Ensemble Kalman Filter and Hybrid methods

Massimo Bonavita

massimo.bonavita@ecmwf.int

Contributions from Mats Hamrud, Mike Fisher

Outline

- Standard Kalman Filter theory
- Kalman Filters for large dimensional systems
- Approximate Kalman Filters: The Ensemble Kalman Filter and 4D-Var
- Hybrid Variational–EnKF algorithms

Standard Kalman Filter

- In a previous lecture it was shown that the linear, unbiased analysis equation had the form:

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}\,(\mathbf{y} - \mathbf{H}(\mathbf{x}^b))$$

$\mathbf{x}^a$ = analysis state; $\mathbf{x}^b$ = background state;

$\mathbf{y}$ = observations; $\mathbf{H}(\mathbf{x}^b)$ = model equivalents of the observations

- It was also shown that the best linear unbiased analysis (BLUE; best here means the analysis that has the minimum error variance) is achieved when the matrix **K** (Kalman Gain Matrix) has the form:

$$\mathbf{K} = \mathbf{P}^b\,\mathbf{H}^T(\mathbf{H}\,\mathbf{P}^b\,\mathbf{H}^T + \mathbf{R})^{-1} = ((\mathbf{P}^b)^{-1} + \mathbf{H}^T\,\mathbf{R}^{-1}\,\mathbf{H})^{-1}\,\mathbf{H}^T\,\mathbf{R}^{-1}$$

$\mathbf{P}^b$ = covariance matrix of the background error

$\mathbf{R}$ = covariance matrix of the observation error

- An expression for the covariance matrix of the analysis error was also found:

Standard Kalman Filter

- An expression for the covariance matrix of the analysis error was also found:

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{KH})\mathbf{P}^b (\mathbf{I} - \mathbf{KH})^T + \mathbf{KRK}^T$$

- In NWP applications of data assimilation we want to update our estimate of the state and its uncertainty at later times, as new observations come in: we want to cycle the analysis

- For each analysis in this cycle we require a background $\mathbf{x}^b_t$ (i.e. a prior estimate of the state valid at time t)

- Usually, our best prior estimate of the state at time t is given by a forecast from the preceding analysis at t-1 (the "background"):

$$\mathbf{x}^b_t = \mathbf{M}(\mathbf{x}^a_{t-1})$$

- What is the error covariance matrix (=> the uncertainty) associated with this background?

Standard Kalman Filter

- What is the error covariance matrix associated with this background?

$$\mathbf{x}^b = \mathbf{M}(\mathbf{x}^a_{t-1})$$

- Subtract the true state $\mathbf{x}^*$ from both sides of the equation:

$$\mathbf{x}^b - \mathbf{x}^*_t = \boldsymbol{\varepsilon}^b_t = \mathbf{M}(\mathbf{x}^a_{t-1}) - \mathbf{x}^*_t$$

- Since $\mathbf{x}^a_{t-1} = \mathbf{x}^*_{t-1} + \boldsymbol{\varepsilon}^a_{t-1}$ we have:

$$\boldsymbol{\varepsilon}^b_t = \mathbf{M}(\mathbf{x}^*_{t-1} + \boldsymbol{\varepsilon}^a_{t-1}) - \mathbf{x}^*_t \approx$$

$$\mathbf{M}(\mathbf{x}^*_{t-1}) + \mathbf{M}\boldsymbol{\varepsilon}^a_{t-1} - \mathbf{x}^*_t =$$

$$\mathbf{M}\boldsymbol{\varepsilon}^a_{k-1} + \eta_k$$

- Here we have defined the model error $\eta_k = \mathbf{M}(\mathbf{x}^*_{t-1}) - \mathbf{x}^*_t$

- We will also assume that no systematic errors are present in our system (!):
  $< \boldsymbol{\varepsilon}^a > = < \eta > = 0 \Rightarrow < \boldsymbol{\varepsilon}^b > = 0$

- The background error covariance matrix will then be given by:

Standard Kalman Filter

$$<\boldsymbol{\varepsilon}^b_t\,(\boldsymbol{\varepsilon}^b_t)^T> = \mathbf{P}^b_t = <(\mathbf{M}\boldsymbol{\varepsilon}^a_{t-1} + \eta_k)\,(\mathbf{M}\boldsymbol{\varepsilon}^a_{t-1} + \eta_k)^T> =$$

$$\mathbf{M}<\boldsymbol{\varepsilon}^a_{t-1}\,(\boldsymbol{\varepsilon}^a_{t-1})^T>\,\mathbf{M}^T + <\eta_k\,(\eta_k)^T> =$$

$$\mathbf{M}\,\mathbf{P}^a_{t-1}\,\mathbf{M}^T + \mathbf{Q}_t$$

- Here we have assumed $< \boldsymbol{\varepsilon}^a_{t-1}\,(\eta_t)^T> = 0$ and defined the model error covariance matrix $\mathbf{Q}_t = <\eta_t\,(\eta_t)^T>$
- We now have all the equations necessary to propagate and update both the state and its error estimates:

$$\mathbf{x}^b_t = \mathbf{M}(\mathbf{x}^a_{t-1})$$

$$\mathbf{P}^b_t = \mathbf{M}\,\mathbf{P}^a_{t-1}\,\mathbf{M}^T + \mathbf{Q}_t$$

Propagation

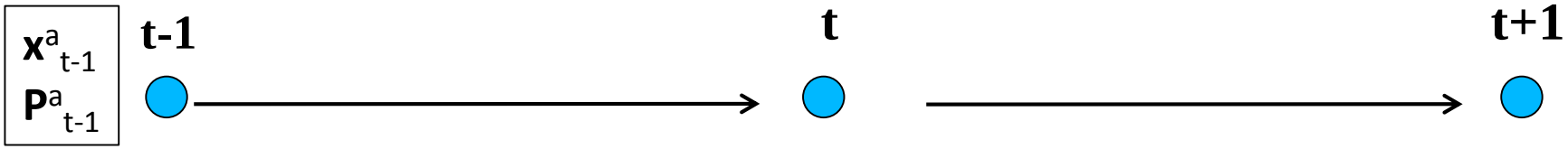
$$\mathbf{K} = \mathbf{P}^b\,\mathbf{H}^T(\mathbf{H}\,\mathbf{P}^b\,\mathbf{H}^T + \mathbf{R})^{-1}$$

$$\mathbf{x}^a_t = \mathbf{x}^b_t + \mathbf{K}\,(\mathbf{y} - \mathbf{H}(\mathbf{x}^b_t))$$

$$\mathbf{P}^a_t = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^b_t\,(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{R}\mathbf{K}^T$$

Update

Standard Kalman Filter



1. Predict the state ahead

$$\mathbf{x}^b_t = \mathbf{M}(\mathbf{x}^a_{t-1})$$

2. Predict the state error cov.

$$\mathbf{P}^b_t = \mathbf{M}\,\mathbf{P}^a_{t-1}\,\mathbf{M}^T + \mathbf{Q}_t$$

Propagation

3. Compute the Kalman Gain

$$\mathbf{K} = \mathbf{P}^b\,\mathbf{H}^T(\mathbf{H}\,\mathbf{P}^b\,\mathbf{H}^T + \mathbf{R})^{-1}$$

4. Update state estimate

$$\mathbf{x}^a_t = \mathbf{x}^b_t + \mathbf{K}\,(\mathbf{y} - \mathbf{H}(\mathbf{x}^b_t))$$

5. Update state error estimate

$$\mathbf{P}^a_t = (\mathbf{I} - \mathbf{KH})\mathbf{P}^b_t\,(\mathbf{I} - \mathbf{KH})^T + \mathbf{KRK}^T$$

Update

New Observations

1. Predict the state ahead

$$\mathbf{x}^b_{t+1} = \mathbf{M}(\mathbf{x}^a_t)$$

2. Predict the state error cov.

$$\mathbf{P}^b_{t+1} = \mathbf{M}\,\mathbf{P}^a_t\,\mathbf{M}^T + \mathbf{Q}_{t+1}$$

Propagation

Standard Kalman Filter

- Under the assumption that the model **M** and the observation operator **H** are linear operators (i.e., they do not depend on $\mathbf{x}^b$), the Kalman Filter produces an optimal sequence of analysis

- The KF analysis $\mathbf{x}^a_t$ is the best (minimum variance) estimate of the state at time t, given $\mathbf{x}^b_0$ and all observations up to time t ($\mathbf{y}_0,\mathbf{y}_1,\ldots,\mathbf{y}_t$).

