ECMWF

Data Assimilation Training Course - Background Error Covariance Modelling

Elias Holm – slides courtesy Mike Fisher, Massimo Bonavita

Importance of Background Covariances

- The formulation of the J_b term of the cost function is ***crucial*** to the performance of current analysis systems.
- To see why, suppose we have a single observation of the value of a model field at one gridpoint.
- For this simple case, the observation operator is:

$$H = (0,...,0,1,0,...,0) .$$

- The gradient of the 3dVar cost function is:

$$\nabla J = B^{-1}(x-x_b) + H^T R^{-1}(Hx-y) = 0$$

- Multiply through by B and rearrange a bit:

$$x - x_b = B\, H^T R^{-1}(y-Hx)$$

- But, for this simple case, $R^{-1}(y-Hx)$ is a scalar

Importance of Background Covariances

- So, we have: $\mathbf{x} - \mathbf{x}_b \propto \mathbf{B}\mathbf{H}^{\mathrm{T}}$
- But, H = (0,...,0,1,0,...,0)
- => The analysis increment is proportional to a column of B.

- The role of B is:
  1. To spread out the information from the observations.
  2. To provide statistically consistent increments at the neighbouring gridpoints and levels of the model.
  3. To ensure that observations of one model variable (e.g. temperature) produce dynamically consistent increments in the other model variables (e.g. vorticity and divergence).

Main Issues in Covariance Modelling

- There are 2 problems to be addressed in specifying B:

1. We want to describe the statistics of the errors in the background.
   - However, we don't know what the errors in the background are, since we don't know the true state of the atmosphere.
2. The B matrix is enormous ($\sim 10^8 \times 10^8$).
   - We are forced to simplify it just to fit it into the computer.
   - Even if we could fit it into the computer, we don't have enough statistical information to determine all its elements.

Diagnosing Background Error Statistics

- **Problem:**
  - We cannot produce samples of background error. (We don't know the true state.)

- **Instead, we must either:**
  - Disentangle background errors from the information we do have: innovation (observation-minus-background) statistics.

- **Or:**
  - Use a surrogate quantity whose error statistics are similar to those of background error. Two possibilities are:
    - Differences between forecasts that verify at the same time.
    - Differences between background fields from an ensemble of analyses.

Diagnosing Background Error Statistics

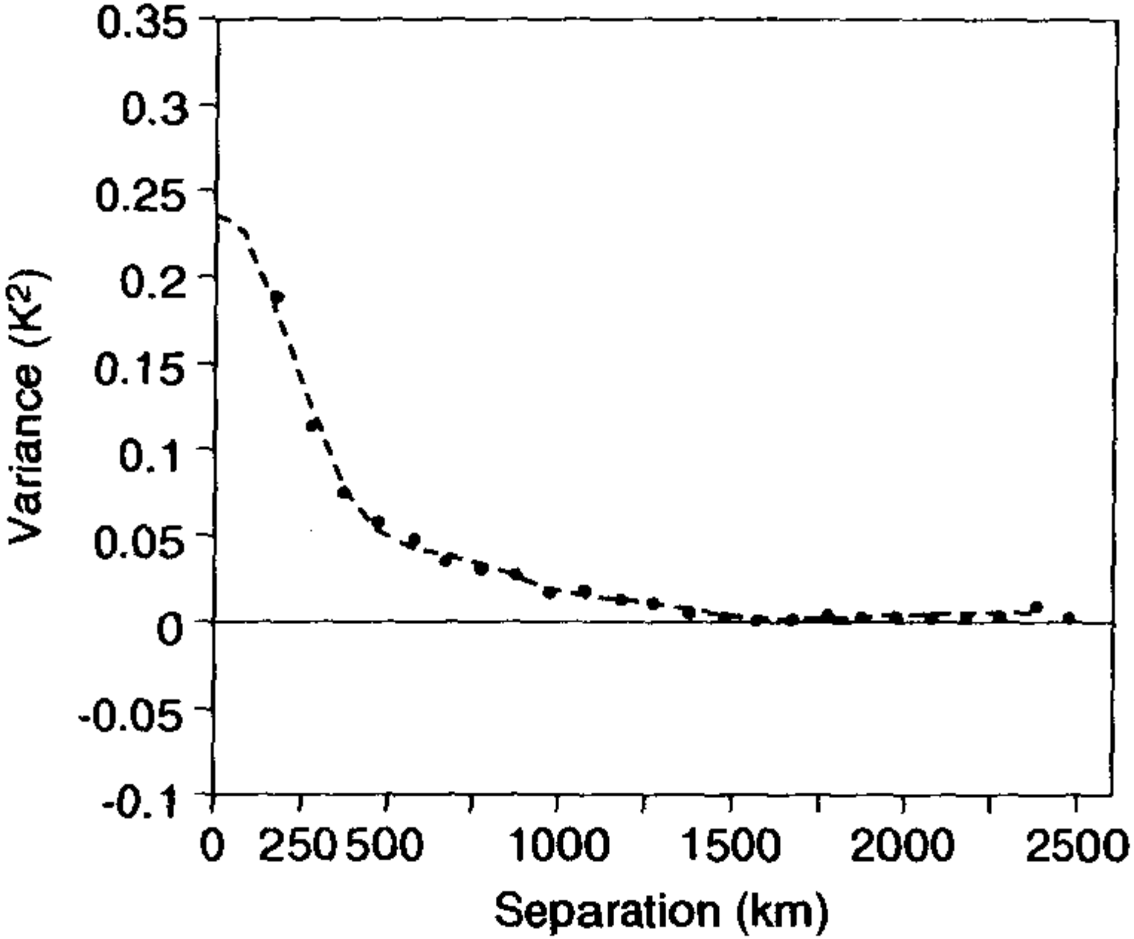
- **Three classic approaches to estimating J_b statistics:**

1. **The Hollingsworth and Lönnberg (1986) method**
   - Differences between observations and the background are a combination of background and observation error.
   - The method tries to partition this error into background errors and observation errors by assuming that the observation errors are spatially uncorrelated.
2. **The NMC method (Parrish and Derber, 1992)**
   - This method assumes that the spatial correlations of background error are similar to the correlations of differences between 48h and 24h forecasts verifying at the same time.
3. **The Analysis-Ensemble method (Fisher, 2003)**
   - This method runs the analysis system several times for the same period with randomly-perturbed observations and models. Differences between background fields for different runs provide a surrogate for a sample of background error.

