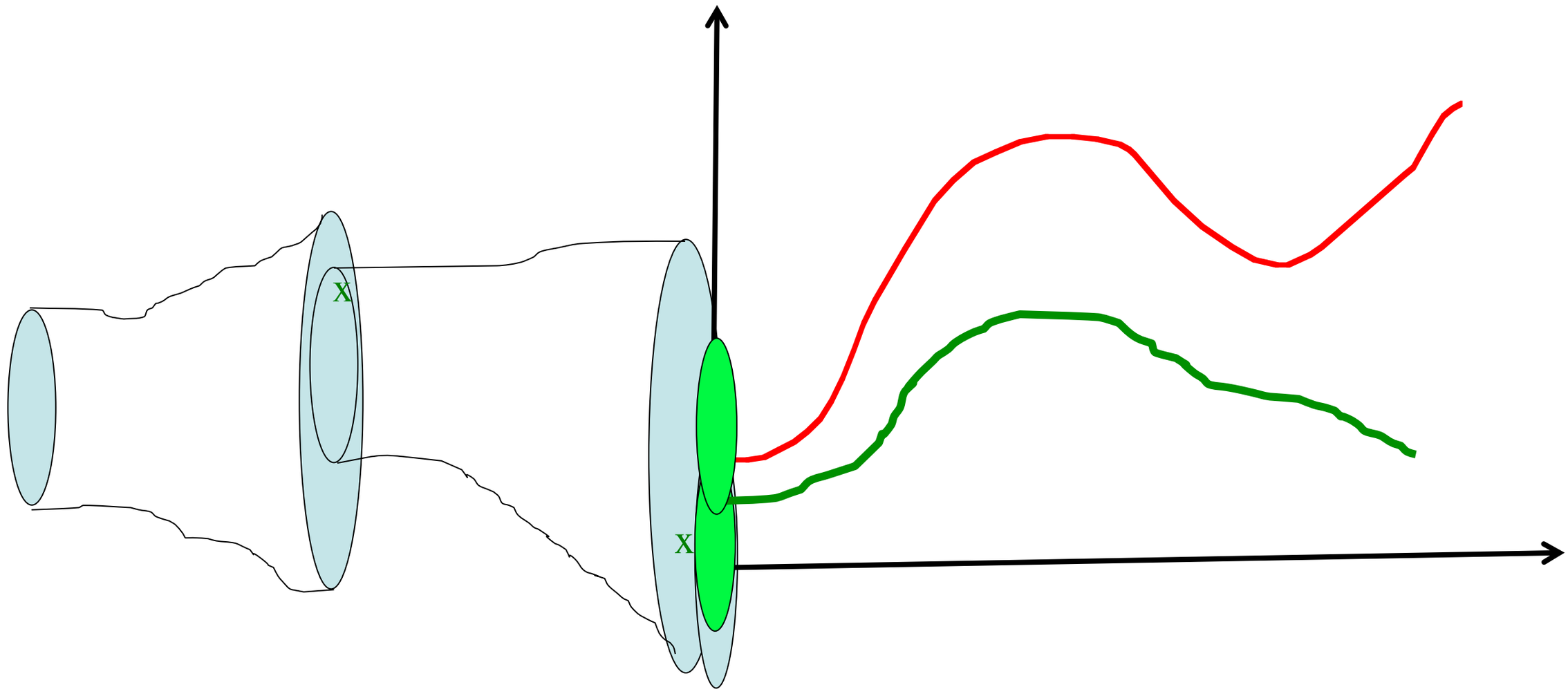
Hybrid Data Assimilation

Peter Jan van Leeuwen
Data Assimilation Research Centre (DARC)
National Centre for Earth Observation
http://www.met.reading.ac.uk/~darc/
University of Reading
p.j.vanleeuwen@reading.ac.uk

