Comparisons between 4DEnVar and 4DVar on the Met Office global model

David Fairbairn

University of Surrey/Met Office

26^{th} June 2013

Joint project by David Fairbairn, Stephen Pring, Andrew Lorenc, Neill Bowler and Ian Roulstone

Funded by EPSRC, Met Office and University of Surrey.

Introduction

- Data assimilation in NWP combines a prior forecast (background) with the latest obs of the atmosphere, to provide the starting point for a weather forecast;
- Estimation of background and observation error covariance matrices are needed to weight the observations and the background, and spread information between variables;
- NWP centres are currently going through a transition point. The climatological background-error covariance is gradually being replaced by a flow-dependent approximation from ensemble forecasts;
- Advances in computing power have only recently made ensemble DA affordable.

Introduction

- Ensemble covariance still not considered good enough to completely replace climatological covariance because of
  1. Sampling error - only order 10-100 ensembles affordable to sample of the order 10^8 model gridpoints;
  2. Model error - difficult to represent flow-dependent model errors.
- Most NWP centres therefore prefer hybrid DA methods, which combine climatological/ensemble covariances;
- Hybrid 4DVar and hybrid 4DEnVar are two competitive DA methods that are not yet fully understood;
- This project aims to improve undestanding of these methods in NWP.

4DVar

- 4DVar [Le-Dimet and Talagrand, 1986] provides a least squares fit between observations and a prior forecast in an assimilation window;
- Based on the incremental formulation of Courtier et al. [1994];
- 4D background state ($\underline{\mathbf{x}}^b$) propagated through the window using the forecast model:

$$\underline{\mathbf{x}}^b = \underline{M}(\mathbf{x}^b(t_0)); \tag{1}$$

- Lower resolution increment:

$$\delta\mathbf{w}(t_0) = S(\mathbf{x}^b(t_0)) - S(\mathbf{x}(t_0)); \tag{2}$$

- Increment propagated using perturbation forecast model:

$$\delta\underline{\mathbf{w}} = \underline{\tilde{\mathbf{M}}}\delta\mathbf{w}(t_0) \tag{3}$$

4DVar cost function

- 4DVar cost function [Rawlins et al., 2007]:

$$
\begin{aligned}
J(\delta\mathbf{w}) &= J_b + J_o + J_c \\
&= \frac{1}{2}\delta\mathbf{w}(t_0)^T\mathbf{B}^{-1}\delta\mathbf{w}(t_0) \\
&+ \frac{1}{2}(\underline{\mathbf{y}} - \underline{\mathbf{y}}^o)^T\underline{\mathbf{R}}^{-1}(\underline{\mathbf{y}} - \underline{\mathbf{y}}^o) \\
&+ J_c,
\end{aligned}
\tag{4}
$$

where

$$
\underline{\mathbf{y}} = \underline{H}(\underline{\mathbf{x}}) + \underline{\tilde{\mathbf{H}}}\delta\underline{\mathbf{w}}. \tag{5}
$$

Control variable transform

- Control variable transform used to pre-condition 4DVar cost function:

$$\delta \mathbf{w} = \mathbf{U}\mathbf{v}, \tag{6}$$

where $\mathbf{U}\mathbf{U}^T = \mathbf{B}$.

- J_b can then be expressed in terms of $\mathbf{v}$:

$$\begin{aligned} J_b &= \frac{1}{2}\mathbf{v}^T\mathbf{U}^T(\mathbf{U}\mathbf{U}^T)^{-1}\mathbf{U}\mathbf{v} \\ &= \frac{1}{2}\mathbf{v}^T\mathbf{v}. \end{aligned} \tag{7}$$

- Cost function gradient:

$$\left[\frac{\partial J}{\partial \mathbf{v}}\right] = \mathbf{v} + \mathbf{U}^T \underline{\tilde{\mathbf{M}}}^T \underline{\tilde{\mathbf{H}}}^T \underline{\mathbf{R}}^{-1}(\underline{\mathbf{y}} - \underline{\mathbf{y}}^o) \tag{8}$$

4DVar-Ben

- 4DVar-Ben uses 3D $\mathbf{P}^b$ instead of $\mathbf{B}$;
- At the Met Office, $\mathbf{P}^b$ currently comes from the ETKF [Bishop et al., 2001];
- 4DVar-Ben initial increment expressed in terms of alpha control variable [Lorenc, 2003];
- Locally weighted linear combination of m ensemble perturbation trajectories:

$$\delta \mathbf{w}(t_0) = \sum_{j=1}^{m} \frac{1}{\sqrt{m-1}} \delta \mathbf{w}_j^b(t_0) \circ \boldsymbol{\alpha}_j, \tag{9}$$

where $\boldsymbol{\alpha}_j$ are the 3D fields of weights for each perturbation, and the fields are modified by the Gaspari-Cohn localization matrix.

