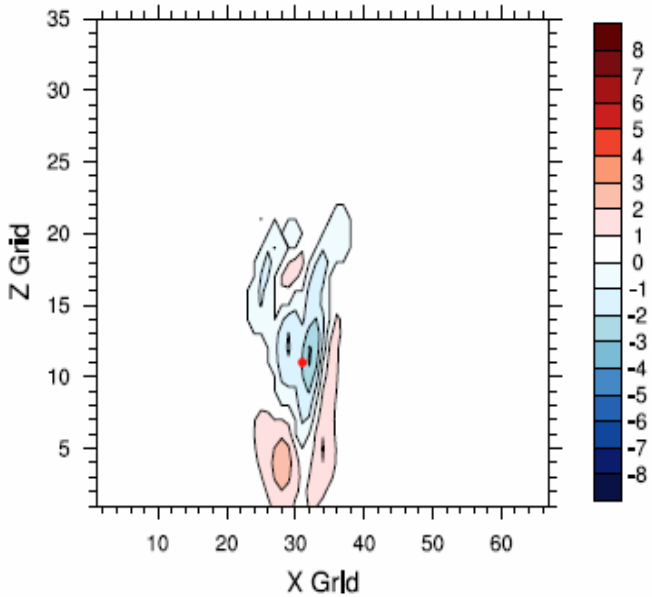
A-B (3DVar)



A-B (EnKF)

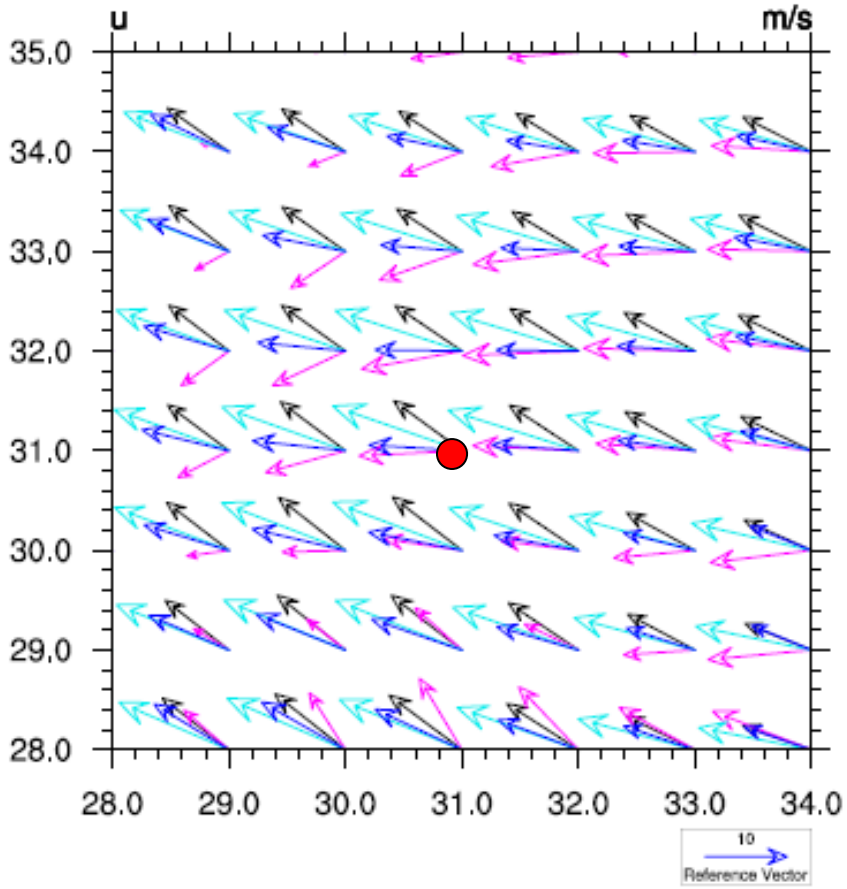


A-B (En3DVar)