Estimating Background Error Statistics from Innovation Statistics

- **Assume:**
  1. Background errors are independent of observation errors.
  2. Observations have spatially uncorrelated errors (for some observation types).
- Let $d_i = y_i - H_i(x_b)$ be the innovation (obs-bg) for the i^{th} observation.
- Then, denoting background error by ε, observation error by η, and neglecting representativeness error, we have $d_i = \eta_i - H_i(\varepsilon)$.
  1. $\Rightarrow Var(d_i) = Var(\eta_i) + Var(H_i(\varepsilon))$
  2. $\Rightarrow Cov(d_i , d_k) = Cov(H_i(\varepsilon) , H_k(\varepsilon))$ (i and k not co-located)
- We can extract a lot of useful information by plotting $Cov(d_i , d_k)$ as a function of the distance between pairs of observations.

Estimating Background Error Statistics from Innovation Statistics

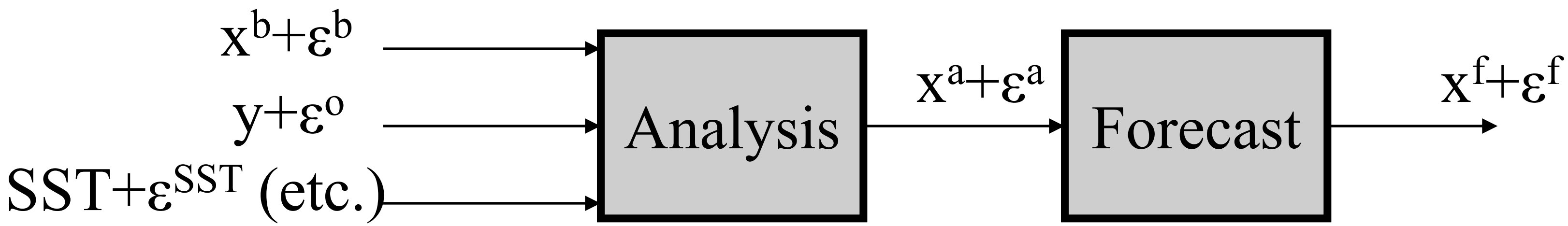


Covariance of $d=y-H(x_b)$ for AIREP temperatures over USA, binned as a function of observation separation.

(from Järvinen, 2001)

Estimating Background Error Statistics from Ensembles of Analyses

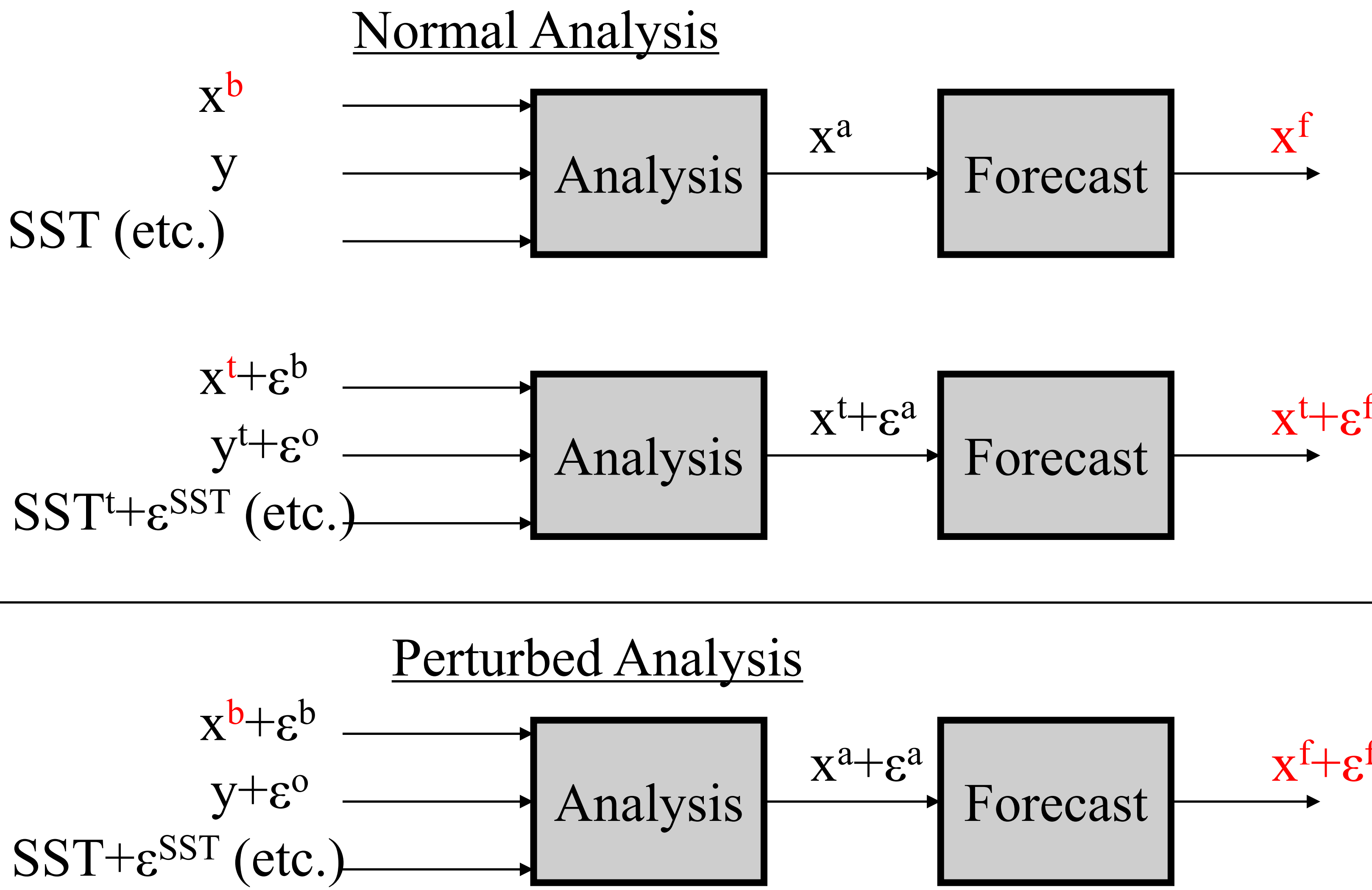
- **Suppose we perturb all the inputs to the analysis/forecast system with random perturbations, drawn from the relevant distributions:**



