Satellite data information content, channel selection and density

Élisabeth Gérard

Météo-France/CNRM/GMAP

Toulouse, France

[EG2], 2nd ENVISAT summer school
Frascati, Italy, 16-26 Aug 2004

Overview

A. Information content
   - Theory
   - Case study
   - Cf. other example in EG3 (WALES)
B. Channel selection
   - 4 methods
   - Illustration for IASI
C. Influence of observation resolution
   - Optimal observation density

A. Information content

- Introduction of useful concepts
  - No. of Degrees of Freedom for Signal
  - Entropy reduction
  - Vertical resolution
- A THORPEX case study

Degrees of Freedom for Signal

The number of degrees of freedom can be smaller than the number of measurements

Notations: $y=Hx+\varepsilon$, $x_a = x_b + K(y-Hx_b)$, $K= A H^T R^{-1}$

Analysis error covariance matrix: $A^{-1}=B^{-1}+H^T R^{-1} H$

Normalisation $x'=B^{-1/2}x$ $y'= R^{-1/2}y$ $\varepsilon'= R^{-1/2} \varepsilon$

Normalised Jacobian matrix: $H'=R^{-1/2}H B^{1/2}$

Thus: $y'=H'x'+ \varepsilon'$ and $Cov(y')=H'H'^T + I$

1st term (variability of the atmosphere)
2nd term (noise)

Degrees of Freedom for Signal

SVD (singular vector decomposition) for H':

H'=U? V^T with U^TU=I, V^TV=I

y ''= U^Ty =U^T(U? V^T x'+ e')=? V^Tx''+e''

With x''= V^Tx', e''= U^Te'

Cov(x'')=V^TV=I, Cov(e'')= U^TU=I

Cov(y'')= ?^{2}+ I where ?^{2} is due to variability, I to noise

The elements of y'' which varie more than the noise are the ones for which λ_i^2 is greater than 1

The number of effective independent measurements is the number of singular values of $R^{-1/2}H\,B^{1/2}$ that are greater than 1

Degrees of Freedom for Signal

Formalisation of the concept of
« degrees of freedom for signal » (DFS):

$x_a = x_b + K(y - Hx_b)$,

$K = A H^T R^{-1} = (B^{-1} + H^T R^{-1} H)^{-1} H^T R^{-1} = BH^T (HBH^T + R)^{-1}$

$x_a - x_b = K(y - Hx_b)$,

$DFS = E((x_a - x_b)^T B^{-1} (x_a - x_b))$

The DFS quantifies what is brought in by the analysis

$DFS = E(tr((x_a - x_b)(x_a - x_b)^T B^{-1}))$

$DFS = tr(E((x_a - x_b)(x_a - x_b)^T) B^{-1})$

Degrees of Freedom for Signal

$E((x_a - x_b)(x_a - x_b)^T) = E(K(y - Hx_b)(y - Hx_b)^T K^T)$

$= K(R + HBH^T)K^T = BH^T(HBH^T + R)^{-1}HB$

$DFS = tr(BH^T(HBH^T + R)^{-1}H) = tr(KH) = tr(HK)$

Knowing that $AB^{-1} + KH = A(B^{-1} + H^T R^{-1} H) = I$

Thus $DFS = tr(KH) = tr(I - AB^{-1})$

$DFS = tr(KH) = tr(G)$ with G the

« Model Resolution Matrix »: $x_a - x_b = G(x_t - x_b)$ if $e = 0$

G shows to which extent the analysis represents the reality

$DFS = tr(I - AB^{-1})$

The more information you put into the system, the more A is different from B

Entropy reduction

- The entropy is the gaussian distribution of the covariance C: $E(C) = cst + 1/2 \log_2 |C|$

  E measures the volume of the space occupied by the probability law that describes the knowledge of the system

- When a measurement is performed, this « uncertainty volume » decreases and the entropy reduction is:

  $$ER = 1/2 \log_2 |B| - 1/2 \log_2 |A|$$

  $$ER = -1/2 \log_2 |AB^{-1}|$$

Vertical resolution

The inversion can be characterised by:

- The matrix A = $Cov(x_a - x_t)$
- The matrix G: $x_a - x_b = G\,(x_t - x_b)$ if $e=0$
  G indicates to which extent the analysis represents the reality. In particular, the vertical resolution of G indicates how the analysis smoothes the reality.

Vertical resolution = dz_i / G_{ii}

where dz_i is the depth of layer i

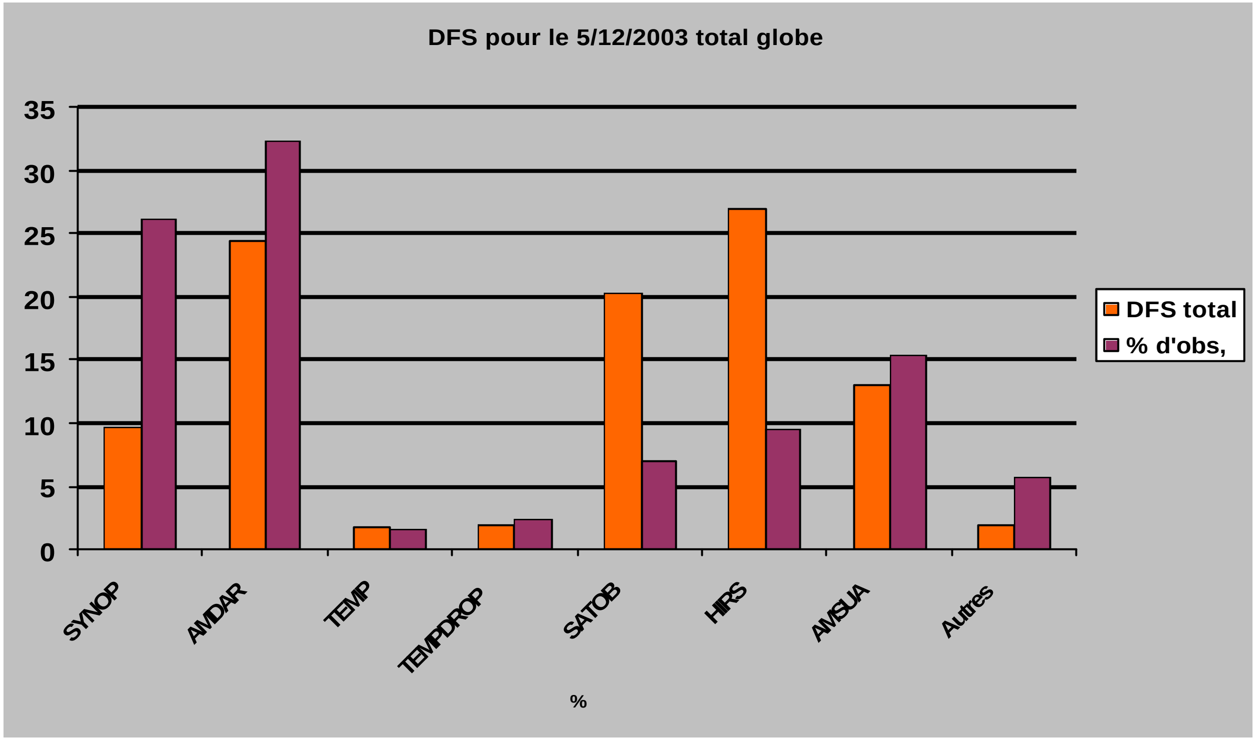
and G_{ii} is the corresponding diagonal element of G