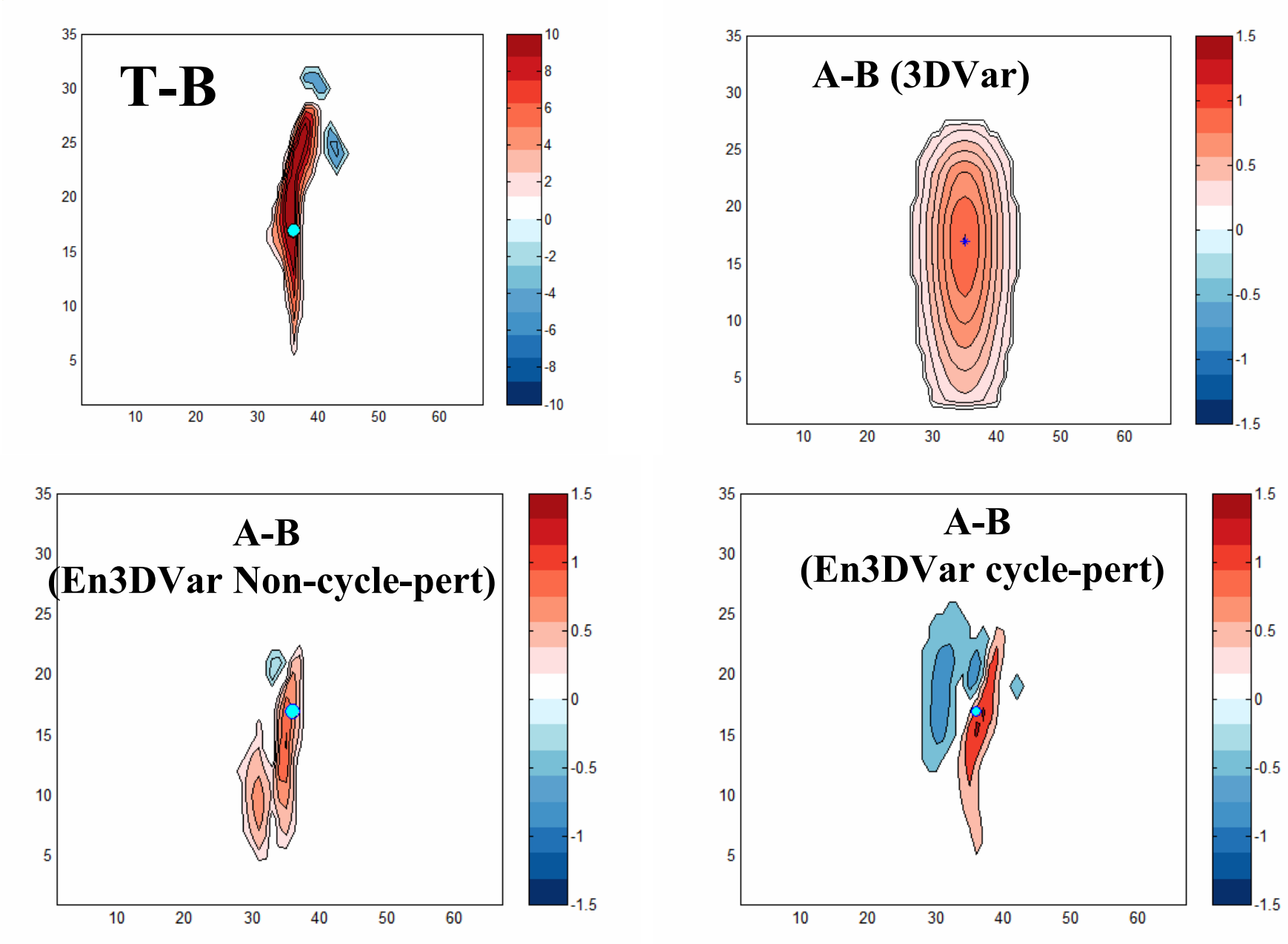
wind vector near observation (Vr)