Control variable transform

- Control variable transform to condition J_α (like 4DVar):

$$\boldsymbol{\alpha}_j = \mathbf{U}^{\boldsymbol{\alpha}} \mathbf{v}_j^{\boldsymbol{\alpha}} \text{ for } j = 1, ..., m, \tag{10}$$

where

$$(\mathbf{U}^{\boldsymbol{\alpha}})^T \mathbf{U}^{\boldsymbol{\alpha}} = \mathbf{C} \tag{11}$$

and **C** is the Gaspari-Cohn localization matrix;

- Sequence of control vectors $\mathbf{v}_j^{\boldsymbol{\alpha}}$ concatenated to make $\mathbf{v}^{\boldsymbol{\alpha}\mathbf{s}}$;
- New operator $\mathbf{U}^{\boldsymbol{\alpha}\mathbf{s}}$ to represent (9) and (10):

$$\delta \mathbf{w}(t_0) = \mathbf{U}^{\boldsymbol{\alpha}\mathbf{s}} \mathbf{v}^{\boldsymbol{\alpha}\mathbf{s}} \tag{12}$$

- Increment then propagated using perturbation forecast model:

$$\delta \underline{\mathbf{w}} = \underline{\tilde{\mathbf{M}}} \delta \mathbf{w}(t_0) \tag{13}$$

Cost function

- Cost function:

$$J(\mathbf{v}^{\alpha\mathbf{s}}) = \frac{1}{2}(\mathbf{v}^{\alpha\mathbf{s}})^T\mathbf{v}^{\alpha\mathbf{s}} + J_o + J_c \tag{14}$$

- Cost function gradient:

$$\left[\frac{\partial J}{\partial \mathbf{v}^{\alpha\mathbf{s}}}\right] = \mathbf{v}^{\alpha\mathbf{s}} + (\mathbf{U}^{\alpha\mathbf{s}})^T \underline{\tilde{\mathbf{M}}}^T \underline{\tilde{\mathbf{H}}}^T \underline{\mathbf{R}}^{-1}(\underline{\mathbf{y}} - \underline{\mathbf{y}}^o). \tag{15}$$

Hybrid 4DVar

- Hybrid 4DVar uses linearly weighted combination of climatological ($\mathbf{v}$) and ensemble ($\mathbf{v}^{\alpha \mathbf{s}}$) background terms:

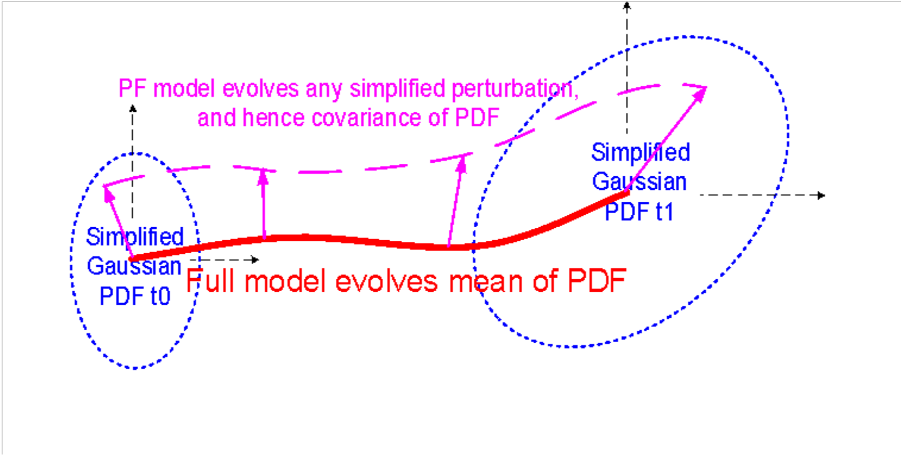
$$\delta \mathbf{w}(t_0) = \sqrt{\beta_c}\mathbf{U}\mathbf{v} + \sqrt{\beta_e}\mathbf{U}^{\alpha \mathbf{s}}\mathbf{v}^{\alpha \mathbf{s}}; \tag{16}$$

- Hybrid covariance ($\mathbf{B}_h$) equivalent to weighting climatological and ensemble covariances [Wang et al., 2007]: $\mathbf{B}_h = \beta_c \mathbf{B} + \beta_e \mathbf{P}^b$.
- Cost function:

$$J(\mathbf{v}, \mathbf{v}^{\alpha \mathbf{s}}) = \frac{1}{2}\mathbf{v}^T\mathbf{v} + \frac{1}{2}(\mathbf{v}^{\alpha \mathbf{s}})^T\mathbf{v}^{\alpha \mathbf{s}} + J_o + J_c \tag{17}$$

- Cost function gradient calculated by concatenating climatological $\left[\frac{\partial J}{\partial \mathbf{v}}\right]$ and ensemble $\left[\frac{\partial J}{\partial \mathbf{v}^{\alpha \mathbf{s}}}\right]$ parts.

Diagram of 4DVar



Strengths and weaknesses

Strengths:

- Linear model propagation of background-error covariance ($(\tilde{\mathbf{M}})^T B_h \tilde{\mathbf{M}}$) is accurate provided that $\tilde{\mathbf{M}}$ is accurate;
- Hybrid background-error covariance performs better than pure ensemble or climatological covariances in NWP (e.g. Clayton et al. [2012]);

Weaknesses:

- Linear models expensive to maintain in terms of staff/computing;
- Linear model approximation less accurate than full nonlinear model.