DFS for a case study (5 Dec 2003)

GLOBAL

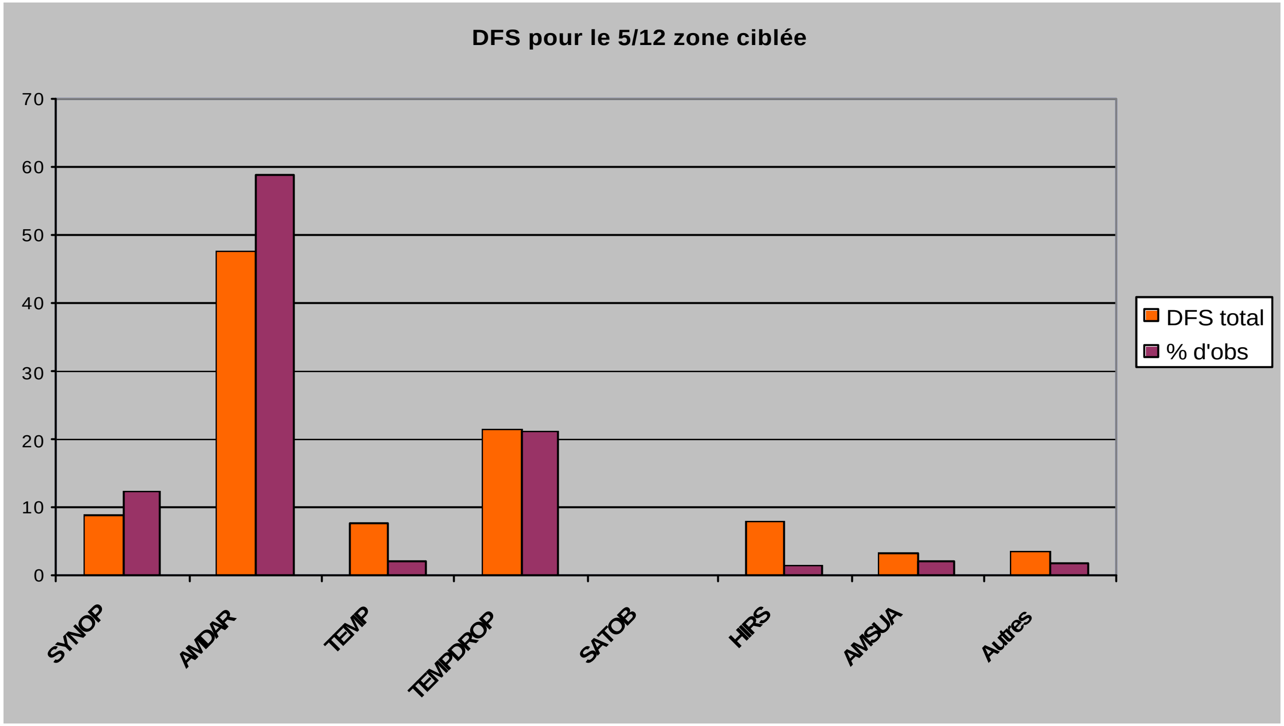
of the North Atlantic TReC (THORPEX Regional Campaign)

15 Oct 2003 – 14 Dec 2003



Marchal, 2004

DFS for a case study (5 Dec 2003) **TARGETED AREA**
of the North Atlantic TReC (THORPEX Regional Campaign)
15 Oct 2003 – 14 Dec 2003



Marchal, 2004

B. Channel selection

- How to choose a subset of channels to perform an inversion when thousands of channels are available ?
- 4 methods
  - DRM method
  - SVD method
  - Iterative method
  - Jacobian method
- Illustration for IASI

DRM method

(Menke)

- « Data Resolution Matrix »: DRM=HK
- As y_a-y_b = DRM (y-y_b), the diagonal elements of DRM indicate how much weight an observation has in its own analysis
- These diagonal elements measure the « importance » of the different channels
- The method implies the computation of matrix A

SVD method

(Prunet)

- SVD of H
- $G = R^{-1/2} H B^{1/2} = U ? V^T$
- Truncation in $?^2$ so that the eigenvalues of $G^T G = B^{1/2} H^T R^{-1} H B^{1/2}$, equivalent to s_b^2 / s_o^2, represent 10% of the contribution of the observations in the analysis
- $G = R^{-1/2} H B^{1/2} => U_p ?_p V_p^T$
- DRM = $V_p V_p^T$. Its diagonal elements are used in terms of « importance » of the channels

Iterative method

(Rodgers)

- This method enables an iterative selection of the channels. At each step, the most interesting channel is chosen and the matrix $B_i=A_{i-1}$ is updated
- After normalising the Jacobian with R
  - $A_i^{-1}=B_i^{-1}+h^Th$
  - where $B_0=B$ and h is a line of H
- The selection criterion is the DFS or ER:
  - $DFS(h)_i=Tr(I-AB_i^{-1})=h^TB_ih/(1+ h^TB_ih)$
  - $ER (h)_i=-1/2 \log_2 det(AB_i^{-1})=1/2 \log_2(1+ h^TB_ih)$

Jacobian method

(Goldberg, Aires)