Operational Hybrid methods

1. ETKF-3/4DVar with control variable transform
2. Ens4DVar ‘ensemble of data assimilations’ EDA
3. 4DEnsVar

1. ETKF-3DVar



$$B = B_c$$

$$B = \beta_c^2 B_c + \beta_e^2 P^b \circ C_{loc}$$

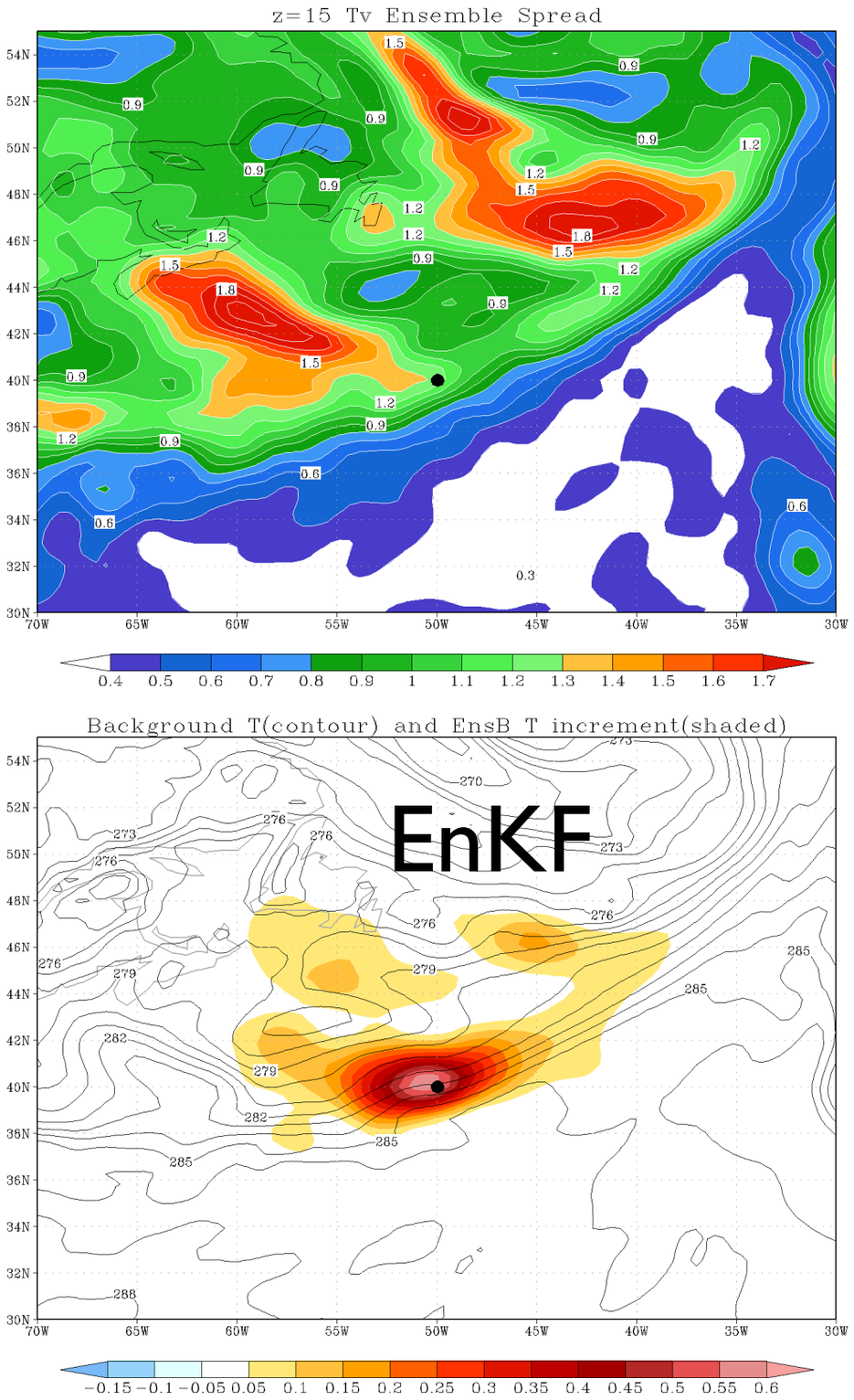
ETKF-3/4DVar algorithm

