AQ Data Assimilation – Methods 2

Richard Ménard
with exercices prepared by Martin Deshaies-Jacques

Air Quality Research Division

Advances in Air Quality Analysis and Prediction
NCAR/ASP Summer Colloquium,
Boulder, August 2, 2016

Outline

1. Covariance modelling (motivation)

Improvement in error covariances

- spectral error covariances
- tuning of the error variances

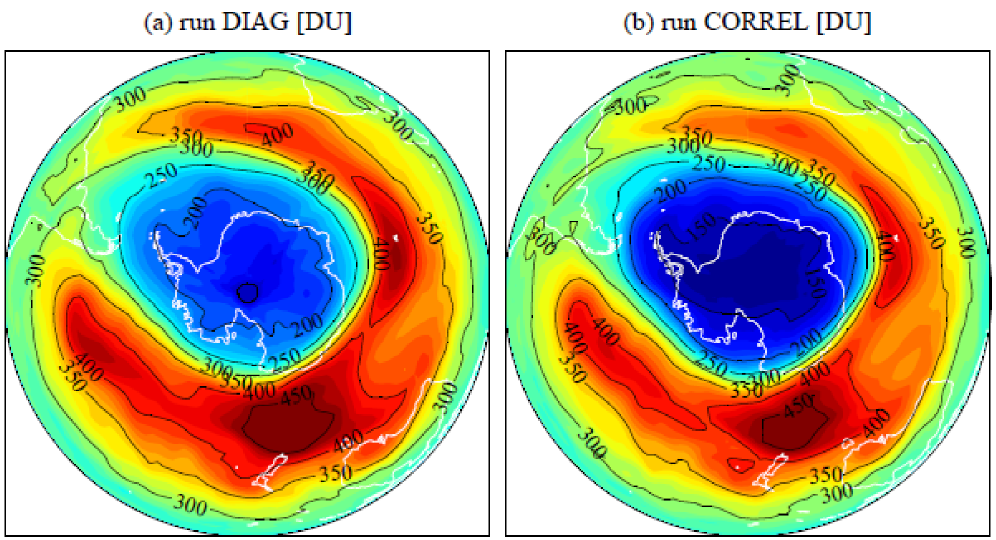


Fig. 6. Total ozone in the Southern Hemisphere on 1 October 2003 obtained by the runs D
Total Ozone Measurement Satellite (TOMS) (c), in Dobson unit (DU).

Errera and Ménard (2012) ACP

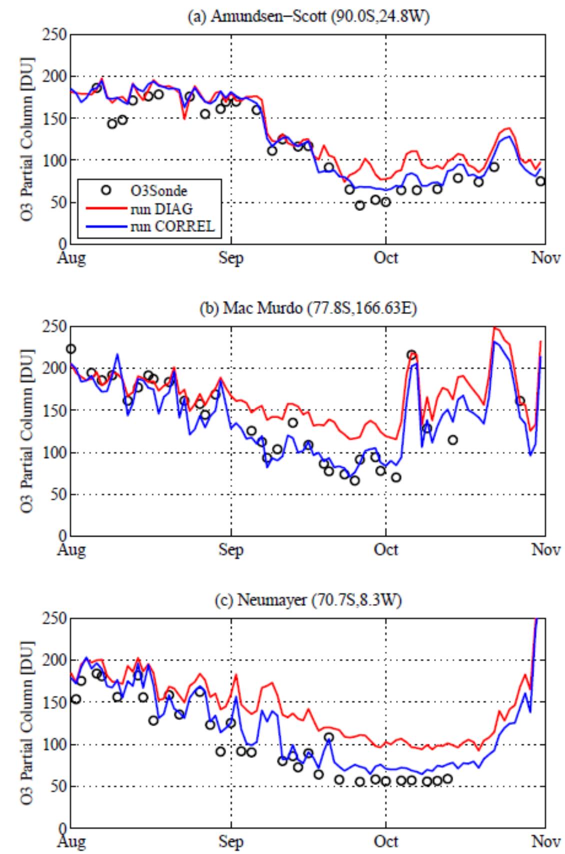
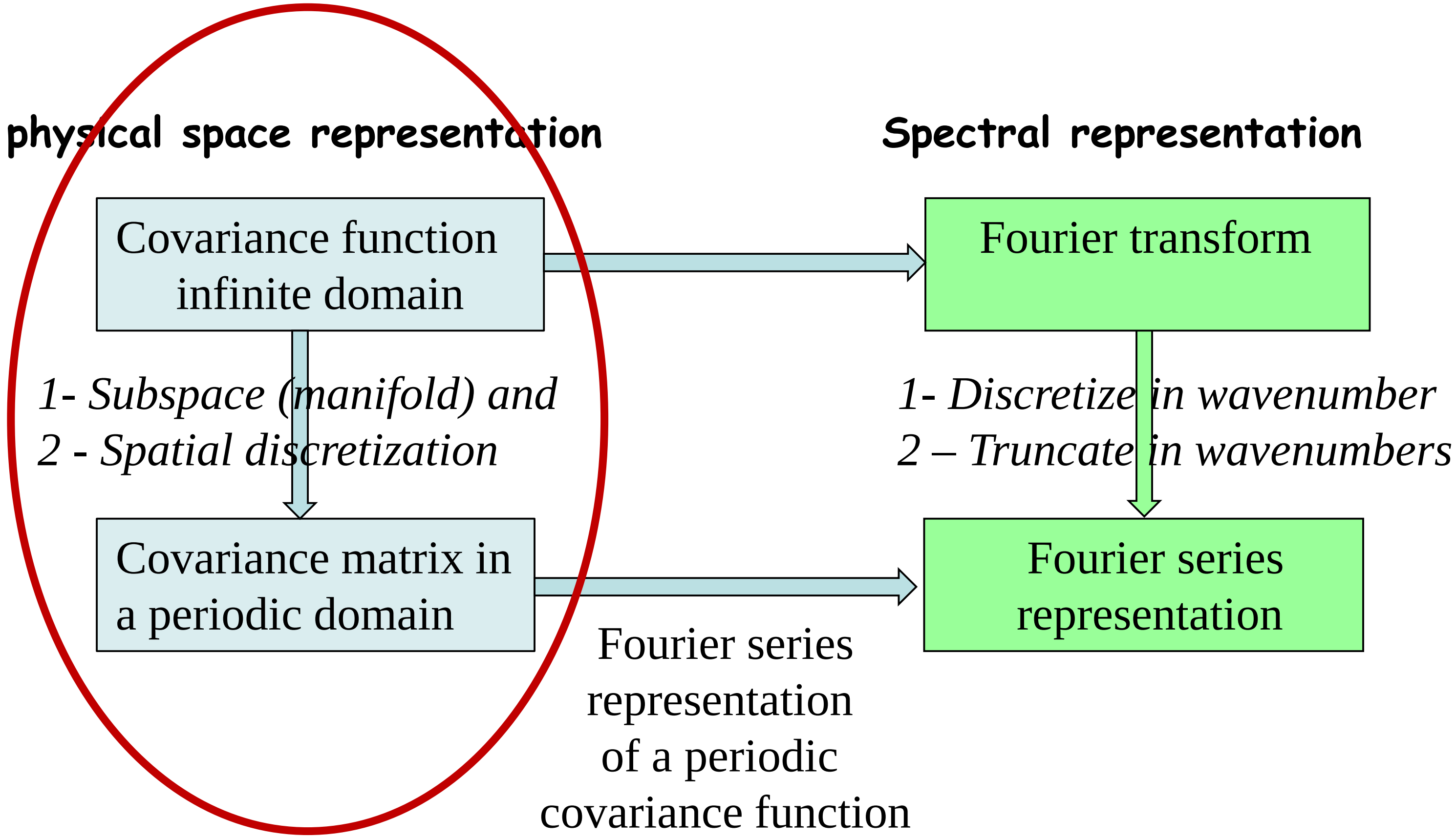


Fig. 5. Time series of the ozone partial column (10–100 hPa) between August and October 2003 obtained above three NDACC stations in Antarctica by the ozone sondes (black circles), and the runs DIAG (red line) and CORREL (blue line), in Dobson unit (DU).

Ways to construct positive-definite covariance matrices



1.1 Covariance functions (Gaspari and Cohn 1999, QJRMS)

Definition 1: A function $P(\mathbf{r},\mathbf{r}')$ is a covariance function of a random field X if

$$P(\mathbf{r},\mathbf{r}') = \left\langle [X(\mathbf{r}) - \langle X(\mathbf{r})\rangle][X(\mathbf{r}') - \langle X(\mathbf{r}')\rangle] \right\rangle$$

Definition 2: A covariance function $P(\mathbf{r},\mathbf{r}')$ is a function that defines positive semi-definite matrices when evaluated on any grid.
That is, letting $\mathbf{r}_i$ and $\mathbf{r}_j$ be any two grid points, the matrix $\mathbf{P}$ whose elements are $\mathbf{P}_{i,j} = P(\mathbf{r}_i,\mathbf{r}_j)$ is defines a covariance matrix, when P is a covariance function

The equivalence between definition 1 and 2 is a result of the *reproducing kernel* property of covariance functions (Rasmussen and Williams, 2006: *Gaussian Processes of Machine Learning*.)

<u>Remark</u> Suppose a covariance function is defined in a 3D space, $\mathbf{r} \in R^3$.
Restricting the value of $\mathbf{r}$ to remain on an manifold (e.g. the surface of a unit sphere) will also define a covariance function, and a covariance matrix (e.g. a covariance matrix on the surface of a sphere)

<u>Correlation function</u> A correlation function $C(\mathbf{r},\mathbf{r}')$ is a covariance function $P(\mathbf{r},\mathbf{r}')$ normalized by the standard deviation at the points **r** and **r**'

$$C(\mathbf{r},\mathbf{r}') = \frac{P(\mathbf{r},\mathbf{r}')}{\sqrt{P(\mathbf{r},\mathbf{r})}\,\sqrt{P(\mathbf{r}',\mathbf{r}')}}$$

<u>Homogeneous and isotropic correlation function</u> If a correlation function is invariant under all translation and all orthogonal transformation, then the correlation function become only a function of the distance between the two points,

$$C(\mathbf{r},\mathbf{r}') = C_0(\|\mathbf{r}-\mathbf{r}'\|)$$

Smoothness properties

- The continuity at the origin determines the continuity allowed on the rest of the domain. For example, if the first derivative is discontinuous at the origin, then first derivative discontinuity is allowed elsewhere (see example with triangle)

Examples of correlation functions (infinite domain)

1. Spatially uncorrelated model (black)

$$C_0(\|\mathbf{r}-\mathbf{r}'\|) = \begin{cases} 1 \text{ if } \mathbf{r}=\mathbf{r}' \\ 0 \text{ if } \mathbf{r}\neq\mathbf{r}' \end{cases}$$

2. First-order auto-regressive model (FOAR) (blue)

$$C_0(\|\mathbf{r}-\mathbf{r}'\|) = \exp\left(-\frac{\|\mathbf{r}-\mathbf{r}'\|}{p_{FOAR}}\right)$$

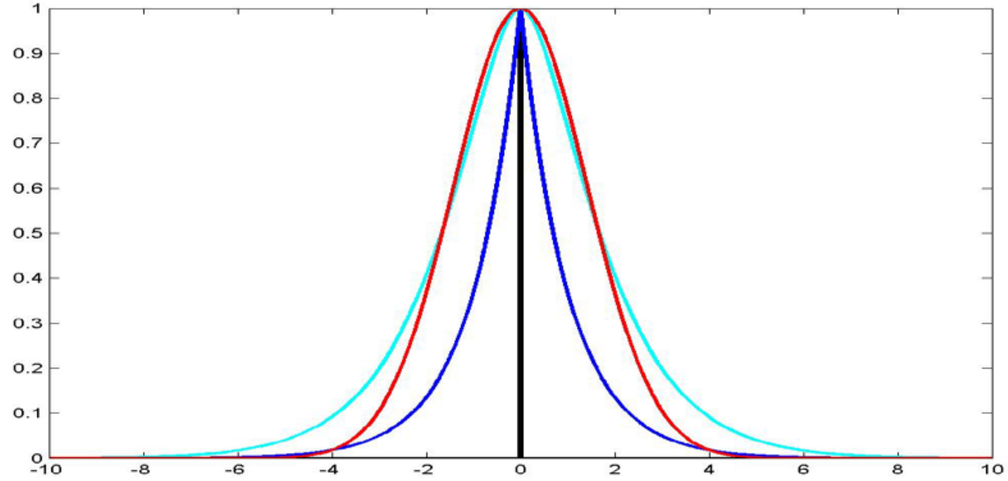
where L is the correlation length scale

3. Second-order auto-regressive model (SOAR) (cyan)

$$C_0(\|\mathbf{r}-\mathbf{r}'\|) = \left(1+\frac{\|\mathbf{r}-\mathbf{r}'\|}{p_{SOAR}}\right)\exp\left(-\frac{\|\mathbf{r}-\mathbf{r}'\|}{p_{SOAR}}\right)$$

4. Gaussian model (red)

$$C_0(\|\mathbf{r}-\mathbf{r}'\|) = \exp\left(-\frac{\|\mathbf{r}-\mathbf{r}'\|^2}{2\,p_G^2}\right)$$



1.2 Covariance matrices

Positive definite matrix (Horn and Johnson ,1985: *Matrix Analysis*, Chap 7)

A real *n* x *n* symmetric matrix **A** is positive definite if

$$\mathbf{c}^T \mathbf{A} \mathbf{c} > 0$$

for any nonzero vector **c**. **A** is said to be positive semi-definite if

$$\mathbf{c}^T \mathbf{A} \mathbf{c} \geq 0$$

Properties

- The sum of any positive definite matrices of the same size is also positive definite
- Each eigenvalue of a positive definite matrix is a positive real number
- The trace and determinant are positive real numbers.

Covariance matrix

The covariance matrix **P** of a random vector $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$ is the matrix $\mathbf{P} = [P_{ij}]$ in which $P_{ij} = \mathbf{E}\left[(X_i - \overline{X}_i)(X_j - \overline{X}_j)\right]$ where $\overline{X}_i = \mathbf{E}\left[X_i\right]$ and **E** is the mathematical expectation.

Property: A covariance matrix is positive semi-definite

$$\mathbf{E}\left[\left(c_1(X_1-\overline{X}_1)+\cdots+c_n(X_n-\overline{X}_n)\right)^2\right]=\mathbf{E}\left[\sum_{i,j=1}^{n}c_i(X_i-\overline{X}_i)c_j(X_j-\overline{X}_j)\right]$$

$$=\sum_{i,j=1}^{n}c_i\mathbf{E}\left[(X_i-\overline{X}_i)(X_j-\overline{X}_j)\right]c_j \;=\mathbf{c}^T\mathbf{P}\,\mathbf{c}\geq 0$$

Remarks

1 - It is often necessary in data assimilation to invert the covariance matrices, and thus we need to have *positive definite* covariances

2 – The positive definite property is global property of a matrix, and it is not trivial to obtain

Correlation matrix

Create a diagonal matrix $\mathbf{\Sigma}$ of error standard deviation $\mathbf{\Sigma}=\sqrt{\mathrm{diag}\,(\mathbf{P})}$

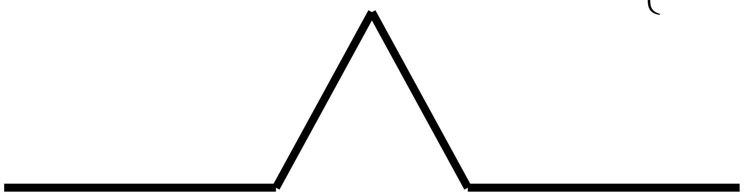
then a correlation matrix $\mathbf{C}$ is related to $\mathbf{P}$ as follows

$$\mathbf{P}=\mathbf{\Sigma}\mathbf{C}\mathbf{\Sigma}$$

$$\mathbf{C}=\mathbf{\Sigma}^{-1}\mathbf{P}\,\mathbf{\Sigma}^{-1}$$

b) Triangle

$$C(i,j) = \begin{cases} 1 - \dfrac{|i-j|}{d} & for\ |i-j| \le n \\ 0 & otherwise \end{cases}$$



for $n=4$

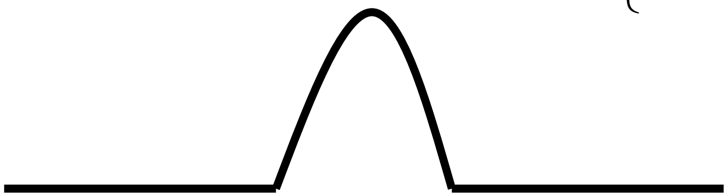
$$\begin{bmatrix} 1.000 & 0.750 & 0.500 & 0.250 & 0.000 \\ 0.750 & 1.000 & 0.750 & 0.500 & 0.250 \\ 0.500 & 0.750 & 1.000 & 0.750 & 0.500 \\ 0.250 & 0.500 & 0.750 & 1.000 & 0.750 \\ 0.000 & 0.250 & 0.500 & 0.750 & 1.000 \end{bmatrix}$$

eigenvalues

3.0646
1.3090
0.2989
0.1910
0.1365

Examples:

a) Truncated parabola $C(i, j) = \begin{cases} 1 - \dfrac{(i-j)^2}{d^2} & for\ |i-j| \le n \\ 0 & otherwise \end{cases}$



for $n=4$

$$\begin{bmatrix} 1.000 & 0.937 & 0.750 & 0.437 & 0.000 \\ 0.937 & 1.000 & 0.937 & 0.750 & 0.437 \\ 0.750 & 0.937 & 1.000 & 0.937 & 0.750 \\ 0.437 & 0.750 & 0.937 & 1.000 & 0.937 \\ 0.000 & 0.437 & 0.750 & 0.937 & 1.000 \end{bmatrix}$$

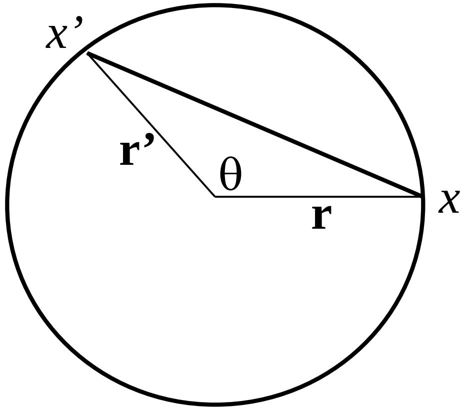
eigenvalues

$$\begin{matrix} 3.8216 \\ 1.2500 \\ 0.0000 \\ 0.0000 \\ -0.0716 \end{matrix}$$

Correlation function on a periodic domain (sub-domain / manifold approach)

Consider a Gaussian model on a 2D plane

$$C(\mathbf{r},\mathbf{r}') = \exp\left(-\frac{\|\mathbf{r}-\mathbf{r}'\|^2}{2\,p_G^2} \right) \quad \text{where } \mathbf{r} \in R^2$$



Any subdomain or manifold of R^2 can also be a domain to define a covariance function.

Consider as a manifold of R^2, a circle of radius a

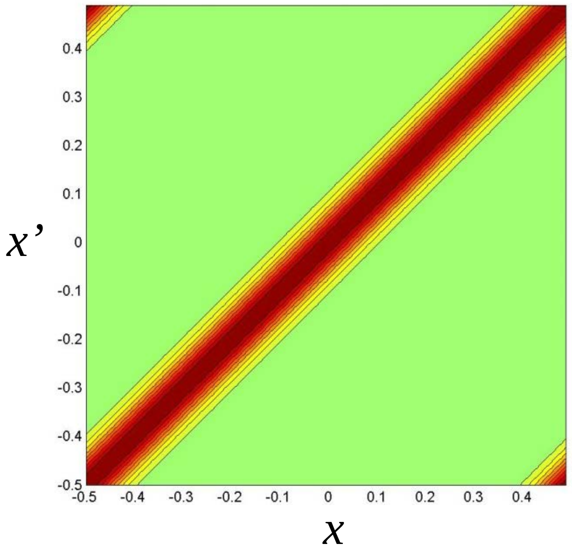
$$\|\mathbf{r}-\mathbf{r}'\|^2 = 2a^2(1-\cos\theta)$$

Now define a coordinate x along the circle ,

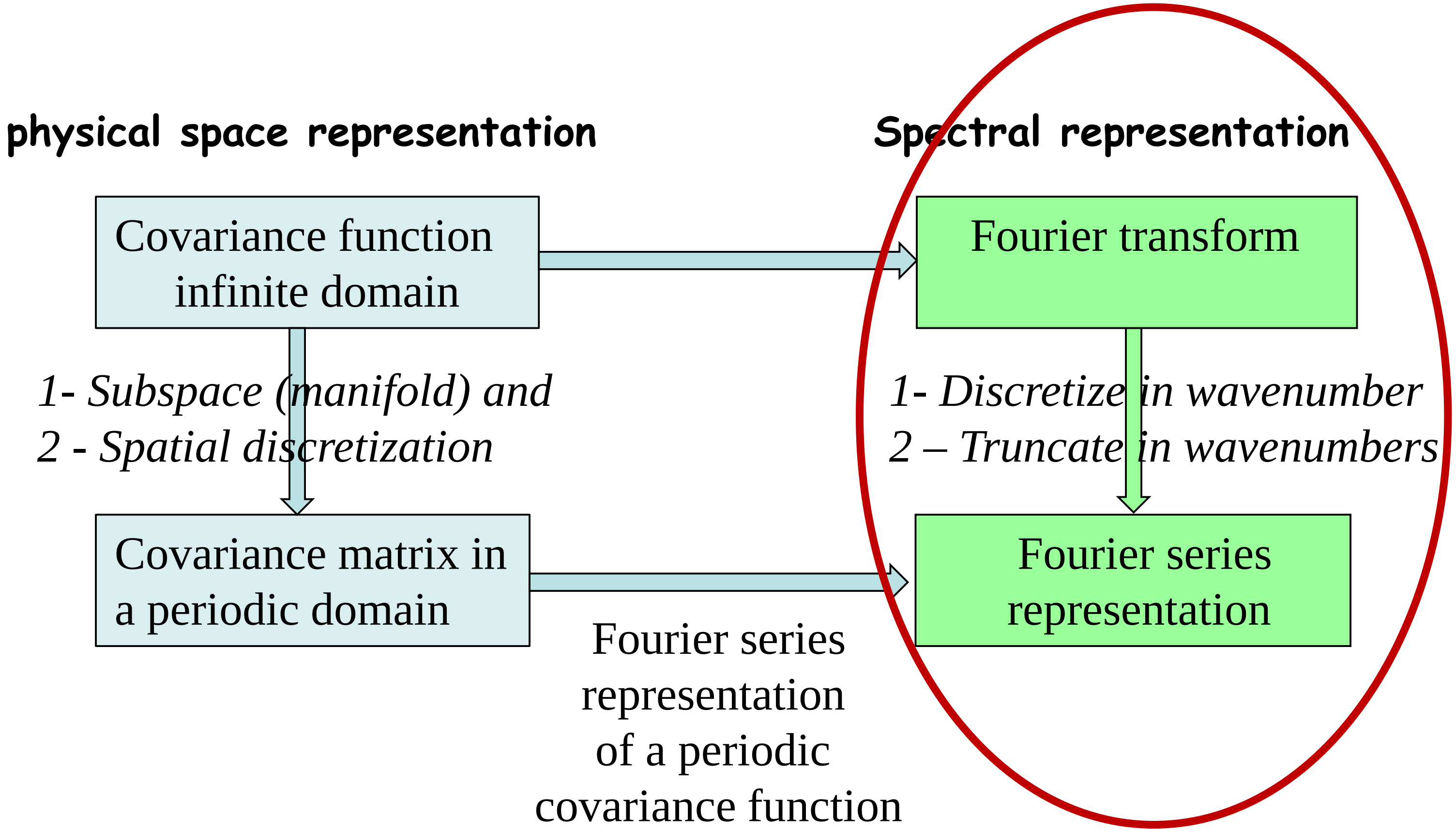
$$x' - x = \frac{a\theta}{2\pi} \quad \text{for } -a \le x, x' \le a$$

then we get

$$C(x,x') = \exp\left(-\frac{(1-\cos[2\pi\,(x-x')/a])}{(p_G/a)^2} \right)$$



Ways to construct positive-definite covariance matrices



1.3 Spectral representation infinite domain

$$C(x) = \sqrt{\frac{2}{\pi}} \int_0^\infty \hat{C}(m) \cos(mx)\, dm$$

$$\hat{C}(m) = \sqrt{\frac{2}{\pi}} \int_0^\infty C(m) \cos(mx)\, dx$$

Gaussian model

$$C(x) = \exp\left(-\frac{x^2}{2p_G^2}\right) \qquad \hat{C}(m) = c_1 \exp\left(-\frac{m^2 p_G^2}{2}\right)$$

First order autoregressive model (FOAR)

$$C(x) = \exp\left(-\frac{|x|}{p_{FOAR}}\right) \qquad \hat{C}(m) = \frac{c_2}{1 + m^2 p_{FOAR}^2}$$

Second order autoregressive model (SOAR)

$$C(x) = \left(1 + \frac{|x|}{p_{SOAR}}\right) \exp\left(-\frac{|x|}{p_{SOAR}}\right) \qquad \hat{C}(m) = \frac{c_3}{\left(1 + m^2 p_{SOAR}^2\right)^2}$$

To obtain a spectra for finite domain we set

$$\gamma_k = \hat{C}\left(\frac{2\pi k}{L} = m\right)$$

← Discretize in wavenumber space for a Fourier series representation

1.4 Spectral representation over a periodic domain

Spectral representation of homogeneous isotropic correlations

- On a unit circle

$$C(\|\mathbf{r}-\mathbf{r}'\|) = \sum_{m=0}^{\infty} a_m \cos(m\theta)$$

where θ is the angle between the two position vectors, and where and all the Fourier coefficients a_m are nonnegative

- On a unit sphere

$$C(\|\mathbf{r}-\mathbf{r}'\|) = \sum_{m=0}^{\infty} b_m P_m(\cos\theta)$$

where all the Legendre coefficients b_m are nonnegative.