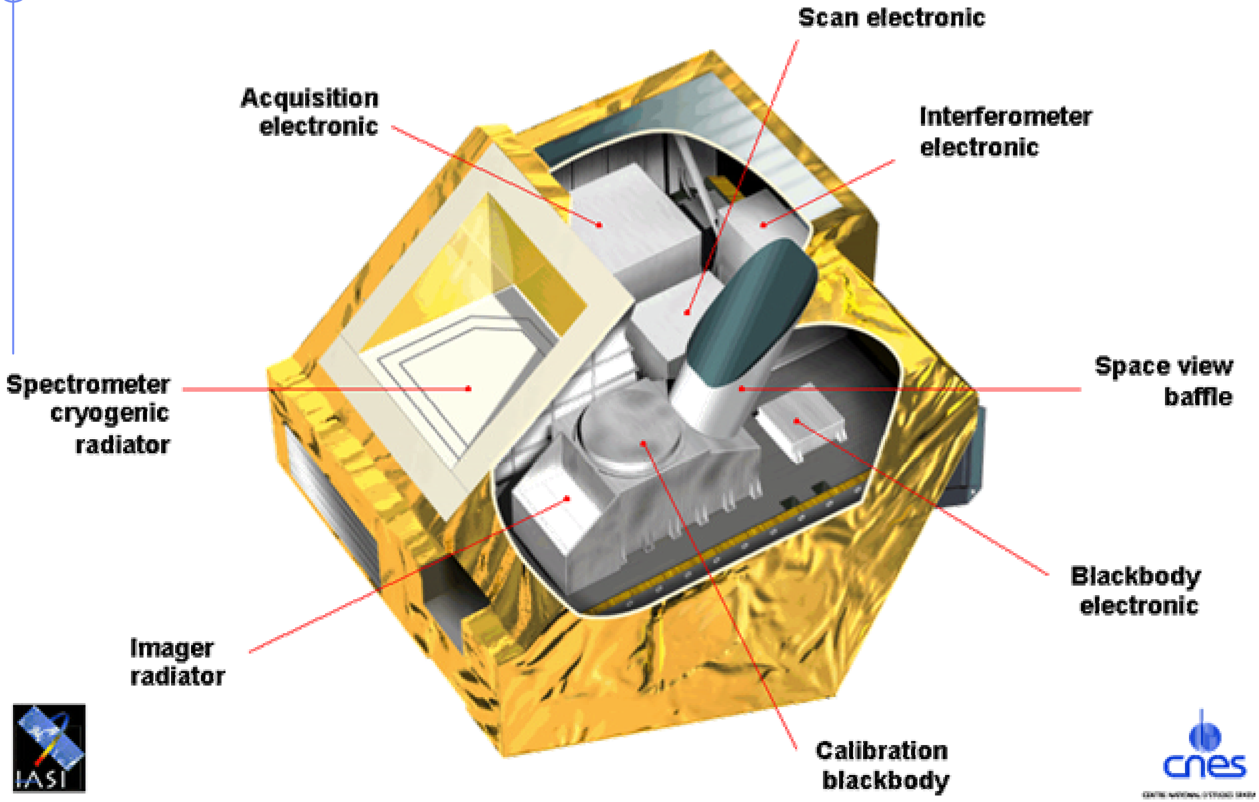
- Method based on weighting functions that describe the channel sensitivity to the atmosphere parameters
- Normalisation of H: $R^{-1/2}HB^{1/2}$
- For each parameter, at each vertical level, one channel is chosen:
  - Amongst the ones which maximum is at this level
  - With the greatest ratio:

  Intensity of the max / width of the weighting function

Illustration for IASI

- Infrared Atmospheric Sounding Interferometer (Michelson interferometer)
- IASI = 8461 radiances in each pixel
- The 4 different methods have been compared for 3 stations representative of the midlatitudes, the Tropics and the polar regions
- For each station, 24 profiles (T,Q,O_3) observed in a year