- **The result will be a perturbed analysis and forecast, with perturbations characteristic of analysis and forecast error.**
- **The perturbed forecast may be used as the background for the next (perturbed) cycle.**
- **After a few cycles, the system will have forgotten the original initial background perturbations.**

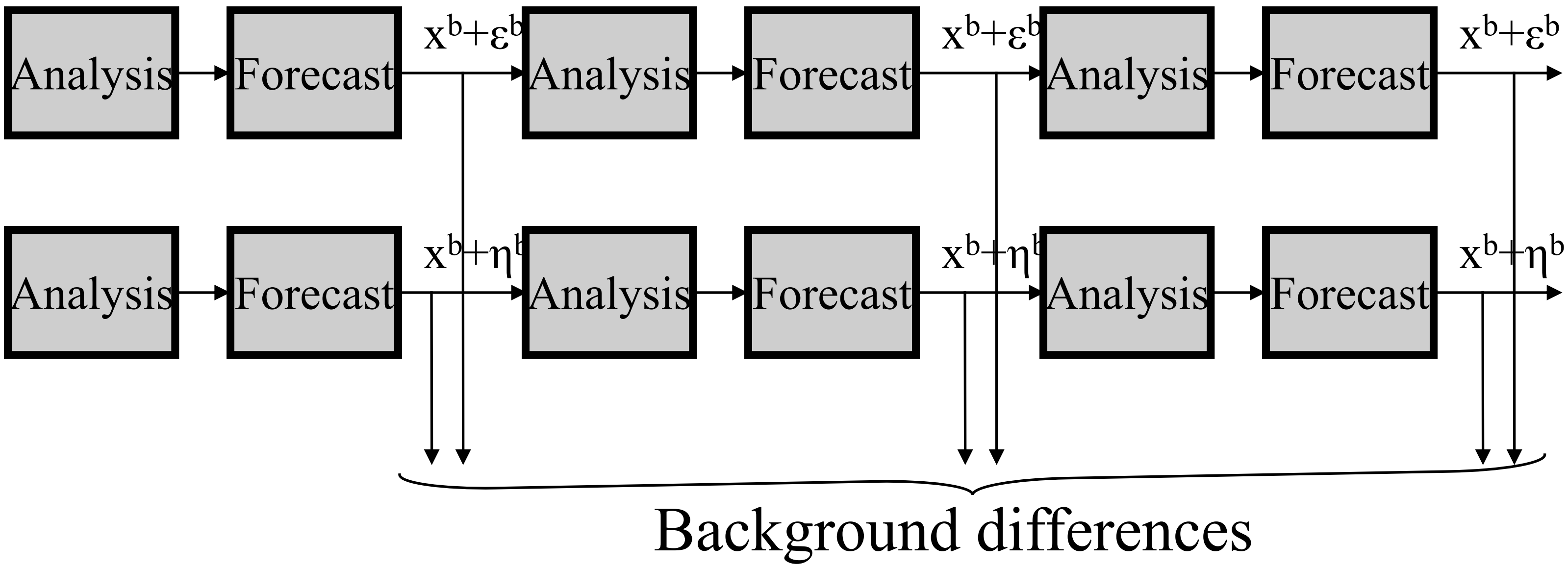
ECMWF

Estimating Background Error Statistics from Ensembles of Analyses

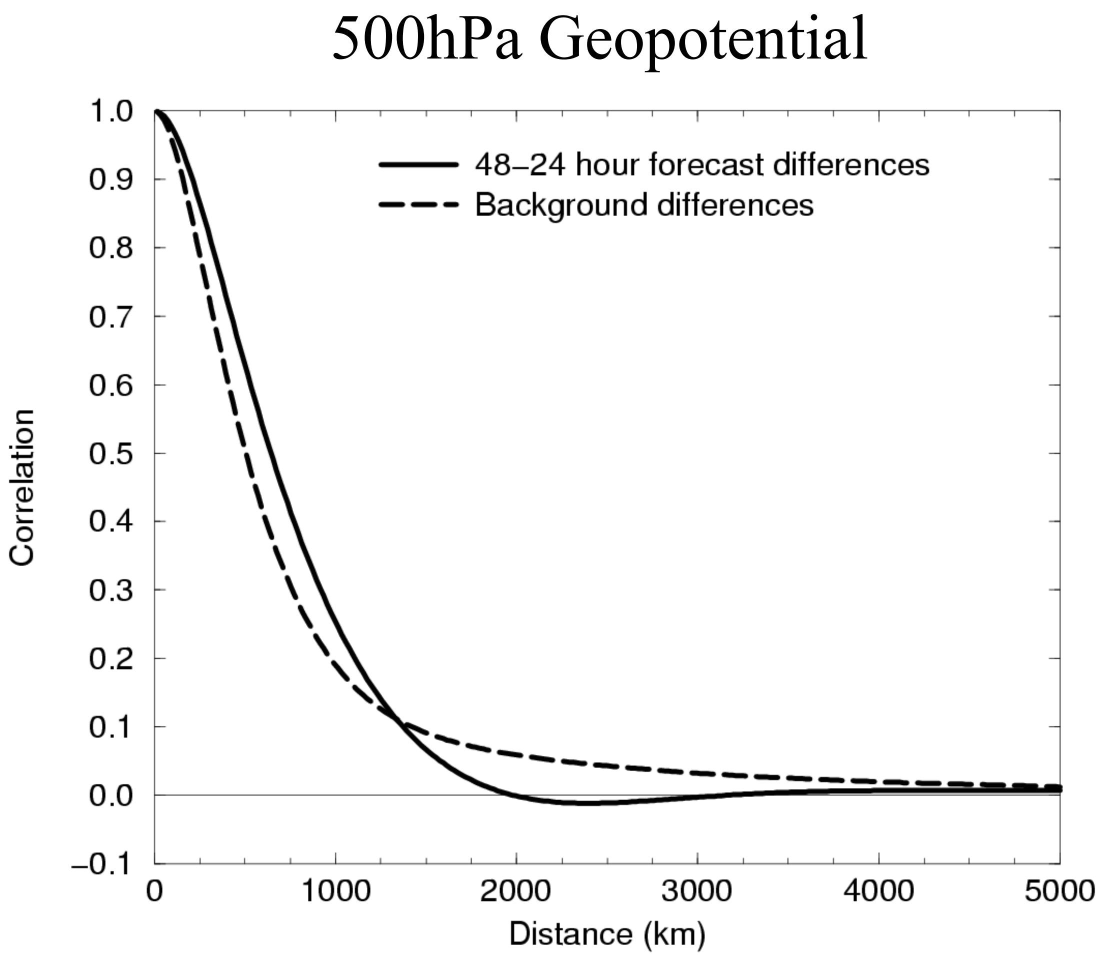


Estimating Background Error Statistics from Ensembles of Analyses

- Run the analysis system **several times with different perturbations**, and form differences between pairs of background fields.
- These differences will have the statistical characteristics of background error (but **twice the variance**).