Consider the complex notation of Fourier series

$$\varphi(x_j) = \sum_k c_k\, e^{2\pi i k x_j / L}$$

The spatial covariance between x_j and $x_{j-h} = x_j - h\,\Delta x$

$$C(x_j, x_{j-h}) = \mathbf{E}[\varphi(x_j)\,\overline{\varphi(x_{j-h})}]$$

$$= \sum_{k,k'} \mathbf{E}[c_k \bar{c}_{k'}]\, e^{2\pi i[k x_j - k'(x_j - h\Delta x)]} = \sum_{k,k'} \mathbf{E}[c_k \bar{c}_{k'}]\, e^{2\pi i[(k-k')x_j + k' h\Delta x)]}$$

For the covariance to depend only on $x_j - x_{j-h} = h\,\Delta x$ we need $\mathbf{E}[c_k \bar{c}_{k'}] = \delta_{k,k'} \mathbf{E}[|c_{k'}|^2]$
i.e. *the Fourier modes are uncorrelated*

Real functions – unitary Fourier series

Fourier series representation over a periodic domain L, using $2N$+1 grid points

$$x_j = \frac{L\,j}{2N+1} \qquad (j = -N, \ldots, 0, \ldots, N)$$

Discrete Fourier series, leading to a unitary matrix, i.e. $\mathbf{F}^{-1} = \mathbf{F}^T$

$$\phi(x_j) = \sqrt{\frac{2}{2N+1}}\left\{\frac{a_0}{\sqrt{2}} + \sum_{k=1}^{N} a_k \cos\left(\frac{2\pi k x_j}{L}\right) + \sum_{k=1}^{N} b_k \sin\left(\frac{2\pi k x_j}{L}\right)\right\}$$

and, in matrix form

$$\begin{pmatrix} \phi(x_{-N}) \\ \cdots \\ \phi(x_0) \\ \cdots \\ \phi(x_N) \end{pmatrix} = \begin{pmatrix} C_{-N,0} & C_{-N,1} & S_{-N,1} & \cdots & \cdots & C_{-N,N} & S_{-N.N} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ C_{0,0} & C_{0,1} & S_{0,1} & \cdots & \cdots & C_{0,N} & S_{0,N} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ C_{N,0} & C_{N,1} & S_{N,1} & \cdots & \cdots & C_{N,N} & S_{N,N} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ b_1 \\ \cdots \\ \cdots \\ a_N \\ b_N \end{pmatrix}$$

where $C_{j,0} = \frac{1}{\sqrt{2N+1}}$; $C_{j,k} = \sqrt{\frac{2}{2N+1}}\cos\left(\frac{2\pi k x_j}{L}\right)$; $S_{j,k} = \sqrt{\frac{2}{2N+1}}\sin\left(\frac{2\pi k x_j}{L}\right)$

- **F** is a Fourier matrix
- **F** it is written here as a unitary matrix, $\mathbf{F}^T = \mathbf{F}^{-1}$

$$\boldsymbol{\varphi} = \mathbf{F}\,\hat{\boldsymbol{\varphi}}$$

How to construct an homogeneous isotropic correlation function in spectral space

Assume uncorrelated modes, i.e.

$$\left.\begin{array}{l}\langle a_i a_j \rangle\\ \langle b_i a_j \rangle\\ \langle b_i b_j \rangle\end{array}\right\} = 0 \quad \text{for } i \neq j$$

then the covariance matrix of the spectral coefficient, $\hat{\mathbf{C}} = \langle \hat{\boldsymbol{\varphi}} \hat{\boldsymbol{\varphi}}^T \rangle$ becomes block-diagonal

$$\hat{\mathbf{C}} = \begin{pmatrix} \hat{\mathbf{C}}_0 & 0 & 0 & 0 \\ 0 & \hat{\mathbf{C}}_1 & 0 & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \hat{\mathbf{C}}_N \end{pmatrix}$$

where

$$\hat{\mathbf{C}}_i = \begin{pmatrix} \langle a_i^2 \rangle & \langle a_i b_i \rangle \\ \langle b_i a_i \rangle & \langle b_i^2 \rangle \end{pmatrix}$$

Transforming in physical space

$$C(x_1,x_2)=\frac{2}{2N+1}\left\{\frac{\langle a_0 a_0\rangle}{2}+\sum_{k=1}^{N}\frac{\langle a_k^2\rangle}{2}\left[\cos\left(\frac{2\pi k(x_1+x_2)}{L}\right)+\cos\left(\frac{2\pi k(x_1-x_2)}{L}\right)\right]\right.$$

$$+\sum_{k=1}^{N}\langle a_k b_k\rangle\sin\left(\frac{2\pi k(x_1+x_2)}{L}\right)$$

$$\left.+\sum_{k=1}^{N}\frac{\langle b_k^2\rangle}{2}\left[\cos\left(\frac{2\pi k(x_1-x_2)}{L}\right)-\cos\left(\frac{2\pi k(x_1+x_2)}{L}\right)\right]\right\}$$

A homogeneous covariance only depends on distance $C(x_1,x_2)=C(|x_1-x_2|,0)$
To obtain such a model we assume

$$\langle a_k^2\rangle=\langle b_k^2\rangle=\gamma_k$$

$$\langle a_k b_k\rangle=0$$

$$\hat{\mathbf{C}}=\begin{pmatrix} a_0^2 & & & & & & \\ & a_1^2 & & & & & \\ & & a_1^2 & & & & \\ & & & a_2^2 & & & \\ & & & & a_2^2 & & \\ & & & & & \ddots & \\ & & & & & & \ddots \end{pmatrix}$$

1.5 Length-scales / smoothness / realizations

Correlation length vs correlation length-scale parameter

Correlation length based on the curvature of the correlation function at the origin *x* = 0 (Daley, 1991)

$$C''(x)\big|_{x=0} = -\frac{1}{L_c^2}$$

for a Gaussian correlation model

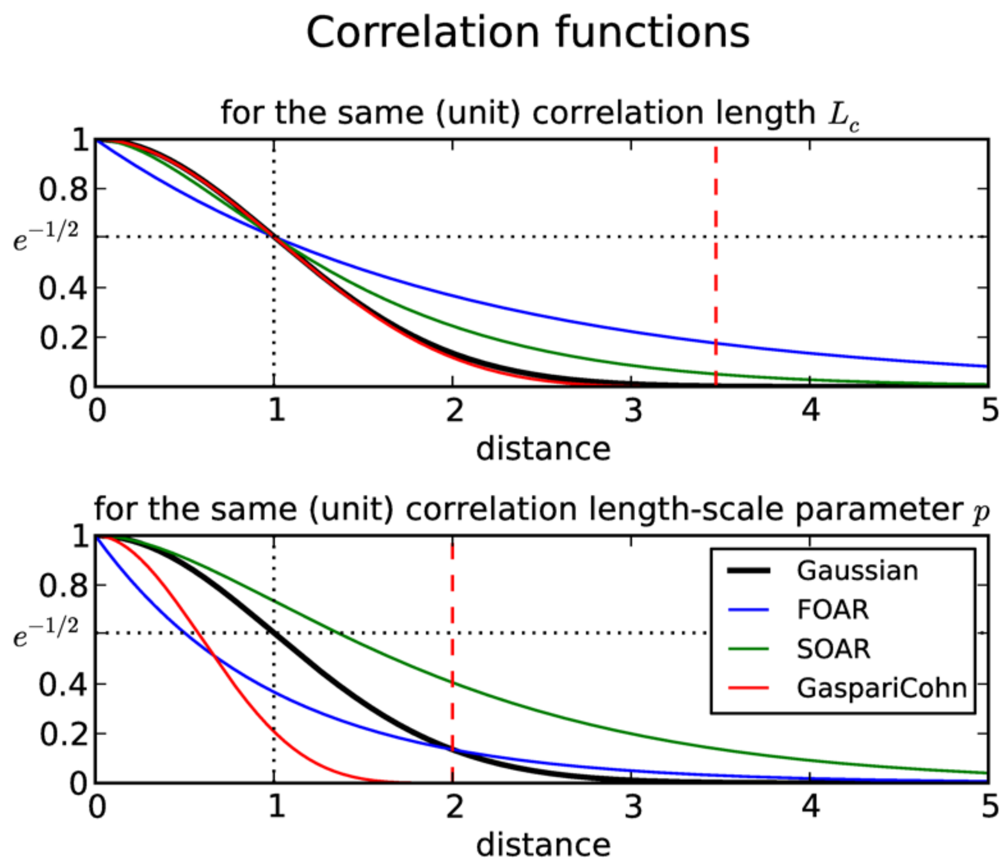
$$C_G(x) = \exp\left(-\frac{x^2}{2p_G^2}\right)$$

we have $p_G = L_c$

so

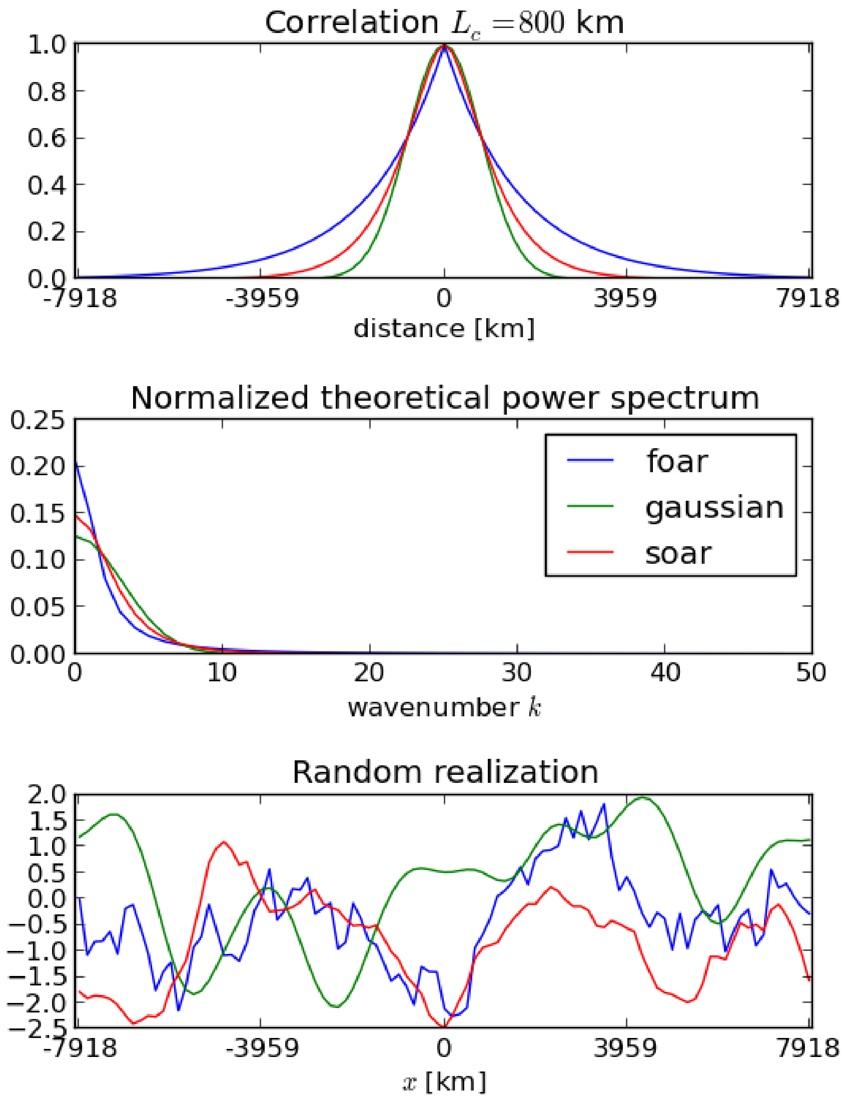
$$C_G(x = L_c) = e^{-1/2} \approx 0.606$$

We define the correlation length of an arbitrary correlation model as the distance for which the correlation reaches a value of 0.606



$$L_c = 0.5005\, p_{FOAR} \quad L_c = 1.3494\, p_{SOAR} \quad L_c = 0.5756\, p_{GC} \quad L_c = p_G$$

Different correlation models means different smoothness of the underlying stochastic realization



Sample correlation from an ensemble of stochastic realization

The spectral decomposition of a correlation matrix $\mathbf{C}$ (in 1D periodic domain) is given by

$$\mathbf{C} = \mathbf{F}\,\hat{\mathbf{C}}\,\mathbf{F}^T$$

where

$$\hat{\mathbf{C}} = \mathrm{diag}(c^2(0), c^2(1), c^2(1), c^2(2), c^2(2), \ldots)$$

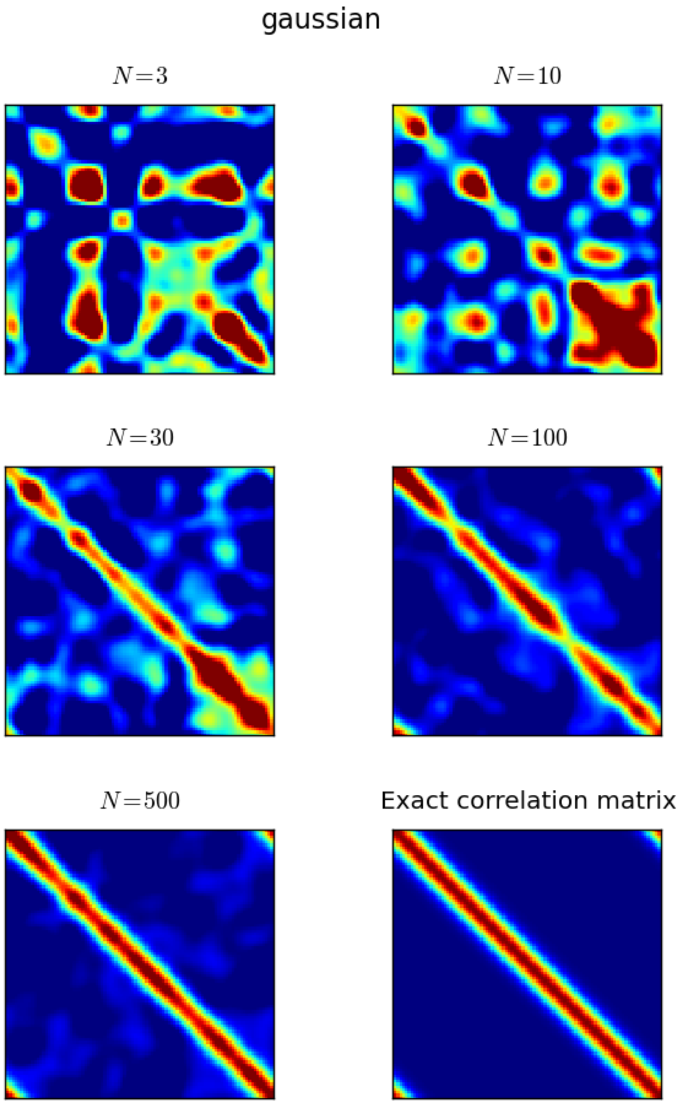
is a diagonal matrix with repeated eigenvalues. Random spatially correlated perturbations $\delta\mathbf{c}$ can be obtained by

$$\delta\mathbf{c} = \sum_{k=1}^{2N+1} d_k c(k)\mathbf{e}_k = \mathbf{F}\hat{\mathbf{C}}^{1/2}\mathbf{d}$$

where $\mathbf{e}_k$ are the columns (or eigenvectors) of $\mathbf{F}$ and $d_k \sim N(0, I)$ are uncorrelated normally distributed random variables, $\mathbf{d} = [d_1, d_2, \ldots, d_{2N+1}]$

So

$$\mathbf{E}[\delta\mathbf{c}(\delta\mathbf{c})^T] = \mathbf{F}\,\hat{\mathbf{C}}^{1/2}\,\mathbf{E}[\mathbf{d}\mathbf{d}^T]\,\hat{\mathbf{C}}^{1/2}\mathbf{F}^T = \mathbf{F}\,\hat{\mathbf{C}}\,\mathbf{F}^T = \mathbf{C}$$



Python script: sampleCorrelations.py

1.6 Spectral representation on a sphere

(Errera and Ménard, 2012, ACP)

Spherical harmonics

$$\psi(\lambda,\varphi) = \sum_{n=0}^{N} \sum_{m=-n}^{n} \hat{\psi}_n^m Y_n^m(\lambda,\varphi) = \sum_{\alpha} \hat{\psi}_\alpha Y_\alpha(S)$$

where $\alpha = (n,m)$ and $S = (\lambda,\varphi)$

$$Y_\alpha(S) = P_n^m(\mu)\exp(im\lambda)$$

$\mu = \sin\varphi$

Orthogonality

$$\frac{1}{2\pi}\int_0^{2\pi}\int_{-1}^{1} Y_n^m(\lambda,\mu) Y_{n'}^{m'}(\lambda,\mu)\, d\lambda\, d\mu = \iint_S Y_\alpha(S) Y_{\alpha'}^*(S)\, dS = \delta_n^{n'}\delta_m^{m'}$$

Addition theorem

$$b_n P_n(x) = \sum_{m=-n}^{n} Y_n^m(S_1) Y_n^{-m}(S_2) \qquad \text{where } b_n = \left[(2n+1)/2\right]^{1/2}$$

$$\iint_S b_{n'} P_{n'}(x) Y_n^m(S_1)\, dS_1 = \delta_n^{n'} Y_n^m(S_2)$$

Covariance

$$C(S_1,S_2) = \sum_{\alpha}\sum_{\alpha'} \left\langle \psi_\alpha \psi_{\alpha'}^* \right\rangle Y_\alpha(S_1) Y_\alpha^*(S_2)$$

$$\left\langle \psi_\alpha \psi_{\alpha'}^* \right\rangle = \iint_{S_1}\iint_{S_2} C(S_1,S_2) Y_\alpha^*(S_1) Y_{\alpha'}(S_2)\, dS_1\, dS_2$$

If a random field is homogeneous and isotropic on a sphere, the covariance function depends only on the geodesic distance between the points, and hence can be expressed as a Legendre series of the form

$$C(S_1, S_2) = C(a) = \sum_n \hat{C}_n P_n(x)$$

where $x = \cos(a)$

$$\langle \psi_\alpha \psi^*_{\alpha'} \rangle = \sum_n \hat{C}_n \iint_{S_1} \iint_{S_2} P_n(x) Y^*_\alpha(S_1) Y_{\alpha'}(S_2)\, dS_1\, dS_2$$

$$= \delta_\alpha^{\alpha'} b_n^{-1} \hat{C}_n = \begin{cases} 0 \text{ if } n \neq n' \; m \neq m' \\ \hat{C}_n / b_n = c_n \end{cases}$$

1 3 5

$$\begin{pmatrix} c_0 &&&&&&&& \\ & c_1 &&&&&&& \\ && c_1 &&&&&& \\ &&& c_1 &&&&& \\ &&&& c_2 &&&& \\ &&&&& c_2 &&& \\ &&&&&& c_2 && \\ &&&&&&& c_2 & \\ &&&&&&&& c_2 \end{pmatrix}$$

1.7 3D covariance models

A horizontally homogeneous isotropic model in 3D takes the form

$$B(S_1, z_1, S_2, z_2) = B(a, z_1, z_2)$$

- A *separable* (horizontal-vertical) model can be obtained by assuming

$$B(a, z_1, z_2) = B(a)\,C(z_1, z_2) = \left(\sum_n \hat{B}_n P_n(x)\right) C(z_1, z_2)$$

- A *non-separable* covariance model (Bartello and Mitchell, 1992, *Tellus*) can be obtained by attaching a different vertical correlation function as a function of the wavenumber n

$$B(a, z_1, z_2) = \sum_n C_n(z_1, z_2)\, P_n(x)$$

C_i are $p \times p$ vertical correlation matrices

1

3

5

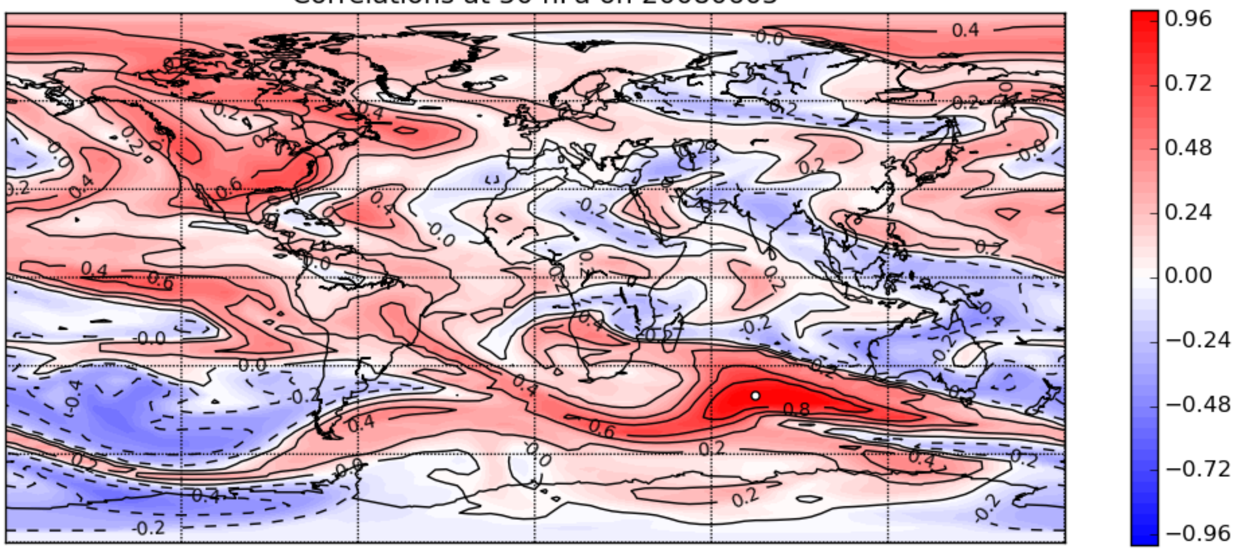
$$\begin{pmatrix} C_0 & & & & & & & & \\ & C_1 & & & & & & & \\ & & C_1 & & & & & & \\ & & & C_1 & & & & & \\ & & & & C_2 & & & & \\ & & & & & C_2 & & & \\ & & & & & & C_2 & & \\ & & & & & & & C_2 & \\ & & & & & & & & C_2 \end{pmatrix}$$

Sample correlation with a 3D chemical transport model

- **Sample from initial homogeneous isotropic correlation model**
- **Correlation after 4 day of transport**