IASI: launch in 2006 on Metop



Remote sounding from IASI

Fig. 1 : Typical atmospheric profiles: water vapour mixing ratio

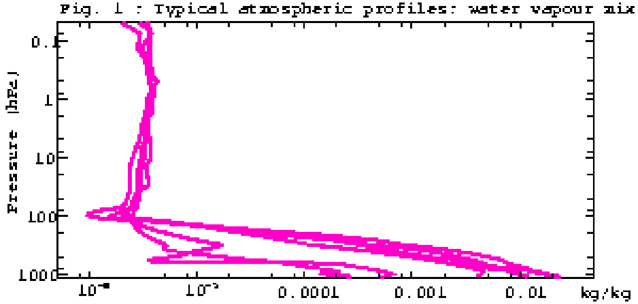


Fig. 2 : Typical atmospheric profiles: temperature

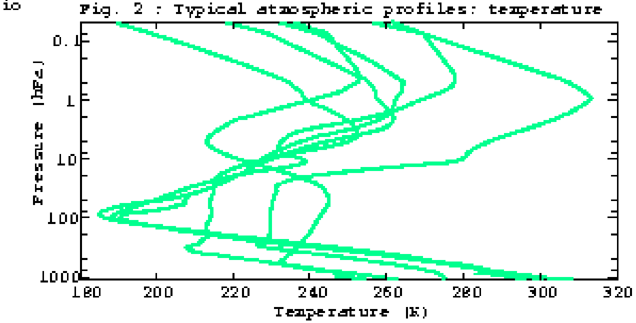
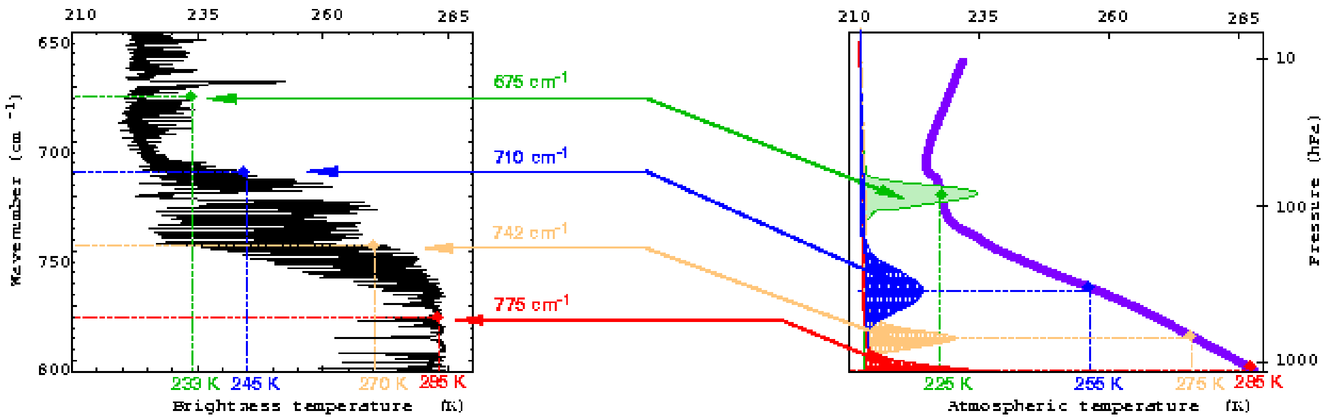
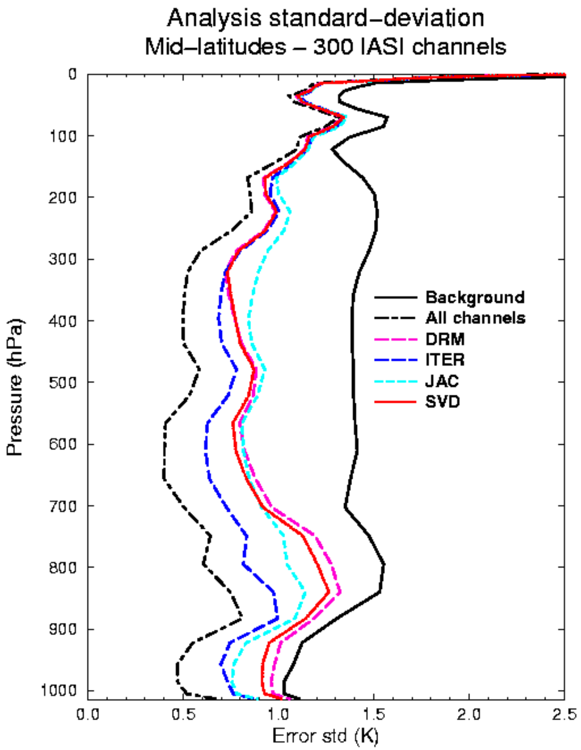
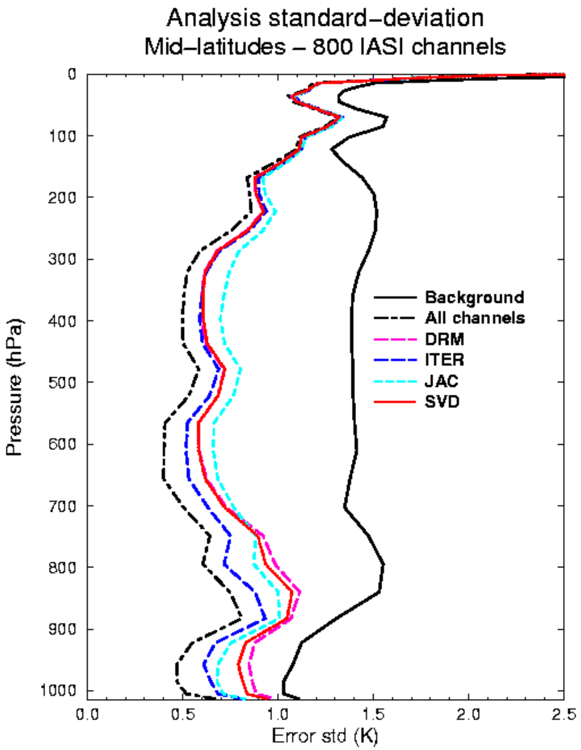


Fig. 3: Correlation between the CO_2 absoption spectrum and the atmospheric temperature profile