- Note that Gaussianity of errors is not required. If errors are Gaussian the Kalman Filter provides the exact conditional probability estimate, i.e. $p(\mathbf{x}^a_t | \mathbf{x}^b_0; \mathbf{y}_0,\mathbf{y}_1,\ldots,\mathbf{y}_t)$. This also implies that if errors are Gaussian then the state estimated with the KF is also the most likely state (the mode of the pdf).

Kalman Filters for Large Dimensional Systems

- The Kalman Filter is unfeasible for large dimensional systems
- The size N of the analysis/background state in the ECMWF 4DVar is $O(10^8)$: the KF requires us to store and evolve in time state covariance matrices ($\mathbf{P}^{a/b}$) of O(NxN)
  - The World's fastest computers can sustain ~ 10^{15} operations per second
  - An efficient implementation of matrix multiplication of two 10^8x10^8 matrices requires ~10^{22} operations ($O(N^{2.8})$): about 4 months on the fastest computer!
  - Evaluating $\mathbf{P}^b_t = \mathbf{M}\,\mathbf{P}^a_{t-1}\,\mathbf{M}^T + \mathbf{Q}_k$ requires $2*N \approx 2*10^8$ model integrations!
- A range of approximate Kalman Filters has been developed for use with large-dimensional systems.
- All of these methods rely on a low-rank approximation of the covariance matrices of background and analysis error.

Kalman Filters for Large Dimensional Systems

- Main assumption: $\mathbf{P}^b_k$ has rank M<<N (e.g. M~100).
- Then we can write $\mathbf{P}^b = \mathbf{X}^b(\mathbf{X}^b)^T$, where $\mathbf{X}^b_k$ is N x M.
- The Kalman Gain then becomes:

$$\mathbf{K} = \mathbf{P}^b \mathbf{H}^T(\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R})^{-1} =$$

$$\mathbf{X}^b(\mathbf{X}^b)^T\mathbf{H}^T(\mathbf{H} \mathbf{X}^b(\mathbf{X}^b)^T \mathbf{H}^T + \mathbf{R})^{-1} =$$

$$\mathbf{X}^b (\mathbf{H}\mathbf{X}^b)^T(\mathbf{H} \mathbf{X}^b(\mathbf{H}\mathbf{X}^b)^T + \mathbf{R})^{-1}$$

- Note that, to evaluate **K**, we apply **H** to the M columns of $\mathbf{X}^b$ rather than to the N columns of $\mathbf{P}^b$.
- The N x N matrices $\mathbf{P}^{a/b}$ have been eliminated from the computation! In their place we have N x M ($\mathbf{X}^b$) and L x M ($\mathbf{H}\mathbf{X}^b$) matrices (L = number of observations)

Kalman Filters for Large Dimensional Systems

- Similar derivation can be done for $\mathbf{P}^a$:

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{KH})\mathbf{P}^b\,(\mathbf{I} - \mathbf{KH})^T + \mathbf{KRK}^T =$$

$$= (\mathbf{I} - \mathbf{KH})\mathbf{P}^b = (\mathbf{I} - \mathbf{KH})\,\mathbf{X}^b(\mathbf{X}^b)^T =$$

$$\mathbf{X}^b(\mathbf{X}^b)^T - \mathbf{KH}\,\mathbf{X}^b(\mathbf{X}^b)^T$$

- Both terms in this expression for $\mathbf{P}^a$ contain an initial $\mathbf{X}^b$ and a final $(\mathbf{X}^b)^T$ so that $\mathbf{P}^a = \mathbf{X}^b\mathbf{W}(\mathbf{X}^b)^T$ for some MxM matrix $\mathbf{W}$

- Finally the background covariance matrix:

$$\mathbf{P}^b = \mathbf{M}\,\mathbf{P}^a\mathbf{M}^T + \mathbf{Q} = \mathbf{M}\,\mathbf{X}^b\mathbf{W}(\mathbf{X}^b)^T\mathbf{M}^T + \mathbf{Q} =$$

$$\mathbf{M}\,\mathbf{X}^b\mathbf{W}\,(\mathbf{M}\mathbf{X}^b)^T + \mathbf{Q}$$

- This requires only(!) M (O(100)) integrations of the linearized model $\mathbf{M}$

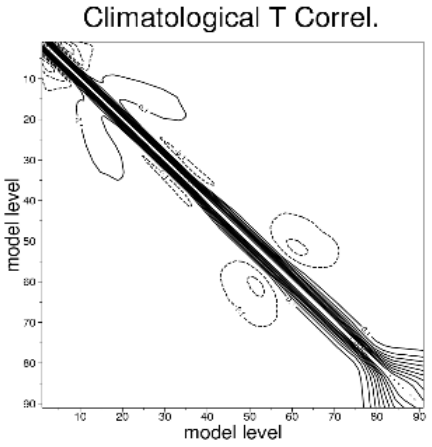
Kalman Filters for Large Dimensional Systems

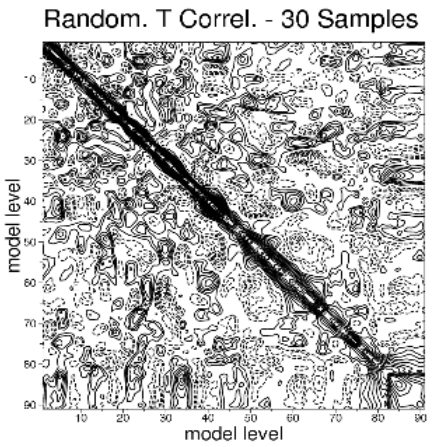
- The algorithm described above is called Reduced-rank Kalman Filter
- These huge gains in computational cost come at a price! ☹
- The analysis increment is a linear combination of the columns of $\mathbf{X}^b_k$:

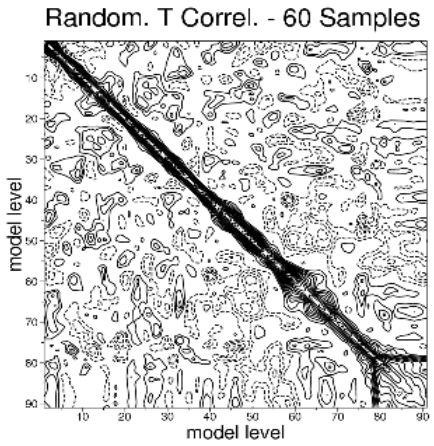
$$\mathbf{x}^a - \mathbf{x}^b = \mathbf{K}\,(\mathbf{y} - \mathcal{H}(\mathbf{x}^b)) = \mathbf{X}^b\,(\mathbf{H}\mathbf{X}^b)^T\,((\mathbf{H}\mathbf{X}^b)(\mathbf{H}\mathbf{X}^b)^T + \mathbf{R})^{-1}\,(\mathbf{y} - \mathcal{H}(\mathbf{x}^b))$$

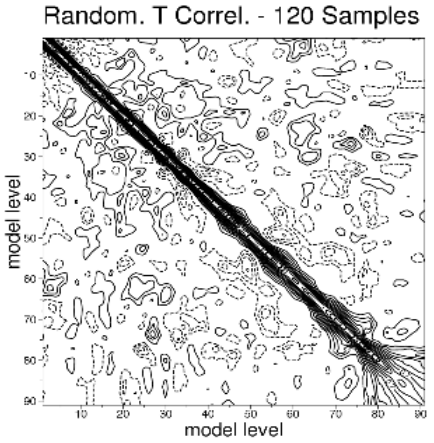
- The analysis increments are formed as a linear combination of the columns of $\mathbf{X}^b$: they are confined to the subspace spanned by $\mathbf{X}^b$, which has at most rank $M << N$.
- This severe reduction in rank has two main effects:
  1. There are too few degrees of freedom available to fit the $\approx 10^6$ observations: the analysis is too "smooth";
  2. The low-rank approximations of the covariance matrices suffer from spurious long-distance correlations.