N = 20

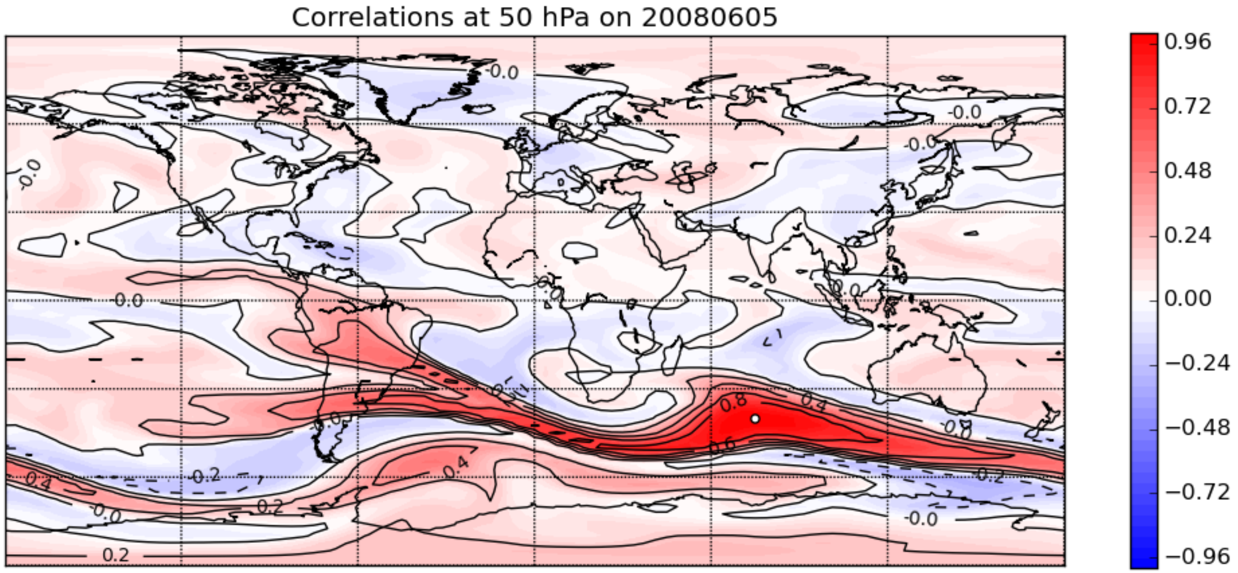
Correlations at 50 hPa on 20080605



Sample correlation with a 3D chemical transport model

- **Sample from initial homogeneous isotropic correlation model**
- **Correlation after 4 day of transport**

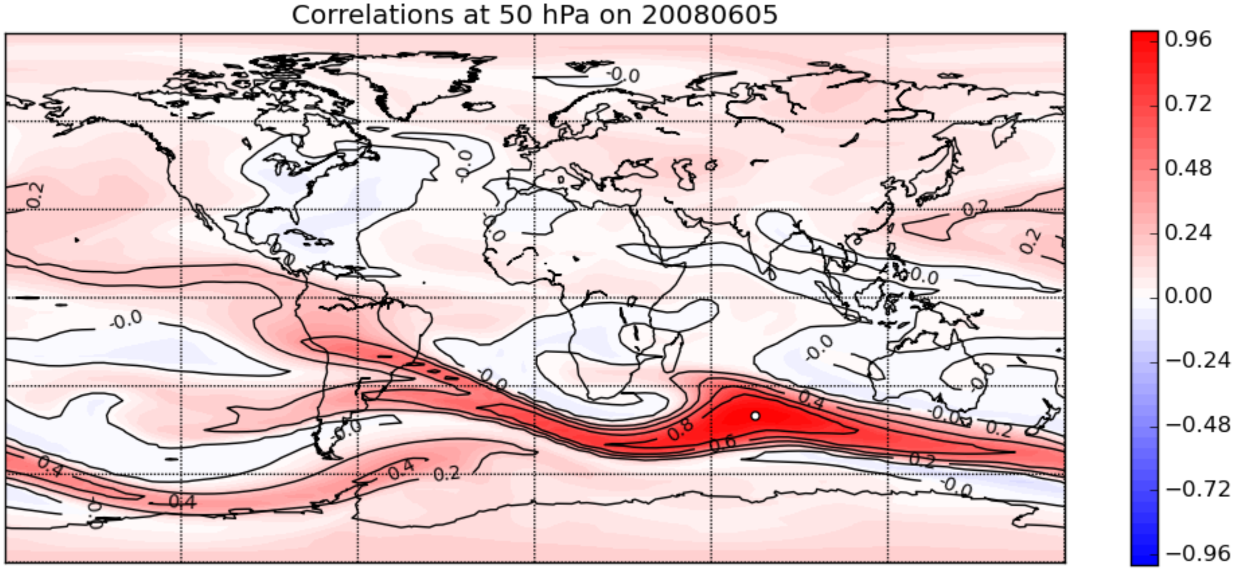
N = 100



Sample correlation with a 3D chemical transport model

- **Sample from initial homogeneous isotropic correlation model**
- **Correlation after 4 day of transport**

N = 500



2. Other ways to look at what is an analysis

2.1 Geometric view (Hilbert space)

Define an inner product of two random vectors **X** , **Y** as

$$\langle \mathbf{X}, \mathbf{Y} \rangle = \mathbf{E}\left[\mathbf{X}^T \mathbf{Y}\right]$$

A distance or 2-*norm* can be defined as

$$\|\mathbf{X}\|_2 = \mathbf{E}\left[\mathbf{X}^T \mathbf{X}\right]$$

2. Other ways to look at what is an analysis

2.1 Geometric view (Hilbert space)

Define an inner product of two random vectors **X** , **Y** as

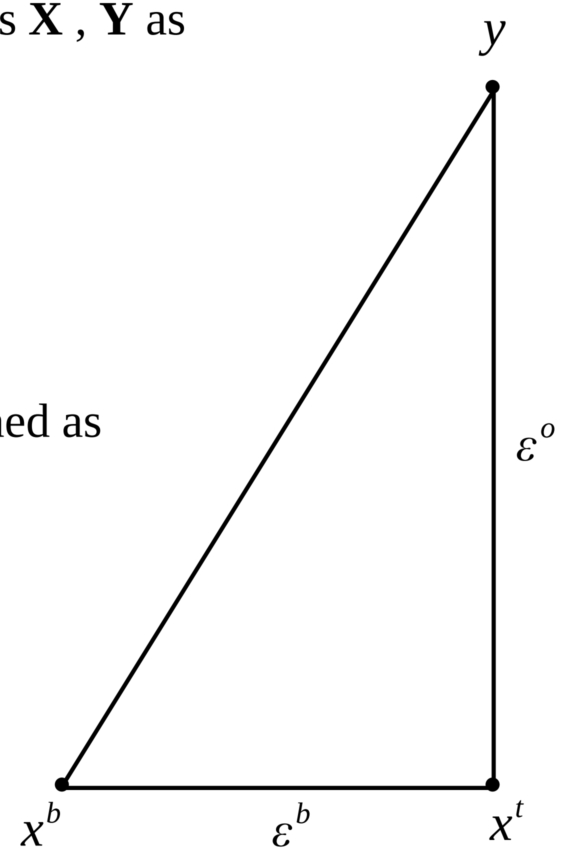
$$\langle \mathbf{X}, \mathbf{Y} \rangle = \mathbf{E}\left[\mathbf{X}^T \mathbf{Y}\right]$$

A distance or 2-*norm* can be defined as

$$\|\mathbf{X}\|_2 = \mathbf{E}\left[\mathbf{X}^T \mathbf{X}\right]$$

The observation and background error is defined as

$$y - x^t = \varepsilon^o$$
$$x^b - x^t = \varepsilon^b$$



2. Other ways to look at what is an analysis

2.1 Geometric view (Hilbert space)

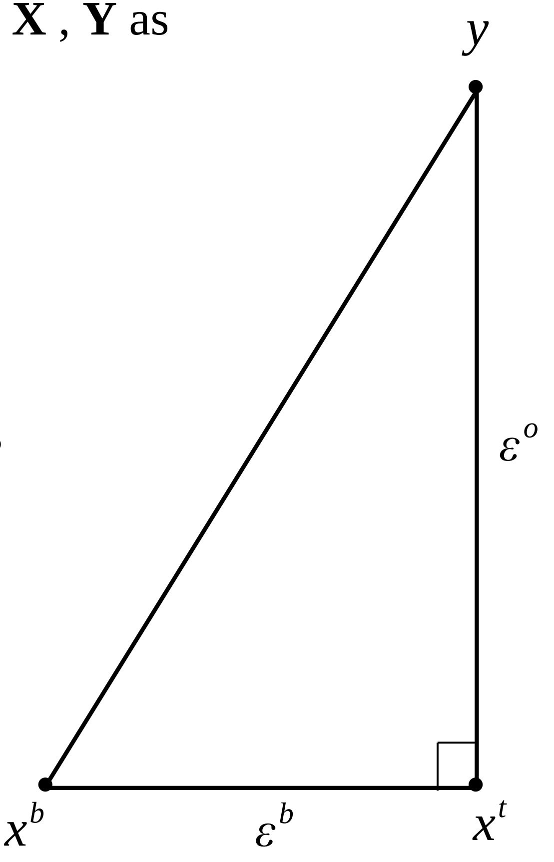
Define an inner product of two random vectors **X** , **Y** as

$$\langle \mathbf{X}, \mathbf{Y} \rangle = \mathbf{E}\left[\mathbf{X}^T \mathbf{Y}\right]$$

A distance or 2-*norm* can be defined as

$$\|\mathbf{X}\|_2 = \mathbf{E}\left[\mathbf{X}^T \mathbf{X}\right]$$

Assume the observation and background errors are uncorrelated $\mathbf{E}\left[\varepsilon^o \varepsilon^b\right] = 0$ then $\varepsilon^o \perp \varepsilon^b$



2. Other ways to look at what is an analysis

2.1 Geometric view (Hilbert space)

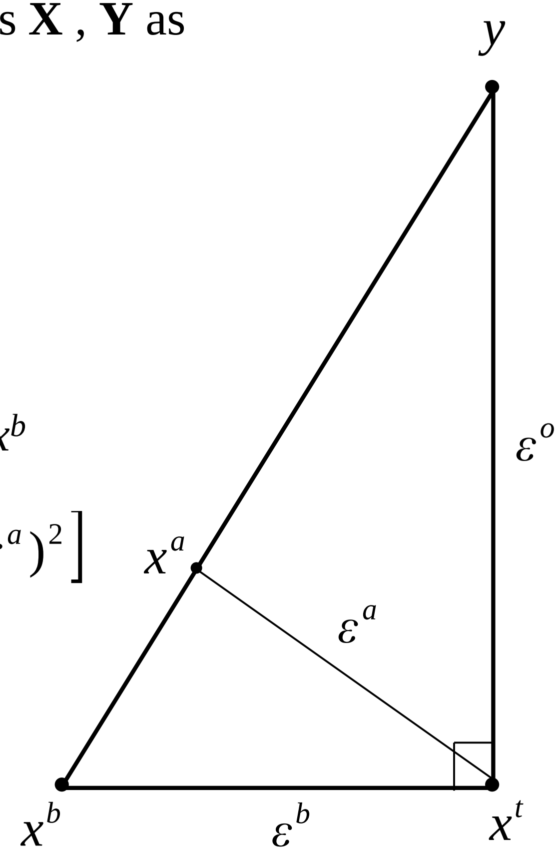
Define an inner product of two random vectors **X** , **Y** as

$$\langle \mathbf{X}, \mathbf{Y} \rangle = \mathbf{E}\left[\mathbf{X}^T \mathbf{Y}\right]$$

A distance or 2-*norm* can be defined as

$$\|\mathbf{X}\|_2 = \mathbf{E}\left[\mathbf{X}^T \mathbf{X}\right]$$

An analysis is a linear combination of y and x^b
It lies on the line y to x^b and
such that it minimizes the analysis error $\mathbf{E}\left[(\varepsilon^a)^2\right]$



2. Other ways to look at what is an analysis

2.1 Geometric view (Hilbert space)

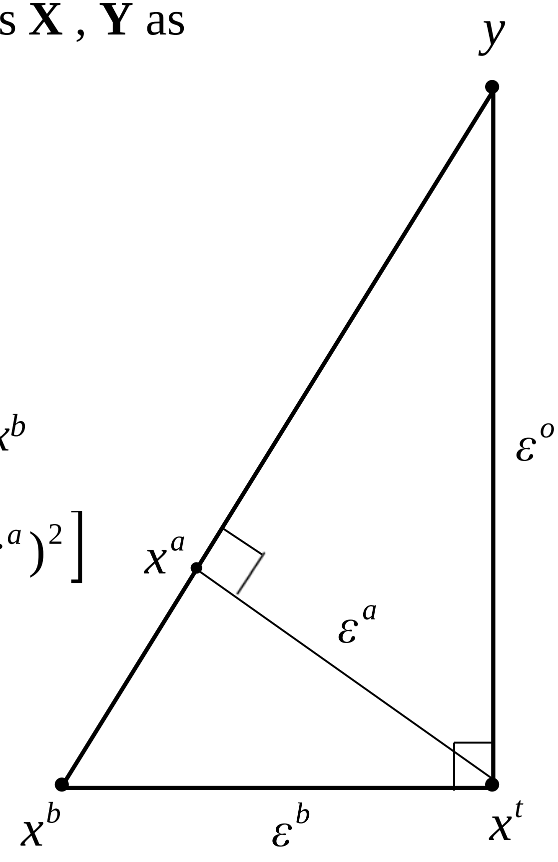
Define an inner product of two random vectors **X** , **Y** as

$$\langle \mathbf{X}, \mathbf{Y} \rangle = \mathbf{E}\left[\mathbf{X}^T \mathbf{Y}\right]$$

A distance or 2-*norm* can be defined as

$$\|\mathbf{X}\|_2 = \mathbf{E}\left[\mathbf{X}^T \mathbf{X}\right]$$

An analysis is a linear combination of y and x^b
It lies on the line y to x^b and
such that it minimizes the analysis error $\mathbf{E}\left[(\varepsilon^a)^2\right]$
so $\mathbf{E}\left[\varepsilon^a (y - x^a)\right] = 0$ or $\varepsilon^a \perp (y - x^a)$



2. Other ways to look at what is an analysis

2.1 Geometric view (Hilbert space)

Define an inner product of two random vectors **X** , **Y** as

$$\langle \mathbf{X}, \mathbf{Y} \rangle = \mathbf{E}\left[\mathbf{X}^T \mathbf{Y}\right]$$

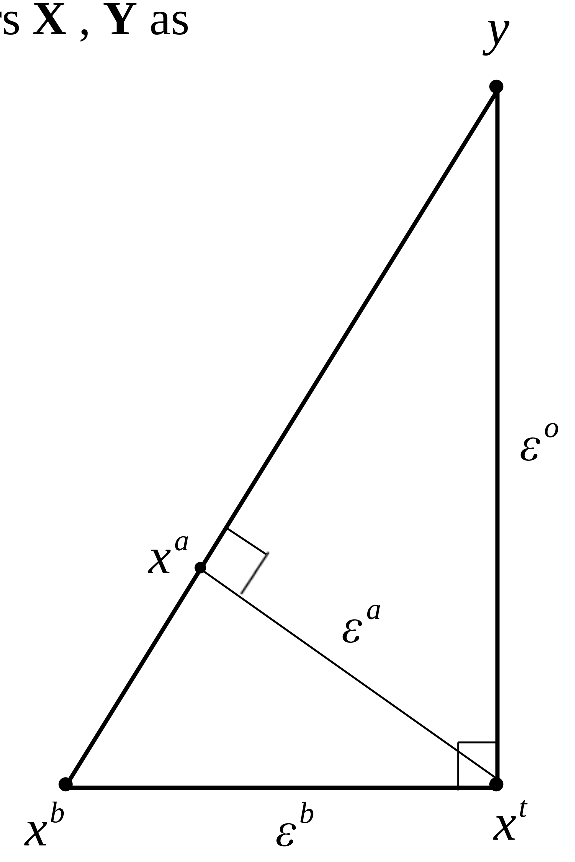
A distance or 2-*norm* can be defined as

$$\left\| \mathbf{X} \right\|_2 = \mathbf{E}\left[\mathbf{X}^T \mathbf{X}\right]$$

Because we assume that observation and background error are uncorrelated then

$$\mathbf{E}\left[(\varepsilon^o)^2\right] + \mathbf{E}\left[(\varepsilon^b)^2\right] = \mathbf{E}\left[(y - x^b)^2\right]$$

$$\mathbf{R} + \mathbf{B} = (O - B)^2$$



2. Other ways to look at what is an analysis

2.1 Geometric view (Hilbert space)

Define an inner product of two random vectors **X** , **Y** as

$$\langle \mathbf{X}, \mathbf{Y} \rangle = \mathbf{E}\left[\mathbf{X}^T \mathbf{Y}\right]$$

A distance or 2-*norm* can be defined as

$$\|\mathbf{X}\|_2 = \mathbf{E}\left[\mathbf{X}^T \mathbf{X}\right]$$

Since the triangles $\Delta\, x^t x^a x^b \sim \Delta\, y\, x^t x^b$ are similar

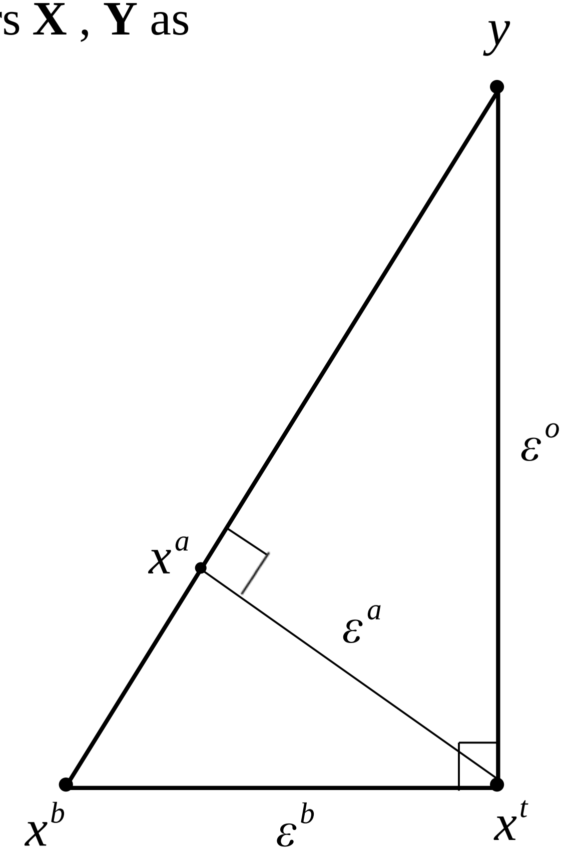
then $$\frac{\varepsilon^b}{x^a - x^b} = \frac{y - x^b}{\varepsilon^b}$$

from which we get

$$\mathbf{E}[(\varepsilon^b)^2] = \mathbf{E}[(x^a - x^b)(y - x^b)]$$

$$\mathbf{B} = \mathbf{E}[(A - B)(O - B)]$$

one of the Desroziers diagnostic (Desroziers et al, 2005, *QJRMS*)



2. Other ways to look at what is an analysis

2.1 Geometric view (Hilbert space)

Define an inner product of two random vectors **X** , **Y** as

$$\langle \mathbf{X}, \mathbf{Y} \rangle = \mathbf{E}\left[\mathbf{X}^T \mathbf{Y}\right]$$

A distance or 2-*norm* can be defined as

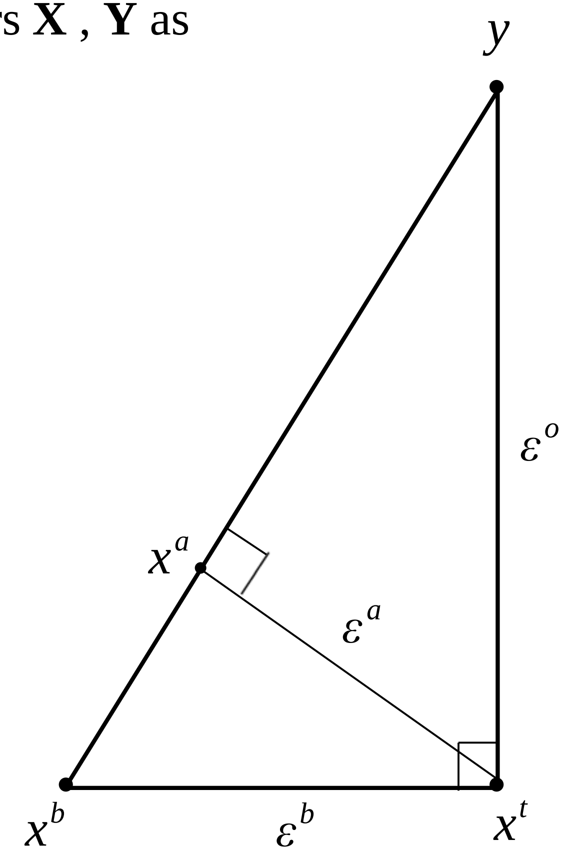
$$\|\mathbf{X}\|_2 = \mathbf{E}\left[\mathbf{X}^T \mathbf{X}\right]$$

The triangles $\Delta\, y\, x^a x^t \sim \Delta\, y\, x^t x^b$
are also similar
from which we get

$$\mathbf{E}[(\varepsilon^o)^2] = \mathbf{E}[(y - x^a)(y - x^b)]$$

$$\mathbf{R} = \mathbf{E}[(O - A)(O - B)]$$

the other Desroziers diagnostic
(Desroziers et al, 2005, *QJRMS*)



2.2 The analysis in spectral space

- Consider a 1D periodic domain

2.2 The analysis in spectral space

- Consider a 1D periodic domain
- And assume observations each grid points, so **H** = **I**

2.2 The analysis in spectral space

- Consider a 1D periodic domain
- And assume observations each grid points, so $\mathbf{H} = \mathbf{I}$
- Variances are uniform and correlations are homogeneous isotropic

$$\mathbf{B} = \mathbf{F}\hat{\mathbf{B}}\mathbf{F}^T$$

$$\hat{\mathbf{B}} = \mathrm{diag}(f^2(0),\ f^2(1),\ f^2(1),\ f^2(2),\ f^2(2),\ ...)$$

$$\mathbf{R} = \mathbf{F}\hat{\mathbf{R}}\mathbf{F}^T$$

$$\hat{\mathbf{R}} = \mathrm{diag}(r^2(0),\ r^2(1),\ r^2(1),\ r^2(2),\ r^2(2),\ ...)$$

2.2 The analysis in spectral space

then ….