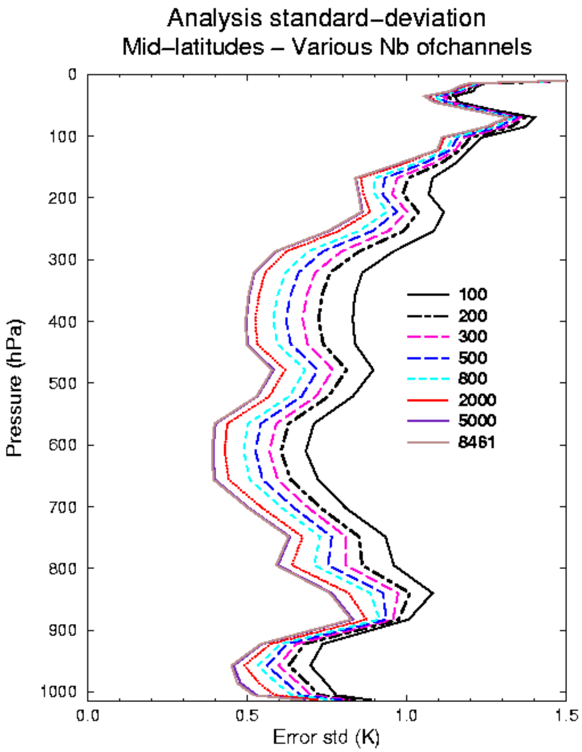


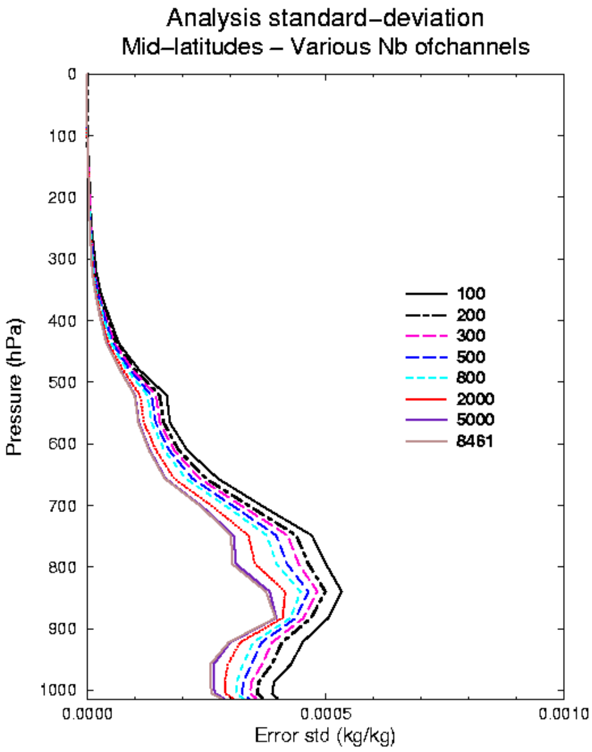
Midlatitudes - T



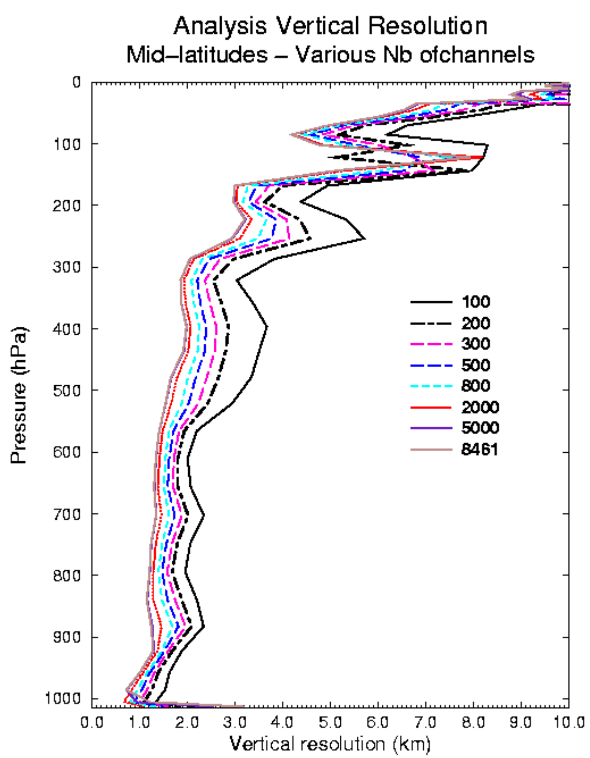


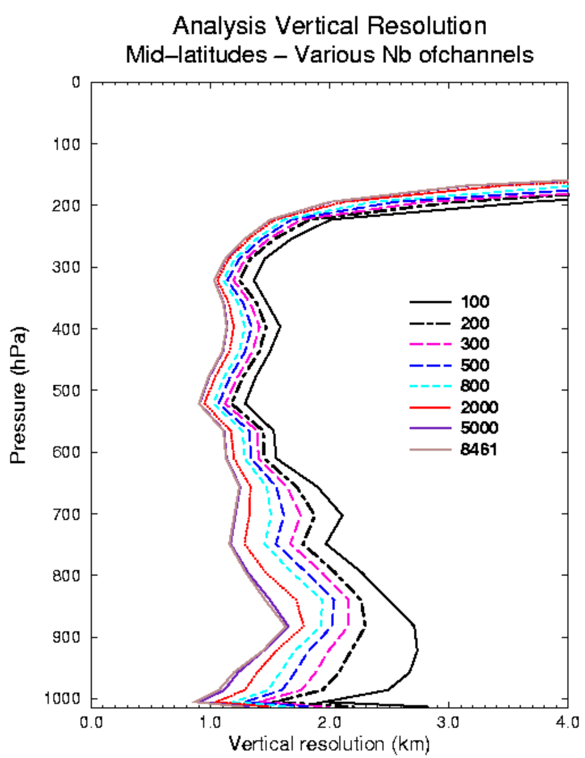
Accuracy / No. of channels



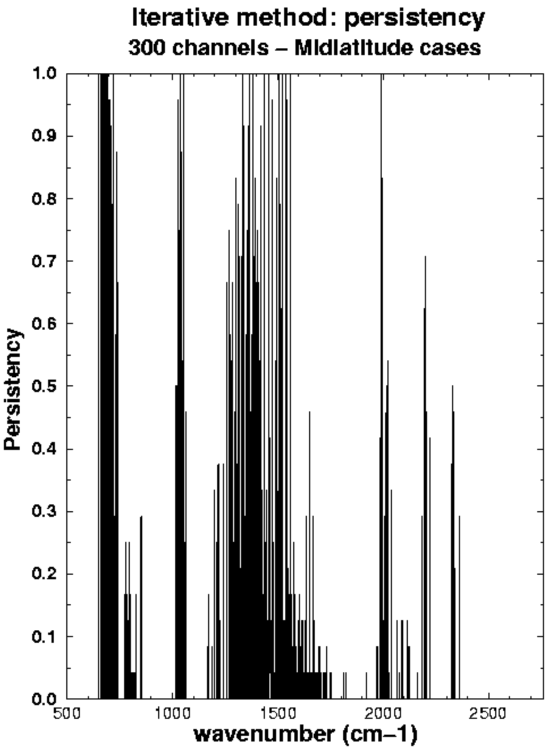


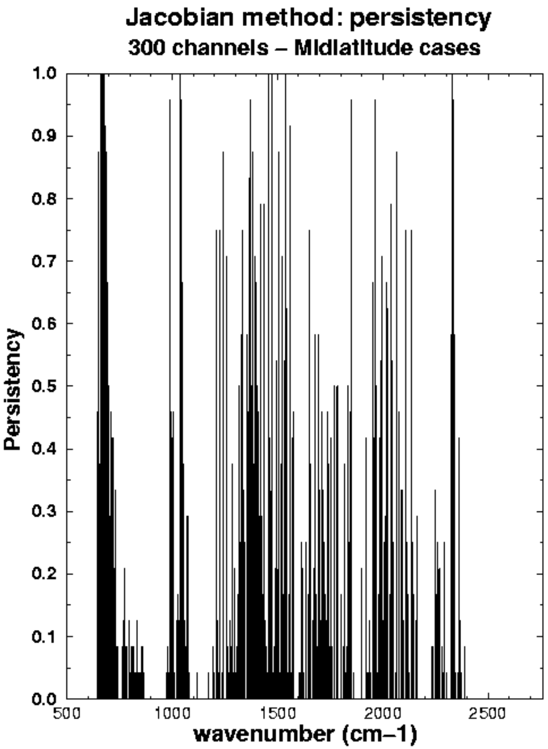
Vert. resolution / No. of channels



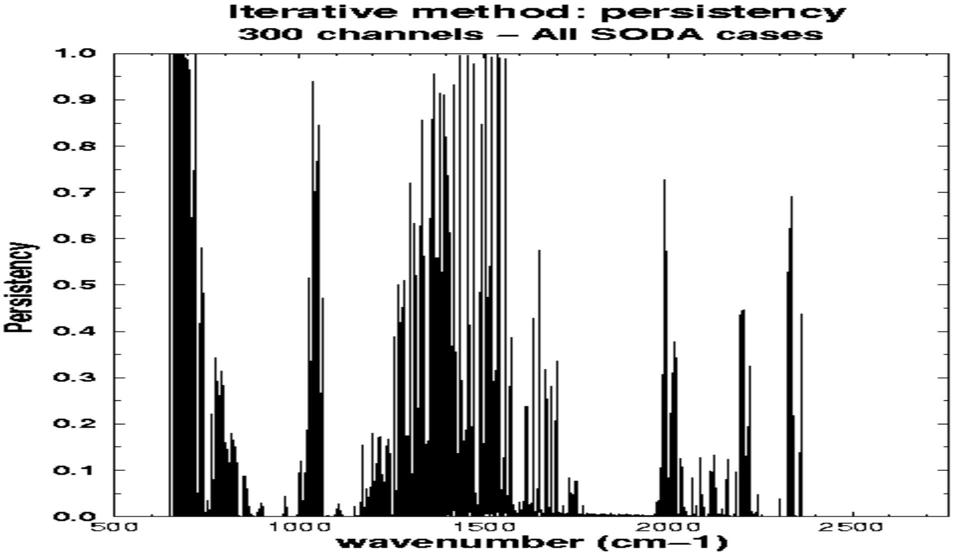


Channel selection – Midlatitudes





Statistics on 518 SODA profiles



Amongst 8461 channels, 28 are used for each selection and 5506 are never used

DFS: 10.7 for T, 9.0 for Q and 1.3 for O_3

Conclusions

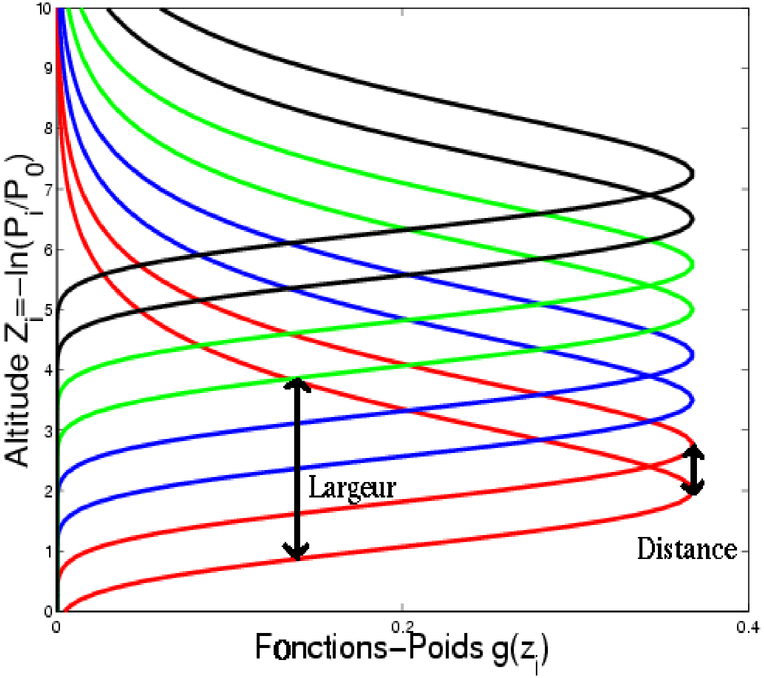
- The iterative method for channel selection gives the best results
- A method based on the Jacobians is efficient in the low troposphere and the method based on SVD is efficient in the high troposphere
- Only 1/3 of the channels are common to selections by ITER et JAC
- The channels are chosen amongst ~3000
- Very few channels are systematically chosen

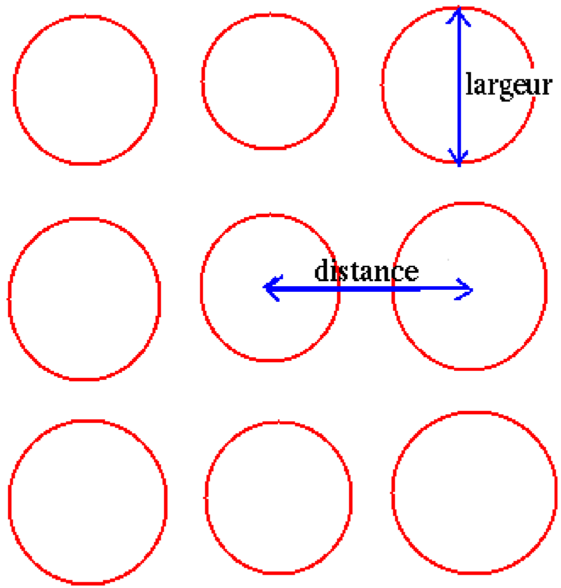