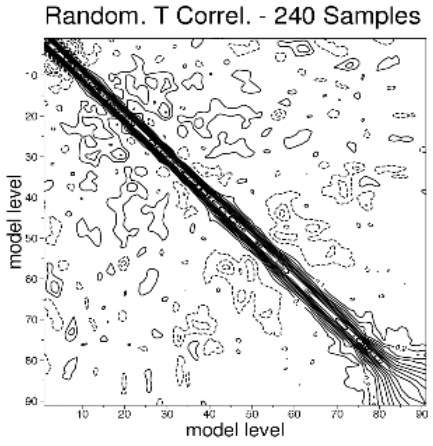
Random estimates of temperature background error correlation matrix for different ensemble sizes

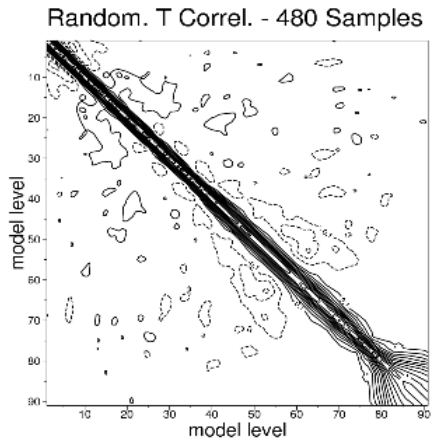












Kalman Filters for Large Dimensional Systems

- There are two ways around the rank deficiency problem:

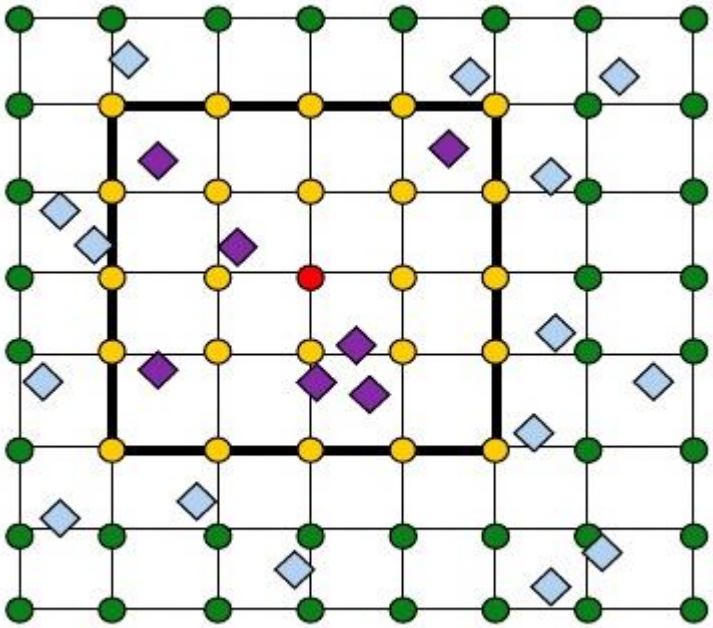
  1. Domain localization (e.g. Houtekamer and Mitchell, 1998; Ott *et al.* 2004);

- Domain localization solves the analysis equations independently for each grid point, or for each of a set of regions covering the domain.
- Each analysis uses only observations that are local to the grid point (or region) and the observations are usually weighted according to their distance from the analysed grid point (e.g., Hunt *et al.*, 2007)
- This guarantees that the analysis at each grid point (or region) is not influenced by distant observations.
- The method acts to vastly increase the dimension of the sub-space in which the analysis increment is constructed because each grid point is updated by a different linear combination of ensemble perturbations
- However, performing independent analyses for each region can lead to difficulties in the analysis of the large scales and in producing balanced analyses.

Kalman Filters for Large Dimensional Systems

1. Domain localization

- Analysed grid point
- Local observations



Kalman Filters for Large Dimensional Systems

- The other way around the rank deficiency problem:

  2. Covariance localization (e.g. Houtekamer and Mitchell, 2001).

- Covariance localization is performed by element wise (Schur) multiplication of the error covariance matrices with a predefined correlation matrix representing a decaying function of distance (vertical and horizontal).

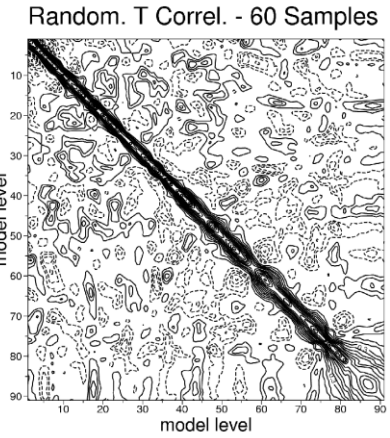
$$\mathbf{P}^b \dashrightarrow \rho_L \circ \mathbf{P}^b$$

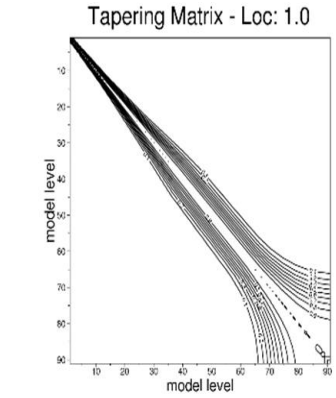
- In this way spurious long range correlations in $\mathbf{P}^b$ are suppressed.

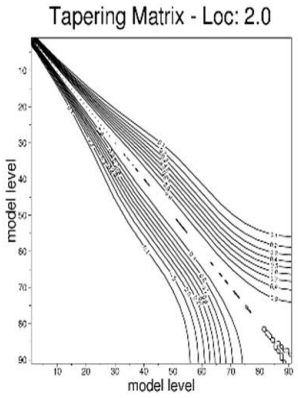
- As for domain localization, the method acts to vastly increase the dimension of the sub-space in which the analysis increment is constructed.

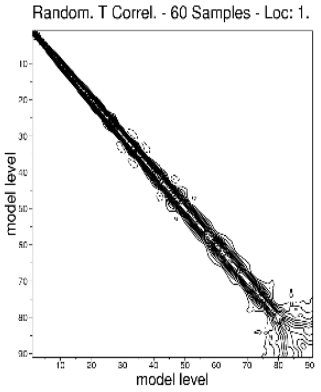
- Choosing the product function is non-trivial. It is easy to modify $\mathbf{P}^b$ in undesirable ways. In particular, balance relationships may be adversely affected.

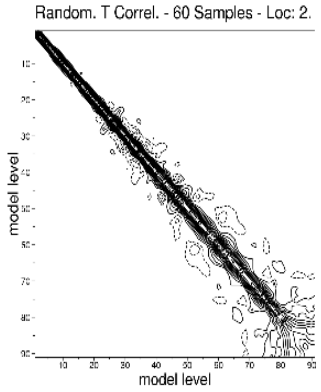
- Standard Error of sample correlation ≈ $(1-\rho^2)/\sqrt{(N_{ens}-1)}$
- for small ρ, N_{ens} it becomes >= ρ;
- since ρ -> 0 for large horiz./vert. distances apply distance based covariance localization on the sample $\mathbf{P^f}$

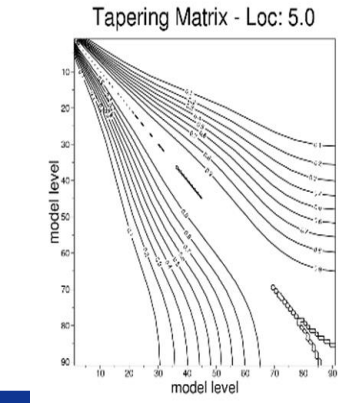


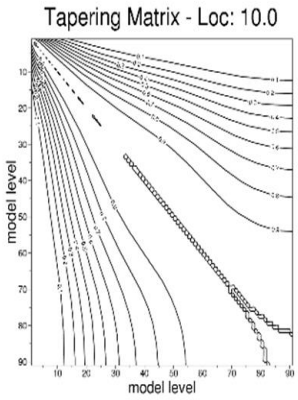


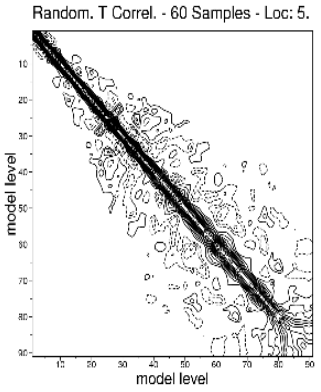


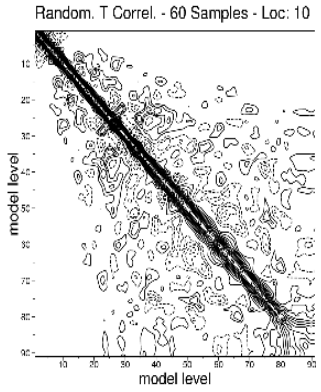












Kalman Filters for Large Dimensional Systems

- Domain/Covariance localization is a practical necessity for using the KF in large dimensional systems
- Finding the right amount of localization is an (expensive) tuning exercise: a good trade-off needs to be found between computational effort, sampling error and imbalance error
- Finding the “optimal” localization scales as functions of the system characteristics is an area of current active research (e.g., Flowerdew, 2015; Periáñez et al., 2014; Menetrier et al., 2014)