4DEnVar

- Least squares fit between ensemble and observations in assimilation window;
- Uses 4D $\underline{\mathbf{P}}^b$ from ETKF, unlike 4DVar-Ben, which uses 3D $\mathbf{P}^b$;
- 4DEnVar equivalent to 4D EnKF, except 4DEnVar correctly localizes in model space [Fairbairn et al., 2013];
- Same algorithm as 4DVar-Ben at initial time;
- Unlike 4DVar-Ben, alpha control variable extended to **all timesteps**, not just initial timestep;

4DEnVar

- Same algorithm as 4DVar-Ben at initial time;
- Unlike 4DVar-Ben, alpha control variable extended to **all timesteps**, not just initial timestep;
- Locally weighted linear combination of m ensemble perturbation trajectories:

$$\delta \underline{\mathbf{w}} = \sum_{j=1}^{m} \frac{1}{\sqrt{m-1}} \delta \underline{\mathbf{w}}_j^b \circ \underline{\boldsymbol{\alpha}}_j. \tag{18}$$

- Most applications (including the Met Office) assume localization is constant in time $\rightarrow \boldsymbol{\alpha}_j$ is constant in time. Control variable $\mathbf{v}^{\boldsymbol{\alpha}}$ is therefore 3D.

Control variable transform

- Control variable transform to condition J_α (like 4DVar):

$$\underline{\boldsymbol{\alpha}}_j = \underline{\mathbf{U}}^{\boldsymbol{\alpha}} \mathbf{v}_j^{\boldsymbol{\alpha}} \text{ for } j = 1, ..., m, \tag{19}$$

where

$$\underline{\mathbf{U}}^{\boldsymbol{\alpha}} (\underline{\mathbf{U}}^{\boldsymbol{\alpha}})^T = \underline{\mathbf{C}} \tag{20}$$

and $\mathbf{C}$ is the same at each timestep.

- Sequence of control vectors $\mathbf{v}_j^{\boldsymbol{\alpha}}$ concatenated to make $\mathbf{v}^{\boldsymbol{\alpha}\mathbf{s}}$;
- New operator $\underline{\mathbf{U}}^{\boldsymbol{\alpha}\mathbf{s}}$ to represent (18) and (19):

$$\delta \underline{\mathbf{w}} = \underline{\mathbf{U}}^{\boldsymbol{\alpha}\mathbf{s}} \mathbf{v}^{\boldsymbol{\alpha}\mathbf{s}}. \tag{21}$$

Cost function

- Cost function:

$$J(\mathbf{v}^{\alpha s}) = \frac{1}{2}(\mathbf{v}^{\alpha s})^T \mathbf{v}^{\alpha s} + J_o \tag{22}$$

- Cost function gradient:

$$\left[\frac{\partial J}{\partial \mathbf{v}^{\alpha s}}\right] = \mathbf{v}^{\alpha s} + (\underline{\mathbf{U}}^{\alpha s})^T \underline{\tilde{\mathbf{H}}}^T \underline{\mathbf{R}}^{-1}(\underline{\mathbf{y}} - \underline{\mathbf{y}}^o). \tag{23}$$

- No J_c term - uses IAU-like initialization instead;
- IAU-like initialization adds increments on gradually through window - prevents fast modes from growing.

Hybrid 4DEnVar

- Hybrid 4DEnVar uses linearly weighted combination of climatological ($\mathbf{v}$) and ensemble ($\mathbf{v}^{\alpha \mathbf{s}}$) background terms:

$$\delta \underline{\mathbf{w}} = \sqrt{\beta_c} \underline{\mathbf{U}} \mathbf{v} + \sqrt{\beta_e} \underline{\mathbf{U}}^{\alpha \mathbf{s}} \mathbf{v}^{\alpha \mathbf{s}}; \tag{24}$$

- Ensemble term is the same as 4DEnVar;
- Climatological term equivalent to 3DVar:

$$\underline{\mathbf{U}} \underline{\mathbf{U}}^T = \underline{\mathbf{B}}, \tag{25}$$

  where $\mathbf{B}$ is the same at each timestep.

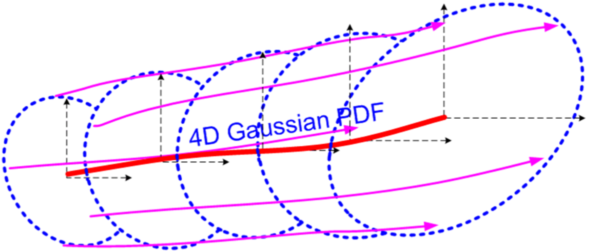
- No flow-dependence in $\mathbf{B}$ at the end of the window, unlike 4DVar ($\tilde{\mathbf{M}} \mathbf{B} \tilde{\mathbf{M}}^T$)!
- Cost function:

$$J(\mathbf{v}, \mathbf{v}^{\alpha \mathbf{s}}) = \frac{1}{2} \mathbf{v}^T \mathbf{v} + \frac{1}{2} (\mathbf{v}^{\alpha \mathbf{s}})^T \mathbf{v}^{\alpha \mathbf{s}} + J_o. \tag{26}$$

- Cost function gradient calculated by concatenating climatological $\left[\frac{\partial J}{\partial \mathbf{v}}\right]$ and ensemble $\left[\frac{\partial J}{\partial \mathbf{v}^{\alpha \mathbf{s}}}\right]$ parts.

Diagram of 4DEnVar

- Full nonlinear model propagation of **ensemble only**;
- Climatological part of covariance is static through window.