Non-cycle-perturbation .VS. cycle-perturbation

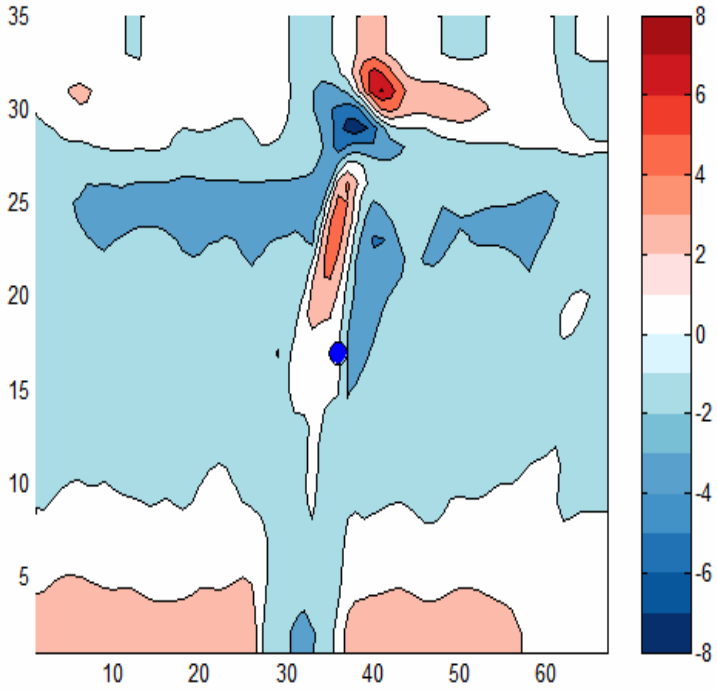


Background, true, En3DVar_NCP , En3DVar_CP

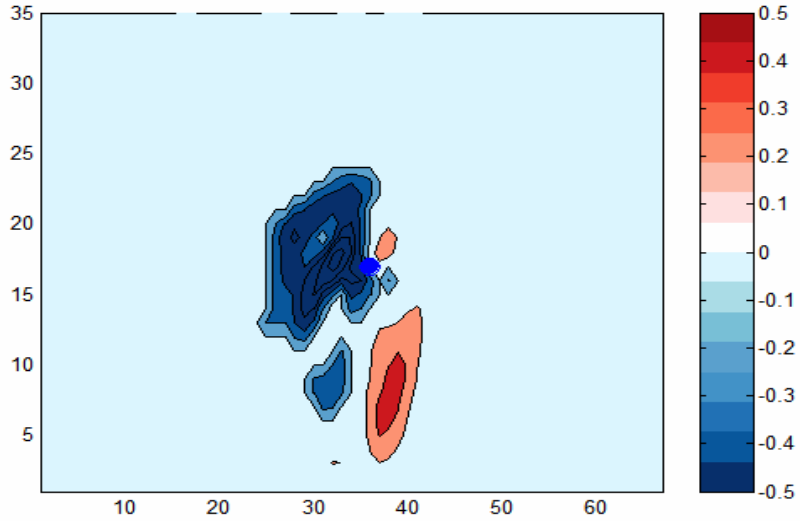