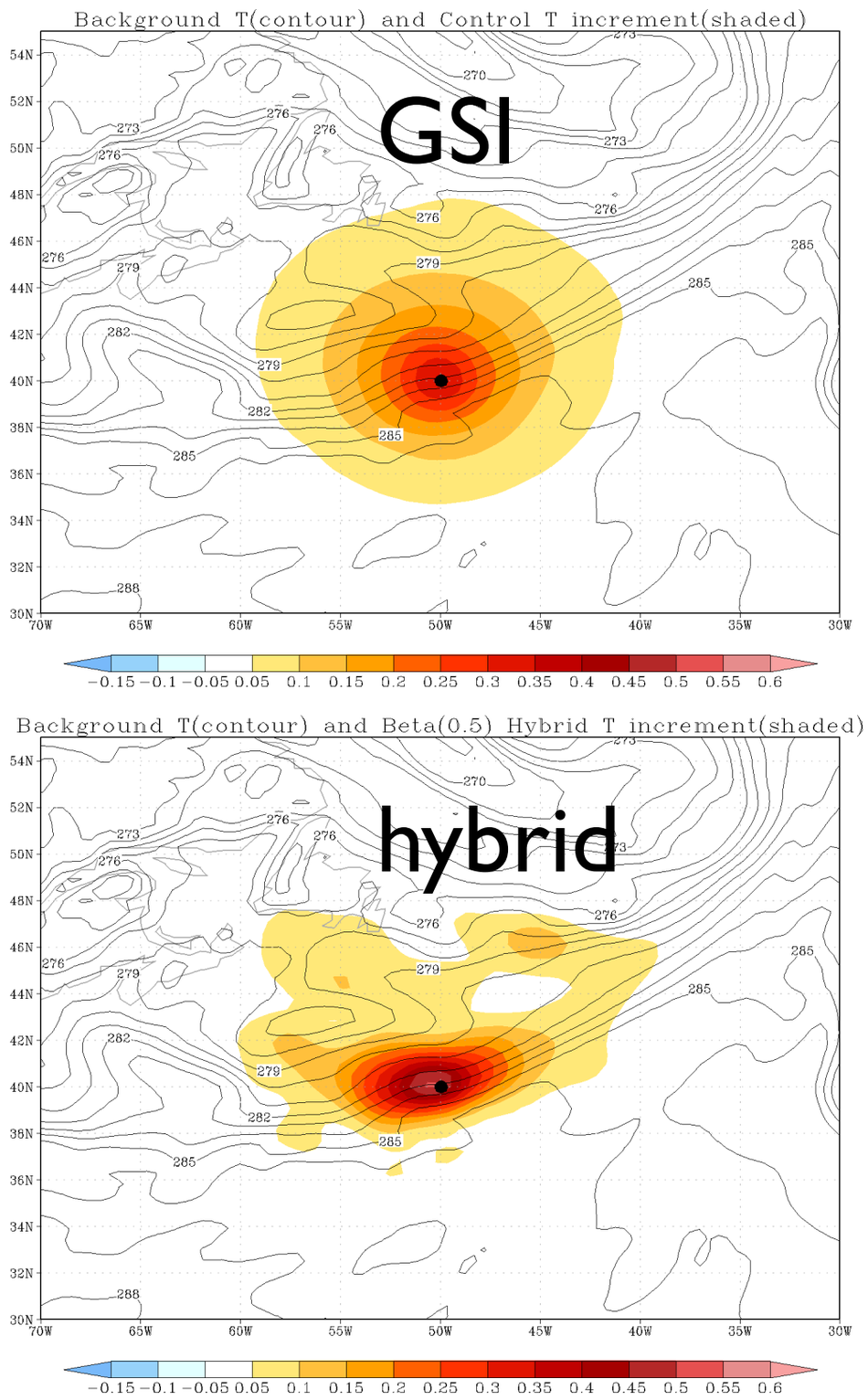
- Run ETKF to observation time (low resolution)
- Form new hybrid B matrix

$$B = \beta_c^2 B_c + \beta_e^2 P^b \circ C_{loc}$$

- Centre around 3/4Dvar forecast
- Precondition
- Calculate 3/4DVar solution(high resolution)
- Run ETKF to next observation time (low resolution)
- Etc.