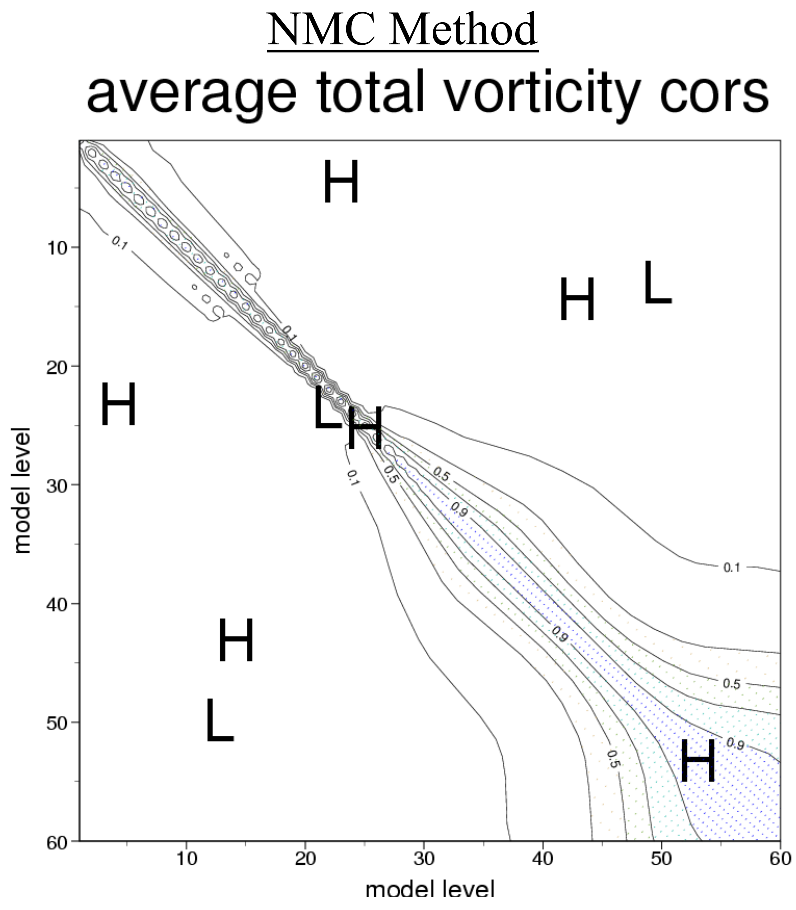


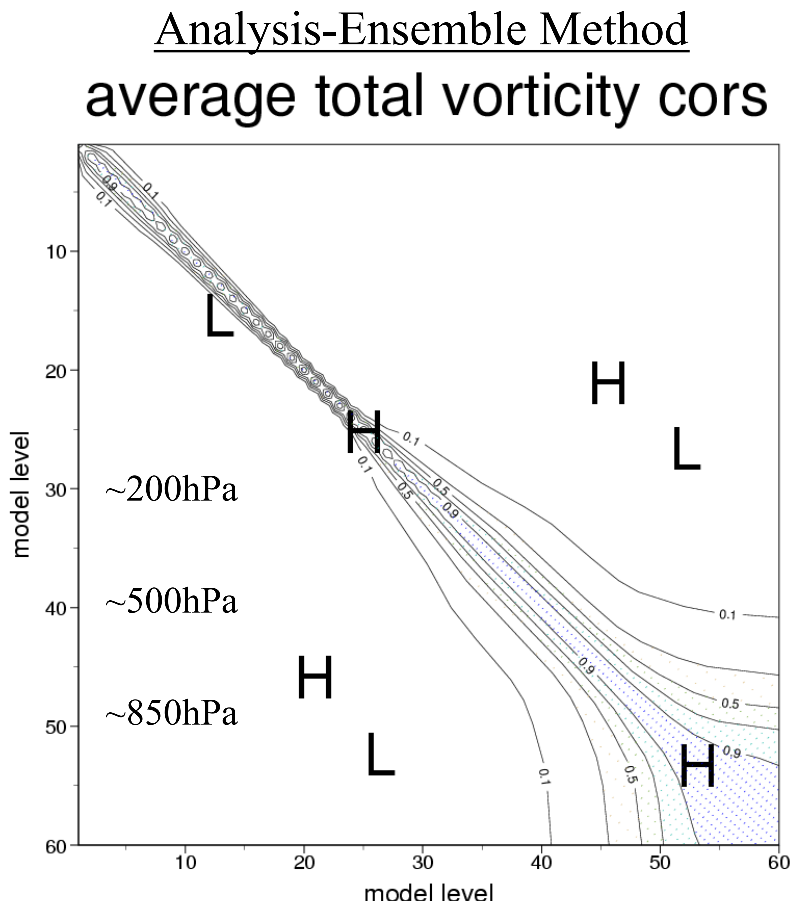
Estimating Background Error Statistics from Ensembles of Analyses



ECMWF

Estimating Background Error Statistics from Ensembles of Analyses





ECMWF

Estimating Background Error Statistics – Pros and Cons of the Various Methods

- **Innovation statistics:**
  - ☺ The only direct method for diagnosing background error statistics.
  - ☹ Provides statistics of background error in observation space.
  - ☹ Statistics are not global, and do not cover all model levels.
  - ☹ Requires a good uniform observing network.
  - ☹ Statistics are biased towards data-dense areas.
- **Forecast Differences:**
  - ☺ Generates global statistics of model variables at all levels.
  - ☺ Inexpensive.
  - ☹ Statistics are a mixture of analysis and background error.
  - ☹ Not good in data-sparse regions.
- **Ensembles of Analyses:**
  - ☹ Assumes statistics of observation error (and SST, etc.) are well known.
  - ☺ Diagnoses the statistics of the actual analysis system.
  - ☹ Danger of feedback. (Noisy analysis system => noisy stats => noisier system.)

Spectral B model