W on vertical cross-section



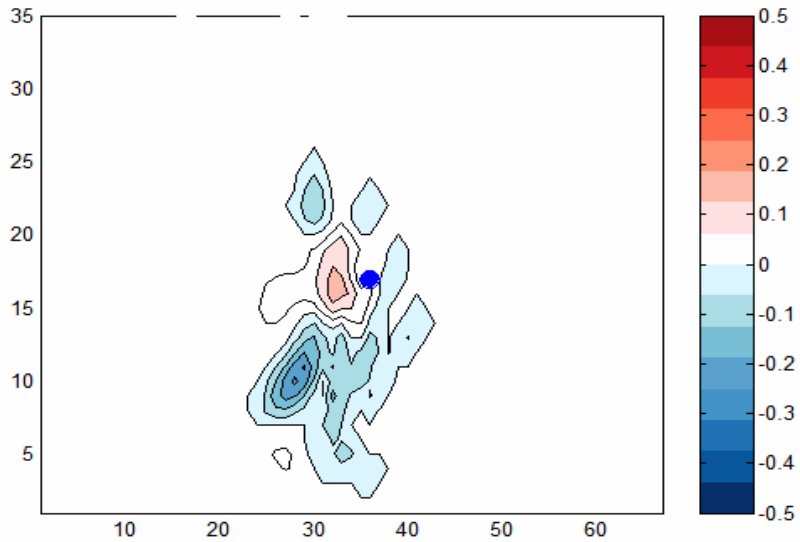
PT on vertical cross-section



T-B

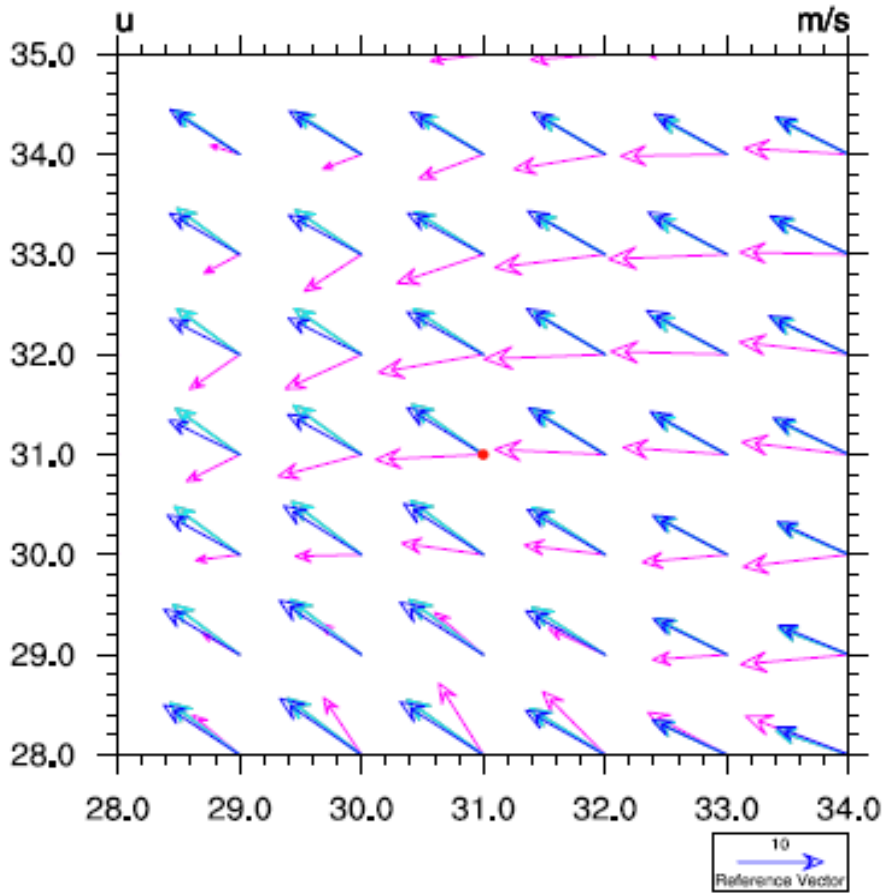


A-B (Non-cycle-pert)



A-B (cycle-pert)