Ensemble Kalman Filters

- Ensemble Kalman Filters (EnKF, Evensen, 1994; Burgers et al., 1998) are Monte Carlo implementations of the reduced rank KF
- In EnKF error covariances are constructed as sample covariances from an ensemble of background/analysis fields, i.e.:

$$\mathbf{P}^{a/b} = \frac{1}{M-1}\sum_m (\mathbf{x}^b{}_m - \langle \mathbf{x}^b{}_m \rangle)\,(\mathbf{x}^b{}_m - \langle \mathbf{x}^b{}_m \rangle)^T =$$

$$= \mathbf{X}^b (\mathbf{X}^b)^T$$

- $\mathbf{X}^b$ is the N x M matrix of background perturbations, i.e.:

$$\mathbf{X}^b = \frac{1}{\sqrt{M-1}} ((\mathbf{x}^b{}_1 - \langle \mathbf{x}^b \rangle), (\mathbf{x}^b{}_2 - \langle \mathbf{x}^b \rangle), .., (\mathbf{x}^b{}_M - \langle \mathbf{x}^b \rangle))$$

- Note that the full covariance matrix is never formed explicitly: The error covariances are usually computed locally for each grid point in the M x M ensemble space

Ensemble Kalman Filters

- In the standard KF the error covariances are explicitly propagated using the tangent linear and adjoint of the model and observation operators, i.e.:

$$\mathbf{K} = \mathbf{P}^b \mathbf{H}^T (\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R})^{-1}$$

$$\mathbf{P}^b = \mathbf{M} \mathbf{P}^a \mathbf{M}^T + \mathbf{Q}$$

- In the EnKF the error covariances are sampled from the ensemble forecasts and the huge matrix $\mathbf{P}^b$ is never explicitly formed:

$$\mathbf{P}^b\mathbf{H}^T = \mathbf{X}^b(\mathbf{X}^b)^T\mathbf{H}^T = \mathbf{X}^b(\mathbf{H}\mathbf{X}^b)^T =$$

$$\frac{1}{M-1}\Sigma_m(\mathbf{x}^b{}_m - \langle \mathbf{x}^b{}_m \rangle)\,(\mathbf{x}^b{}_m - \langle \mathcal{H}(\mathbf{x}^b{}_m) \rangle)^T$$

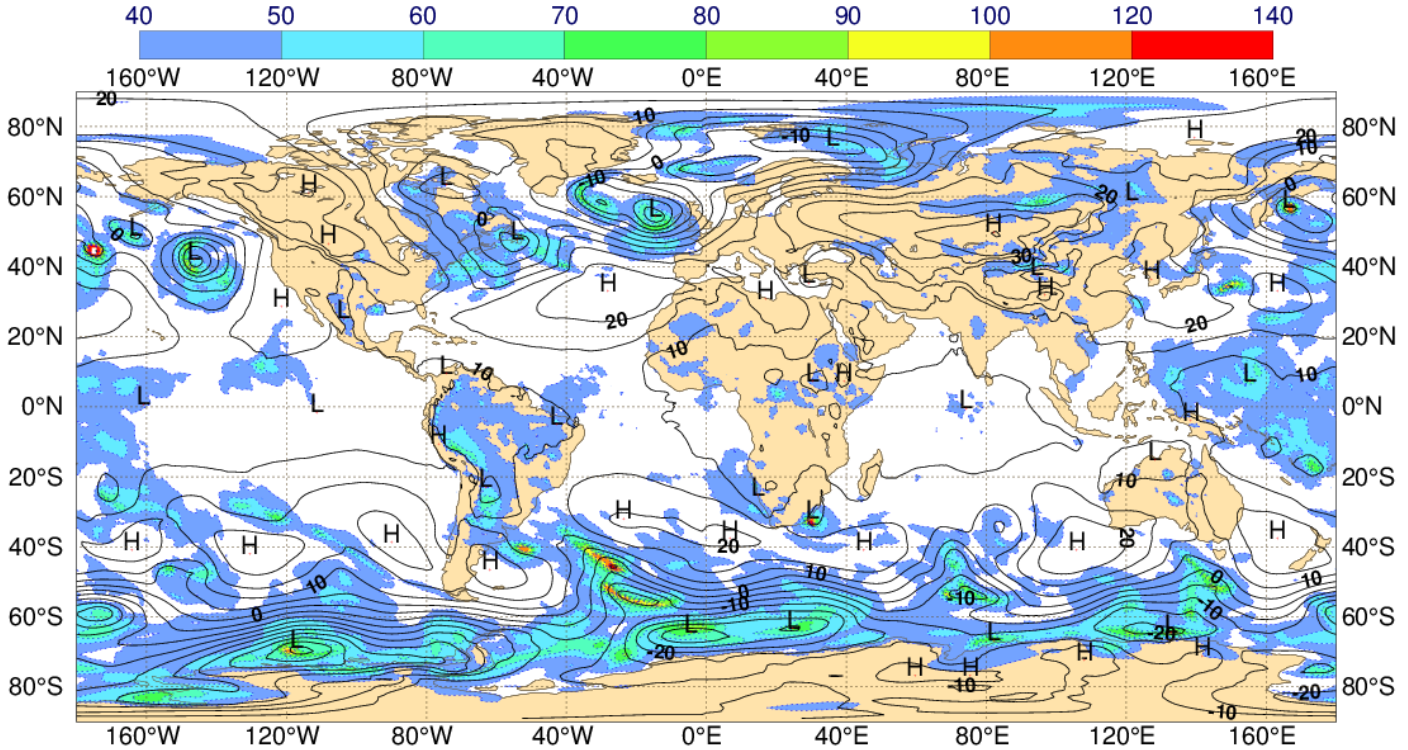
$$\mathbf{H}\mathbf{P}^b\mathbf{H}^T = \mathbf{H}\mathbf{X}^b(\mathbf{H}\mathbf{X}^b)^T =$$

$$\frac{1}{M-1}\Sigma_m(\mathbf{x}^b{}_m - \langle \mathcal{H}(\mathbf{x}^b{}_m) \rangle)\,(\mathbf{x}^b{}_m - \langle \mathcal{H}(\mathbf{x}^b{}_m) \rangle)^T$$

- Not having to code TL and ADJ operators is a distinct advantage!