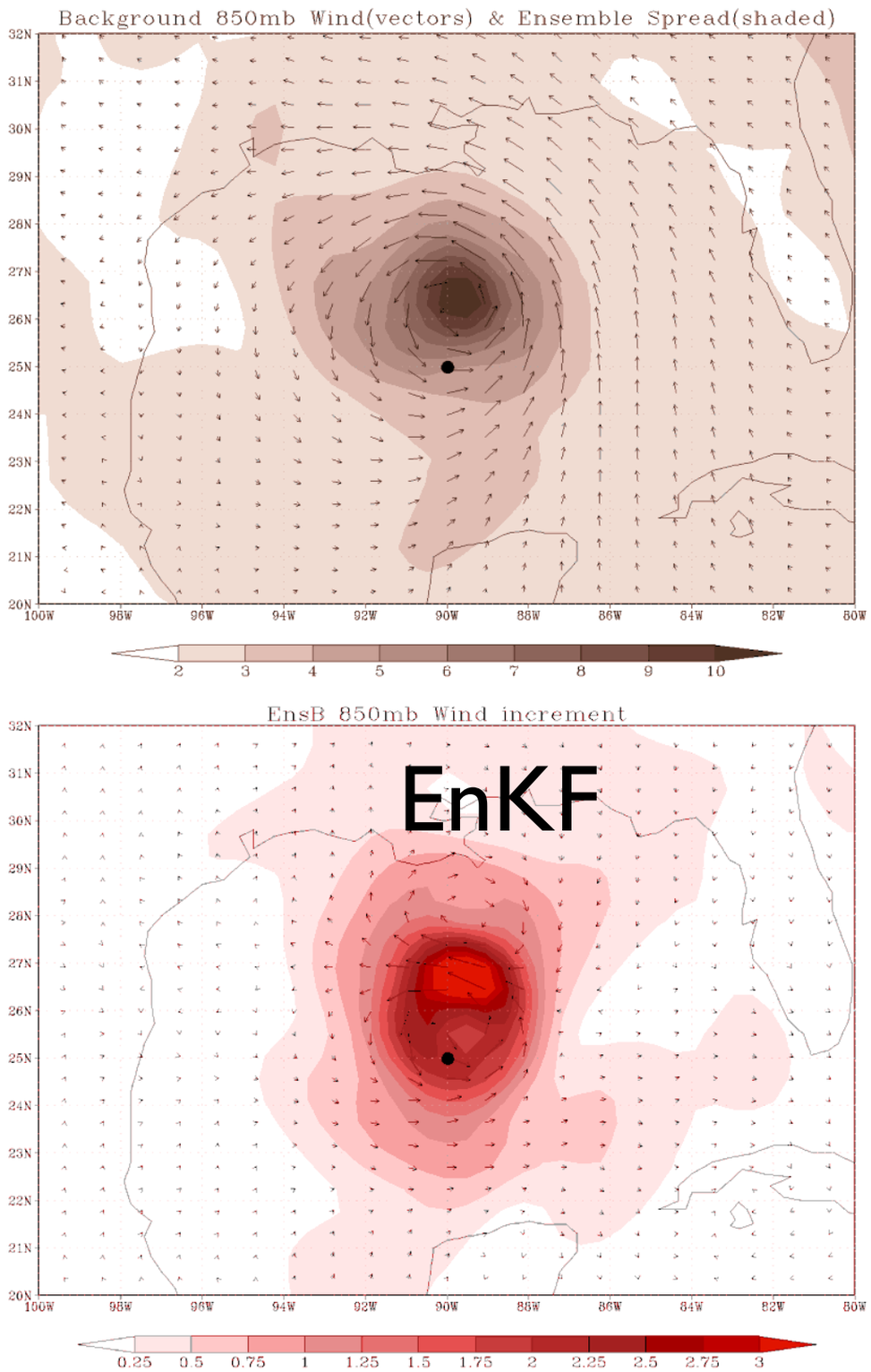
Single-observation increments