The analysis error covariance $\mathbf{A}$ can be obtained as

$$\mathbf{A} = (\mathbf{I} - \mathbf{K})\mathbf{B} \quad \text{where } \mathbf{K} = \mathbf{B}(\mathbf{B}+\mathbf{R})^{-1}$$

$$= \mathbf{R}(\mathbf{B}+\mathbf{R})^{-1}\mathbf{B}$$

and its spectral decomposition $\hat{\mathbf{A}} = \mathbf{F}^T \mathbf{A}\,\mathbf{F}$
gives spectral variances as

$$a^2(k) = \frac{r^2(k)\, f^2(k)}{f^2(k) + r^2(k)} \quad \text{or} \quad \frac{1}{a^2(k)} = \frac{1}{r^2(k)} + \frac{1}{f^2(k)}$$

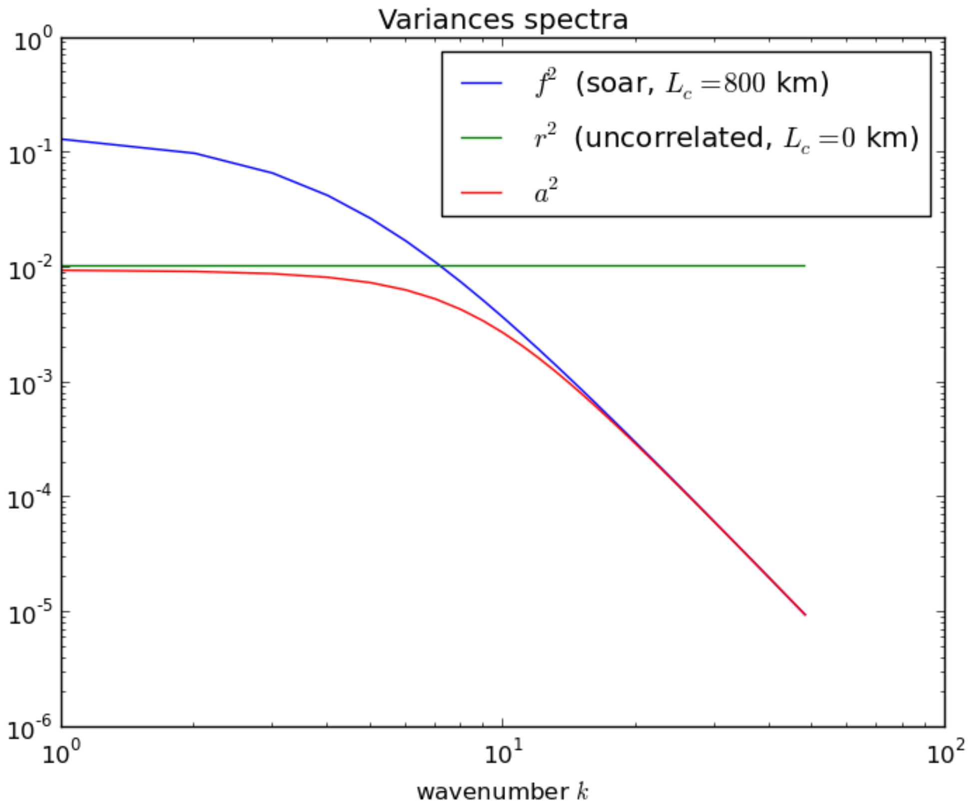
for each wavenumber k

note that ….

$$a^2(k) < \min\left\{ r^2(k),\ f^2(k) \right\}$$

- For any spectra $\left\{ c^2(k) \right\}$ in 1D the variance is given by a sum of the spectral components $\sigma_C^2 = c^2(0) + 2\sum_{k=1}^{N} c^2(k)$

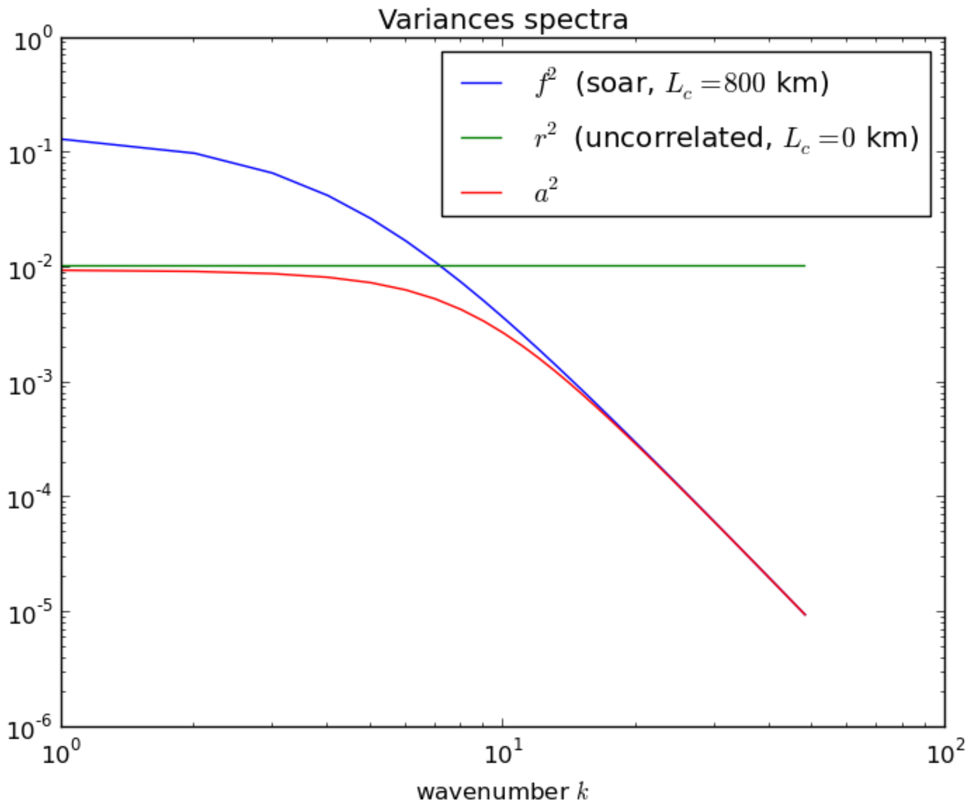
- For any spectra $\{c^2(k)\}$ in 1D the variance is given by $\sigma_C^2 = c^2(0) + 2\sum_{k=1}^{N} c^2(k)$ a sum of the spectral components
- Lets consider the case where the background and observation error variances are identical, i.e. $\sigma_o^2 = \sigma_b^2 = 1$



Python script: spectralVariance.py

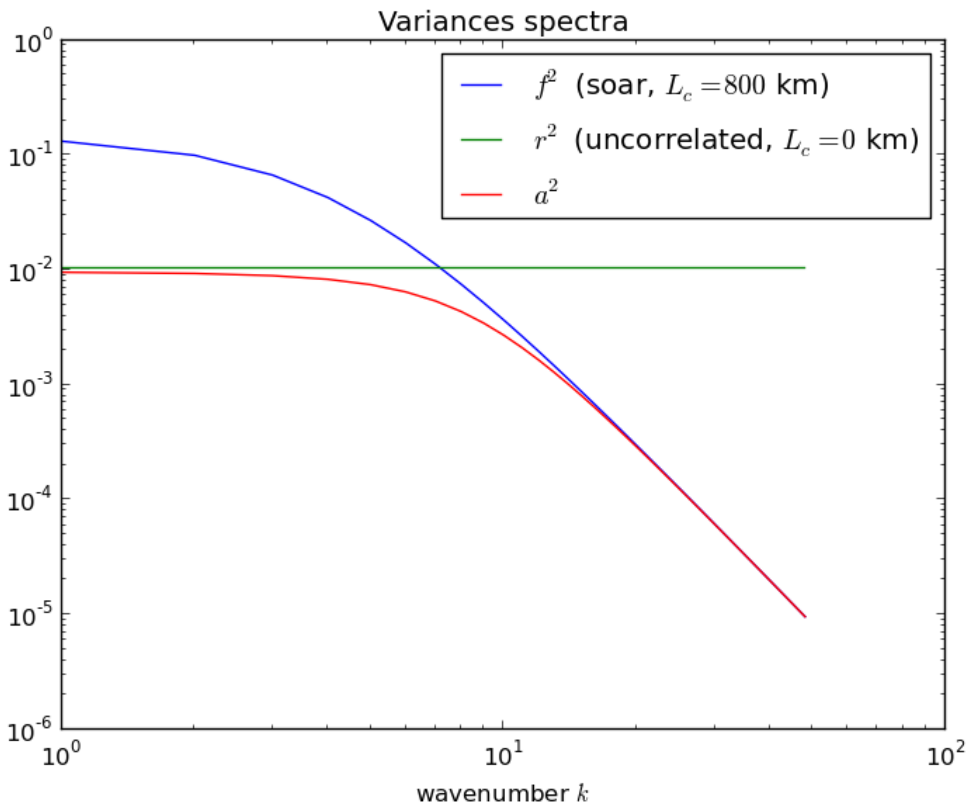
- So for the large scales $f^2 > r^2$ and consequently $a^2 \approx r^2$
- And for the small scales $r^2 > f^2$ and consequently $a^2 \approx f^2$

$$\frac{1}{a^2(k)} = \frac{1}{r^2(k)} + \frac{1}{f^2(k)}$$



Python script: spectralVariance.py

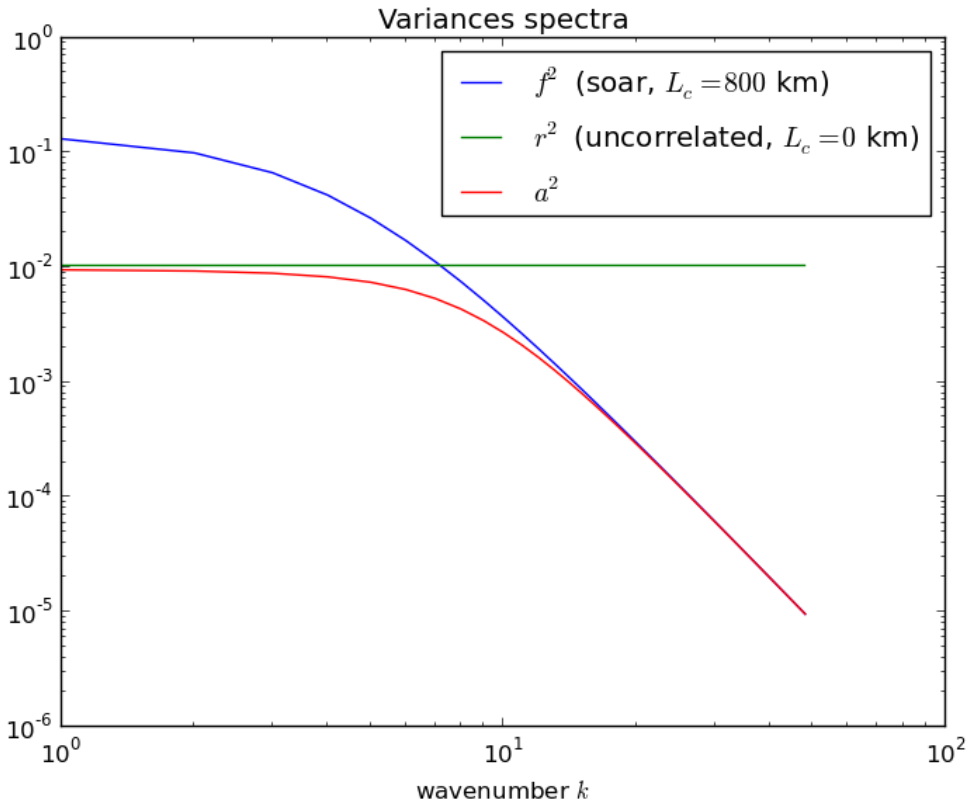
- Most of the analysis correction is done on the large scales
- And there is nearly no analysis correction on small scales



Python script: spectralVariance.py

then ...

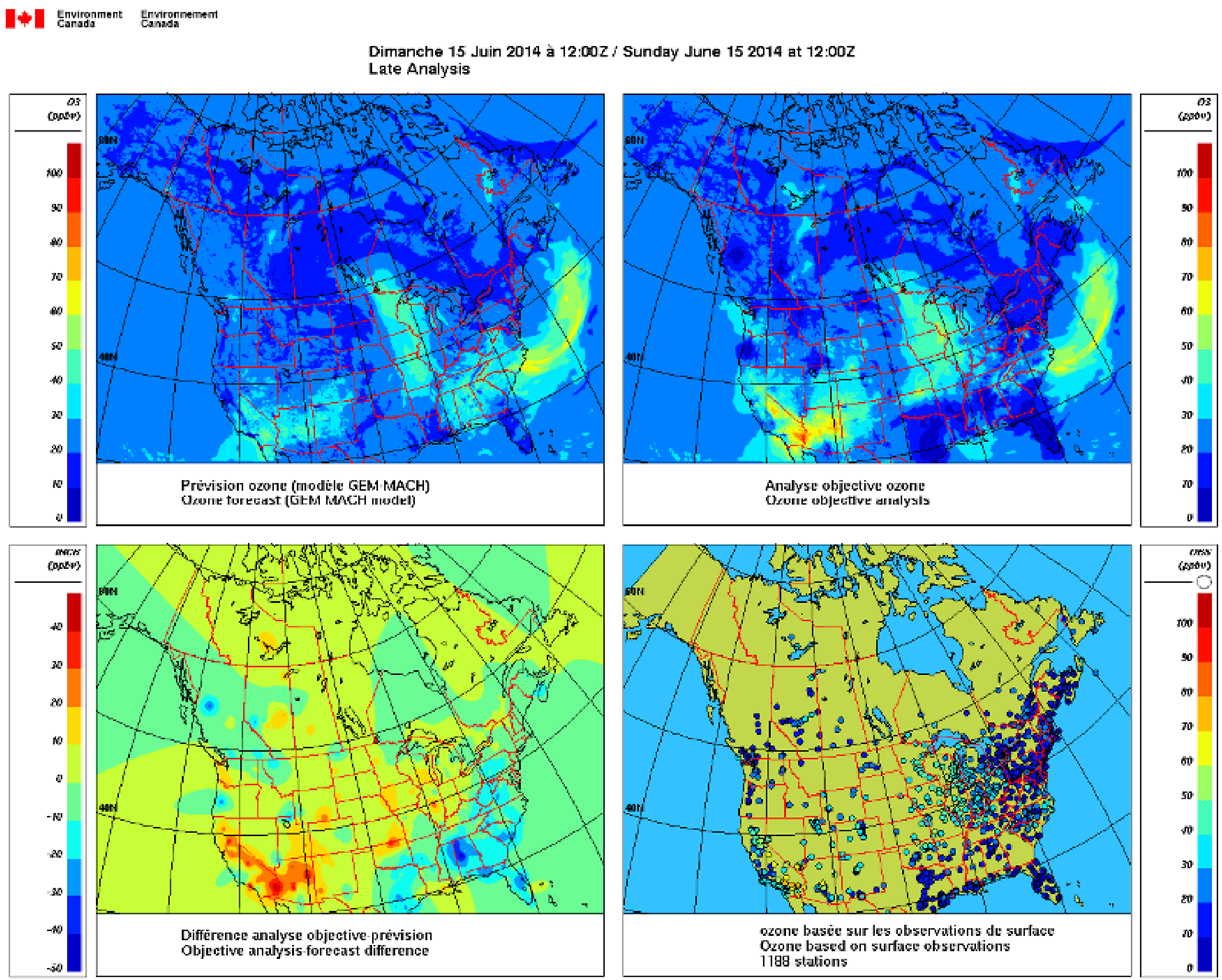
There is a problem for a high resolution models that covers a wide range of scales
a *single correlation model approach is not appropriate to offer an analysis correction on all scales*



Python script: spectralVariance.py

3. Application of AQ analyses

- Optimum interpolation of AirNow observations with GEMMACH
- Operational since 2013 (O3, PM2.5), but running in experimental mode since 2002
  (Ménard and Robichaud 2005: *ECMWF Proceedings*) (Robichaud and Ménard 2014, *ACP*)
- April 2015 we added NO, NO2, SO2, PM10 (Robichaud et al. *2015*, *Air Qual Atmos Health*)



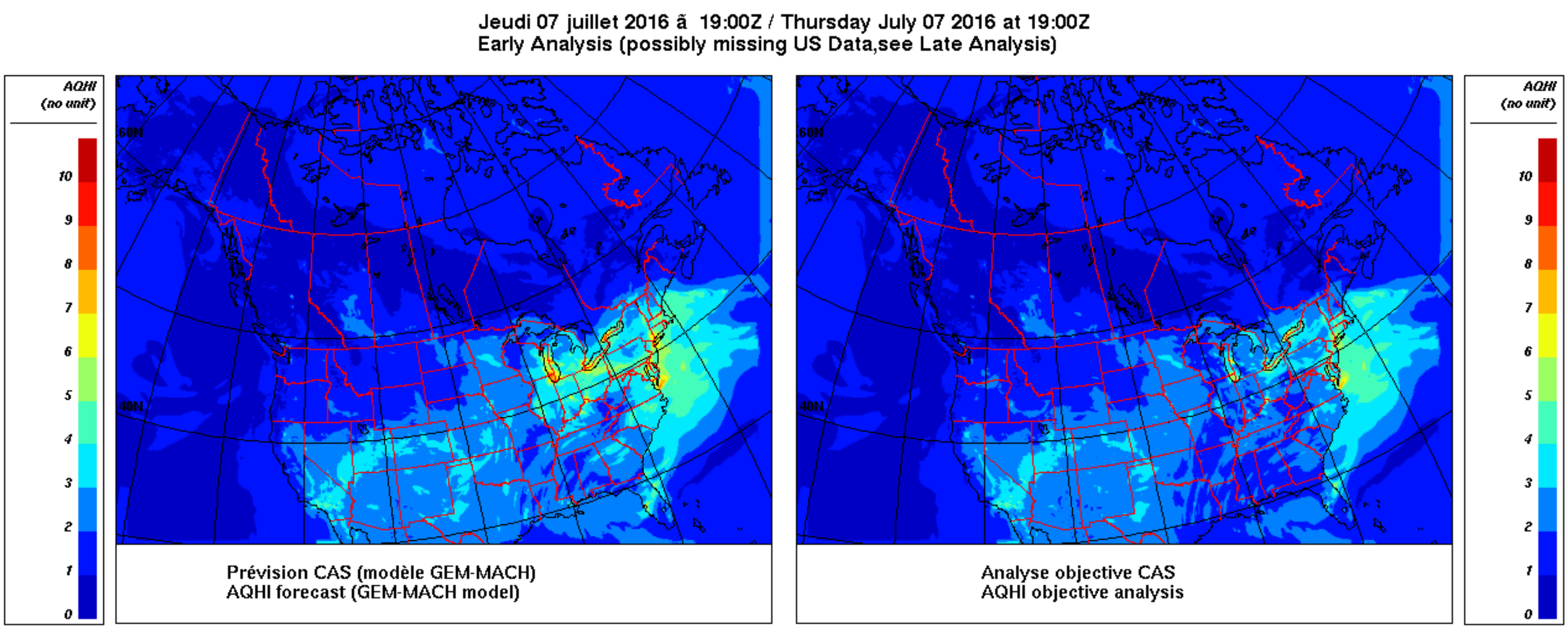
3.1 Air Quality Health Index Maps

Canadian Air Quality Health Index (Stieb et al. 2008, *JA&WMA*)

- Ten year old program that has evolved from an O_3-only forecast in Eastern Canada to a Canada-wide O_3, NO_2, $PM_{2.5}$ forecast program

$$AQHI = 10/10.4 \times 100 \times [(\exp(0.000871[NO_2])-1) + (\exp(0.000537[O_3]) -1)+(\exp(0.000487[PM_{2.5}]) -1)]$$

- A map of AQHI is delivered operationally (each hour)



3.2 Health impact studies

10 year AQ analyses using AirNow CHRONOS and GEM-MACH
Robichaud et Ménard, 2014, *Atmos. Chem. Phys.*, **14**, 1769-1800

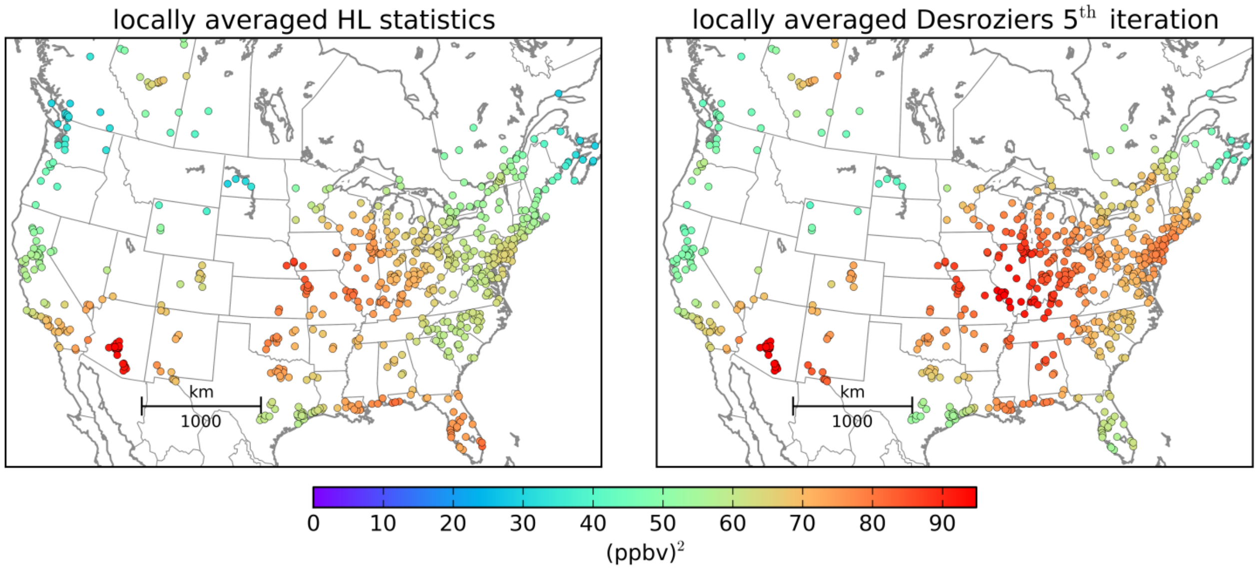
Ambient $PM_{2.5}$, O_3, and NO_2 Exposures and Associations with Mortality over 16 Years of Follow-Up in the Canadian Census Health and Environment Cohort (CanCHEC)
Crouze et al. (2015)**,** *Environ. Health Perspect.*, **123**, 1180-1186

The Canadian Urban Environmental (CANUE) Health Research Consortium
Jeff Brook (PI) (ECCC and UofT) with 15 Canadian Universities, Federal, Provincial and Local Governments.
Develop an easy access geospatial data server (e.g. Google Earth) to support quantitative research on the effect urban environment on health. Data linked to postal codes will contain information on numerous metrics, NDVI, local climatic zones, building density, land use, noise level, air pollution, greenspace, walkability. Data from 1980's up to now.

Some details about the OI used for AQ analysis

- **Optimum interpolation (OI) currently**
  - Uses a local Hollingsworth-Lönnberg fitting to obtain error variance
  - Uses a parametrization of error statistics for isolated stations
  - Has a seasonal bias correction, based on four large regions
- **Next release**
  - Maximum likelihood estimation of correlation length
  - Use compact support correlation models (Ménard et al. 2016, *JA&WMA*)
  - Use hybrid error statistics. Locally averaged H-L or Desroziers in observation space and ensemble of model runs
  - Run in assimilation mode for (at least) verification

σ_b^2 estimation



4. Kalman filtering

4.1 Theory

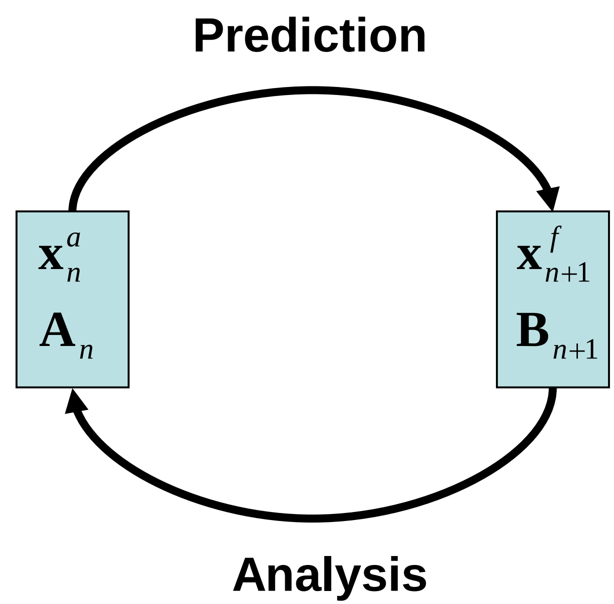
In a Kalman filter, the error covariance is dynamically evolved between the analyses, so that both the state vector and the error covariance are updated by the dynamics and by the observations.

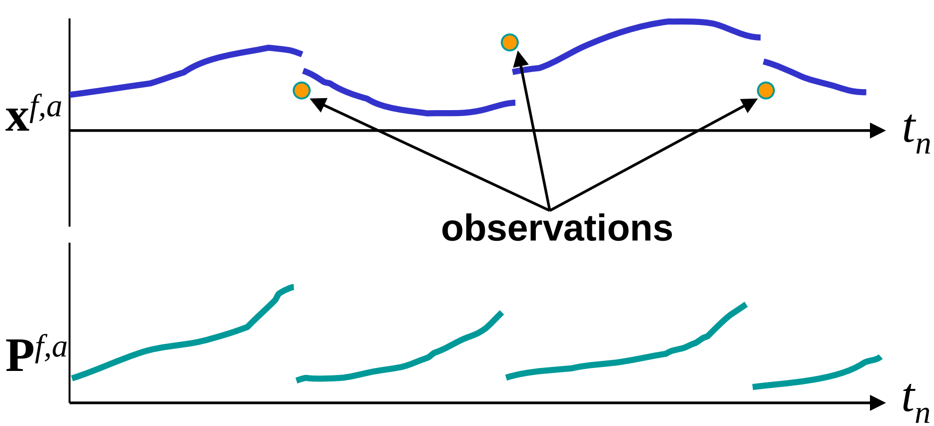
- *Prediction*

$$\mathbf{x}_{n+1}^{f} = \mathbf{M}_n \mathbf{x}_n^{a}$$
$$\mathbf{B}_{n+1} = \mathbf{M}_n \mathbf{A}_n \mathbf{M}_n^{T} + \mathbf{Q}_n$$

- *Analysis*

$$\mathbf{x}_n^{a} = \mathbf{x}_n^{f} + \mathbf{K}_n (\mathbf{y}_n - \mathbf{H}_n \mathbf{x}_n^{f})$$
$$\mathbf{K}_n = \mathbf{B}_n \mathbf{H}_n^{T} \left(\mathbf{H}_n \mathbf{B}_n \mathbf{H}_n^{T} + \mathbf{R}_n \right)^{-1}$$
$$\mathbf{A}_n = \left(\mathbf{I} - \mathbf{K}_n \mathbf{H}_n \right) \mathbf{B}_n$$





The Kalman filter produces the best estimate of the atmospheric state given *all current and past* observations, and yet the algorithm is *sequential in time* in a form of a predictor-corrector scheme.

From a Bayesian point of view the Kalman filter constructs an estimate based on

$$p(\mathbf{x}_n \mid \mathbf{y}_n, \mathbf{y}_{n-1}, \ldots, \mathbf{y}_0)$$

Time sequential property of a Kalman filter is however not easy to show, and this is one of the main result of Kalman (Kalman 1960, *Trans. ASME-J. Basic Eng*.)

Prediction of errors

$$\mathbf{x}_{n+1}^{f} = M_n\left(\mathbf{x}_n^{a}\right)$$

M is a discretized model of the atmosphere which expresses our theoretical understanding but also involves discretization errors

Using the same model to represent the evolution of the *true* state $\mathbf{x}^t$

$$\mathbf{x}_{n+1}^{t} = M_n\left(\mathbf{x}_n^{t}\right) + \boldsymbol{\varepsilon}_n^{q}$$

where $\boldsymbol{\varepsilon}_n^{q}$ is called the model error (or modelling error).