Ensemble Kalman Filters

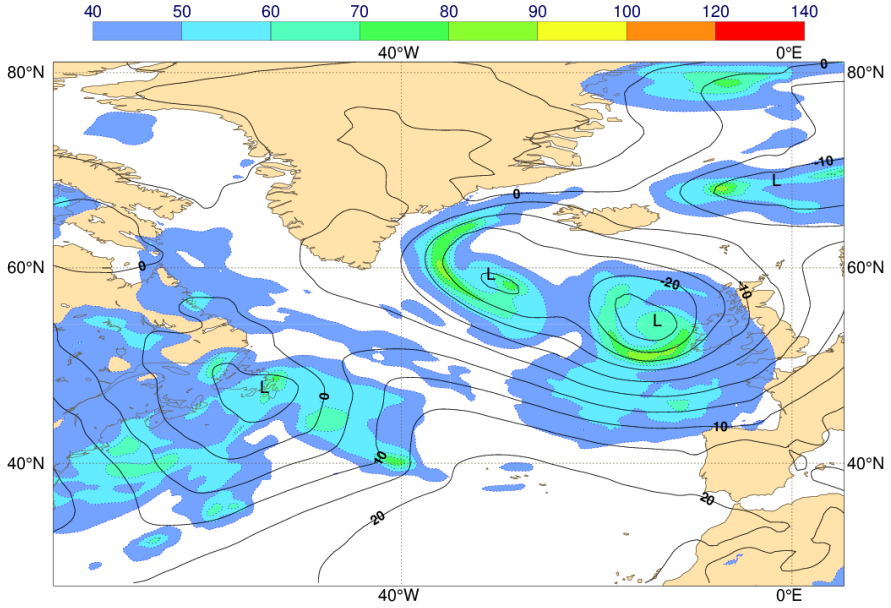
- In the EnKF the error covariances are sampled from the ensemble forecasts. They reflect the current state of the atmospheric flow

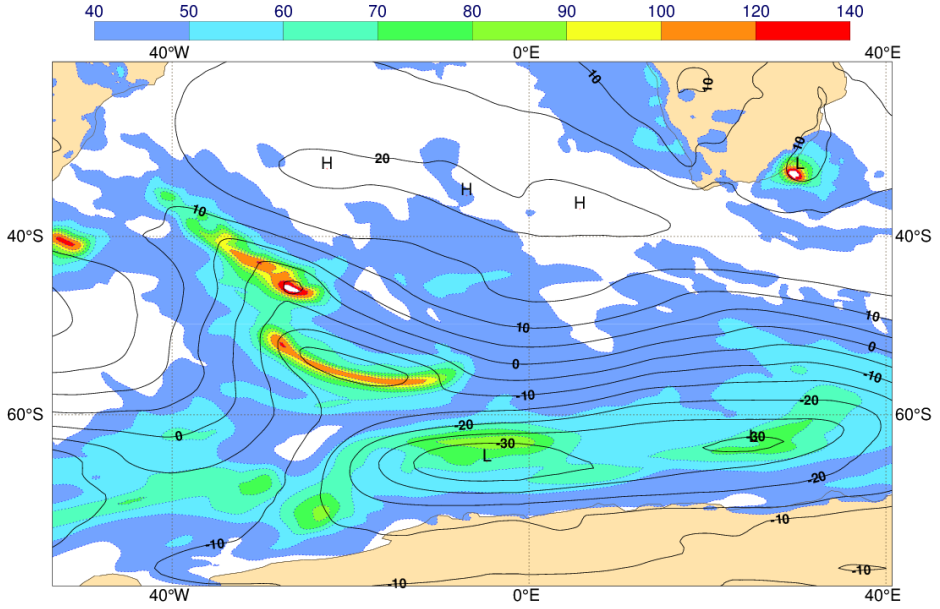


Spread of surface pressure background t+6h fcst (shaded, Pa)
Z1000 background t+6h fcst (black isolines)

Ensemble Kalman Filters

- In the EnKF the error covariances are sampled from the ensemble forecasts. They reflect the current state of the atmospheric flow





Spread of surface pressure background t+6h fcst (shaded, Pa)
Z1000 background t+6h fcst (black isolines)

ECMWF

Ensemble Kalman Filters

- The Ensemble Kalman Filter requires us to generate a sample $\{\mathbf{x}^b_{k,m}; m=1,..,M\}$ drawn from the pdf of background error: how to do this?
- We can generate this from a sample $\{\mathbf{x}^a_{k-1,m}; m=1,..,M\}$ from the pdf of analysis error for the previous cycle:

$$\mathbf{x}^b_m = \mathcal{M}(\mathbf{x}^a_{t-1,m}) + \boldsymbol{\eta}_m$$

where $\boldsymbol{\eta}_m$ is a sample drawn from the pdf of model error.

- The question is then: How do we generate a sample from the analysis pdf? Let us look at the analysis update again:

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}\,(\mathbf{y} - \mathbf{H}(\mathbf{x}^b)) = (\mathbf{I}\text{-}\mathbf{KH})\,\mathbf{x}^b + \mathbf{Ky}$$

- If we subtract the true state $\mathbf{x}^*$ from both sides (and assume $\mathbf{y}^*=\mathbf{Hx}^*$)

$$\mathbf{e}^a = (\mathbf{I}\text{-}\mathbf{KH})\,\mathbf{e}^b + \mathbf{Ke}^o$$

- i.e., the errors have the same update equation as the state

Ensemble Kalman Filters

- Consider now an ensemble of analysis where all the inputs to the analysis (i.e., the background forecast and the observations) have been perturbed according to their errors:

$$\mathbf{x}^{a\prime} = (\mathbf{I}\text{-}\mathbf{KH})\,\mathbf{x}^{b\prime} + \mathbf{K}\mathbf{y}'$$

- If we subtract the unperturbed analysis $\mathbf{x}^a = (\mathbf{I}\text{-}\mathbf{KH})\,\mathbf{x}^b + \mathbf{K}\mathbf{y}$

$$\boldsymbol{\varepsilon}^a = (\mathbf{I}\text{-}\mathbf{KH})\,\boldsymbol{\varepsilon}^b + \mathbf{K}\boldsymbol{\varepsilon}^o$$

- Note that the observations (during the update step) and the model (during the forecast step) are perturbed explicitly.
- The background is implicitly perturbed , i.e.:

$$\mathbf{x}^b = \mathcal{M}(\mathbf{x}^a_{t\text{-}1,m}) + \boldsymbol{\eta}_m$$

- Hence, one way to generate a sample drawn from the pdf of analysis error is to perturb the observations with perturbations characteristic of observation error.
- The EnKF based on this idea is called Perturbed Observations (Stochastic) EnKF (Houtekamer and Mitchell, 1998). It is also the basis of ECMWF EDA

Ensemble Kalman Filters

- Another way to construct the analysis sample without perturbing the observations is to make a linear combination of the background sample:

$$\mathbf{X}^a=\mathbf{X}^b\mathbf{T}$$

where T is a *M* x *M* matrix chosen such that:

$$\mathbf{X}^a(\mathbf{X}^a)^T = (\mathbf{X}^b\mathbf{T})\,(\mathbf{X}^b\mathbf{T})^T = \mathbf{P}^a = (\mathbf{I}\text{-}\mathbf{K}\mathbf{H})\mathbf{P}^b$$

- Note that the choice of **T** is not unique: Any orthonormal transformation **Q** ($\mathbf{Q}\mathbf{Q}^T=\mathbf{Q}^T\mathbf{Q}=\mathbf{I}$) can be applied to **T** and give a valid analysis sample!

- Implementations also differ on the treatment of observations (i.e., local patches, one at a time)

- Consequently there are a number of different, functionally equivalent, implementations of the Deterministic EnKF (ETKF, Bishop *et al.*, 2001; LETKF, Ott *et al.*, 2004, Hunt *et al.*, 2007; EnSRF, Whitaker and Hamill, 2002; EnAF, Anderson, 2001;…)

Ensemble Kalman Filters

- The real question then is:

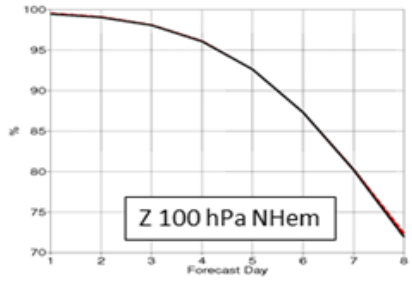
  How does the EnKF compare with standard 4DVar?