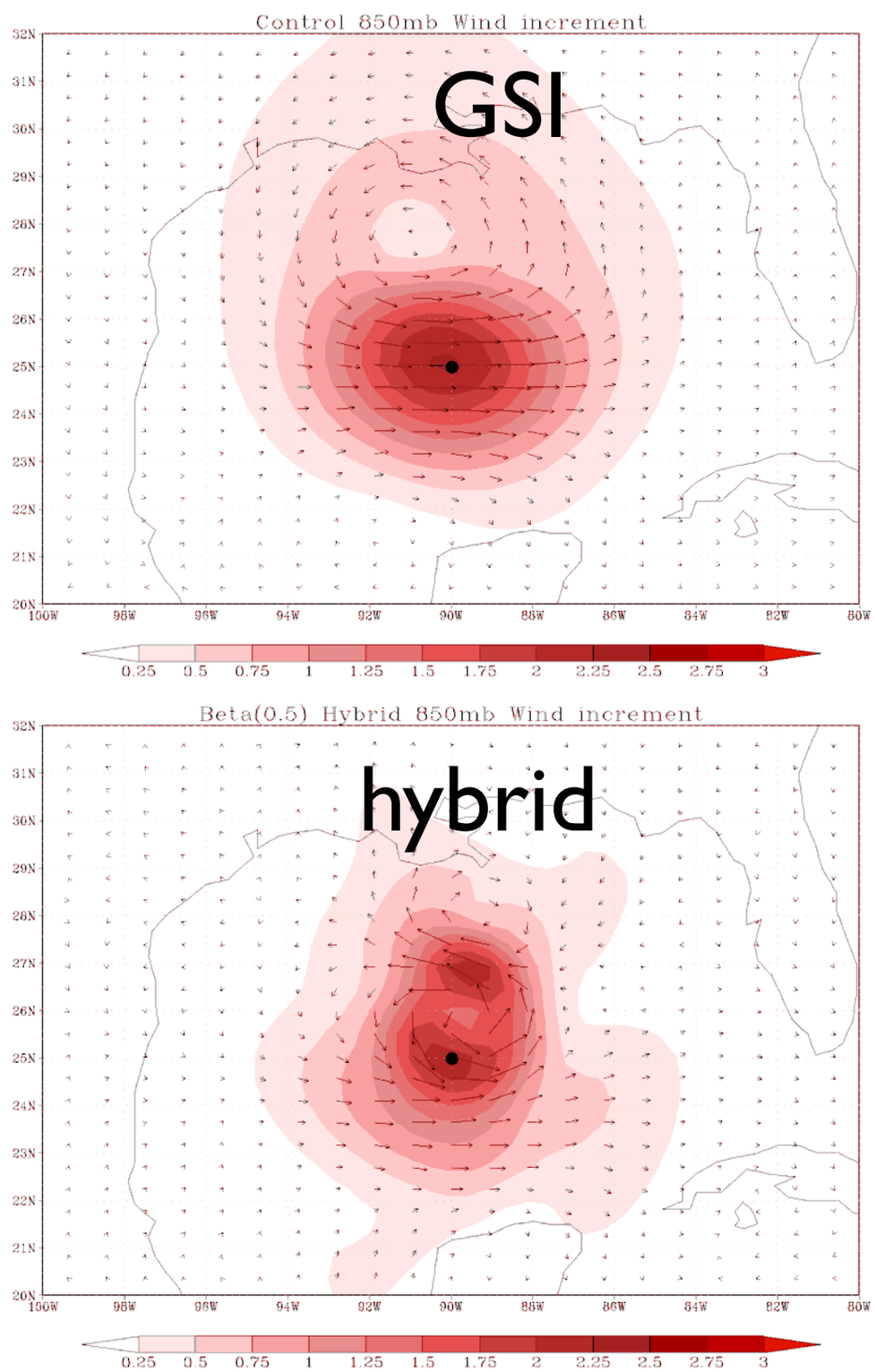




Single 850 hPa T_v observation (1K O-F, 1K error)