Trajectories of perturbations from ensemble mean

Full model evolves mean of PDF

Localised trajectories define 4D PDF of possible increments

Strengths and weaknesses (compared with hybrid 4DVar)

Strengths:

- 4D $\underline{\mathbf{P}}^b$ should be more accurate than 4DVar $\underline{\tilde{\mathbf{M}}}\mathbf{P}^b\underline{\tilde{\mathbf{M}}}^T$;
- No need for expensive Linear and adjoint models;
- Shares many of the same features as 4DVar e.g. minimization algorithm etc...

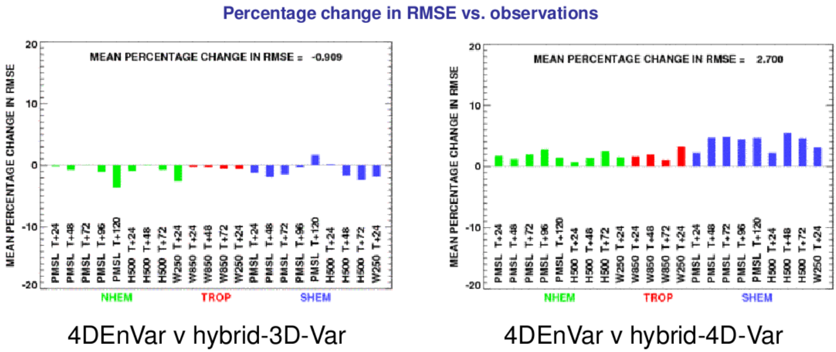
Weaknesses

- 4DEnVar uses 3D representation of climatological $\mathbf{B}$, but 4DVar uses 4D representation $\underline{\tilde{\mathbf{M}}}\mathbf{B}\underline{\tilde{\mathbf{M}}}^T$;
- Localization function and Schur product do not commute - Severe localization can significantly degrade time correlations of $\underline{\mathbf{P}}^b$ [Fairbairn et al., 2013]:

$$\mathbf{C} \circ \mathbf{M}\mathbf{P}^b\mathbf{M}^T \neq \mathbf{M}(\mathbf{C} \circ \mathbf{P}^b)\mathbf{M}^T. \quad (27)$$

Motivation for this research

- The Met Office is thinking of moving from hybrid 4DVar to hybrid 4DEnVar;
- Trial was run in 2012 to compare 44 member hybrid 4DVar against hybrid 4DEnVar (ratio $0.8\beta_c : 0.5\beta_e$);
- Hybrid 4DEnVar beat hybrid 3DVar, but Hybrid 4DEnVar performed worse than hybrid 4DVar, particularly in the Southern hemisphere!
- Single obs experiments can help to explain these results.



Single observation experiments

- Analysis increment for each DA method computed for single observations;
- Pseudo observation generated from real analysis - large increments added to obs to increase impact;
- Two very different extreme weather types selected:
  1. Strong midlatitude jet, where $||\mathbf{P}^b||_2 \approx ||\mathbf{B}||_2$;
  2. Hurricane Sandy, where $||\mathbf{P}^b||_2 >> ||\mathbf{B}||_2$.
- This talk focuses on the jet stream case;

Why single obs experiments?

- Provides test of background-error covariance to spread information;
- Quick and easy to run (unlike trials, which can take months);
- Results can help to direct future trials;

The observation types

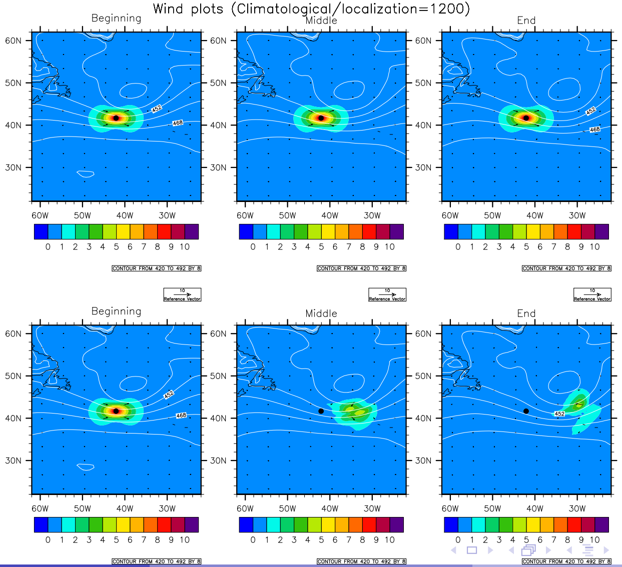
- Single ob located at beginning of window → 4DVar and 4DEnVar should be equivalent at t_0;
- Localization function $exp(\frac{-Z^2}{2L^2})$, where Z is the distance and L is length-scale.
- Met Office currently use $L = 1200km$. When $Z = 1200km$, $Localization = 0.61$, when $Z = 2400km$, $Localization = 0.14$.