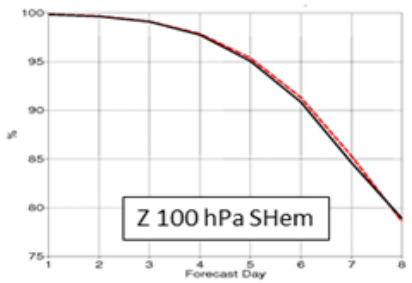
ECMWF EnKF vs 4DVar deterministic forecast skill

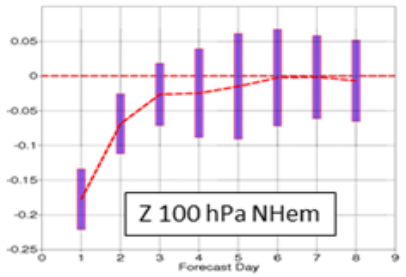
TL399 100 member EnKF

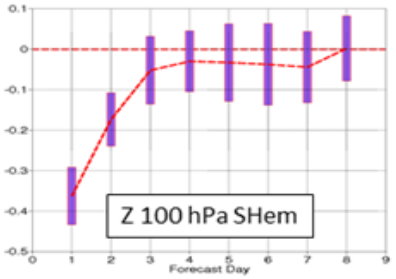
TL399 (95/159) 4DVar with static B

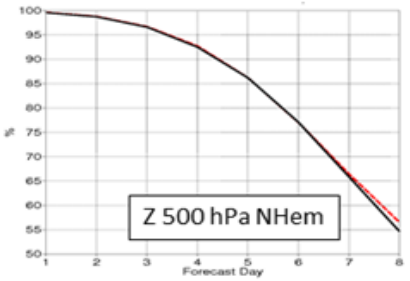
Verification against ECMWF Operations (T1279 4DVar analysis)

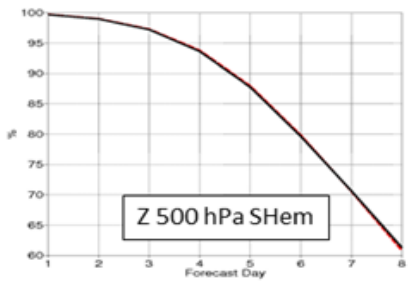


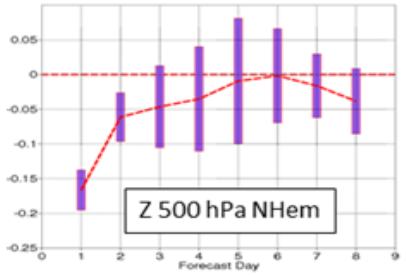


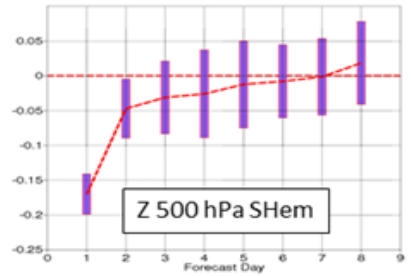












Ensemble Kalman Filters

- Pros:
  1. Background error estimates reflect state of the flow
  2. Provides an ensemble of analyses: can use for Ensemble prediction
  3. EnKF competitive with standard 4DVar at intermediate resolutions
  4. Very good scalability properties

Ensemble Kalman Filters

- Cons:
  1. The basic approximation of the EnKF is to replace the mean and covariances of the KF with sample mean and covariances
  2. As the affordable ensembles are relatively small sampling noise and rank deficiency of the sampled error covariances become performance limiting factor for the EnKF
  3. Careful localization of sampled covariances becomes crucial: This is an on-going research topic for both EnKF and Ensemble Variational hybrid systems
  4. Note how covariance localization becomes conceptually and practically more difficult for observations (satellite radiances) which are non-local, i.e. they sample a layer of the atmosphere (Campbell *et al.*, 2010)

Hybrid Variational–EnKF algorithms

The best of both worlds?

Hybrid Variational–EnKF algorithms

4D-Var

If we neglect model error (perfect model assumption) the problem of finding the model trajectory that best fits the observations over an assimilation interval t=0,1,…,T) given a background state $\mathbf{x_b}$ and its error covariance $\mathbf{P^b}$ can be solved By finding the minimum of the cost function:

$$J(\mathbf{x}_0) = (\mathbf{x}_b - \mathbf{x}_o)^T (\mathbf{P}^b)^{-1} (\mathbf{x}_b - \mathbf{x}_o) + \sum_{t=0}^{T} (\mathbf{y}_t - H_t M_{0\to t}(\mathbf{x}_0))^T \mathbf{R}_t^{-1} (\mathbf{y}_t - H_t M_{0\to t}(\mathbf{x}_0))$$

This is equivalent, for the same $\mathbf{x_b}$, $\mathbf{P^b}$, to the Kalman filter solution at the end of the assimilation window (t=T) (*Fisher et al.,*2005).

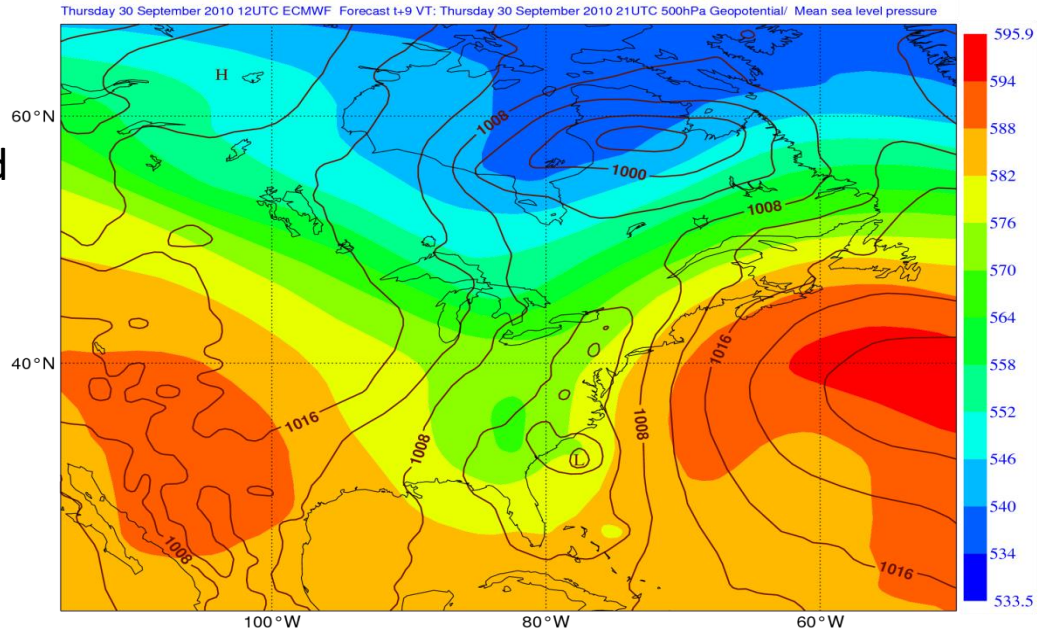
Hybrid Variational–EnKF algorithms

4D Variational methods

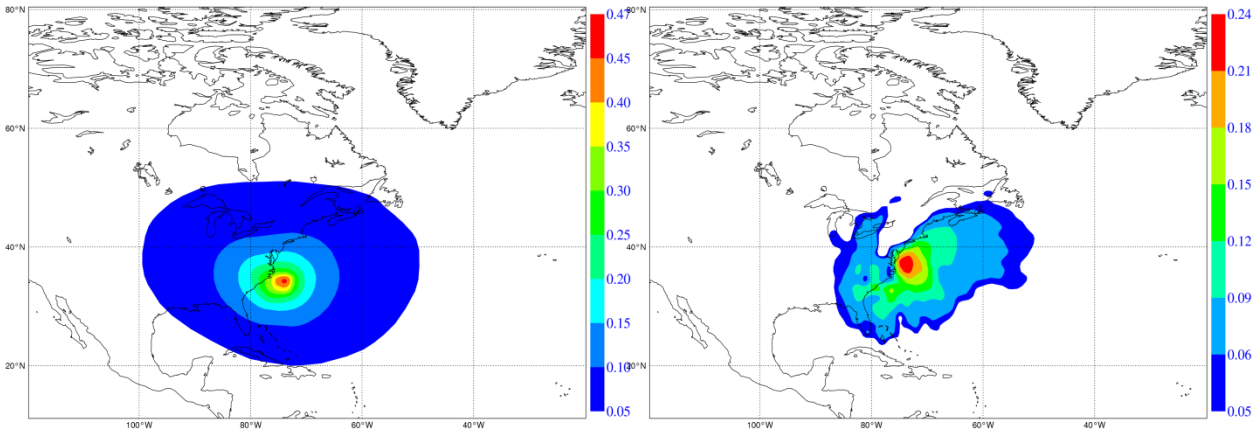
The 4D-Var solution implicitly evolves background error covariances *over the length of the assimilation window* (Thepaut *et al*.,1996) with the tangent linear dynamics:

$$\mathbf{P}^{\mathbf{b}}(\mathrm{t}) \approx \mathbf{M}\mathbf{P}^{\mathbf{b}}\mathbf{M}^{\mathbf{T}}$$