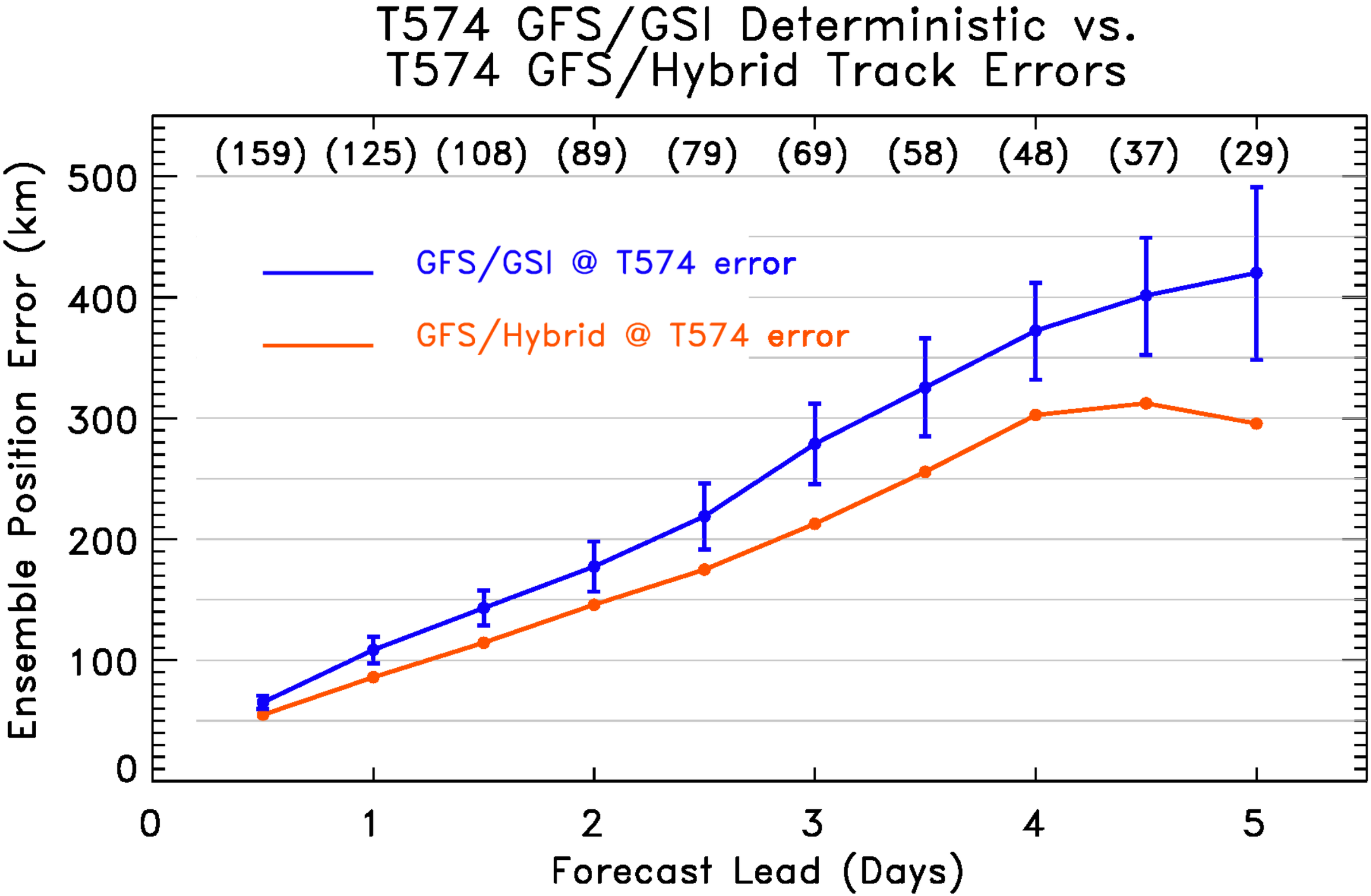
Single-observation increments





Single 850 hPa zonal wind observation (3 m/s O-F, 1m/s error) Hurricane Ike

The excellent news



Practical implementation

The 3DVar costfunction reads:

$$J(x) = \frac{1}{2}(x - x_b)^T B^{-1}(x - x_b) + \frac{1}{2}(y - H(x))^T R^{-1}(y - H(x))$$

with now: $$B = \beta_c^2 B_c + \beta_e^2 P^b \circ C_{loc}$$

Use alpha control variable transform:

$$\delta x = \beta_c B_c^{1/2} v + \beta_e X' \circ \alpha$$

Practical implementation

With this control variable transform

$$\delta x = \beta_c B_c^{1/2} v + \beta_e X' \circ \alpha$$

we find for the background term:

$$\frac{1}{2}\delta x^T B^{-1} \delta x =$$

$$\frac{1}{2}\left(\beta_c B_c^{1/2} v + \beta_e X' \circ \alpha\right)^T \left(\beta_c^2 B_c + \beta_e^2 P^b \circ C_{loc}\right)^{-1} \left(\beta_c B_c^{1/2} v + \beta_e X' \circ \alpha\right) =$$