In variational analysis the B matrix is usually defined implicitly in terms of a transformation from the departure δx in state space to a control variable χ:

$$\delta \mathbf{x} = \mathbf{x} - \mathbf{x}_b = \mathbf{L}\chi$$

where $\mathbf{L}$ verifies $\mathbf{B} = \mathbf{L}\mathbf{L}^T$

In the spectral formulation (Derber and Bouttier, 1999), the change of variable $\mathbf{L}$ has the form:

$$\mathbf{L} = \mathbf{K}\,\mathbf{B}_u^{1/2}$$

where $\mathbf{K}$ is a balance operator going from the set of "unbalanced " variables [ζ, η_u, $(T,ps)_u$, q] (the "control vector") to the set of state variables [ζ, η, (T,ps), q]

There is a degree of flow-dependence in $\mathbf{K}$ as the balance constraints are linearised about the first-guess trajectory

Spectral B model

$$\delta \mathbf{x} = \mathbf{x} - \mathbf{x}_b = \mathbf{L}\chi \qquad \mathbf{L} = \mathbf{K}\,\mathbf{B}_u^{1/2}$$

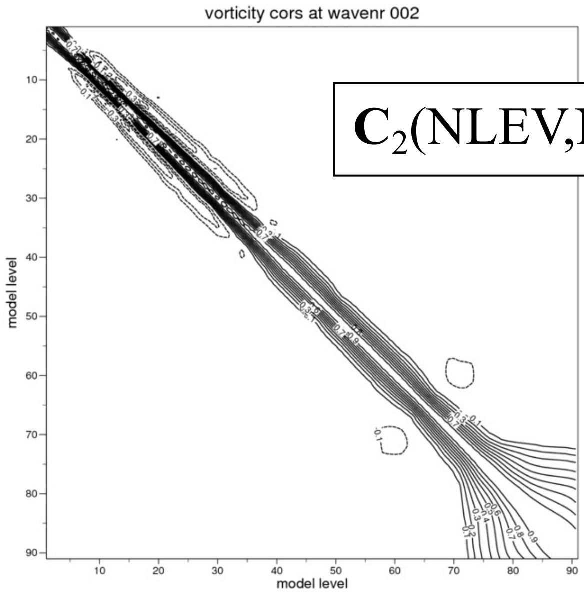
Since we assume that the balance operator accounts for all inter-variable correlations, B_u is block diagonal

$$\mathbf{B}_u = \begin{pmatrix} \mathbf{B}_\zeta & 0 & 0 & 0 \\ 0 & \mathbf{B}_{D_u} & 0 & 0 \\ 0 & 0 & \mathbf{B}_{(T,p_s)_u} & 0 \\ 0 & 0 & 0 & \mathbf{B}_q \end{pmatrix}$$