Jet observation:

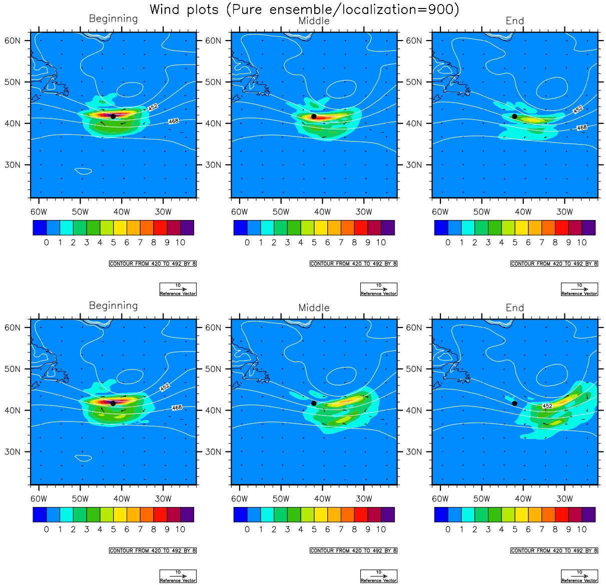
- - Single Westerly wind (u) observation with increment +10m/s;
- - Observation located at level 29 (≈500hPa), at coordinates 41N,41W.

Jet: $\beta_C = 1.0$, $\beta_e = 0.0$

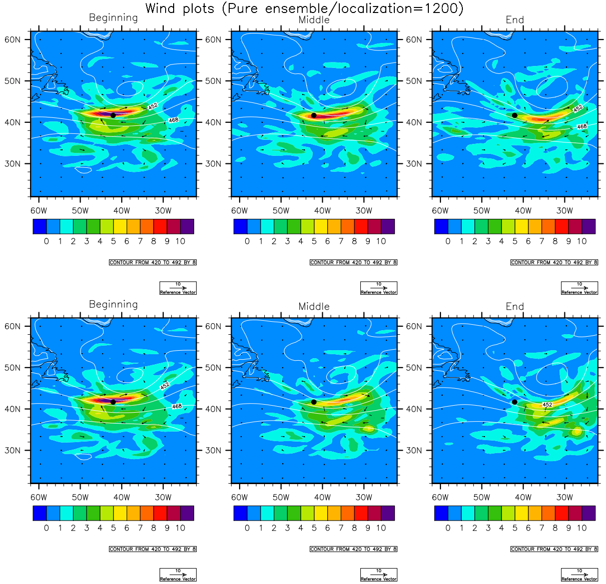
4DEnVar (top) and 4DVar (bottom) wind increments at beginning (left), middle (middle) and end (right) of the assimilation window:



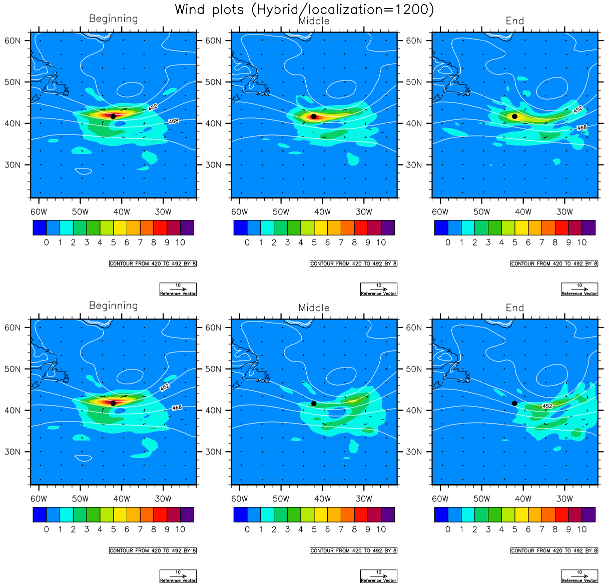
Jet: $\beta_C = 0.0$, $\beta_e = 1.0$, $L = 500km$



Jet: $\beta_C = 0.0$, $\beta_e = 1.0$, $L = 1200km$



Jet: $\beta_C = 0.5$, $\beta_e = 0.5$, $L = 1200km$



“4D” errors

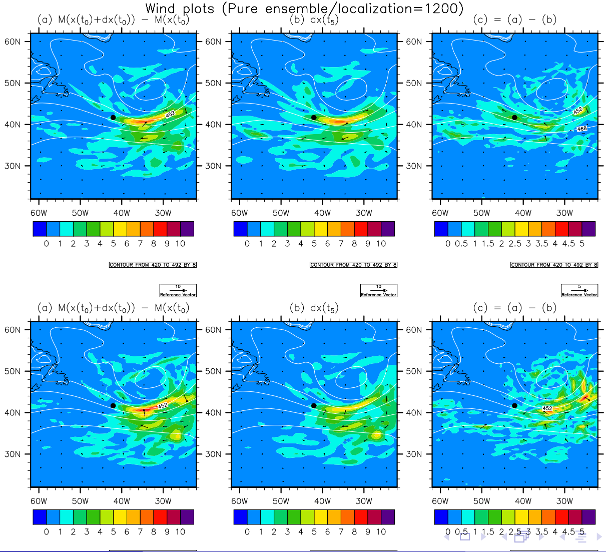
- The “4D” errors are introduced as a result of the ”time” dimension in the assimilation window:

$$M_{0\to5}(\mathbf{x}^b(t_0) + \delta\mathbf{x}^a(t_0)) - M_{0\to5}(\mathbf{x}^b(t_0)) - \delta\mathbf{x}^a(t_5) \tag{28}$$