C. Influence of observation resolution on data assimilation

Liu and Rabier, 2002 [QJRMS]

- More and more satellite data at present and in the future
- Only 10 to 20 % of these data are used !
- A few considerations
  - Analysis grid vs. distance between obs
  - Error correlation length vs. distance between obs
  - Observation error correlations not accounted for
- How to determine the optimal sampling ?
- What is the optimal observation size for a given analysis resolution and observation density ?

Observation resolution





Vertical resolution

Horizontal resolution

General context

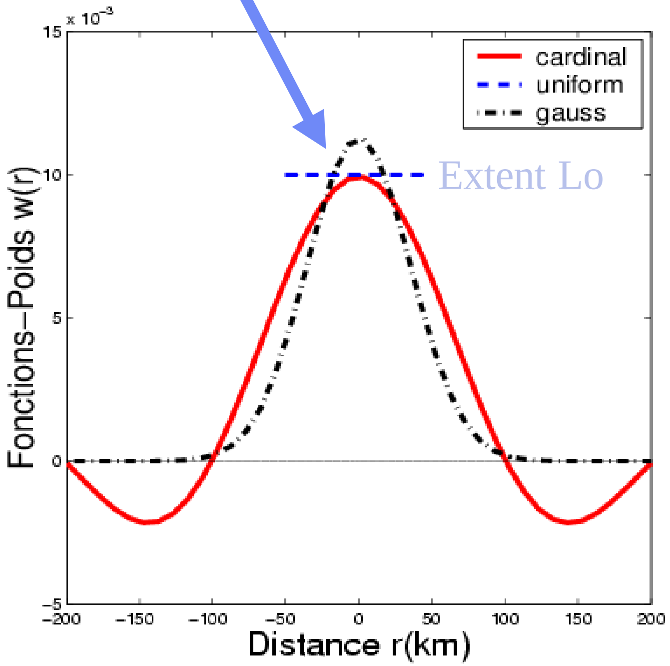
- A periodical domain 1D: L=8000 km
- Model variables are Fourier coefficients. Δx=100 km. No dynamic
- The background error is correlated with a correlation length L_b=208 km, σ_b=1
- The remotely sensed observation has an extent L_o, an interval Δy, the instrument error (uncorrelated) is σ_o=1
- Analysis error covariance matrix

$$\mathbf{A}=(\mathbf{I}-\mathbf{KH})\mathbf{B}(\mathbf{I}-\mathbf{KH})^T+\mathbf{KRK}^T$$