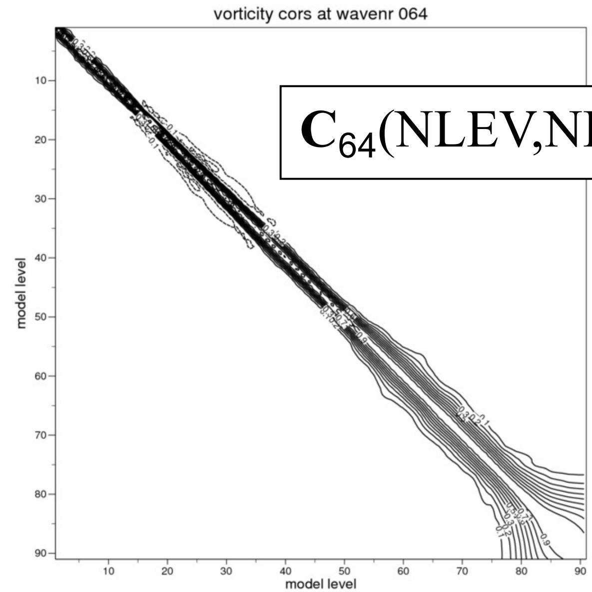
Each block in B_u is of the form $\mathbf{\Sigma}^T\mathbf{C}\mathbf{\Sigma}$.

$\mathbf{\Sigma}$ is the gridpoint standard deviation of background errors.

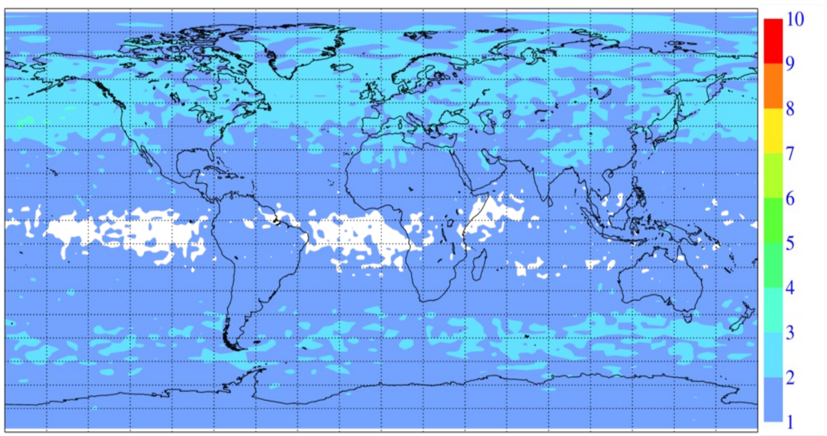
$\mathbf{C}$ models the autocorrelation of the control variables. It is block diagonal with one full vertical correlation matrix for each spectral wavenumber, i.e. C_n(NLEV,NLEV) (non-separable B model)



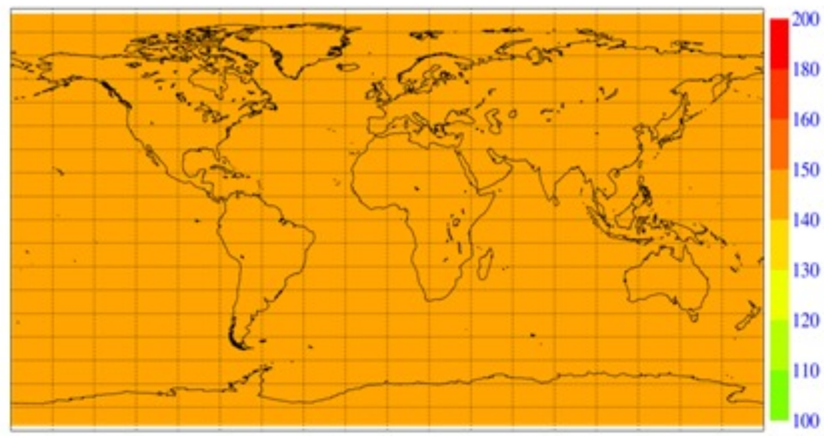
Vorticity correl. wavenum=2



Vorticity correl. wavenum=64



Vorticity bg error stdev, 500hPa



Vorticity bg error corr. Lscale, 500hPa

From Spectral to Wavelet B model

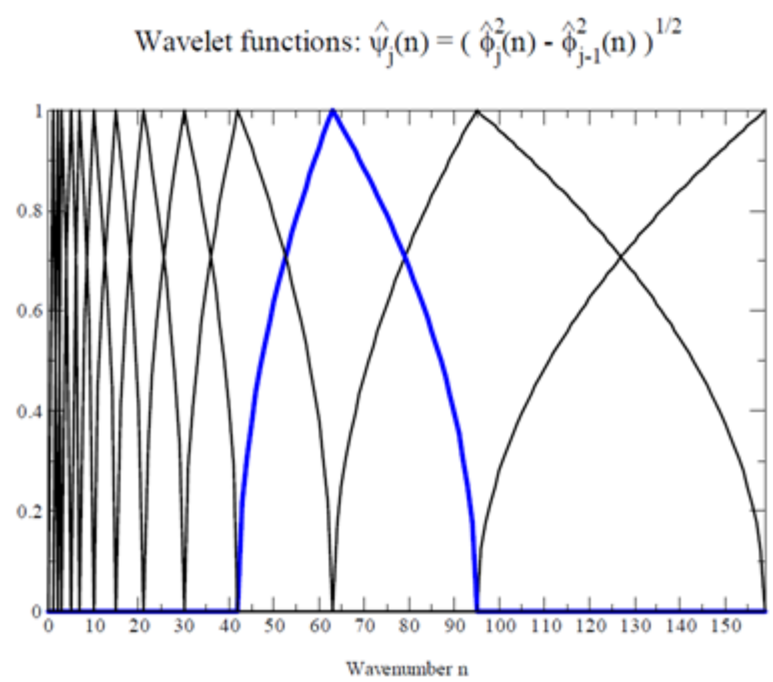
- The spectral B model is one end of the spectrum: full resolution of the variation of vertical correlation with horizontal wavenumber, but it allows no variation with horizontal location.