$$= \frac{1}{2} v^T v + \frac{1}{2} \alpha^T C_{loc}^{-1} \alpha$$

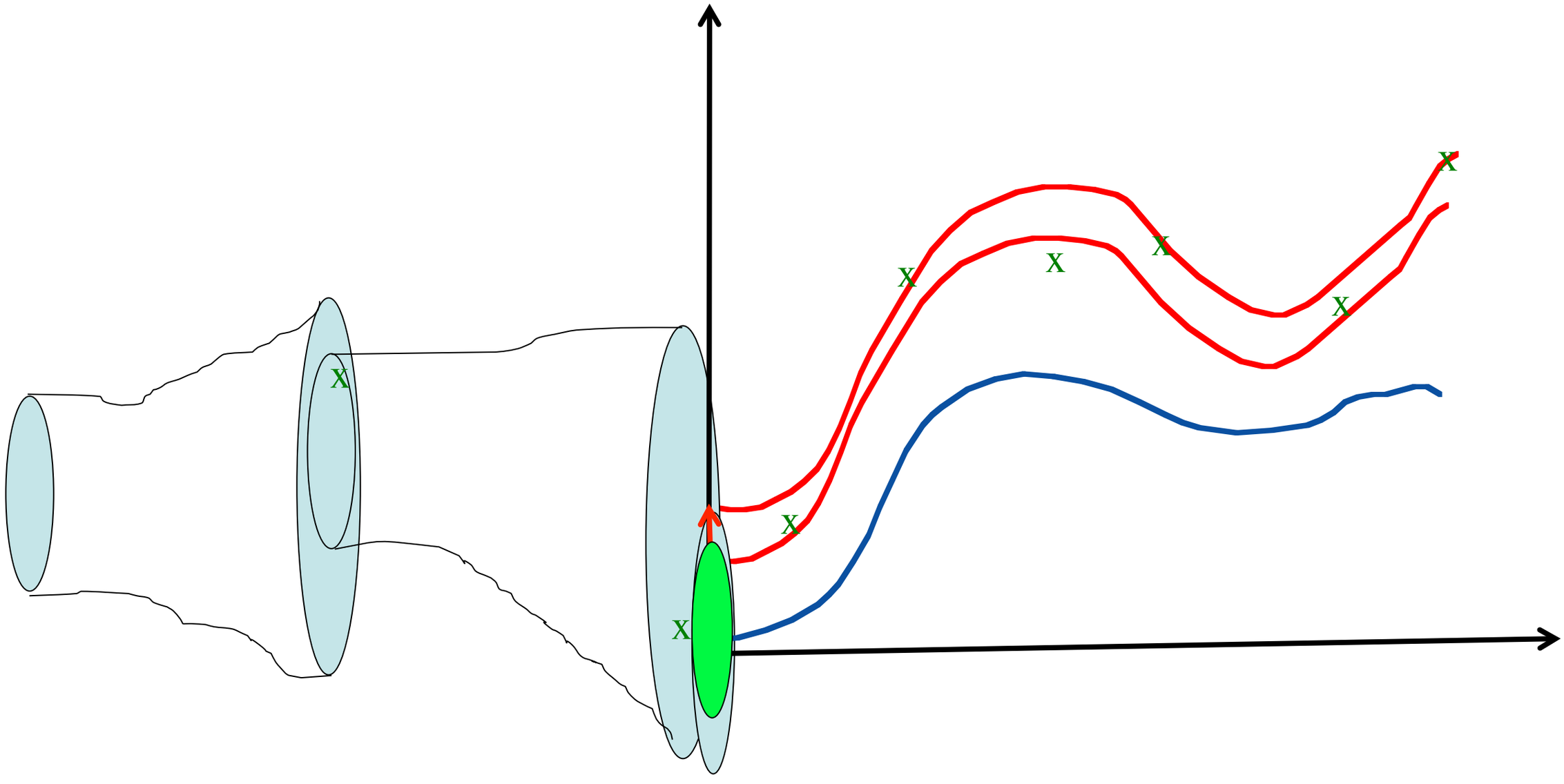
Meaning of control-variable transform

The control-variable transform can be written as:

$$\delta x = \beta_c B_c^{1/2} v + \beta_e X' \circ \alpha = \delta x_c + \delta x_e$$

- The increment is now a weighted sum of the static B_c component and the flow-dependent, ensemble based B_e
- The flow-dependent increment is a linear combination of ensemble perturbations X', modulated by the α fields
- If the α fields were homogeneous δx_e only spans Nens-1 degrees of freedom;
- We allow for varying α fields, which effectively increases the degrees of freedom
- C_{loc} is the localisation matrix for the flow-flow-dependent increments: it controls the spatial variation of α