Reflectivity DA

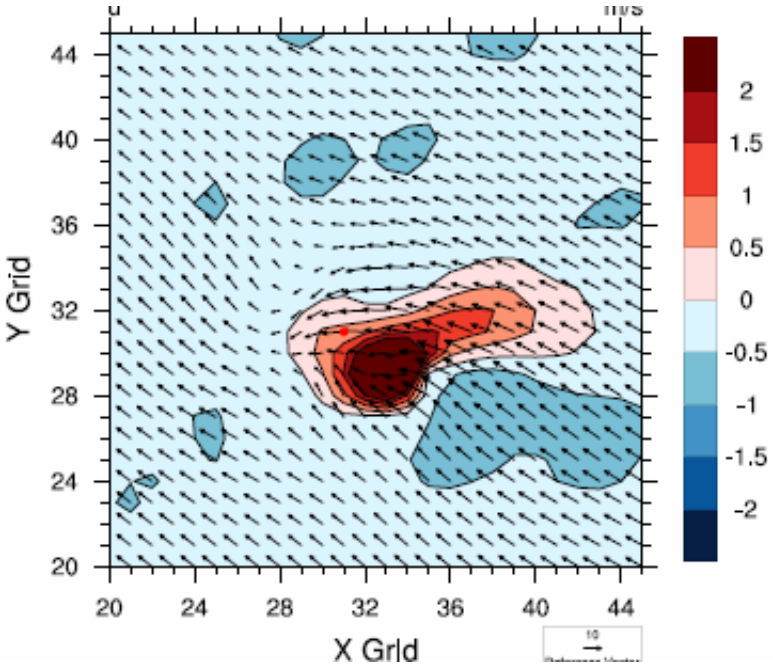


True, background and 3DEnVar

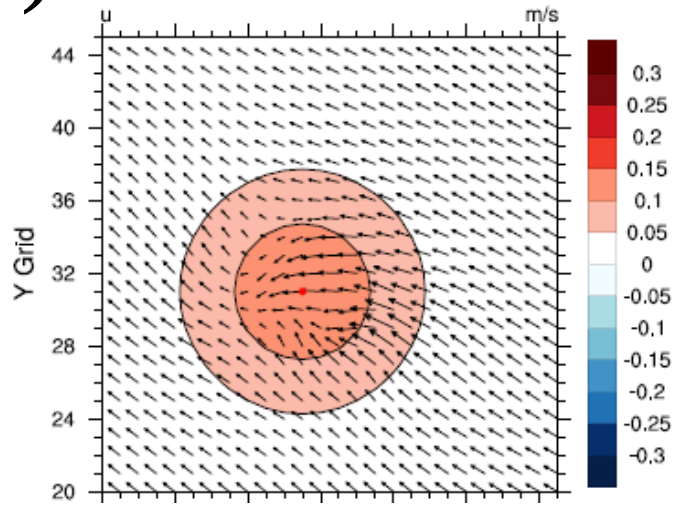
(there are little affect to wind by reflectivity)

Qr (k=12)

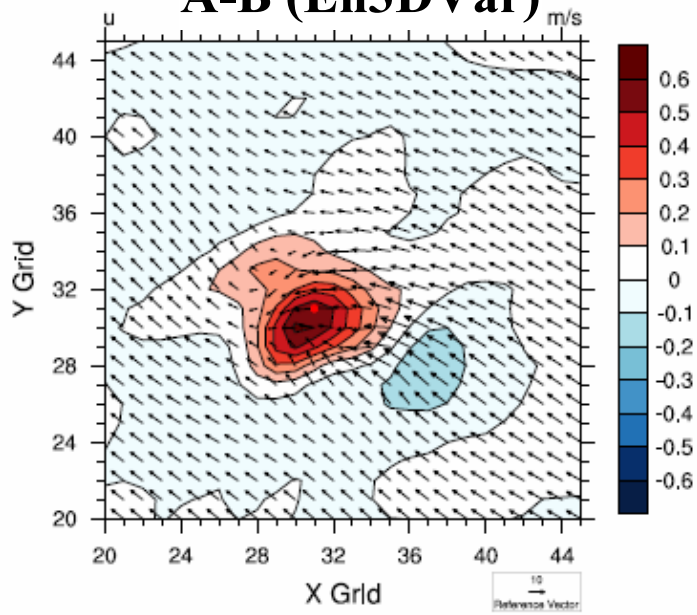
T-B



A-B (3DVar)