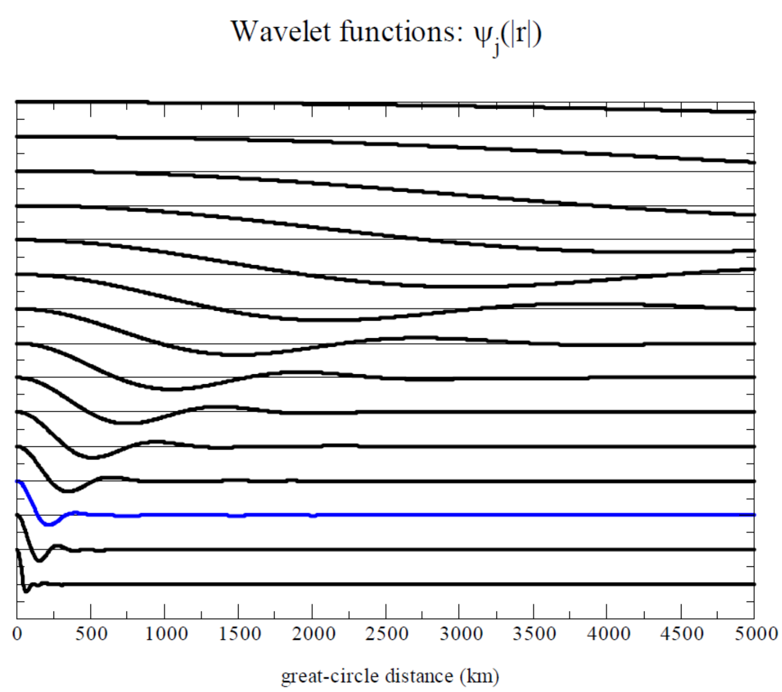
- The other end of the spectrum is represented by the separable formulation which allows full variation of the correlations with horizontal location, but allows no variation of vertical correlation with horizontal wavenumber.

- The wavelet B (Fisher, 2003) is a compromise between these two extremes and allows a degree of variation of correlation with both wavenumber and horizontal location.

Wavelet B model

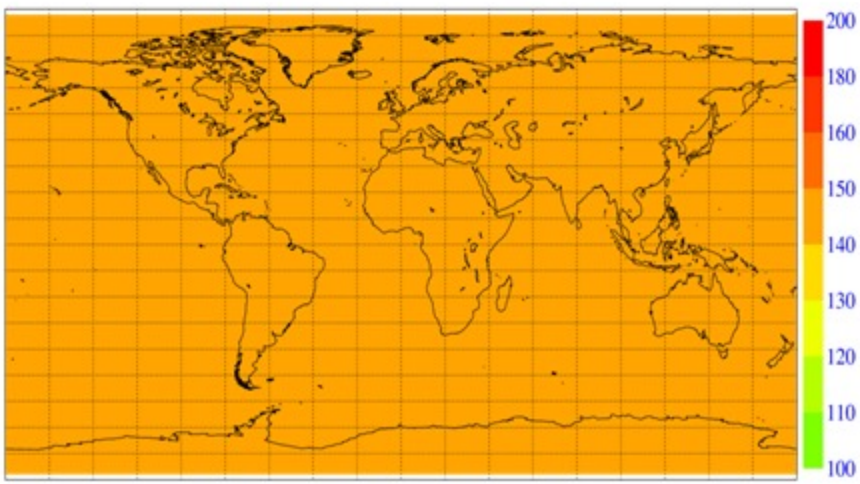
- The wavelet **B** is based on a wavelet expansion on the sphere.
- The basis functions (wavelets) are chosen to be band-limited and, to a good approximation, spatially localized