MSLP and 500 hPa Z (shaded) background fcst



Temperature analysis increments for a single temperature observation at the start of the assimilation window: $x^a(t)-x^b(t) \approx MP^bM^TH^T(y-Hx)/(\sigma_b^2 + \sigma_o^2)$



t=+0h t=+3h t=+9h

Hybrid Variational–EnKF algorithms

4D Variational methods

- The 4D-Var solution implicitly evolves background error covariances *over the length of the assimilation window* with the tangent linear dynamics:

$$\mathbf{P}^{\mathbf{b}}(t) \approx \mathbf{M}\mathbf{P}^{\mathbf{b}}\mathbf{M}^{\mathbf{T}}$$

- But it does not propagate error information from one assimilation cycle to the next: $\mathbf{P}^{\mathbf{b}}$ is not evolved according to KF equations (i.e., $\mathbf{P}^{\mathbf{b}} = \mathbf{M}\mathbf{P}^{\mathbf{a}}\mathbf{M}^{\mathbf{T}} + \mathbf{Q}$) but is reset to a climatological, stationary estimate at the beginning of each assimilation window.

- Only information about the state $(\mathbf{x}_{\mathbf{b}})$ is propagated from one cycle to the next.

Hybrid Variational–EnKF algorithms

a) Kalman Filter is computationally unfeasible for large dimensional systems (e.g., operational NWP);

b) Variational (4D-Var) do not cycle state error estimates: works well for short assimilation windows (6-12h). Longer windows, where $\mathbf{Q}$ is required, have proved more difficult;

c) EnKF cycle reduced-rank estimates of state error covariances: need for spatial localization to combat rank deficiency, with possibly negative impact on dynamical balance and use non-local observations (radiances);

….

Hybrid approach: Use cycled, flow-dependent state error estimates (from an EnKF/Ensemble DA system) in a 3/4D-Var analysis algorithm

Hybrid Variational–EnKF algorithms

Hybrid approach: Use cycled, flow-dependent state error estimates
(from an EnKF/EDA system) in a 3/4D-Var analysis algorithm

This solution would:

1) Integrate flow-dependent state error covariance information into a variational analysis
2) Keep the full rank representation of $\mathbf{P}^{b}$ and its implicit evolution inside the assimilation window
3) More robust than pure EnKF for limited ensemble sizes and large model errors
4) Allow for flow-dependent quality control of observations

Hybrid Variational–EnKF algorithms

In operational use (or under test), there are various approaches to doing hybrid DA in a VAR context:

1. Extended control variable method (Met Office, NCEP/GMAO, CMC)
2. 4D-Ensemble-Var (under active development in all of the above)
3. Ensemble of Data Assimilations method (ECMWF, Meteo France)
4. Hybrid Gain Ensemble Data Assimilation (ECMWF)

Hybrids: extended control variable

1. Extended control variable (Barker, 1999; Lorenc, 2003)

Conceptually add a flow-dependent term to the model of $\mathbf{P}^b$ ($\mathbf{B}$):

$$\mathbf{B} = \beta_c^2 \mathbf{B}_c + \beta_e^2 \mathbf{P}_e \circ \mathbf{C}_{loc}$$

$\mathbf{B}_c$ is the static, climatological covariance
$\mathbf{P}_e \circ \mathbf{C}_{loc}$ is the localised ensemble sample covariance

In practice this is done through augmentation of the control variable:

$$\delta\mathbf{x} = \beta_c \mathbf{B}_c^{1/2} \mathbf{v} + \beta_e \mathbf{X}' \circ \boldsymbol{\alpha}$$

and introducing an additional term in the cost function:

$$J = \frac{1}{2}\boldsymbol{v}^T\boldsymbol{v} + \frac{1}{2}\boldsymbol{\alpha}^T\boldsymbol{C}_{loc}^{-1}\boldsymbol{\alpha} + J_o + J_c$$

from A.Clayton

Hybrids: extended control variable

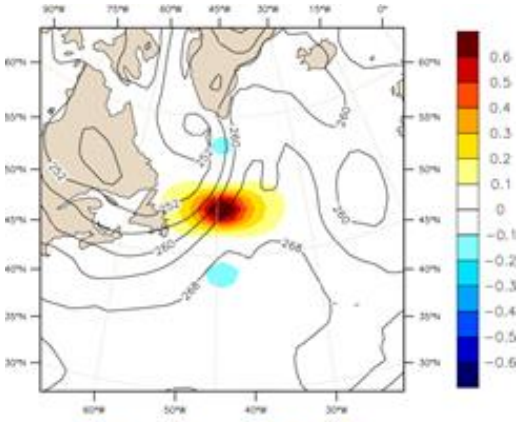
1. Extended control variable method

$$\delta\mathbf{x} = \beta_c \mathbf{B}_c^{1/2}\mathbf{v} + \beta_e \mathbf{X}' \circ \boldsymbol{\alpha} = \delta\mathbf{x}_{c\,\mathrm{lim}} + \delta\mathbf{x}_{ens}$$

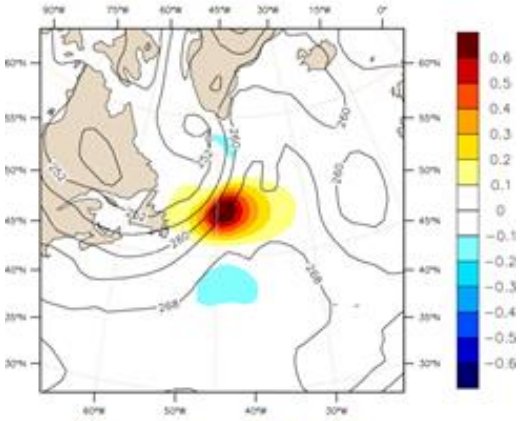
- The increment is now a weighted sum of the static $\mathbf{B}$ component and the flow-dependent, ensemble based $\mathbf{B}$
- The flow-dependent increment is a linear combination of ensemble perturbations $\mathbf{X}'$, modulated by the $\boldsymbol{\alpha}$ fields
- If the $\boldsymbol{\alpha}$ fields were homogeneous $\delta\mathbf{x}_{ens}$ could only span N_{ens}-1 degrees of freedom; instead $\boldsymbol{\alpha}$ fields are smoothly varying, which effectively increases the degrees of freedom
- $\mathbf{C}_{loc}$ is a covariance (localization) model for the flow-dependent increments: it controls the spatial variation of $\boldsymbol{\alpha}$

Hybrids: extended control variable

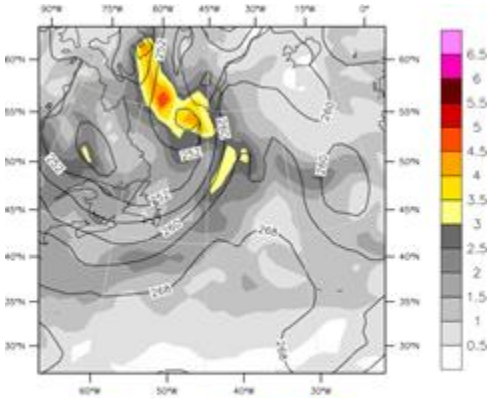
u response to a single u observation at centre of window