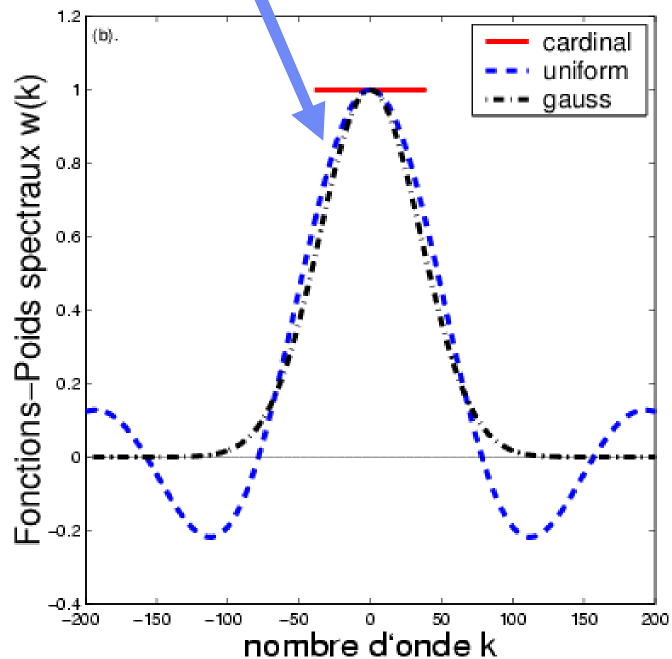
Representation of model and observation weighting functions

$$y(r_i)=\int w(r-r_i)x(r)dr+e_i^w+e_i^m$$

$$y(r_i)=\sum_{k=-\infty}^{+\infty}\hat{w}_k\hat{x}_k\exp(ikr_i)+e_i^w+e_i^m$$



In physical space

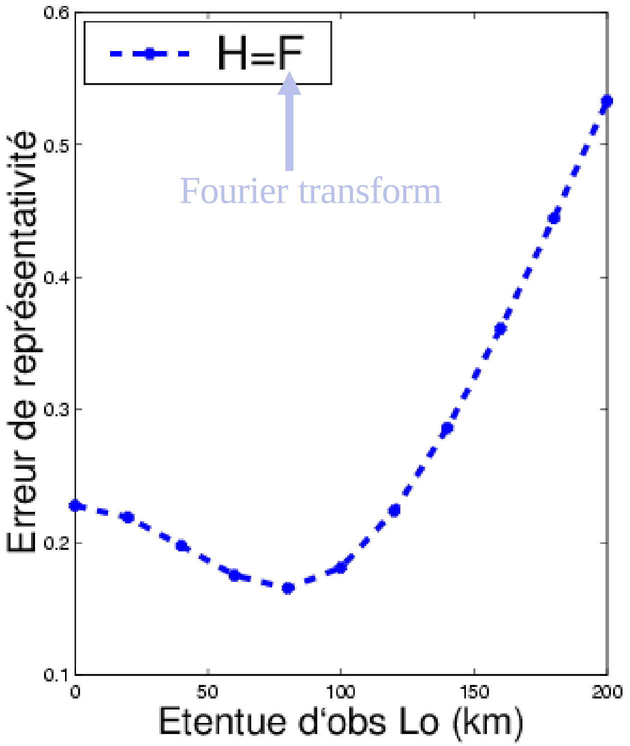


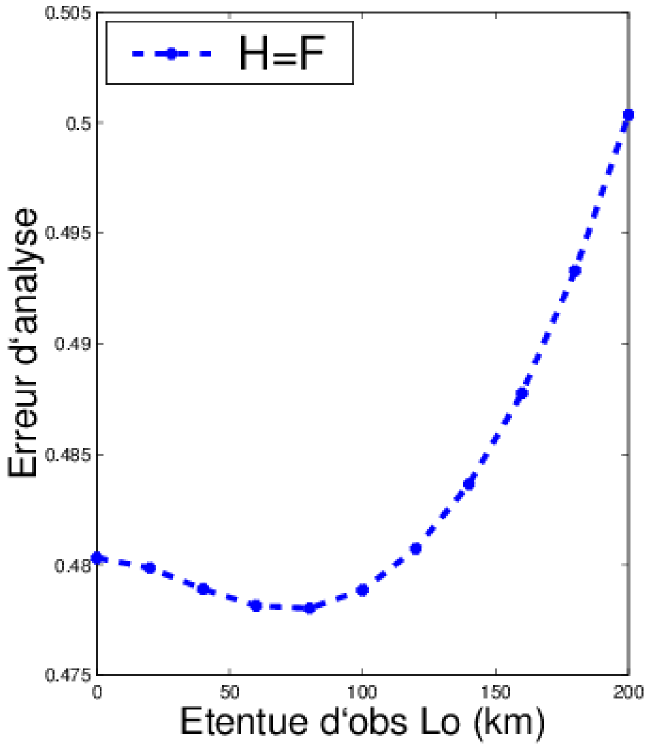
In Fourier space

Representativeness and analysis errors as a function of observation extent

$\sigma_b=\sigma_o=1$

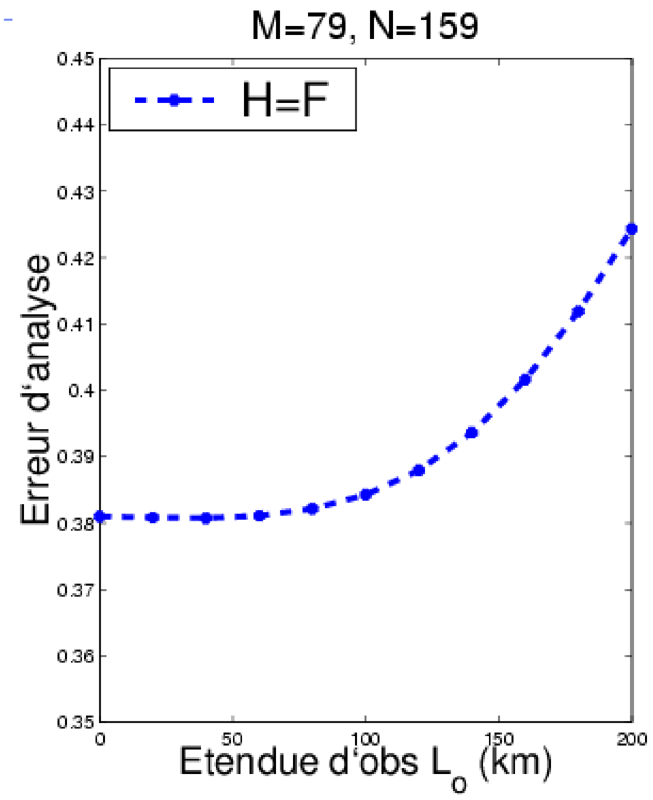
Δx=100km, Dimension of obs and model N=M=79,
Observation weighting function is uniform