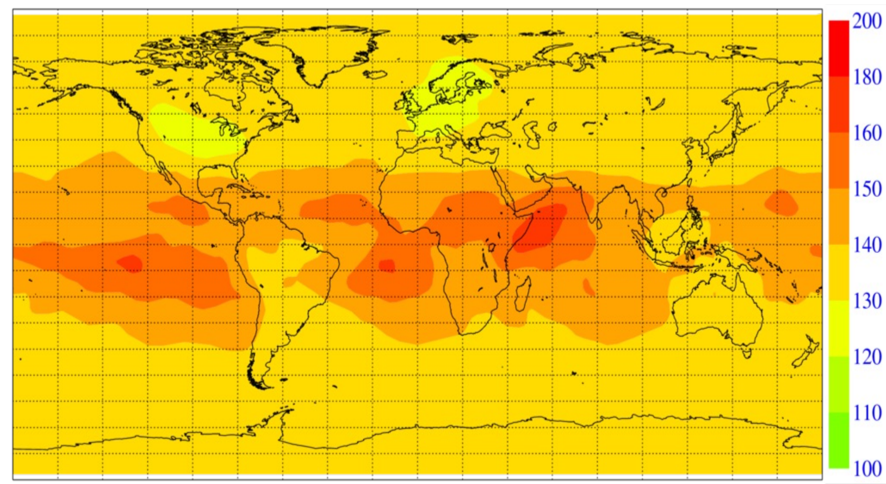


Wavelet B model

- The correlation matrices $\mathbf{C}_n[N_{lev}xN_{lev}]$ are now of the form $\mathbf{C}_j[N_{lev}xN_{lev}](\lambda,\varphi)$, where j is now the index of the wavelet component
- The choice of the wavelet bandwidths $[N_j, N_{j+1}]$ determines the trade-off between spectral and spatial resolution. If the bands are narrow, the corresponding wavelet functions are not spatially localized, and vice versa



Climat. Spectral B
Vorticity bg error corr. Lscale, 500hPa



Climat. Wavelet B
Vorticity bg error corr. Lscale, 500hPa

Flow-dependent wavelet B model

The wavelet **B** formulation:

$$\left(\mathbf{x}-\mathbf{x}_b\right)=\mathbf{L}\boldsymbol{\chi}=\mathbf{K}\boldsymbol{\Sigma}_b^{1/2}\sum_j \psi_j \otimes \left[\mathbf{C}_j^{1/2}(\lambda,\phi)\chi_j\right]$$

can be made flow-dependent by obtaining flow-dependent estimates of the background error variances ($\boldsymbol{\Sigma}_b$) and correlations ($\mathbf{C}_j(\lambda,\varphi)$) from the EDA background perturbations

Diffusion Operators and Digital Filters

- Spectral/wavelet approaches are efficient and convenient for models with regular (e.g. spherical or rectangular) domains.
- Difficult to use if the domain is not regular (e.g. ocean models).
- Because the spectral approach is based on convolutions, it is difficult to incorporate inhomogeneity and anisotropy.
- Diffusion operators and digital filters provide alternatives to the spectral approach that address these difficulties.

Diffusion Operators

- **The 1-dimensional diffusion equation:**

$$\frac{\partial \eta}{\partial t} - \kappa \frac{\partial^2 \eta}{\partial x^2} = 0$$

- **Has solution at time *T*:**

$$\eta(x,T) = \frac{1}{\sqrt{4\pi\kappa T}} \int_{x'} e^{-(x-x')^2/4\kappa T} \eta(x',0)\, dx'$$

- **That is, $\eta(x,T)$ is the result of convolving $\eta(x,0)$ with the Gaussian function:**

$$\frac{1}{\sqrt{4\pi\kappa T}} \exp\left(-x^2/4\kappa T\right)$$

Diffusion Operators

- The one-dimensional result generalizes to more dimensions, and to different geometries (e.g. on the sphere).
- Weaver and Courtier (2001) realized that numerical integration of a diffusion equation could be used to perform convolutions for covariance modelling.
- Irregular boundary conditions (e.g. coastlines) are easily handled.
- More general partial differential equations can be used to generate a large class of correlation functions:

$$\frac{\partial \eta}{\partial t} + \sum_{p=1}^{P} \kappa_p \left(-\nabla^2 \right)^p \eta = 0$$

Digital Filters

- In one-dimension, convolution with a Gaussian may be achieved, to good approximation, using a pair of recursive filters:

$$q_i = \beta p_i + \sum_{j=1}^{n} \alpha_j q_{i-j}$$

$$s_i = \beta q_i + \sum_{j=1}^{n} \alpha_j s_{i+j}$$

- In two dimensions, the Fourier transform of the Gaussian factorizes:

$$\exp\left(-\frac{a^2\left(k^2+l^2\right)}{2}\right) = \exp\left(-\frac{a^2k^2}{2}\right)\exp\left(-\frac{a^2l^2}{2}\right)$$

  - => 2-D convolution may be achieved by 1-D filtering in the x-direction, and then in the y-direction.

- NB: This factorization only works for Gaussians!

Digital Filters

- Non-Gaussian covariance functions may be produced as a superposition of Gaussians.
    - I.e. the filtered field is the weighted sum of convolutions with a set of Gaussians of different widths.
- Inhomogeneous covariances may be synthesized by allowing the filter coefficients to vary with location.
- Simple anisotropic covariances (ellipses), with different north-south and east-west length scales, can be produced by using different filters in the north-south direction.
- However, fully general anisotropy (bananas) requires 3 independent filters (north-south, east-west, and SW-NE) in 2 dimensions and 6 filters in 3 dimensions.

Digital Filters

- There is a close connection between digital filter methods and diffusion operator methods.
  - One timestep of integration of a diffusion operator can be viewed as one application of a digital filter.
- Advantages of Digital Filters:
  - Computational Efficiency
  - Generality
- Disadvantages:
  - Filter coefficients are difficult to determine from data.
  - Grid geometry, polar singularities and boundary conditions must be handled carefully.

Summary

- A good B matrix is vitally important in any (current) data assimilation system.
- In a large-dimension system, covariances must be modelled: The matrix is too big to specify every element.
- Innovation Statistics are the only real data we have to diagnose background error statistics, but they are difficult to use.
- Analysis ensembles allow us to generate a good surrogate for samples of background error.
- Spectral representations work well for simple geometries (spherical or rectangular domains) but anisotropic and/or inhomogeneous covariances are tricky!
- Wavelet formulation allows inhomogeneous covariances.
- Diffusion operators and digital filters have fewer limitations, but calculating the diffusion/filter coefficients is non-trivial.