ETKF-4DVar, e.g. Met Office



$$B = \beta_c^2 B_c + \beta_e^2 P^b \circ C_{loc}$$

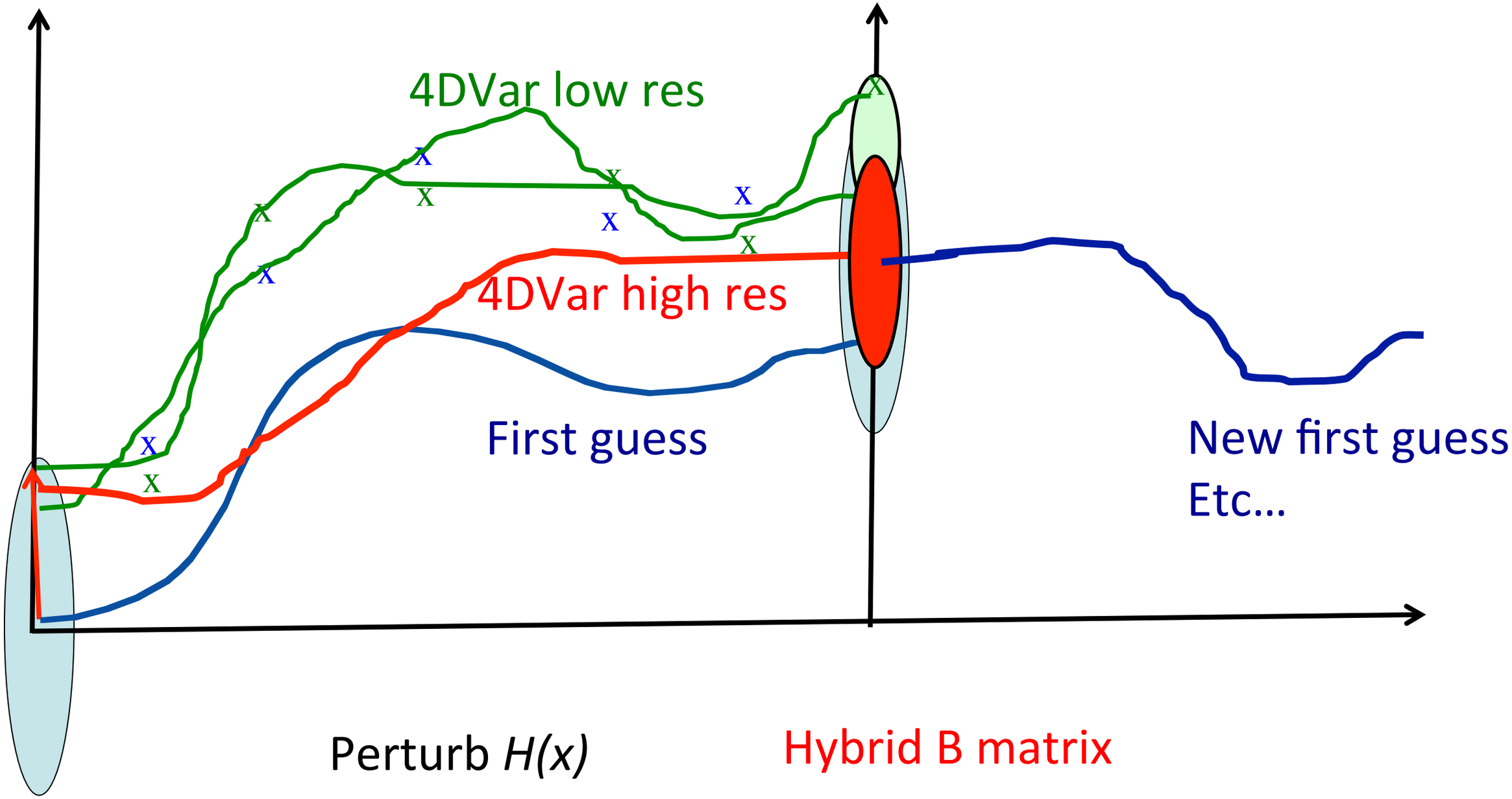
Practical implementation ETKF-4DVar

- Up to now this has only been implemented in strong-constraint 4Dvar
- Implementation similar to EnKF-3DVar.
- All usual tricks like preconditioning and incremental 4DVar are fully explored.

ETKF-4DVar

1) Prior is assumed Gaussian
2) B matrix is informed by previous observations, flow dependent B matrix
3) Error estimate from the EnKF ensemble
4) Natural ensemble prediction system
5) Possibility of getting stuck in local minima
6) Extra linearity by replacing ensemble mean by 4Dvar solution

2. EnsVars
‘ensemble of data-assimilations’ EDA



EnsVars algorithm

- Perturb *H(x)* in 4Dvar costfunction N_e times ('perturb observations')
- Solve 4DVar for each of these costfunctions (low resolution)
- Solve high-res 4Dvar over same window
- Use ensemble of 4Dvar ensemble forecasts to form hybrid B matrix for 4DVars of next time window.
- All usual tricks like preconditioning and incremental 4DVar are fully explored.

Characteristics of EnsVars

1) Prior is assumed Gaussian
2) B matrix is informed by previous observations, flow dependent B matrix
3) Error estimate from the low-res 4DVar ensemble
4) Natural ensemble prediction system
5) Possibility of getting stuck in local minima
6) Extra linearity by replacing ensemble mean by high-res 4Dvar solution

Conclusions hybrid methods

Variational methods are the most popular data-assimilation scheme, but use static B matrix and need adjoint.

Ensemble-Kalman Filters are used to make B flow dependent: ETKF-4DVar, EnsVars (EDA), 4DEnsVar. The latter doesn't need an adjoint.

Strong-constraint 4DEnsVar, in which the space-time covariances are generated by an ensemble run has issues with localisation.

Weak-constraint 4DEnsVar partially solves localisation issue.

MORE WORK IS NEEDED !