To simply we assume here that the model is linear in $\mathbf{x}$

$$\boldsymbol{\varepsilon}_{n+1}^{f} = \mathbf{M}_n \boldsymbol{\varepsilon}_n^{a} - \boldsymbol{\varepsilon}_n^{q} \quad \text{where } \boldsymbol{\varepsilon}_n^{f,a} = \mathbf{x}_n^{f,a} - \mathbf{x}_n^{t}$$

Assuming that model error at time t_n is uncorrelated with the analysis error at time t_n then

$$\mathbf{B}_{n+1} = \mathbf{M}_n \mathbf{A}_n \mathbf{M}_n^{T} + \mathbf{Q}_n$$

where $\mathbf{B}_n = \mathbf{E}\left[\boldsymbol{\varepsilon}_n^{f}\left(\boldsymbol{\varepsilon}_n^{f}\right)^{T}\right]$; $\mathbf{A}_n = \mathbf{E}\left[\boldsymbol{\varepsilon}_n^{a}\left(\boldsymbol{\varepsilon}_n^{a}\right)^{T}\right]$

4.2 Advection-diffusion transport in 1D

$$\frac{\partial c}{\partial t} + U\frac{\partial c}{\partial x} = \nu\frac{\partial^2 c}{\partial x^2} \qquad \text{....(1)}$$

with a uniform wind U. Both the *concentration* and *concentration error* obey (1). Fourier series representation over a periodic domain L, using $2N+1$ grid points

$$x_j = \frac{L\,j}{2N+1} \qquad (j = -N,\dots,0,\dots,N)$$

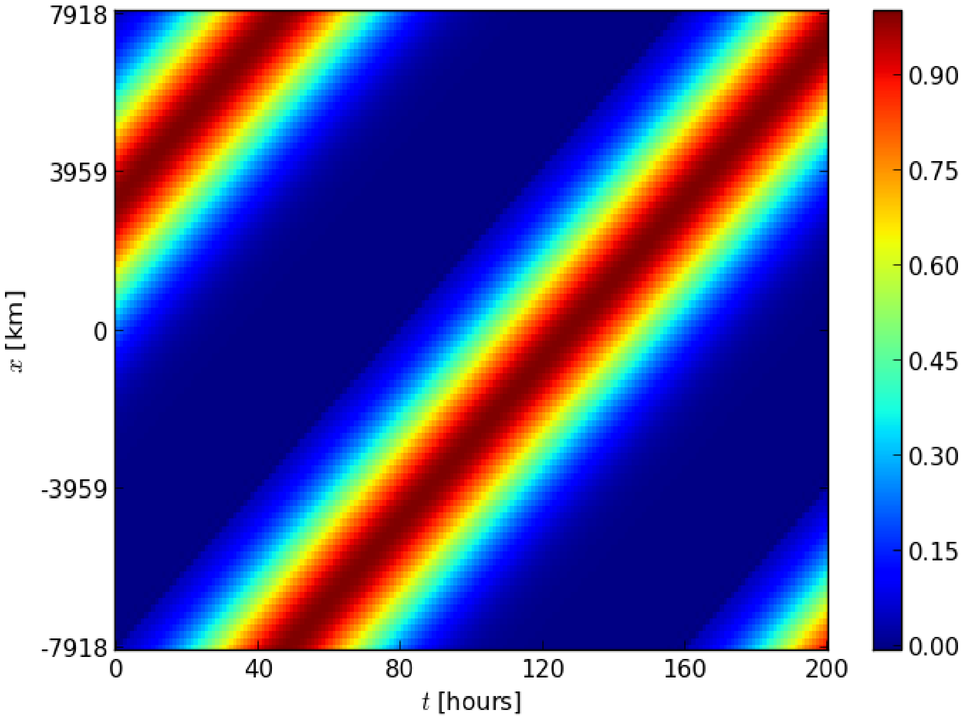
Discrete Fourier series, leading to a unitary matrix, i.e. $\mathbf{F}^{-1} = \mathbf{F}^T$

$$c(x_j) = \sqrt{\frac{2}{2N+1}}\left\{\frac{a_0}{\sqrt{2}} + \sum_{k=1}^{N} a_k \cos\left(\frac{2\pi k x_j}{L}\right) + \sum_{k=1}^{N} b_k \sin\left(\frac{2\pi k x_j}{L}\right)\right\}$$

The transport model in matrix $\mathbf{M}$ can also be transformed in spectral space by a transformation of the form $\hat{\mathbf{M}} = \mathbf{F}^T\,\mathbf{M}\,\mathbf{F}$, where $\hat{\mathbf{M}}$ is a block-diagonal matrix, with blocks

$$\exp\left(-\frac{4\pi^2\nu\,\Delta t\,k^2}{L^2}\right)\begin{bmatrix} \cos\left(\frac{2\pi k\,U\Delta t}{L}\right) & -\sin\left(\frac{2\pi k\,U\Delta t}{L}\right) \\ \sin\left(\frac{2\pi k\,U\Delta t}{L}\right) & \cos\left(\frac{2\pi k\,U\Delta t}{L}\right) \end{bmatrix}$$

Example of advection of gaussian hill (no diffusion)



Python script: propagation.py

4.3 Kalman filter for the advection-diffusion in 1D and H=I

Observing at each grid points, i.e. $\mathbf{H} = \mathbf{I}$, and having homogeneous isotropic model and observation error covariances, **Q** and **R**,

⇨ *the whole Kalman filter equation system can be diagonalized* (Daley and Ménard, 1993, *MWR*)

$$f_{n+1}^2(k) = m^2(k)\, a_n^2(k) + q(k)$$

$$a_n^2(k) = \frac{r^2(k)\, f_n^2(k)}{f_n^2(k) + r^2(k)}$$

where $m^2(k) = \exp\left(-\frac{8\pi^2 \Delta t k^2}{L^2}\right)$

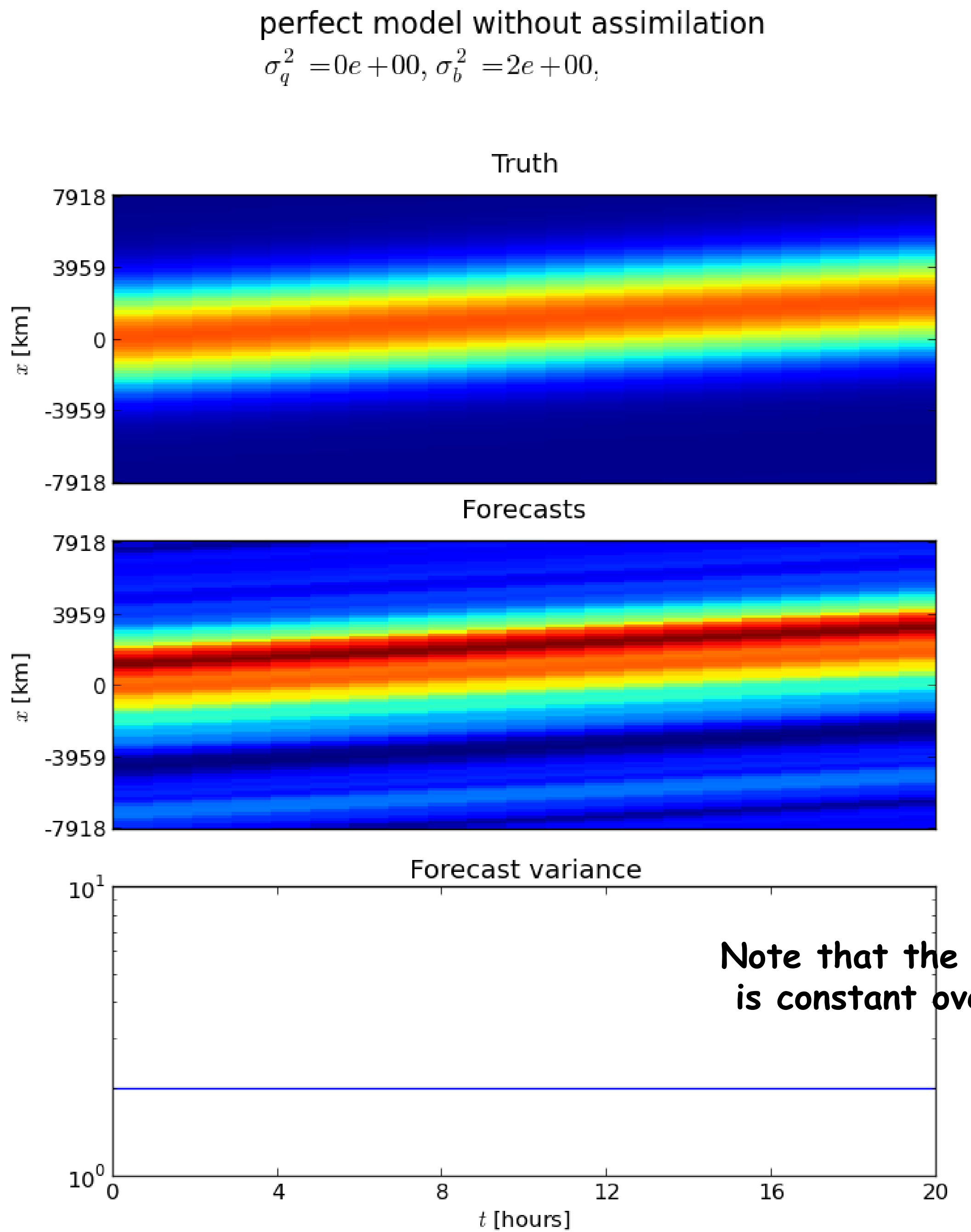
and we recall that

$$\hat{\mathbf{B}}_n = \mathrm{diag}(f_n^2(0),\ f_n^2(1),\ f_n^2(1),\ f_n^2(2),\ f_n^2(2),\ \ldots)$$

$$\hat{\mathbf{A}}_n = \mathrm{diag}(a_n^2(0),\ a_n^2(1),\ a_n^2(1),\ a_n^2(2),\ a_n^2(2),\ \ldots)$$

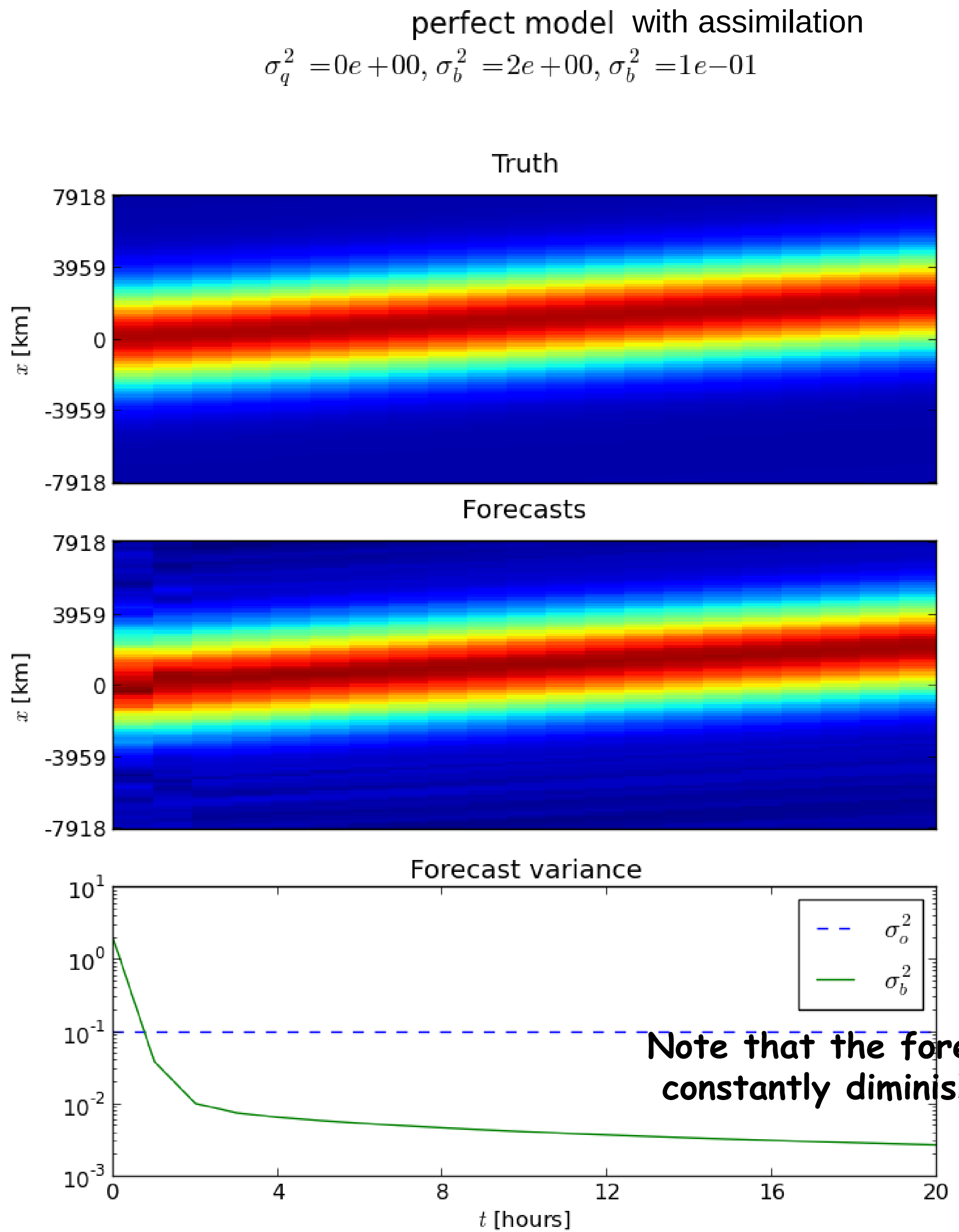
$$\hat{\mathbf{Q}} = \mathrm{diag}(q^2(0),\ q^2(1),\ q^2(1),\ q^2(2),\ q^2(2),\ \ldots)$$

$$\hat{\mathbf{R}} = \mathrm{diag}(r^2(0),\ r^2(1),\ r^2(1),\ r^2(2),\ r^2(2),\ \ldots)$$



Note that the forecast error variance is constant over time

Python script: kalmanFilter.py

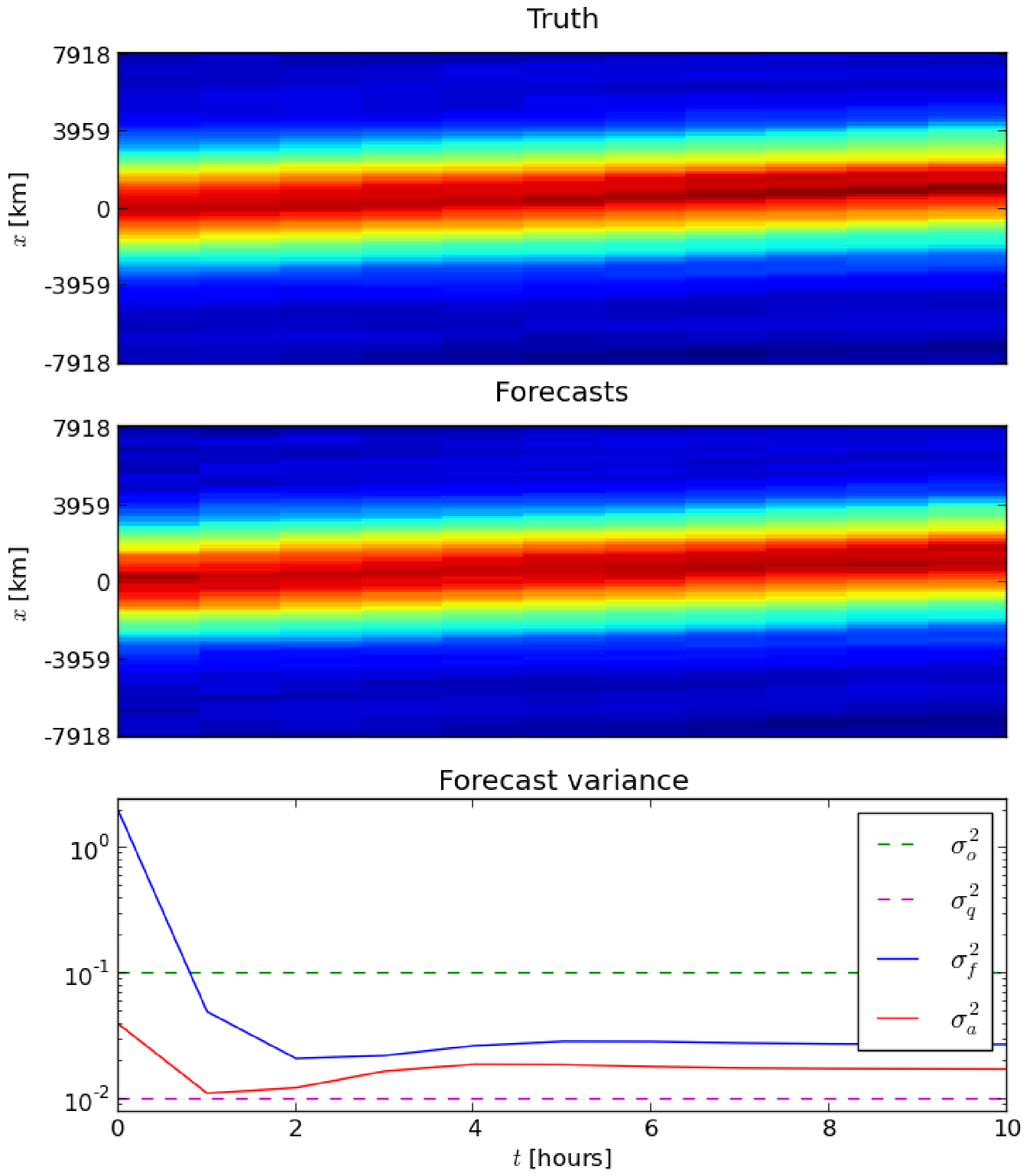


Note that the forecast error variance constantly diminishes ⟹ filter divergence

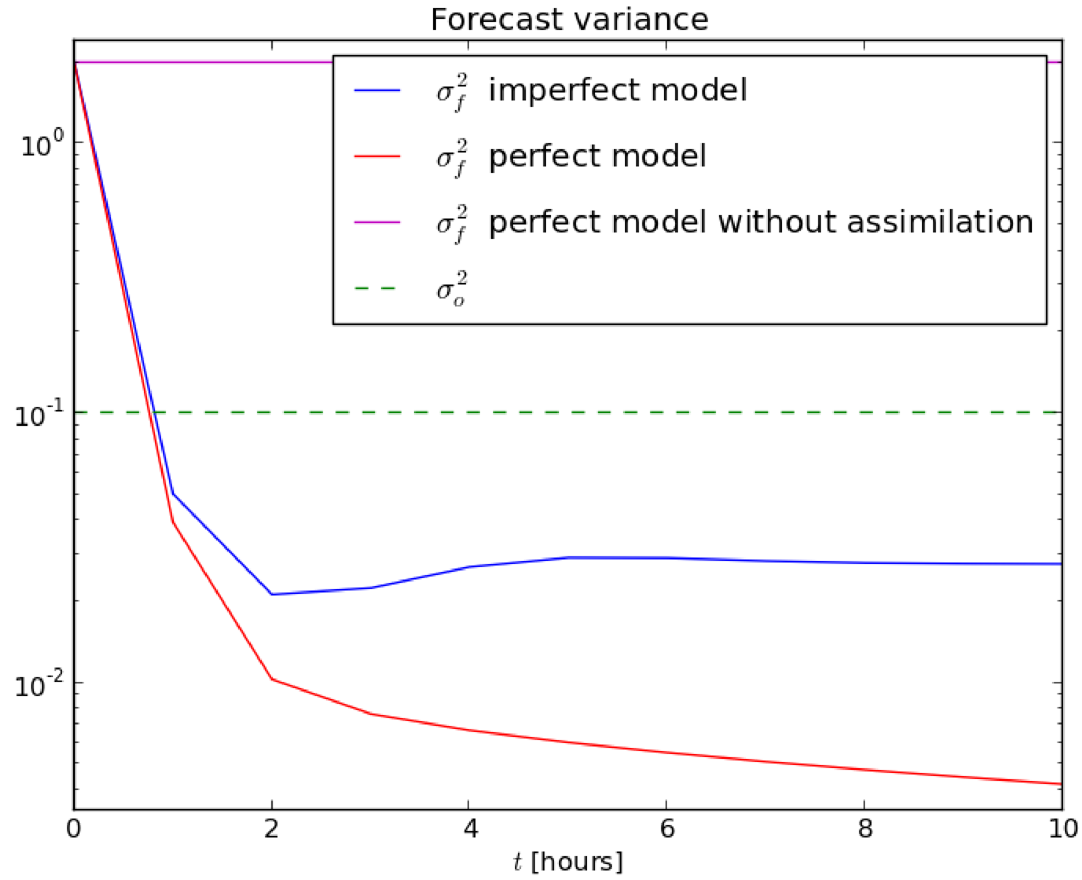
Python script: kalmanFilter.py

$$\sigma_q^2 = 1e-02,\ \sigma_b^2 = 2e+00,\ \sigma_o^2 = 1e-01$$

And when we add model error (and wind was made a bit smaller)



Python script: kalmanFilter.py



stationary solution

Python script: filterDivergence.py

4.4 Stationary solution

- After a short time (days or less) most Kalman filter reaches a nearly stationary regime where the initial conditions have been forgotten
- This simple model actually can provide properties of the stationary solution in spectral space