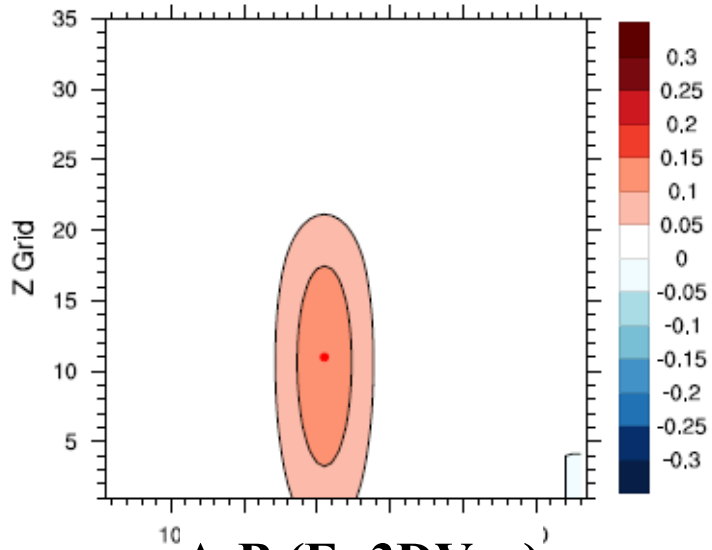


A-B (En3DVar)

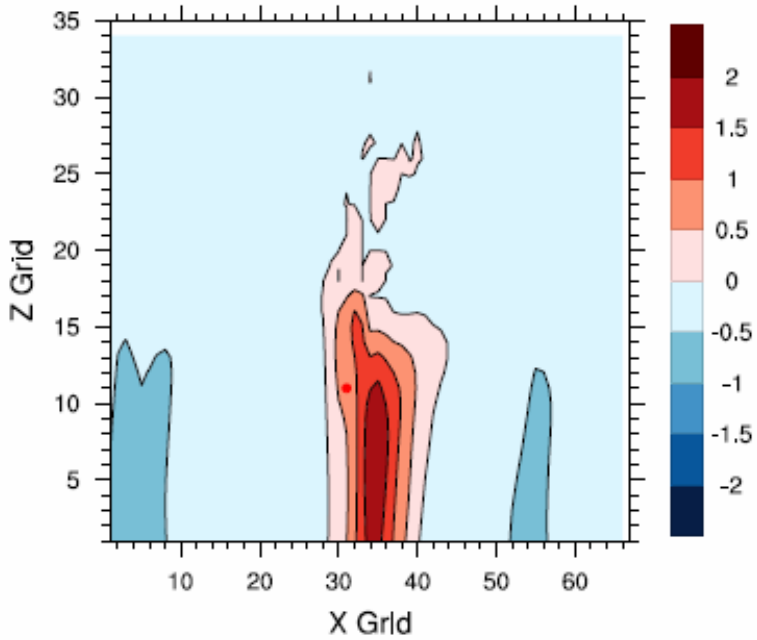


Qr on vertical cross-section

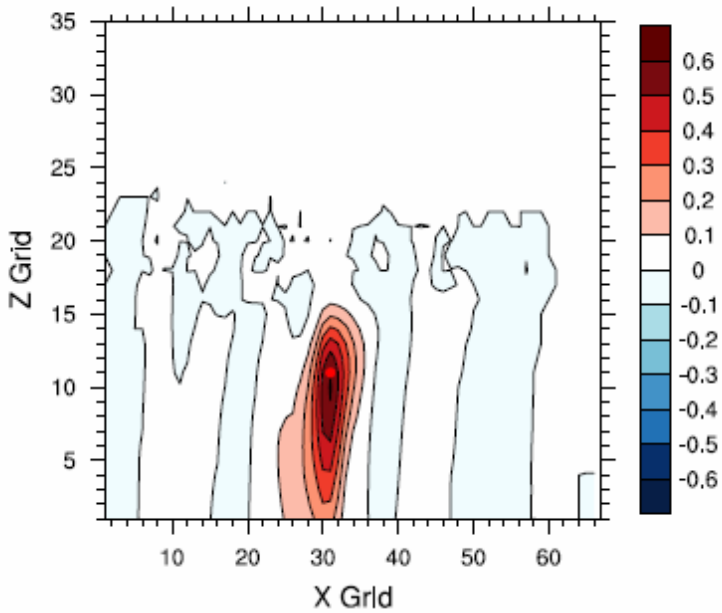
A-B (3DVar)



T-B



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 de-correlation 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