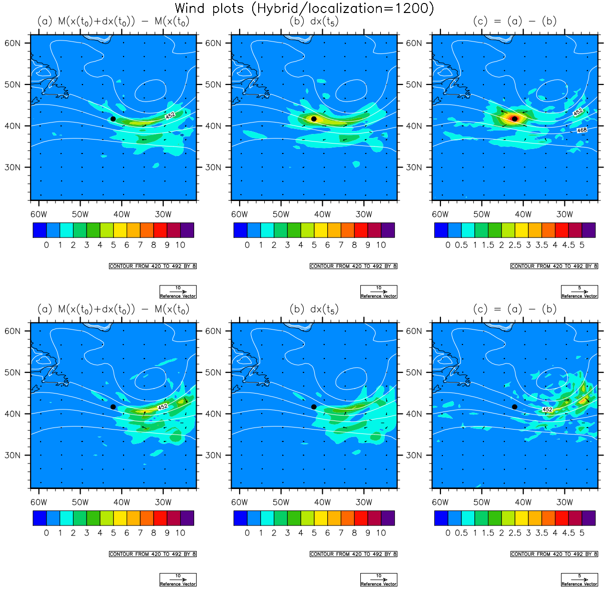
- For 4DVar, this measures the errors in the TL hypothesis ($\delta\mathbf{x}^a(t_5) = \tilde{\mathbf{M}}_{0\to5}\delta\mathbf{x}^a(t_0)$);
- For 4DEnVar, the “4D” error includes two sources:
  1. Errors from 3D approximation of climatological **B** throughout assimilation window;
  2. Errors from the localization not moving with the flow (degrading the time correlations of $\mathbf{P}^b$);
- Following plots show absolute errors.

Jet "4D" errors - $\beta_C = 0.0$, $\beta_e = 1.0$, $L = 1200km$

4DEnVar (top) and 4DVar (bottom) showing $M_{0\to5}(\mathbf{x}^b(t_0) + \delta\mathbf{x}^a(t_0)) - M_{0\to5}(\mathbf{x}^b(t_0))$ (left), $\delta\mathbf{x}^a(t_5)$ (middle) and "4D" error (right):



Jet "4D" errors - $\beta_C = 0.5$, $\beta_e = 0.5$, $L = 1200km$



Relative errors

- The relative errors are proportional to the size of the increment:

$$R.E. = \frac{M_{0\to5}(\mathbf{x}^b(t_0) + \delta\mathbf{x}^a(t_0)) - M_{0\to5}(\mathbf{x}^b(t_0)) - \delta\mathbf{x}^a(t_5)}{M_{0\to5}(\mathbf{x}^b(t_0) + \delta\mathbf{x}^a(t_0)) - M_{0\to5}(\mathbf{x}^b(t_0))} \tag{29}$$

- Global averaged relative errors calculated as RMS of gridpoints;
- With a pure ensemble, 4DVar R.E. = 0.54, 4DEnVar R.E. = 0.51;
- With a 50:50 hybrid, 4DVar R.E. = 0.66, 4DEnVar R.E. = 0.78.

Conclusion

- Single observation experiments used to compare 4DVar with 4DEnVar;
- Jet stream used as example;
- Methods differ only in their "4D" assimilation of observations;
- Calculation of "4D" errors used to compare the methods;
- With a pure ensemble covariance, 4DVar and 4DEnVar have similar errors;
- With a hybrid covariance, 4DVar performs much better than 4DEnVar;
- Superior performance of hybrid 4DVar based on 4D representation of climatological **B**;
- 4D **B** important for jet because $||\mathbf{P}^b||_2 \approx ||\mathbf{B}||_2$ and fast flow.
- Jet stream case fairly typical and effects a large area $\rightarrow$ it could explain the trials, but more evidence is needed!

Limitations

- Only 2 case studies selected;
- Results do not show the effect of multiple observations, which would cause further sampling error issues for the ensemble covariance;
- Trials are needed to gain statistically significant results.

Future work

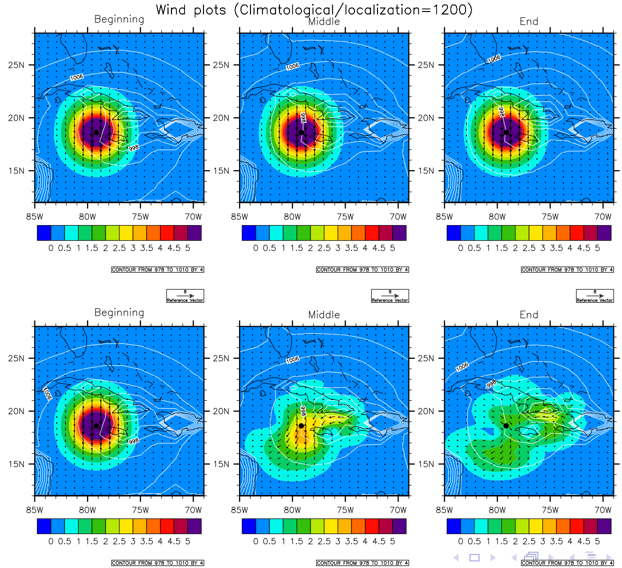
- Run trial of pure ensemble 4DVar vs pure ensemble 4DEnVar. If the results are similar, then this would suggest that the 3D climatological covariance is the main problem for 4DEnVar;
- Investigate ways to reduce the dependence of these methods on the climatological **B**. Some possible ways:
  1. Increase ensemble size (44 members is not enough!);
  2. Ensemble of 4DEnVars [Fairbairn et al., 2013];
  3. Improve flow-dependent representation of model error in ensemble (e.g. stochastic physics);
  4. Waveband localization (More severe localization of high frequency waves than low frequency waves).