Combining the first and second equation of

$$f_{n+1}^2(k) = m^2(k)\,a_n^2(k) + q(k)$$

$$a_n^2(k) = \frac{r^2(k)\,f_n^2(k)}{f_n^2(k) + r^2(k)}$$

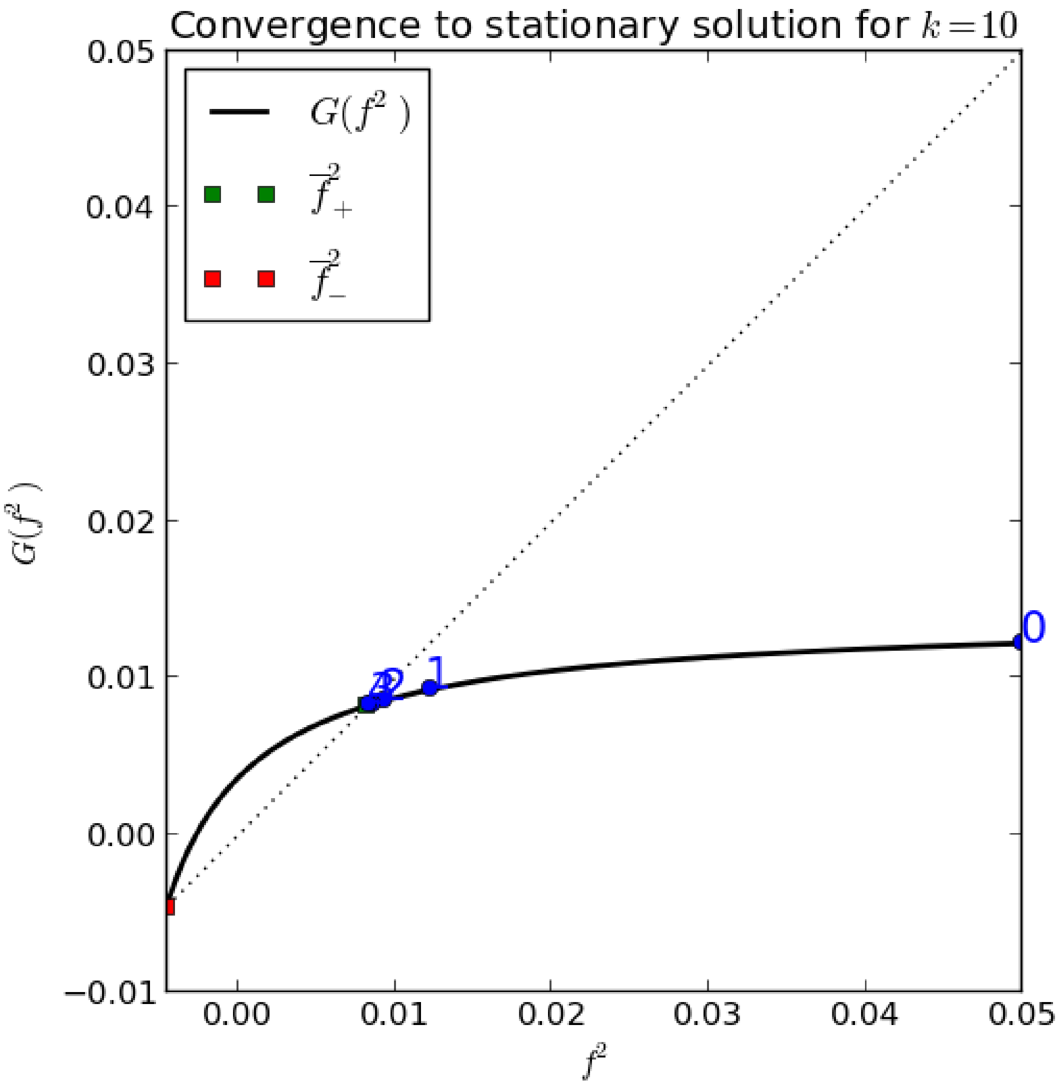
we get a mapping

$$f_{n+1}^2 = G\left(f_n^2\right)$$

for each wavenumber k , and where the mapping function G is of the form

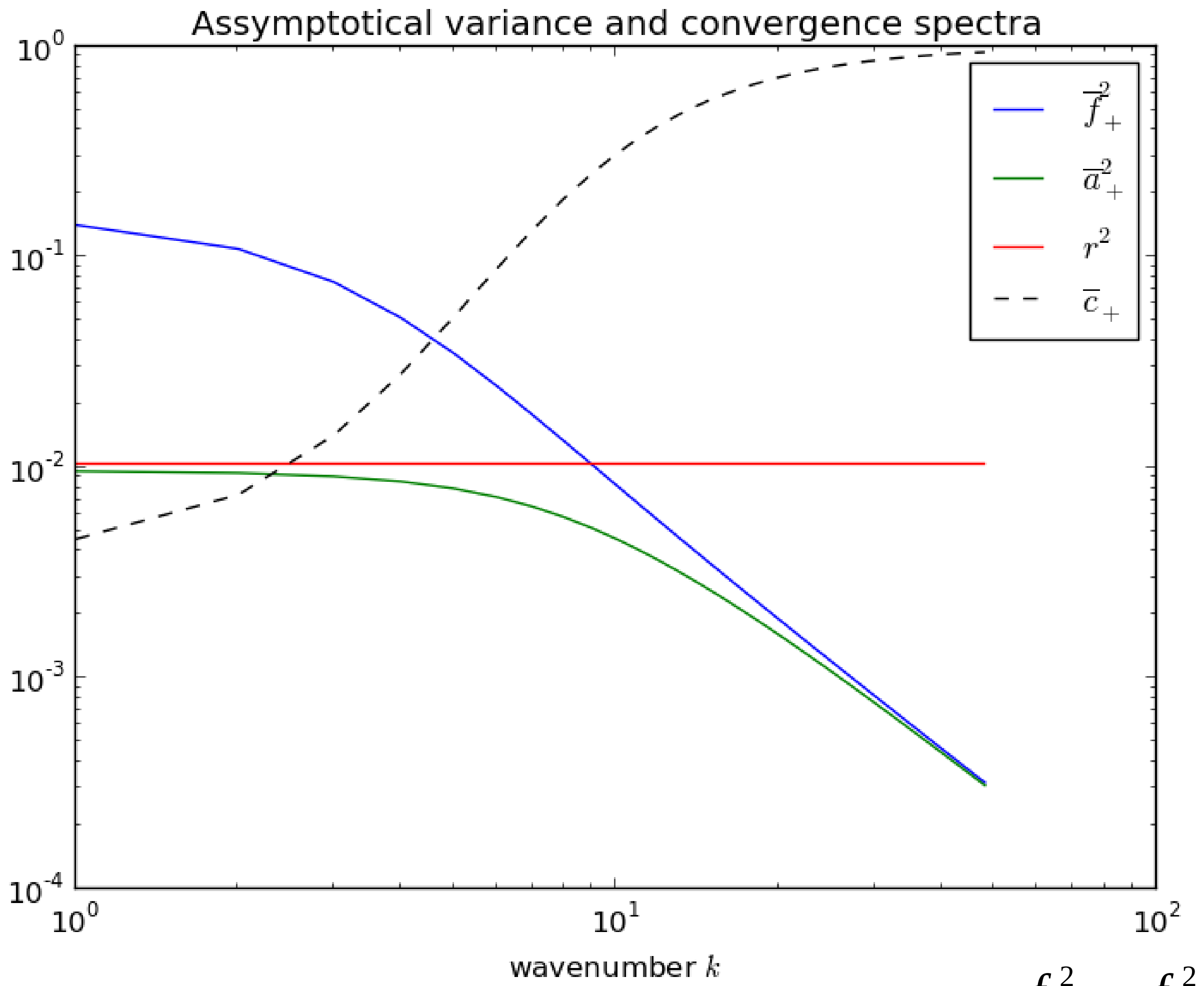
$$G\left(f^2\right) = \frac{m^2 r^2 f^2}{f^2 + r^2} + q^2$$

- This mapping has two fixed-point solutions.
- One unstable fixed-point with negative variance (red square)
- One stable fixed-point with positive variance (green square)



Python script: stationarySolutions.py

Stationary solution in wavenumber space
(no diffusion)



The rate of convergence to the stationary solution $c_n = \dfrac{f_{n+1}^2 - f_n^2}{f_n^2 - f_{n-1}^2}$

Python script: spectralVariance.py

$$f_{n+1}^2(k) = m^2(k)\,a_n^2(k) + q(k)$$

$$a_n^2(k) = \frac{r^2(k)\, f_n^2(k)}{f_n^2(k) + r^2(k)}$$

Remarks for the inviscid case

- $r^2(k) \to 0$ (perfect obs) then $f^2(k) \to q^2(k)$

  and $a^2(k) \to r^2(k)$

The analysis error goes to zero and so the forecast error is simply the model error

- $q^2(k) \to 0$ (perfect model) then $a^2(k) = f^2(k) \to \sqrt{r^2(k)q^2(k)}$

Because the error variance is conserved by transport, in a perfect model there is no growth of error and the analysis error and forecast error are identical

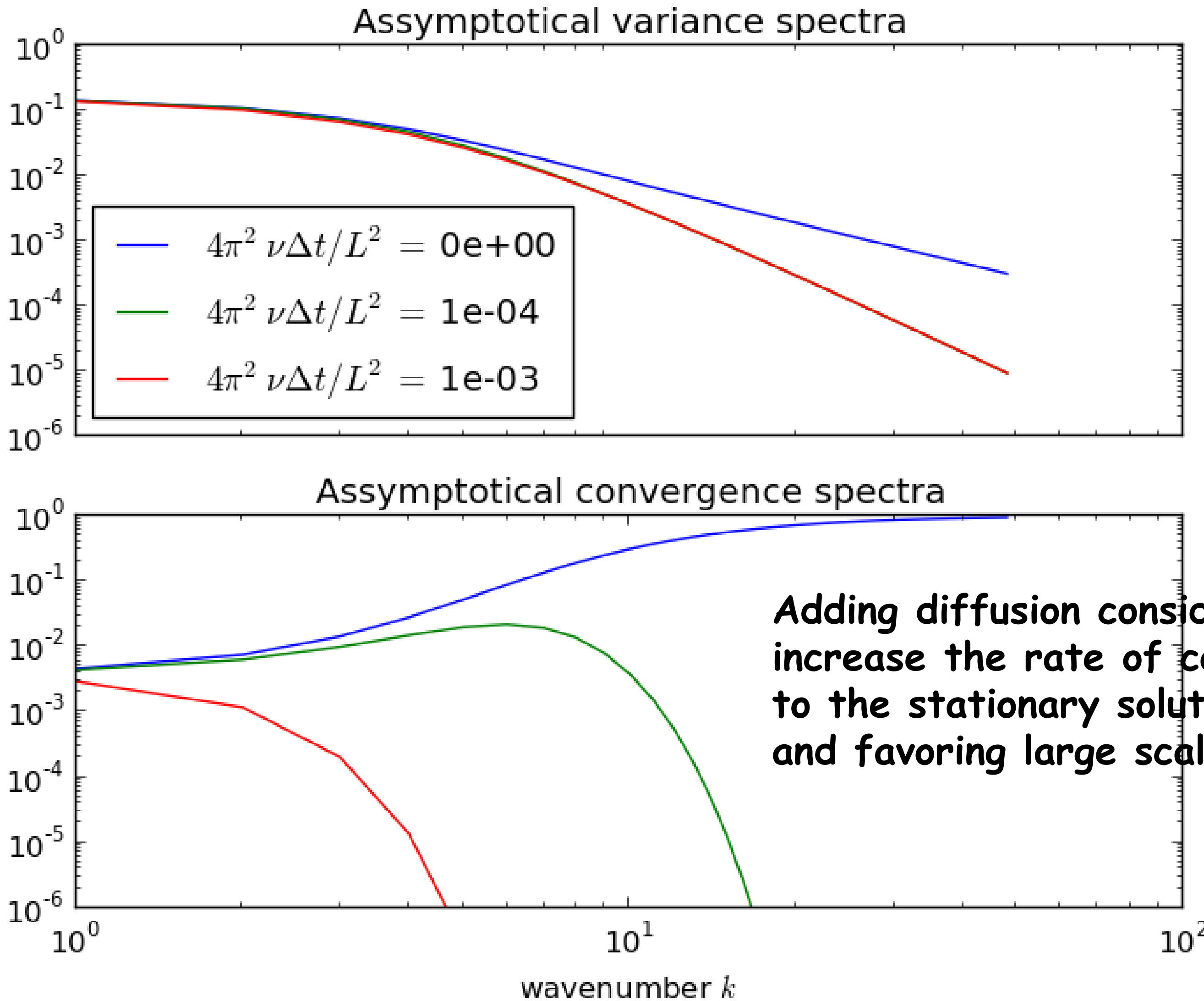
- rate of convergence

$$\text{as } r^2(k) \to 0 \text{ then } c(k) \to 0$$

$$\text{as } q^2(k) \to 0 \text{ then } c(k) \to 1$$

The slow convergence of perfect models is such that a model takes a very long time to forget the initial condition.

Stationary solution and convergence rate: model with diffusion



Adding diffusion considerably increase the rate of convergence to the stationary solution, and favoring large scales

Python script: viscosity.py

Tutorial

Package
This bundle contains a module (DM93) and a collection of python scripts illustrating important characteristics of the Kalman Filter using a simple spectral advection model

Dependencies
Python 2, Numpy, Matplotlib
These packages are readilly available on all major Linux distributions

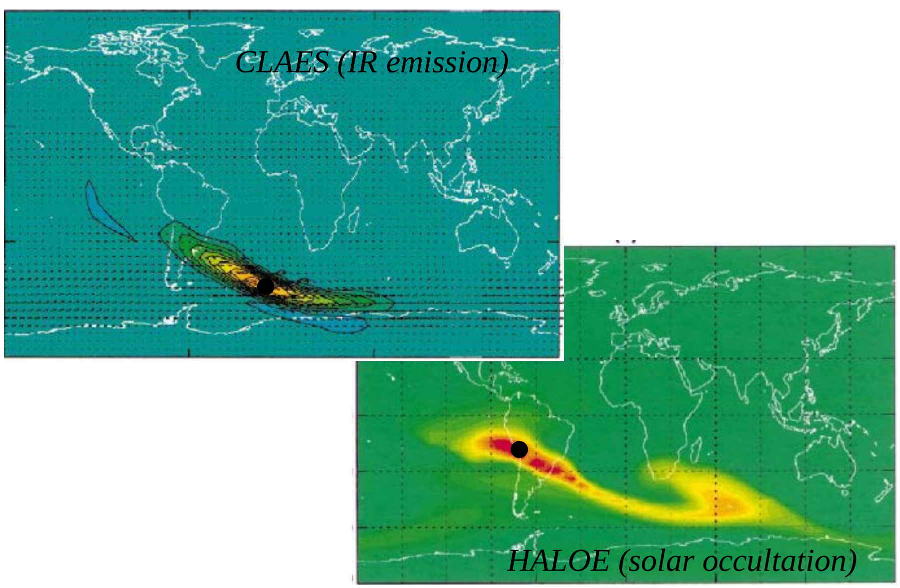
Installation
To obtain the bundle, you can either download a zip file from
github.com/martndj/DaleyMenard1993

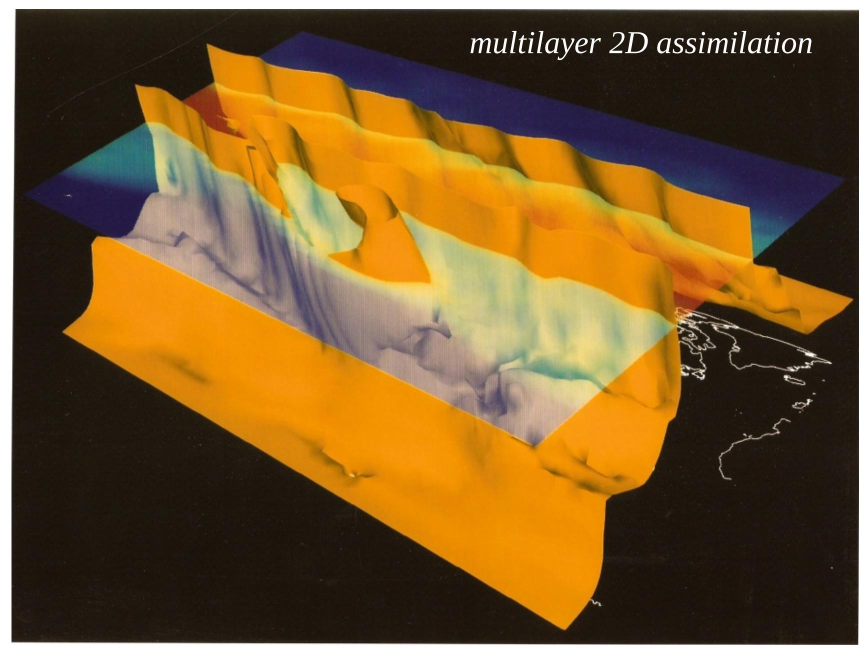
or use git in command line:
git clone https://github.com/martndj/DaleyMenard1993.git

5. Implementation of different flavors of Kalman filtering

5.1 Eulerian KF (Lyster et al. 1997, Ménard et al. 2000, Ménard and Chang 2000)

- 2D advection of long-lived species on isentropic surfaces in the stratosphere
- Limb sounding observations (UARS observations CH_4, N_2O, HNO_3,…)
- 2D isentropic assimilation decoupling
- Implementation of KF with no approximations





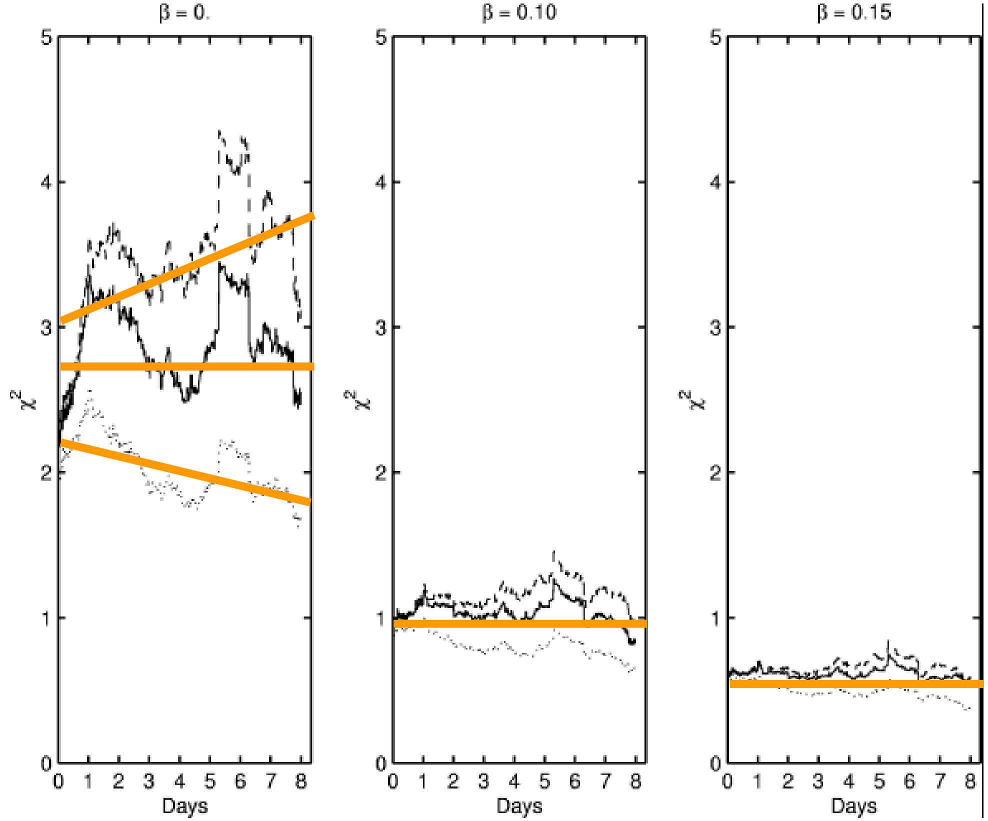
Chi-square diagnostic: Tuning of observation and model error variance parameter

Chi-square diagnostic

$$\langle \chi_k^2 \rangle = \left\langle \nu^T \left[\mathbf{H}\left(\mathbf{H}\mathbf{P}^f\right)^T + \mathbf{R} \right]^{-1} \nu \right\rangle = p$$

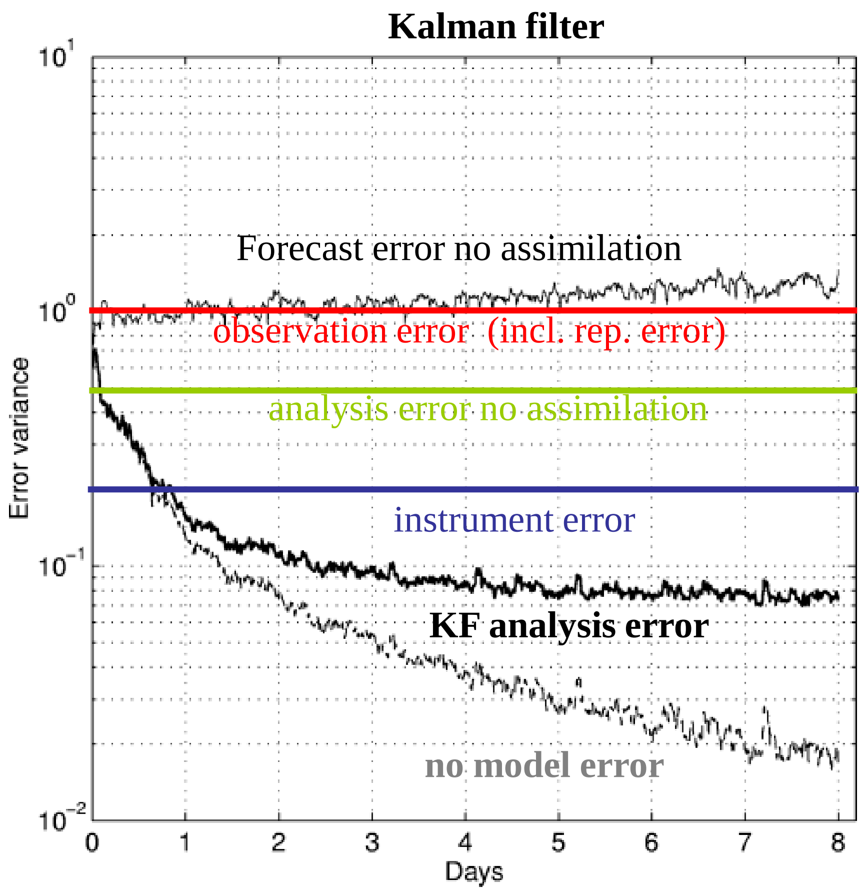
ν is the innovation: OmF
p is the number of observations

When error variance are evolved the tendency of the innovation variance provides information about model error variance



- β is the value of the observation error (rep. error)
- Each panel has three curves corresponding to different value of model error

Error variance

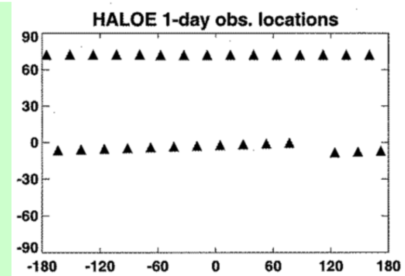


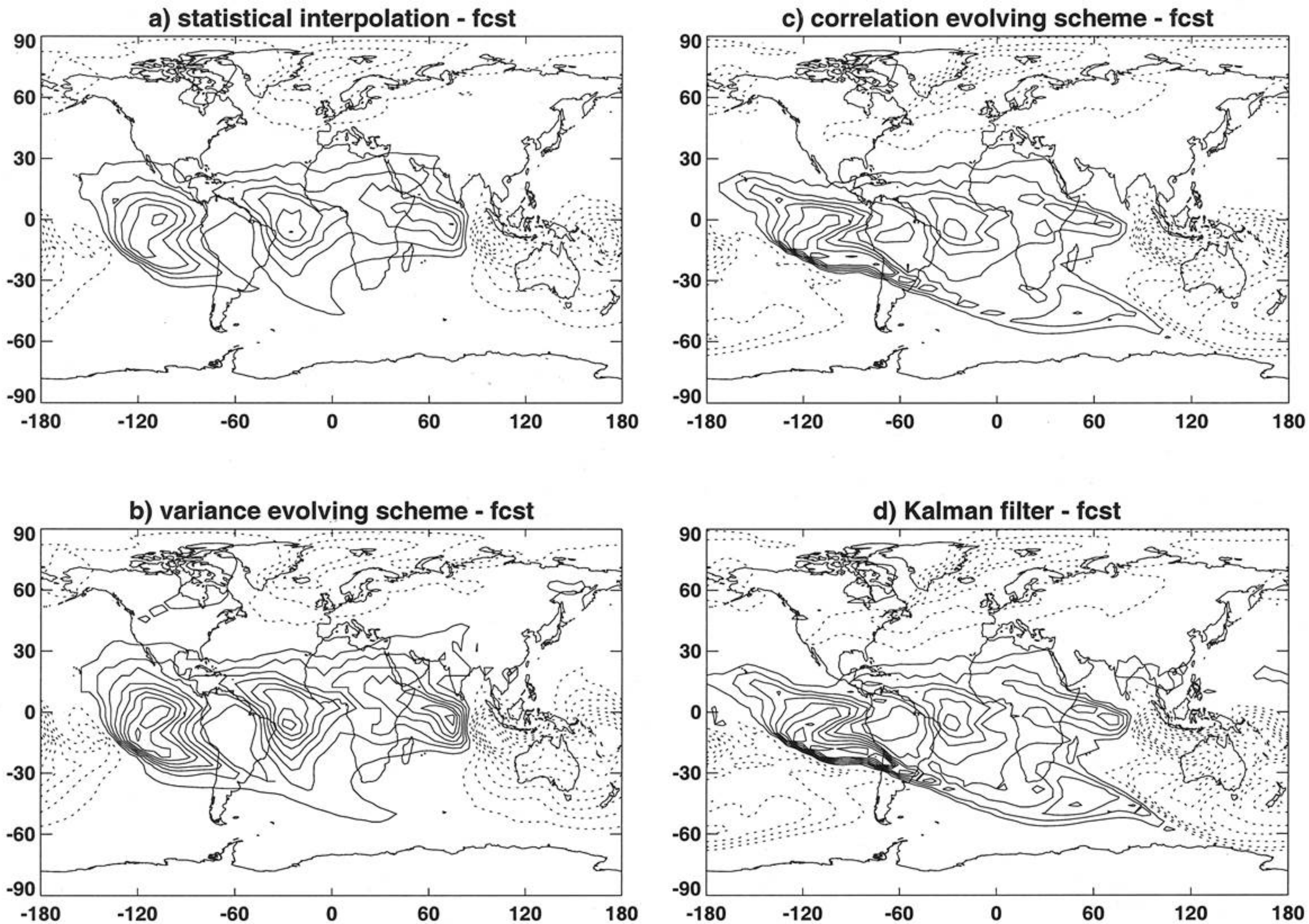
all experiments are χ^2 tunned

1 - Static error covariance (e.g. OI)
2 – Evolving the error variance only
3 – Evolving the error correlation only
4 – Full KF

Analysis increments for different schemes: HALOE CH_4 data

Sparse observations
HALOE is a solar occultation measurement





HALOE CH4 in ppmv

5.2 Ensemble Kalman filter and comparison with 4D-Var

- No chemistry (Skachko et al. 2014, *GMD*)
- With full chemistry (Skachko et al. 2016, *GMD*)

Stochastic EnKF – observation perturbation , with careful tuning of the model and observation error variance for optimal assimilation

BASCOE CTM

- Tracer transport or full chemistry
- 3.75° x 2.5° horizontal
- 37 vertical hybrid-pressure levels
- ECMWF ERA-Interim wind and temperature