Standard 3D-Var

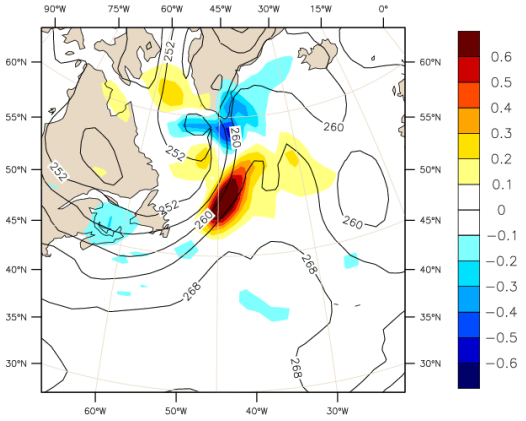


Standard 4D-Var

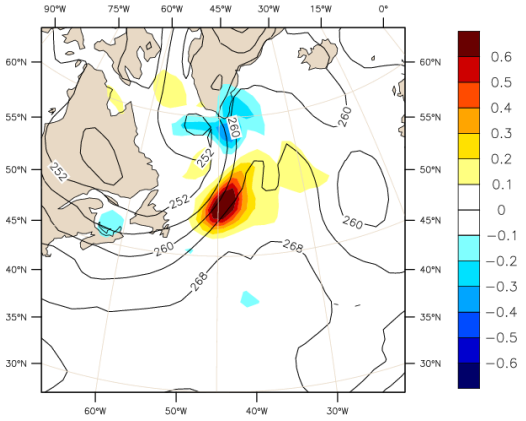


Ensemble RMS

from A.Clayton



Pure ensemble 3D-Var



50/50 hybrid 3D-Var

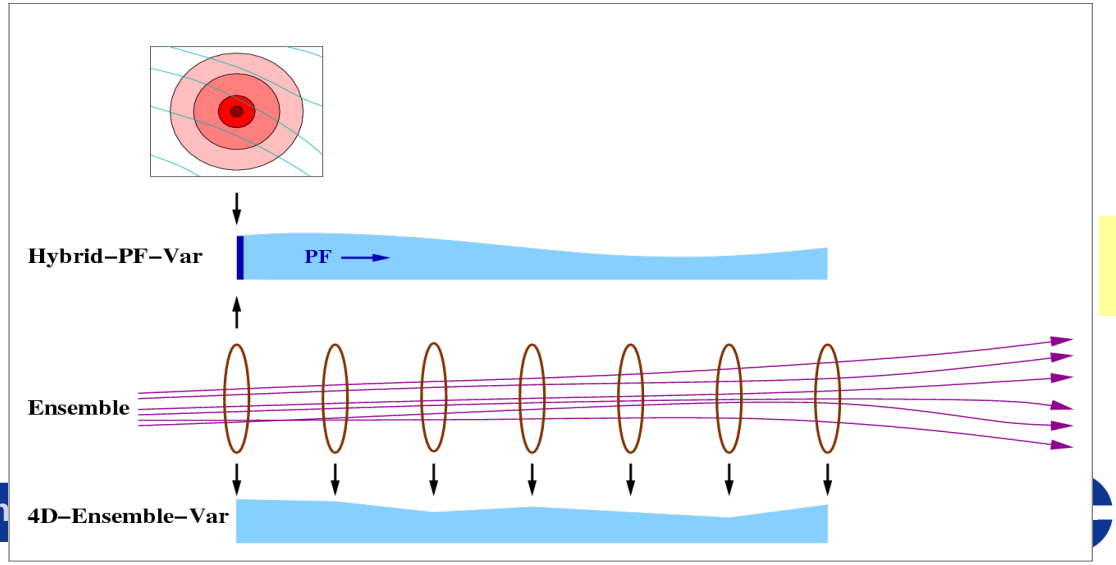
ECMWF

Hybrids: 4D-En-Var

2. 4D-Ensemble-Var method (Liu et al., 2008)

- In the extended control variable method one uses the ensemble perturbations to estimate $\mathbf{P}^b$ only at the start of the 4DVar assimilation window: the evolution of $\mathbf{P}^b$ inside the window is done by the tangent linear dynamics ($\mathbf{P}^b(t) \approx \mathbf{M}\mathbf{P}^b\mathbf{M}^T$)

- In 4D-En-Var $\mathbf{P}^b$ is sampled from ensemble trajectories throughout the assimilation window:



from D. Barker

Hybrids: 4D-En-Var

- The 4D-Ens-Var analysis increment is thus a localised linear combination of ensemble trajectories perturbations:

$$\delta\mathbf{x} = \sum_{k=1,N} \boldsymbol{\alpha}_k \circ \mathbf{x}_k^{'}(t)$$

$$\mathbf{x}_k^{'}(t) = \mathbf{x_k}(t) - \overline{\mathbf{x_k}(t)}$$

- This is fundamentally the same state update procedure of the LETKF version of EnKF (Hunt et al., 2007): It is difficult to imagine that 4D-En-Var should work better than an EnKF.
- While traditional 4DVar requires repeated, **sequential** runs of $\mathbf{M}$, $\mathbf{M^T}$, ensemble trajectories from the previous assimilation time can be pre-computed in parallel
- As in the EnKF, 4D-Ens-Var does not require developing and maintaining the TL and Adjoint models

Hybrids

3. Ensemble of Data Assimilations
4. Hybrid Gain Ensemble Data Assimilation

- To be continued…

References

1. Anderson, J. L., 2001. An ensemble adjustment Kalman filter for data assimilation. Mon. Wea. Rev. 129, 2884–2903.

2. Bishop, C. H., Etherton, B. J. and Majumdar, S. J., 2001. Adaptive sampling with ensemble transform Kalman filter. Part I: theoretical aspects. Mon. Wea. Rev. 129, 420–436.

3. Burgers, G., Van Leeuwen, P. J. and Evensen, G., 1998. On the analysis scheme in the ensemble Kalman filter. Mon. Wea. Rev. 126, 1719–1724.

4. Campbell, W. F., C. H. Bishop, and D. Hodyss, 2010: Vertical covariance localization for satellite radiances in ensemble Kalman Filters. Mon. Wea. Rev.,138,282–290.

5. Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res. 99(C5), 10 143–10 162.

6. Evensen, G . 2004 . Sampling strategies and square root analysis schemes for the EnKF . No.: 54 . p.: 539-560 . Ocean Dynamics

7. Fisher, M., Leutbecher, M. and Kelly, G. A. 2005. On the equivalence between Kalman smoothing and weak-constraint four-dimensional variational data assimilation. Q.J.R. Meteorol. Soc., 131: 3235–3246.

8. Flowerdew, J. 2015. Towards a theory of optimal localisation. Tellus A 2015, 67, 25257, http://dx.doi.org/10.3402/tellusa.v67.25257

References

9. Houtekamer, P. L. and Mitchell, H. L., 1998. Data assimilation using an ensemble Kalman filter technique. Mon. Wea. Rev. 126, 796–811.

10. Houtekamer, P. L. and Mitchell, H. L., 2001. A sequential ensemble Kalman filter for atmospheric data assimilation. Mon. Wea. Rev. 129, 123–137.

11. Hunt, B. R., Kostelich, E. J. and Szunyogh, I., 2007. Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. Physica D, 230, 112–126.

12. Liu C, Xiao Q, Wang B. 2008. An ensemble-based four-dimensional variational data assimilation scheme. part i: Technical formulation and preliminary test. Mon. Weather Rev. 136: 3363–3373.

13. Lorenc, A.C.,2003: The potential of the ensemble Kalman filter for NWP—A comparison with 4D-VAR. Q. J. R. Meteorol. Soc., 129: 3183–3203.

14. Ménétrier, B., T. Montmerle, Y. Michel and L. Berre, 2014: Linear Filtering of Sample Covariances for Ensemble-Based Data Assimilation. Part I: Optimality Criteria and Application to Variance Filtering and Covariance Localization. Mon. Wea. Rev, doi: http://dx.doi.org/10.1175/MWR-D-14-00157.1

15. Ott, E., Hunt, B. H., Szunyogh, I., Zimin, A. V., Kostelich, E. J. and co-authors. 2004. A local ensemble Kalman filter for atmospheric data assimilation. Tellus 56A, 415–428.

References

16. Periáñez, Á., Reich, H. and Potthast, R. 2014. Optimal localization for ensemble Kalman filter systems. J. Met. Soc. Japan. 62, 585 597.

17. Thépaut, J.-N., Courtier, P., Belaud, G. and Lemaître, G. 1996. Dynamical structure functions in a four-dimensional variational assimilation: A case-study. Q. J. R. Meteorol. Soc., 122, 535–561

18. Whitaker, J. S. and Hamill, T. M., 2002. Ensemble data assimilation without perturbed observations. Mon. Wea. Rev. 130, 1913–1924.