Sandy observation

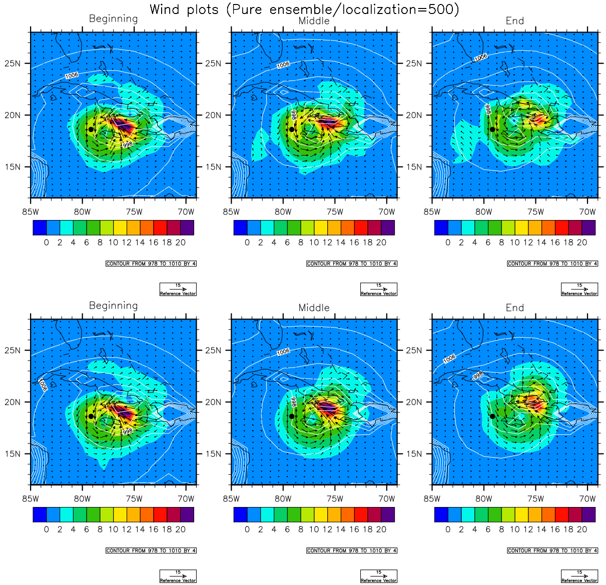
- - Single Southerly wind (v) observation with increment $+10m/s$;
- - Observation located at level 1 (surface), at coordinates 18N,79W.

Sandy: $\beta_C = 1.0$, $\beta_e = 0.0$

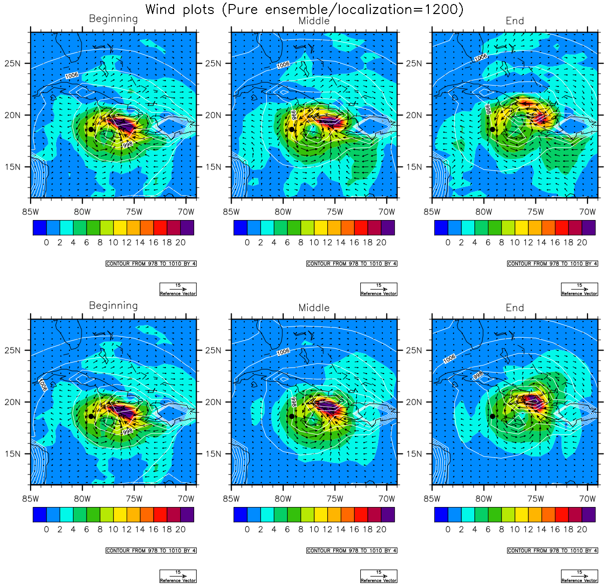
4DEnVar (top) and 4DVar (bottom) wind increments at beginning, middle and end of the assimilation window:



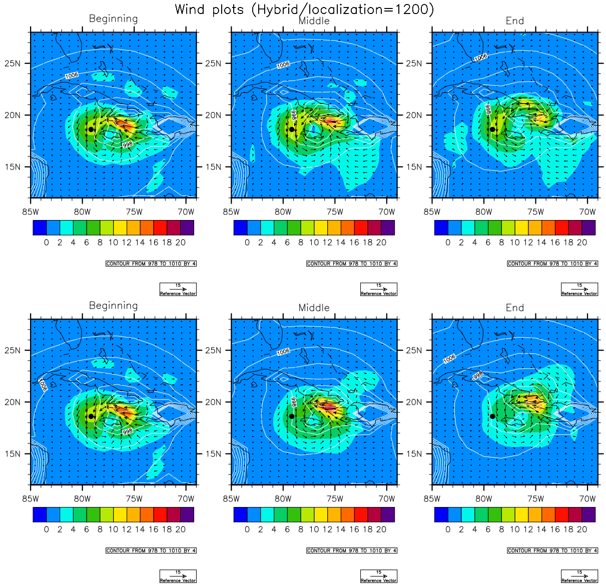
Sandy: $\beta_C = 0.0$, $\beta_e = 1.0$, $L = 500km$



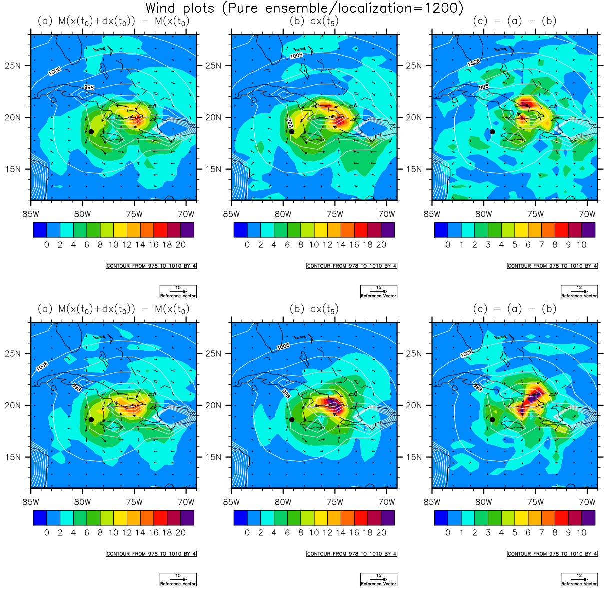
Sandy: $\beta_C = 0.0$, $\beta_e = 1.0$, $L = 1200km$