Assimilation aspects related to CTM

- Offline of the meteorology
- No changes to winds and temp
- Winds and temp. considered perfect

Observations

- MLS EOS-Aura ozone profiles

Spatial correlations – spectral

- Follows Errera and Ménard 2012, ACP
- $\mathbf{B} = \mathbf{L}\mathbf{L}^T$ where $\mathbf{L} = \boldsymbol{\Sigma} \mathbf{S} \boldsymbol{\Lambda}^{1/2}$
- $\tilde{\mathbf{x}} = \mathbf{L}\boldsymbol{\theta}$ where $\boldsymbol{\theta} \sim N(0, \mathbf{I})$

Covariance models

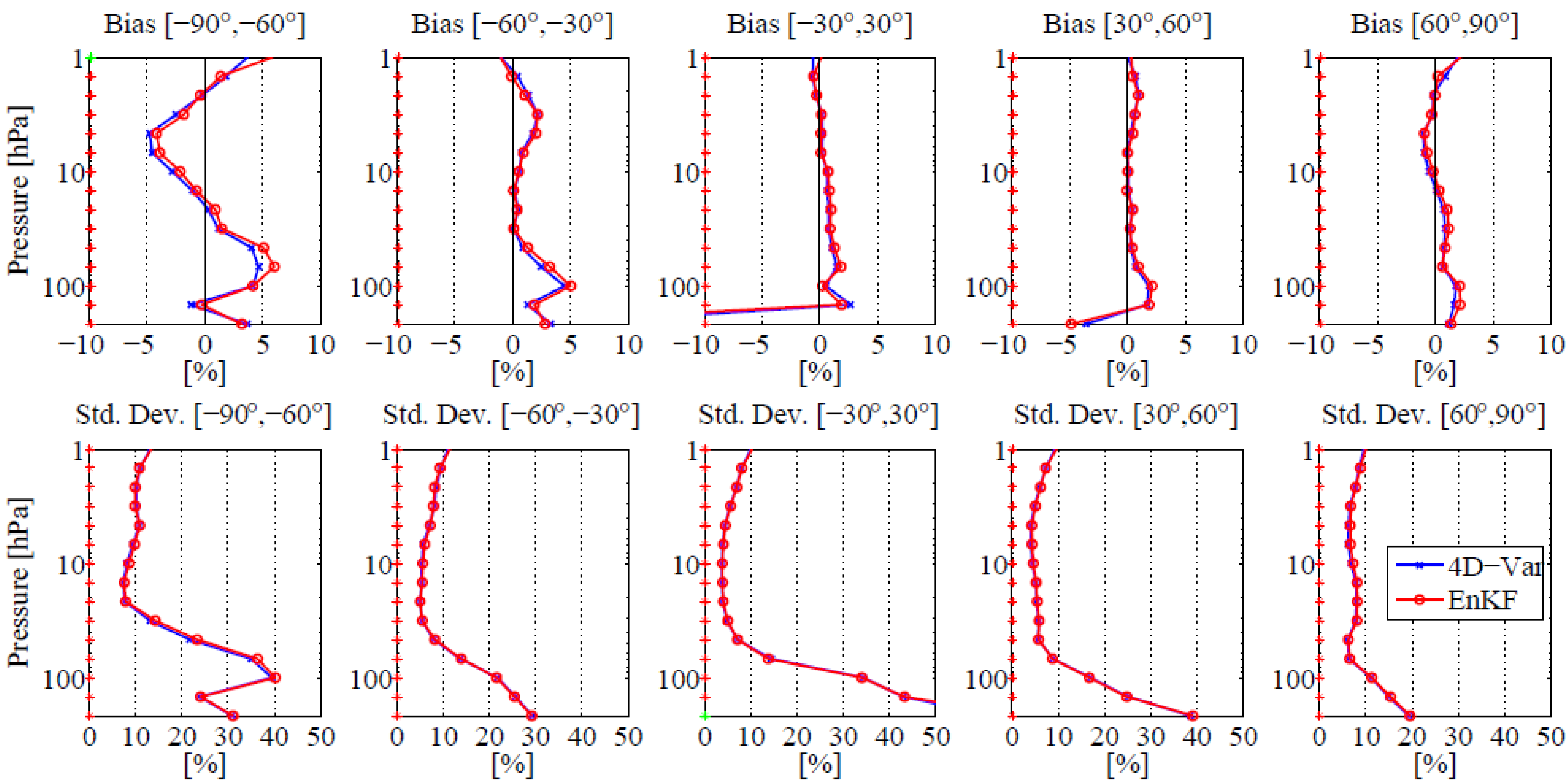
- $\mathbf{B}_0$ Gaussian model; horizontal and vertical
- $\mathbf{R}$ horizontally and vertically uncorrelated
- $\mathbf{Q}$ additive model error, with same correlation structure as $\mathbf{B}_0$

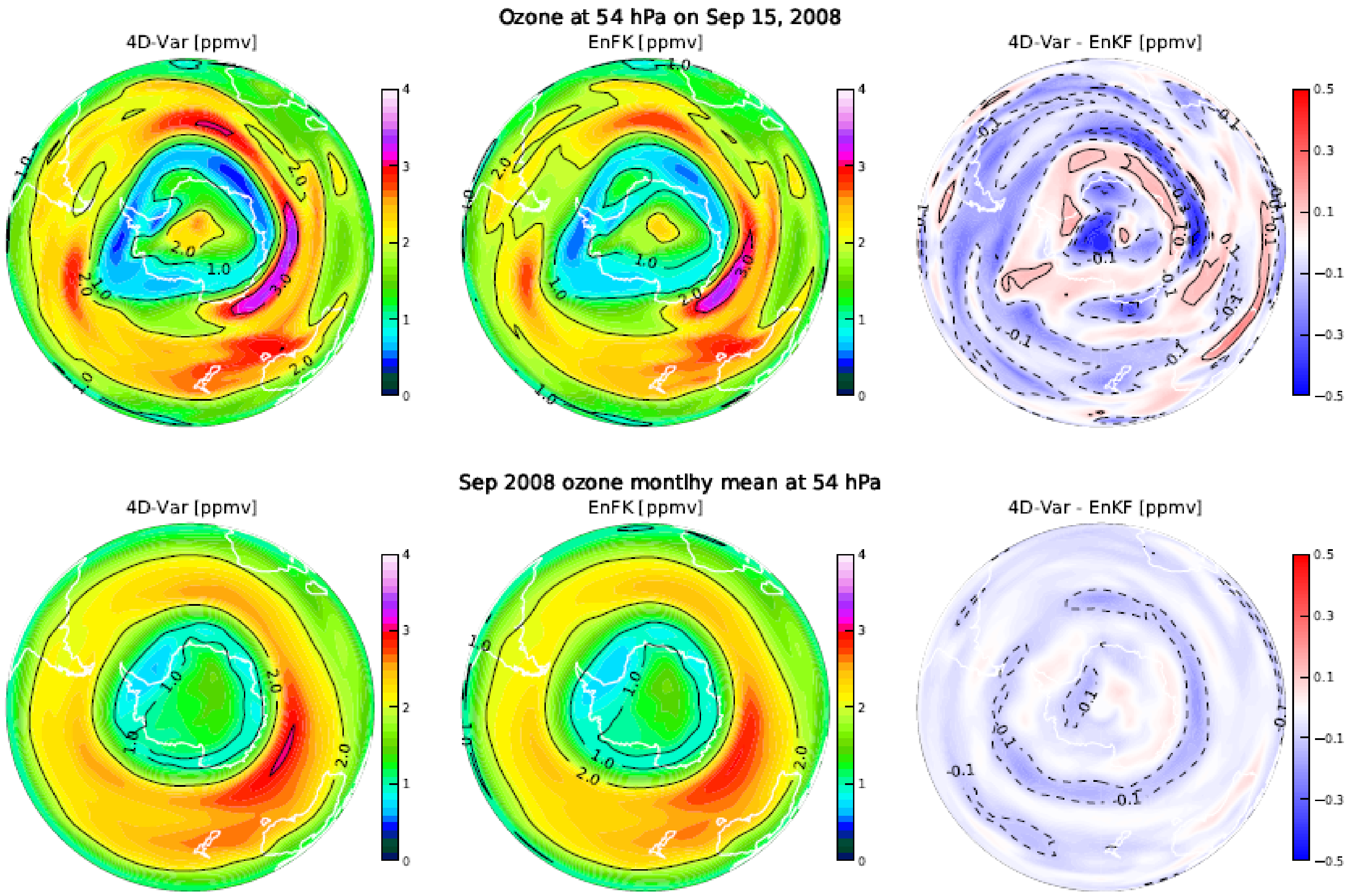
Relative error formulation

- $\boldsymbol{\Sigma} = diag(\beta \mathbf{x}^b)$
- $\mathbf{R}(i,i) = r\, (\mathbf{y}^o \circ \boldsymbol{\sigma}_{rel})^2$
- $\boldsymbol{\eta} = diag(\alpha \mathbf{x}^f)\mathbf{L}\boldsymbol{\theta}$ where $\mathbf{Q} = \langle \boldsymbol{\eta}\boldsymbol{\eta}^T \rangle$

No chemistry – assimilation of ozone

Bias and standard deviation of O-P (September – October 2008)

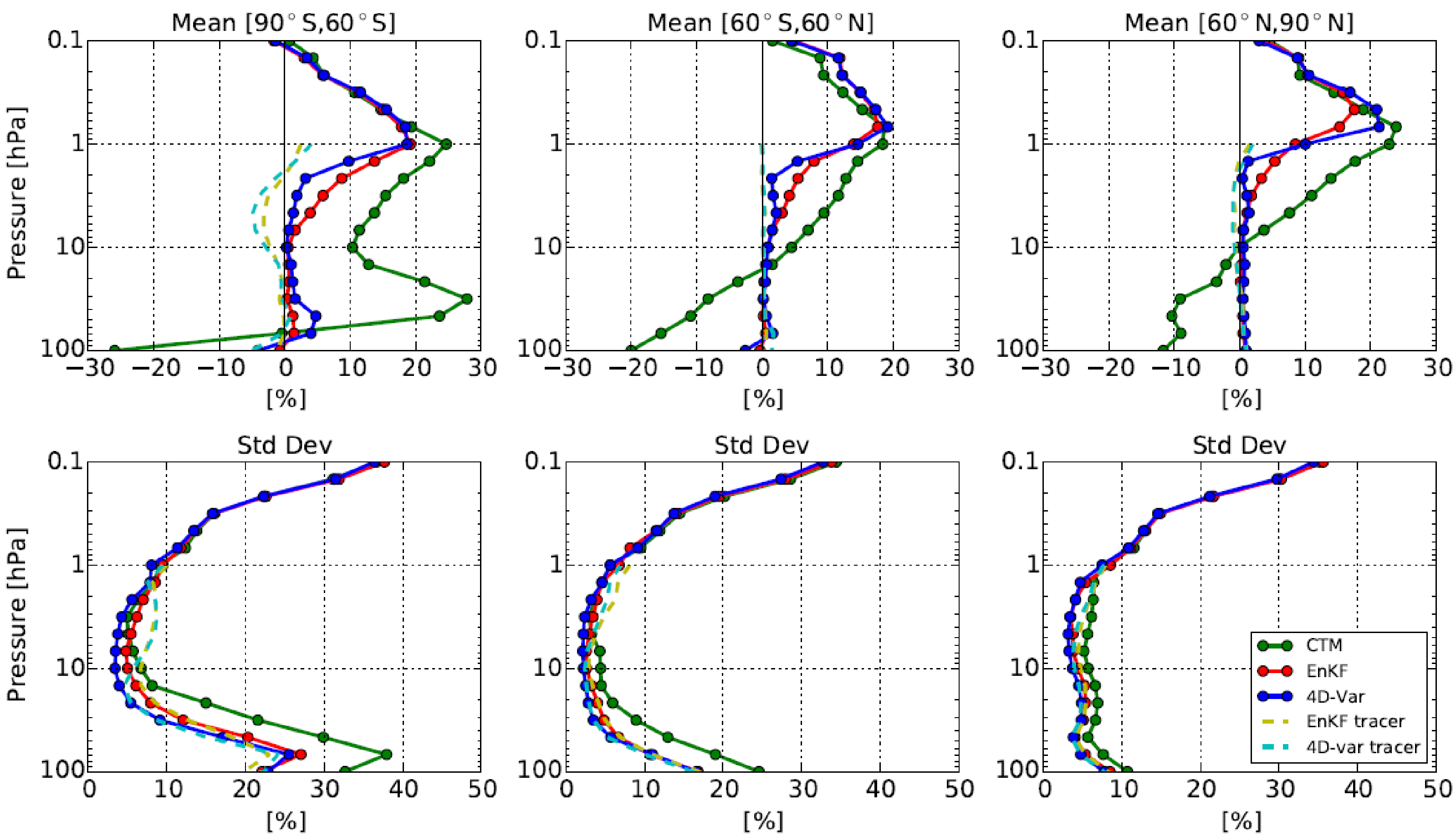




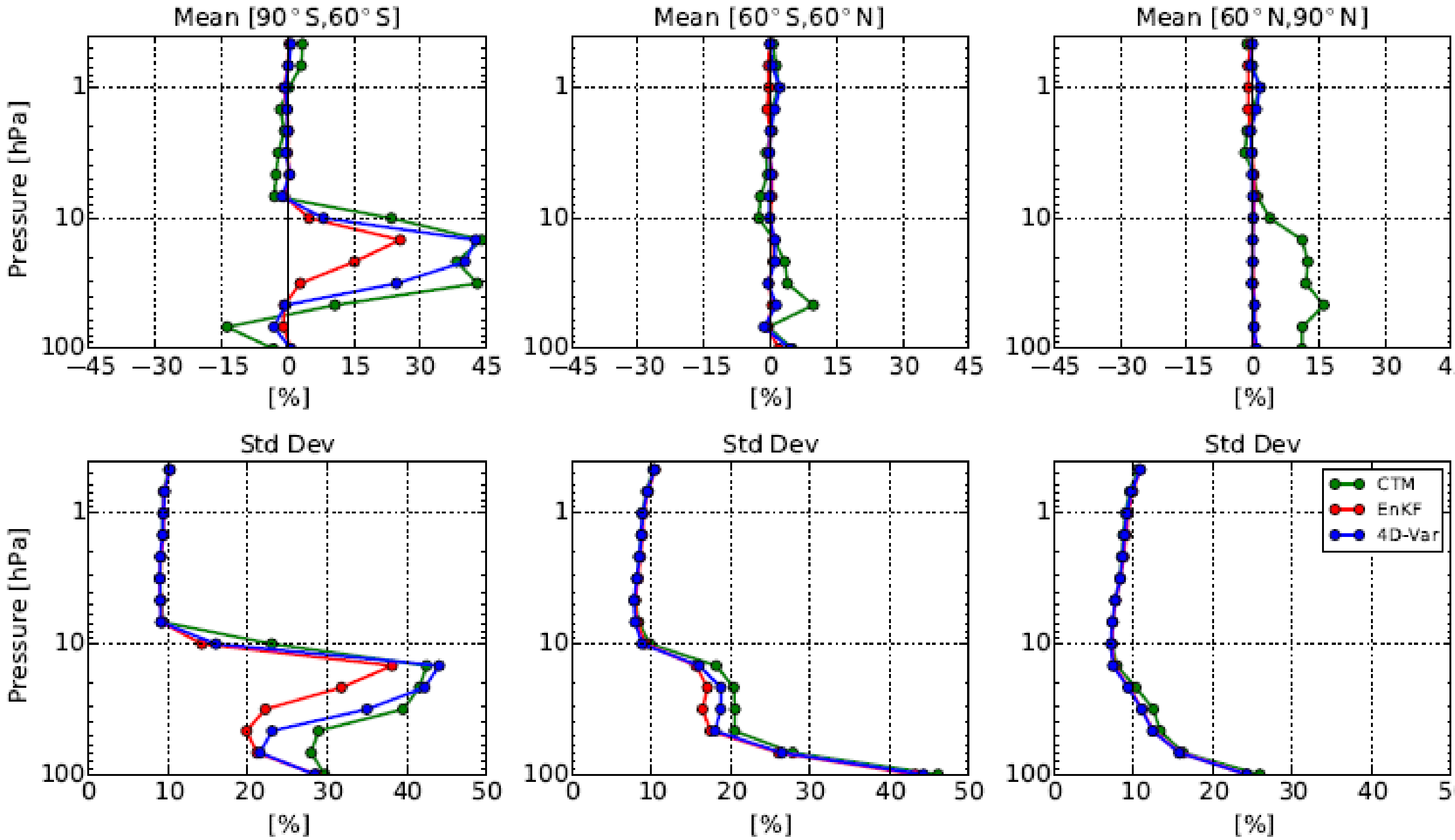
Assimilation of ozone as passive tracer transport, using the same input errors and with model error the EnKF and 4D-Var solutions gives nearly identical O-P zonal statistics, but the EnKF analyses are somewhat smoother than the 4D-Var analyses

With chemistry – assimilation of O3, N2O, H2O, HCl, HNO3

Bias and standard deviation of **O_3** (O-P) (September – October 2008)



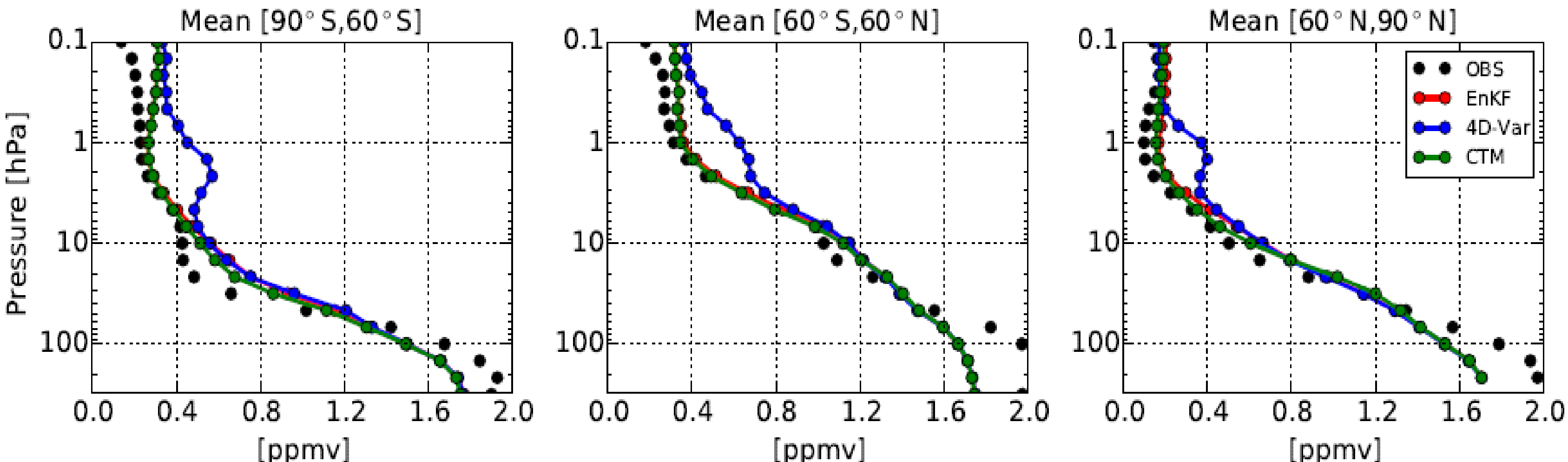
Bias and standard deviation of **HCl** (O-P) (May – June 2008)



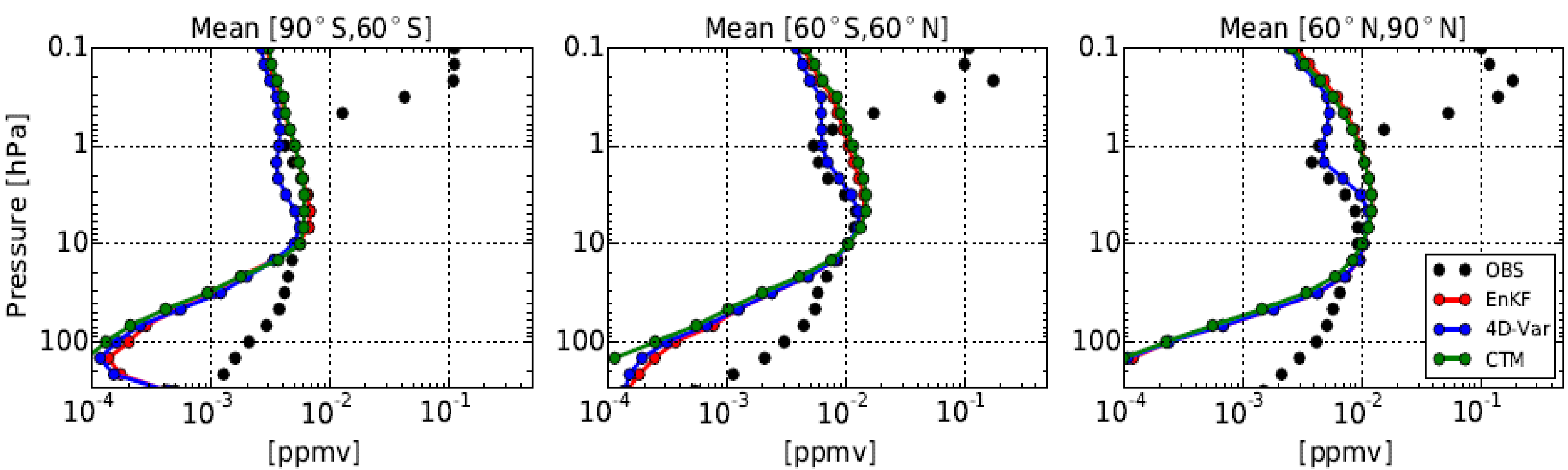
During this period the chemical lifetime of HCl (in polar vortex) is much shorter than at other latitudes, because the heterogeneous removal due to the formation of PSC has already started. This loss process is overestimated in the BASCOE CTM, due to a crude cold-point temperature parametrization. The CTM underestimates HCl by up to 45% at 30 hPa in the Antarctic polar vortex region

Impact on non-observed species (September – October 2008)

CH4

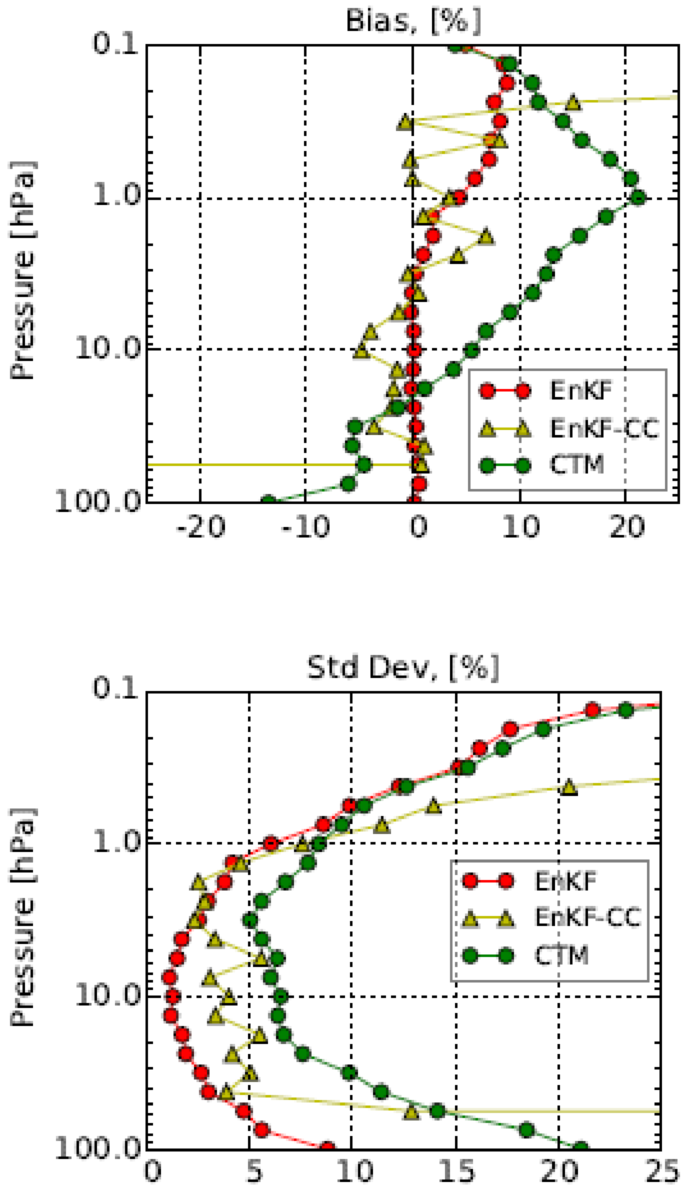


NOx



Outstanding issues with multi-specie assimilation

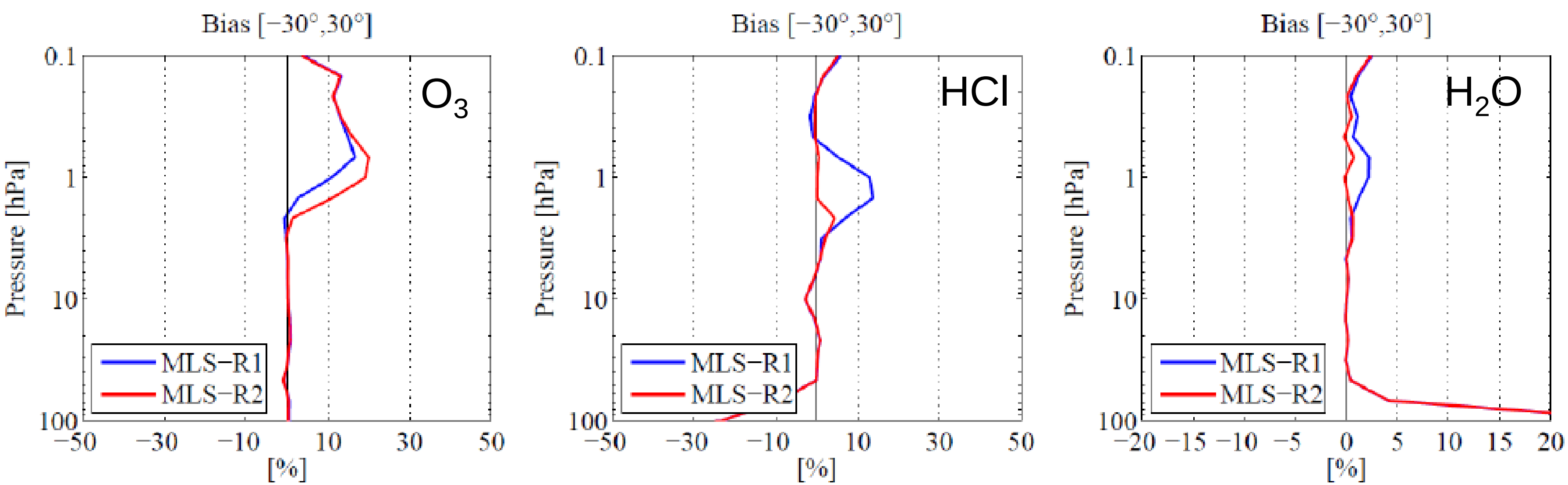
EnKF – O3 and N2O assimilation. Problem of specie localization



Outstanding issues with multi-specie assimilation

4D-Var – O3 assimilation. Problem due to model error

- BASCOE model suffers from an "ozone deficit" : overestimation of MLS around 1hPa by 20%
- "For ozone below 70 km, we continue to report a photochemical model deficit relative to observations… in the 10-50% range…" (Siskind et al., JGR, 2013)
- Rejecting O3 obs. above 3 hPa in the assimilation removes the bias in HCl (and H2O)



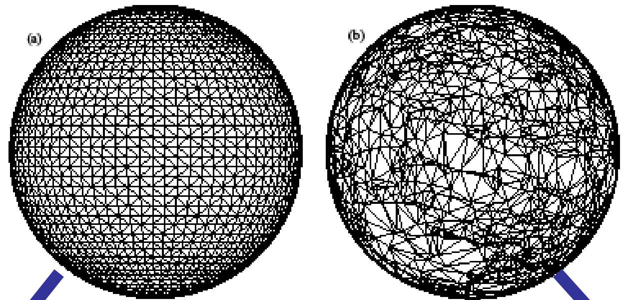
R1: As REAN01 but MLS O3 assimilated up to 0.1 hPa
R2=REAN01: Idem with O3 observations rejected above 3 hPa

5.3 Lagrangian KF

$P^f(\mathbf{x}_1(t;\mathbf{X}_1),\ \mathbf{x}_2(t;\mathbf{X}_2),\ t) = P^f(\mathbf{X}_1,\ \mathbf{X}_2, 0)$ with no model error

(Lyster et al. 2004)

- One set of trajectories
- Delaunay triangulation for **H**. comp ~ *N log N*
- Remapping each 2 to 3 days

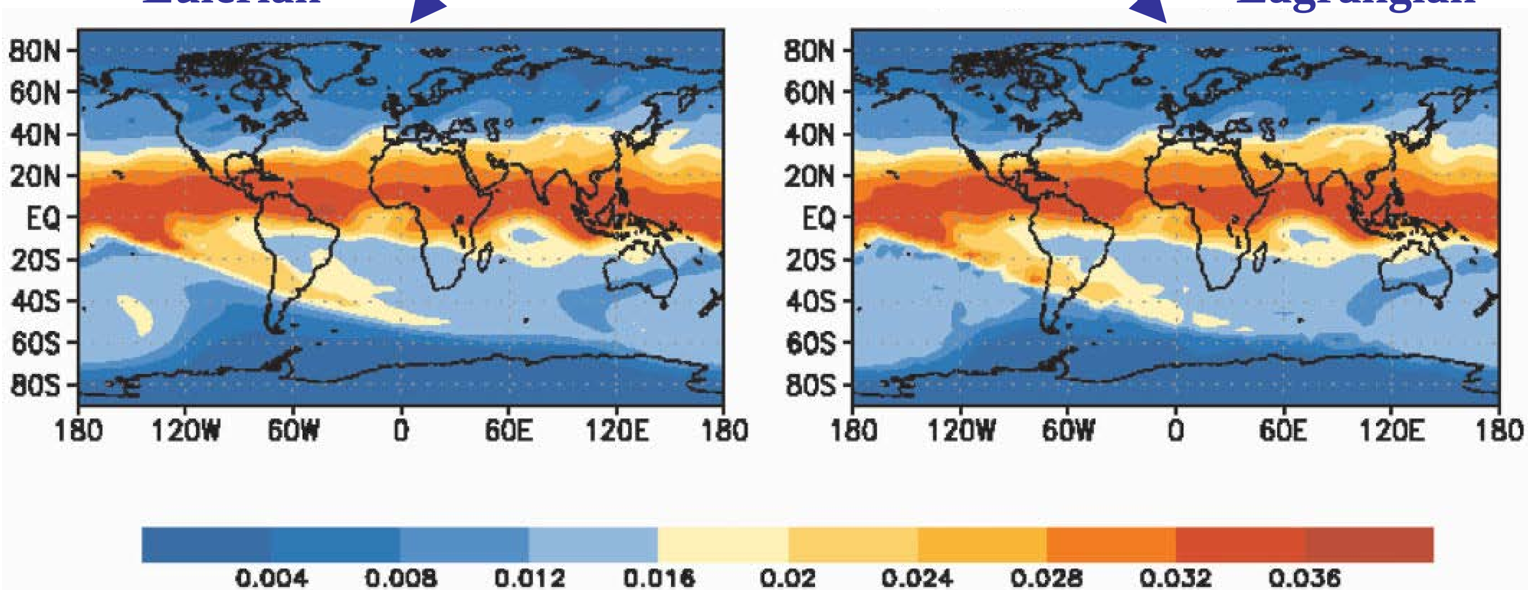


covariance evolution ~ $O(\mathrm{N}^2)$

covariance evolution ~ $O(\mathrm{N})$

Eulerian

Lagrangian



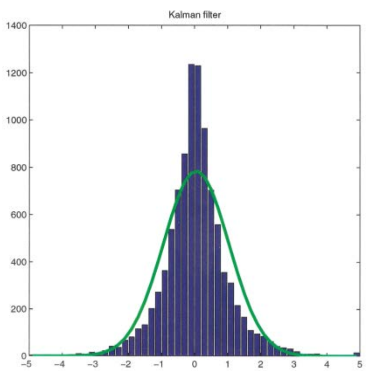
- Lagrangian analysis are much more noisy
- Remapping (field and cov.) is needed each 2-3 days because of trajectories clumping

5.4 Lognormal KF

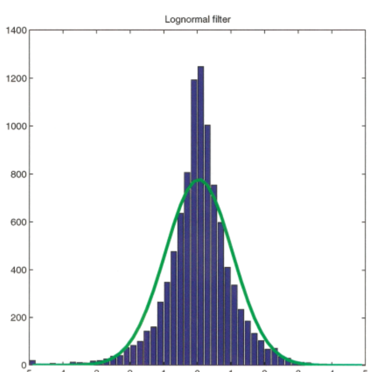
- KF relative error formulation

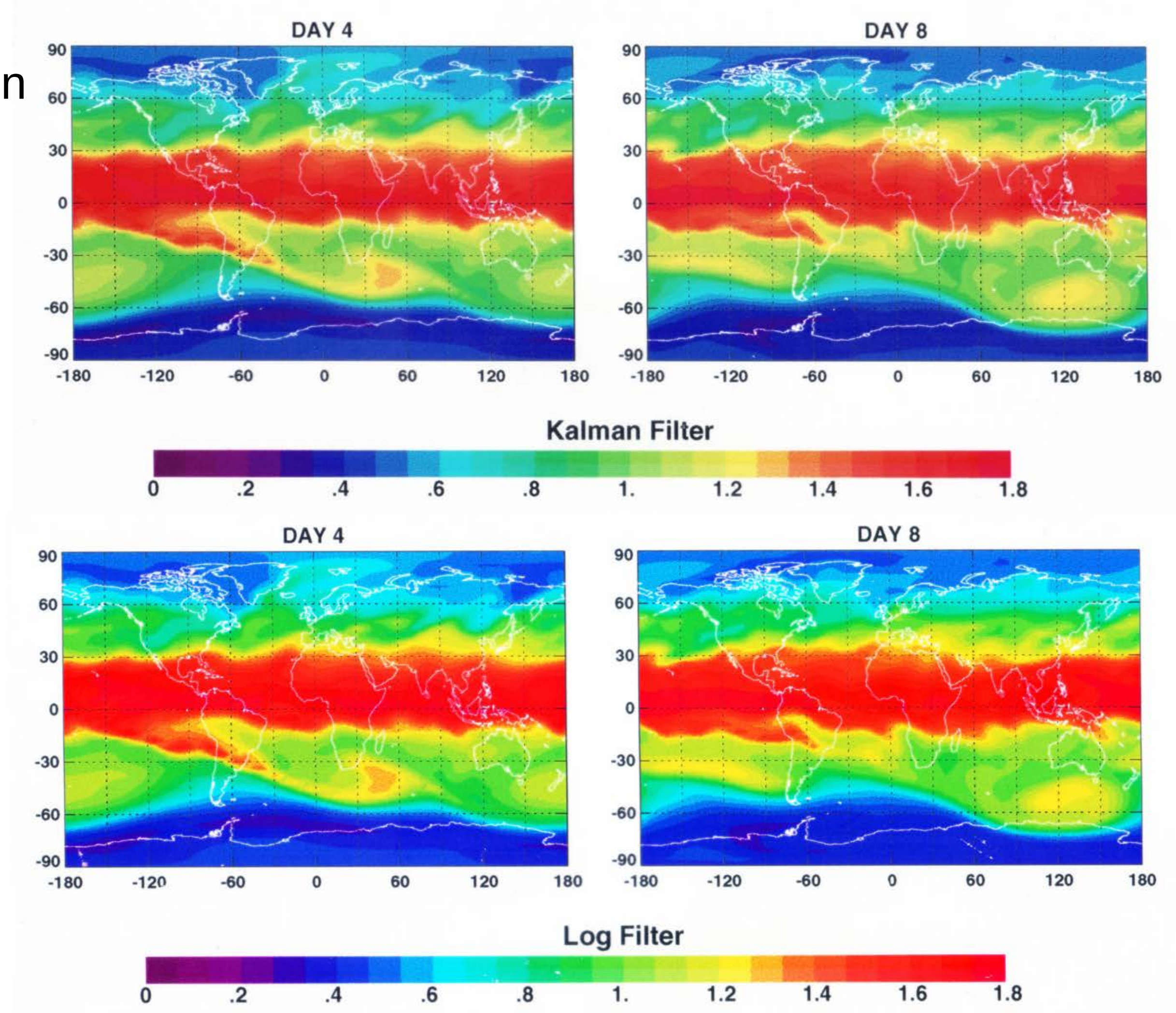
$$\boldsymbol{\varepsilon}^o = \boldsymbol{\mu}^f \circ \tilde{\varepsilon}^o$$

$$\boldsymbol{\varepsilon}^q = \boldsymbol{\mu}^a \circ \tilde{\varepsilon}^q$$



- Lognormal filter





So a lognormal formulation doesn't seem to resolve the kurtosis of the OmF distributions

5.5 Sequential filter

(Khatattov et al. 2000, Dee 2003, Eskes et al. 2003, Rösevall et al. 2007, van der A et al. 2010)

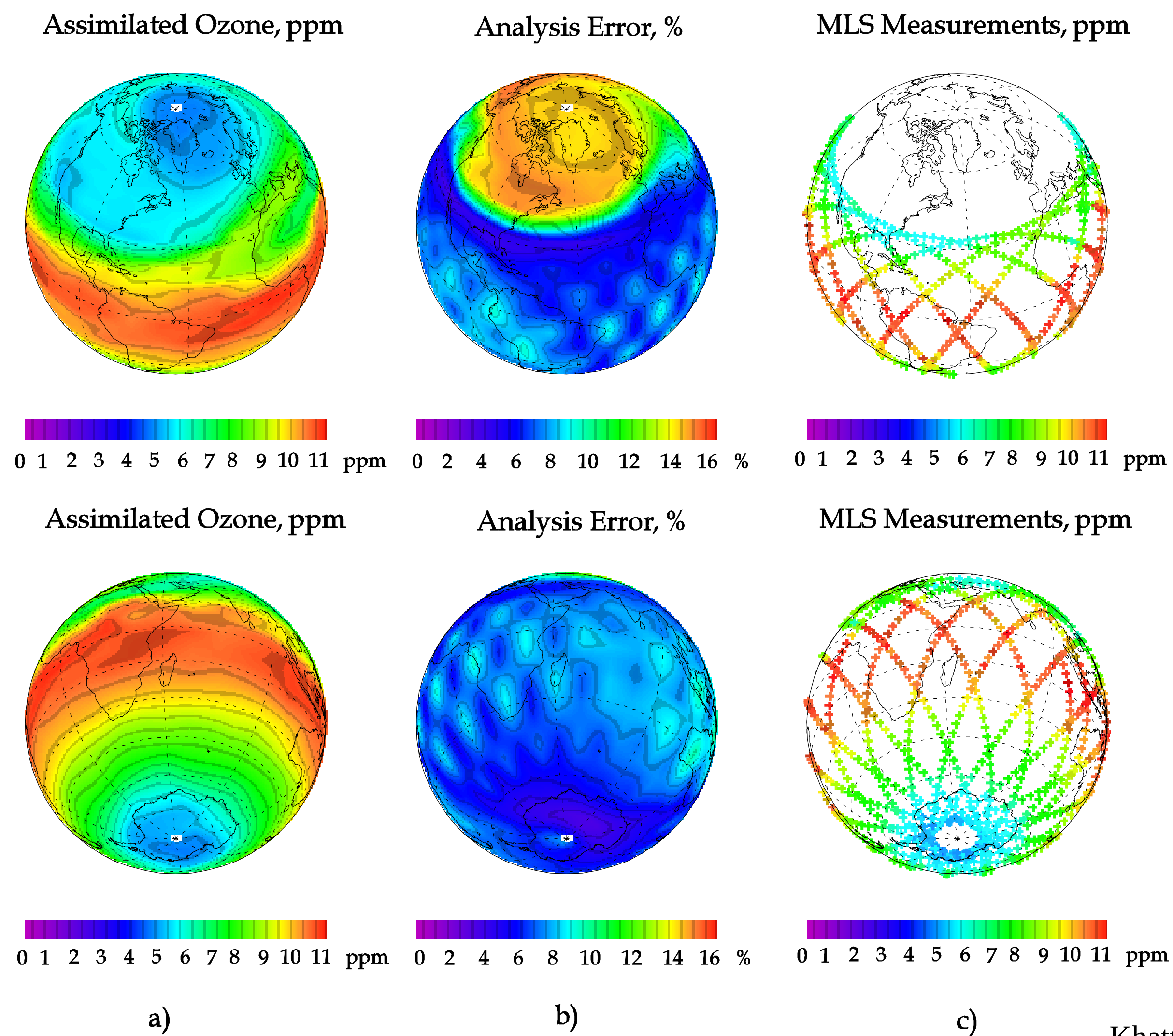
- Error variance evolution using the method of characteristics
- Analysis error variance computed using Chloleski decomposition or sequential variance update (Dee 2003)
- Error correlation kept fixed
- Model error variance estimated by innovation statistics

Has been applied to 3D CTM of long-lived species in
- Stratosphere (UARS, GOME, Flight planning for measurement campaign)
- Troposphere (MOPITT)

also to multispecies, and to
- humidity in the troposphere

Using a Choleski decomposition (for small matrices ~ 2000 or less) and a prescribed error correlation we can calculate the analysis error variance

$\mathbf{v}^a(\mathbf{x}_i) = \mathbf{v}^f(\mathbf{x}_i) - \mathbf{p}_i^T \left(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R}\right)^{-1} \mathbf{p}_i$ where $\mathbf{p}_i$ is the column of $\mathbf{P}^f$ associated with $\mathbf{x}_i$



Khattatov et al. 2000

Summary

1. We have discussed how construct covariance matrices for DA from standard correlation functions constructed in an infinite domain
2. Characteristics of correlation functions such as smoothness and correlation length was also discussed
3. We examined the spectral and orthogonality properties of the analysis
4. We discussed how AQ analysis (alone) is useful of health studies
5. We presented a full analytical solution (in spectral form) of the Kalman filter with H=I
6. We presented real applications of KF and discussed
   - the importance of error correlation
   - compared the EnKF with 4D-Var with and without chemistry, and discuss some outstanding issues
   - discussed a Lagrangian KF that reduces considerably the cost
   - show that the kurtosis of OmF is not an effect of Gaussian/Lognormal distributed errors
   - put into context what is a sequential filter

Thank you

References

Bartello, P. and Mitchell, H. L., 1992: A continuous three-dimensional model of short-range forecast error covariances, Tellus A, 44, 217, doi:10.1034/j.1600-0870.1992.t01-2-00002.x

Crouze, D.L., P.A. Peters, P. Hystad, J.R. Brook, A. van Donkelaar, R.V. Martin, P.J. Villeneuve, M. Jerrett, M.S. Goldberg, C.A. Pope III, M. Brauer, R.D. Brook, A. Robichaud, R. Ménard, R.T. Burnett. 2014: Ambient PM2.5, O3, and NO2 exposures and association with mortality over 16 years of follow-up in the Canadian Census Health and Environment Cohort (CanCHEC). Environ. Health Perspect., 123:1180-1186, http://dx.doi.org/10.1289/ehp.1409276 (accessed 2 November 2015)

Daley, R. 1991: Atmospheric Data Analysis. New York: Cambridge University Press. 457 pp.

Daley, R., and R. Ménard, 1993: Spectral characteristics of Kalman filter systems for atmospheric data assimilation. Mon. Wea. Rev., 121:1554-1565.

Dee, D., 2003: An adaptive scheme for cycling background error variances during data assimilation. ECMWF/GEWEX workshop on humidity analysis pp.15 (July 2002).

Desroziers, G., L. Berre, B. Chapnik, and P. Poli. 2005. Diagnostic of observation, background and analysis-error statistics in observation space. Q. J. R. Meteorol. Soc., 131:3385-3396. doi:10.1256/qj.05.108.

Errera, Q, and R. Ménard, 2012: Technical Note: Spectral representation of spatial correlations in variational assimilation with grid point models and application to the Belgian Assimilation System for Chemical Observations (BASCOE). Atmos. Chem. Phys., 12:10015-10031. www.atmos-chem-phys.net/12/10015/2012/ doi:10.5194/acp-12-10015-2012

Eskes, H. J., P. F. J. Van Velthoven, P. J. M. Valks, and H. M. Kelder, 2003: Assimilation of Gome total-ozone satellite observations in a three-dimensional tracer-transport model. Q. J. R. Meteorol. Soc., 129:1663–1681.

Gaspari, G. and S.E. Cohn, 1999. Construction of correlation functions in two and three dimensions. Q. J. R. Meteorol. Soc., 125:723-757.

Kalman, R. E., 1960: A new approach to linear filtering and prediction problems. Journal Basic Engineering, 82:35–45.

Khattatov, B., J.-F. Lamarque, L. Lyjak, R. Ménard, P. Levelt, X. Tie, G. Brasseur, and J. Gille, 2000: Assimilation of satellite observations of long-lived chemical species in global chemistry transport models. J. Geophys. Res., 105:29–135.

Hollingsworth,A., and P. Lönnberg. 1986. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. Tellus A, 38:111-136. doi:10.1111/j.1600-0870.1986.tb00460.x

Horn R.A, and C.R. Johnson, 1985. Matrix Analysis. Cambridge University Press. 607 pp.

Lyster, P.M., S.E. Cohn, R. Ménard, L.-P. Chang, S.-J. Lin and R.G. Olsen, 1997: Parallel implementation of a Kalman filter for constituent data assimilation. Mon. Wea. Rev., 125:1674-1686.

Lyster, P., S. Cohn, B. Zhang, L.-P. Chang, R. Menard, K. Olson, and R. Renka, 2004: A lagrangian trajectory filter for constituent data assimilation. Q. J. R. Meteorol. Soc., 130 (601):2315–2334, 10.1256/qj.02.234.

Ménard, R., S. Cohn, L.-P. Chang, and P.M. Lyster, 2000: Assimilation of stratospheric chemical tracer observations using a Kalman filter. Part I: Formulation. Mon. Wea. Rev., 128:2654-2671.

Ménard, R. and L.P. Chang. 2000. Assimilation of stratospheric chemical tracer observations using a Kalman filter. Part II: -validated results and analysis of variance and correlation dynamics. Mon. Wea. Rev., 128:2672-2686.

Ménard, R., and A. Robichaud. 2005. The chemistry-forecast system at the Meteorological Service of Canada. In ECMWF Seminar Proceedings on Global Earth-System Monitoring. 5-9 Sept 2005, pp 297-308. Reading, UK.

Ménard, R., M. Deshaies-Jacques, and N. Gasset, 2016 : A comparison of correlation-length estimation methods for the objective analysis of surface pollutants at Environment and Climate Change Canada. Journal of the Air & Waste Management Association, DOI: 10.1080/10962247.2016.1177620

Rasmussen C.E. and C.K.I. Williams, 2006: Gaussian processes of machine learning. MIT Press. 248 pp., www.GaussianProcess.org/gpml

Robichaud, A. and R. Ménard. 2014. Multi-year objective analysis of warm season ground-level ozone and PM2.5 over North-America using real-time observations and Canadian operational air quality models. Atmos. Chem. Phys., 14:1769-1800. doi: 10.5194/acp-14-1769-2014.

Robichaud, A., R. Ménard, Y. Zaïtseva, and D. Anselmo. 2015. Multi-pollutant surface objective analyses and mapping of air quality health index over North America. Air Qual. Atmos. Health. DOI 10.1007/s11869-015-0385-9.

Rösevall, J.D., D.P. Murtagh, J. Urban and A.K. Jones. 2007. A study of polar ozone depletion based on sequential assimilation of satellite data from the ENVISAT/MIPAS and Odin/SMR instruments. Atmos. Chem. Phys., 7:899-911. www.atmos-chem-phys.net/7/899/2007/ (accessed 16 February 2007).

Skachko, S., Q. Errera., R. Ménard, Y. Christophe, and S. Chabrillat, 2014: Comparison of the ensemble Kalman filter and 4D-Var assimilation methods using a stratospheric tracer model. Geosci. Model Dev., 7:1451-1465. www.geosci-model-dev.net/7/1451/2014/ doi:10.5194/gmd-7-1451-2014

Skachko, S., R. Ménard, Q. Errera, Y. Christophe, and S. Chabrillat, 2016 : EnKF and 4D-Var data assimilation with a chemistry transport model. Geosci. Model Dev., doi:10.5194/gmd-2016-95, 2016

Stieb, D.M., R.T. Burnett, M. Smith-Dorion, O. Brion, H.H. Shin, and V. Economou. 2008. A new multipollutant, no-threshold air quality health index based on short-term associations observed in a daily time-series analyses. J. Air Waste Manage. Assoc., 58:435-450. doi:10.3155/1047-3289,58.3,3,435.

van der A, R. J., M. A. F. Allaart, and H. J. Eskes, 2010: Multi sensor reanalysis of total ozone. Atmos. Chem. Phys., 10 (22):11 277–11 294, 10.5194/acp-10-11277-2010.