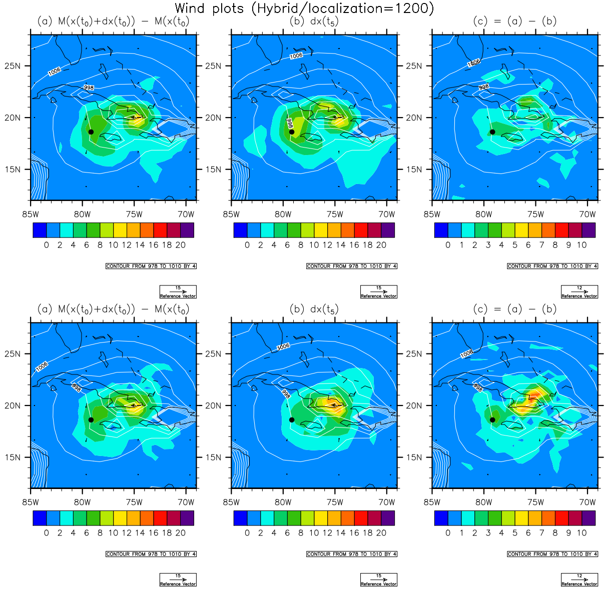
Sandy: $\beta_C = 0.5$, $\beta_e = 0.5$, $L = 1200km$



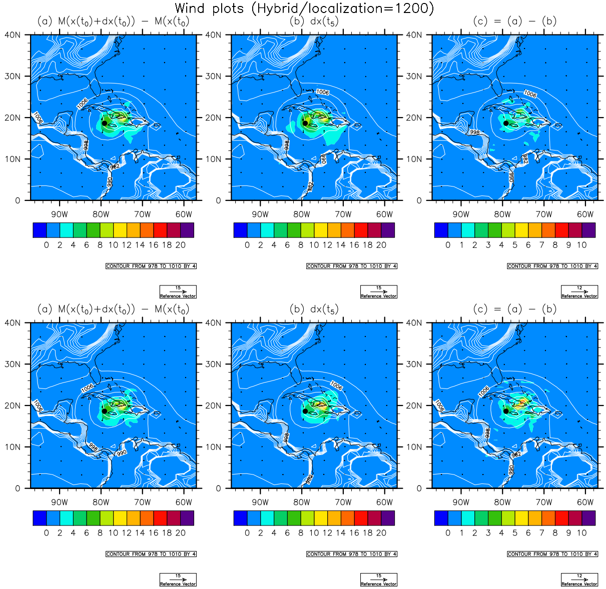
Sandy "4D" errors - $\beta_C = 0.0$, $\beta_e = 1.0$, $L = 1200km$



Sandy “4D” errors - $\beta_C = 0.5$, $\beta_e = 0.5$, $L = 1200km$



Sandy "4D" errors - $\beta_C = 0.5$, $\beta_e = 0.5$, $L = 1200km$



Relative errors

- The relative errors are proportional to the size of the increment:

$$R.E. = \frac{M_{0\to 5}(\mathbf{x}^b(t_0) + \delta\mathbf{x}^a(t_0)) - M_{0\to 5}(\mathbf{x}^b(t_0)) - \delta\mathbf{x}^a(t_5)}{M_{0\to 5}(\mathbf{x}^b(t_0) + \delta\mathbf{x}^a(t_0))} \tag{30}$$

- Global averaged relative errors calculated as RMS of gridpoints;
- With a pure ensemble, 4DVar R.E. $= 0.57$, 4DEnVar R.E. $= 0.69$;
- With a 50:50 hybrid, 4DVar R.E. $= 0.66$, 4DEnVar R.E. $= 0.75$;
- Hurricane Sandy case important but much more localized/more rare than jet stream case.

References

C. H. Bishop, B. Etherton, and S. Majumdar. Adaptive sampling with the ensemble transform Kalman filter. Part I: theoretical aspects. *Mon. Wea. Review*, 42:420–436, 2001.

A. M. Clayton, A. C. Lorenc, and D. M. Barker. Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Q.J.R. Meteorol. Soc*, submitted:16pp, 2012.

P. Courtier, J. Thepaut, and A. Hollingsworth. A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.*, 120:1389–1408, 1994.

D. Fairbairn, S. Pring, A.C.Lorenc, and I. Roulstone. A comparison of 4DVar with ensemble data assimilation methods. *Q. J. R. Meteorol. Soc.*, Accepted: 14pp, 2013. doi: 10.1002/qj.2135.

F. X. Le-Dimet and O. Talagrand. Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, 38A: 97–110, 1986.

A. Lorenc. Modelling of error covariances by 4D-Var data assimilation. *Q. J. R. Meteorol. Soc*, 129:3167–3182, 2003.

F. Rawlins, S. Ballard, K. Bovis, A. Clayton, D. Li, G. Inverarity, A. Lorenc, and T. Payne. The Met Office four-dimensional variational data assimilation scheme. *Q. J. R. Meteorol. Soc*, 133:347–362, 2007.

X. Wang, C. Snyder, and T. M. Hamill. On the theoretical equivalence of differently proposed ensemble–3DVAR hybrid analysis schemes. *Mon. Weather Rev.*, 135:222–227, 2007. doi: 10.1175/MWR3282.1.