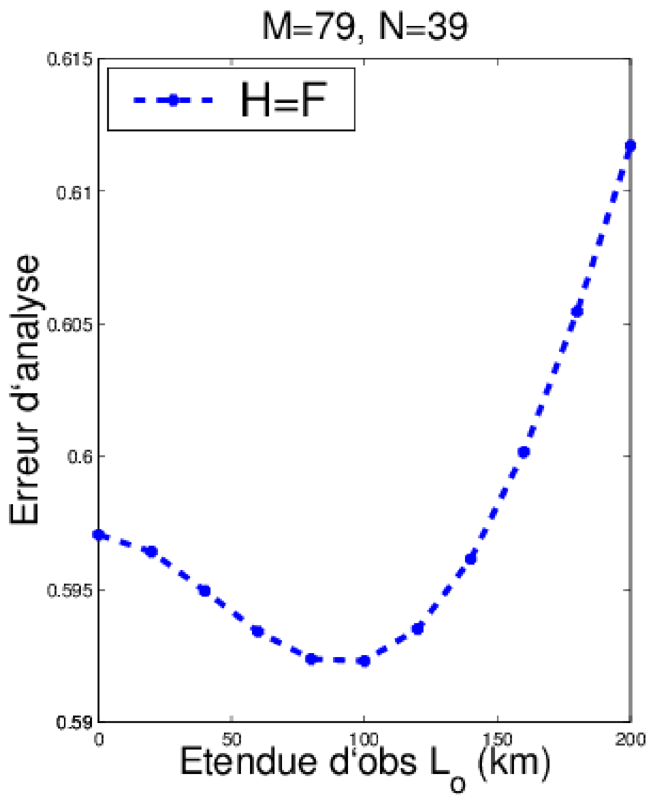


Analysis error as a function of observation extent

Obs no. Nx 2



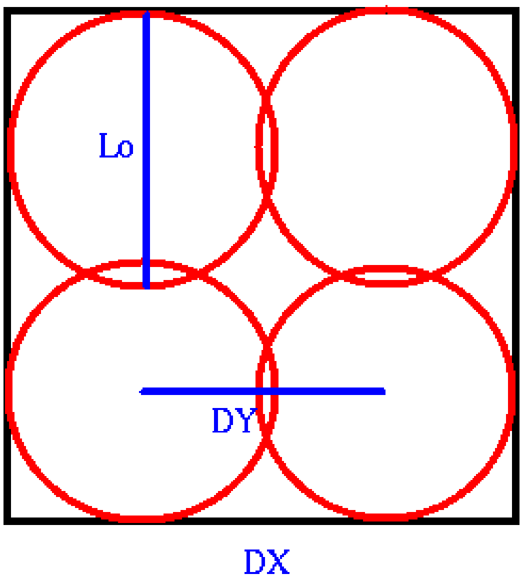
Obs no. N/2

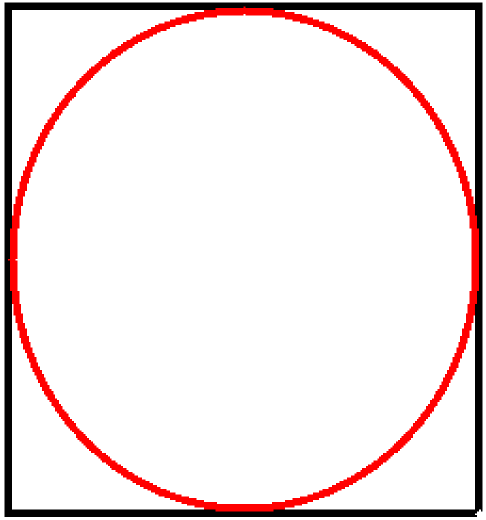


Optimal observation extent

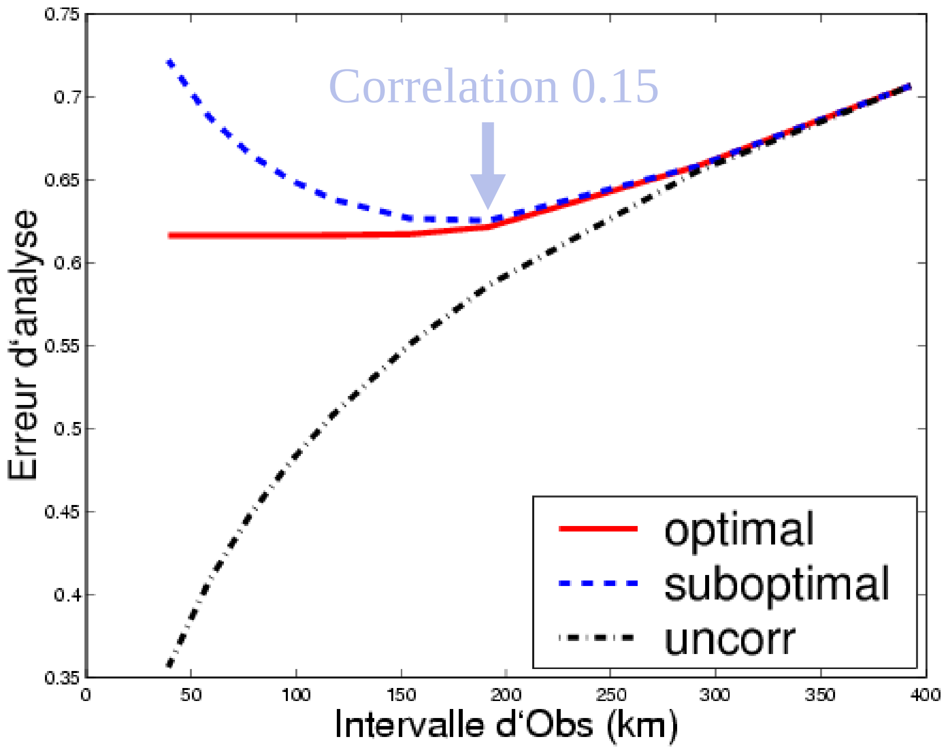
Optimal Lo $\approx$ min(Δx,Δy)

More observations   Less observations





Optimal observation sampling



Configuration:

Grid box
Δx=100km

Background correlation
L_b=208km

Observation correlation
L=100km

$\sigma_b=\sigma_o=1$

Conclusions

- The optimal observation extent is equal to min(Δx,Δy)
- For uncorrelated observations, increasing the observation density improves the analysis
- For correlated data
  - Increasing the density beyond a threshold leads to a slight improvement even with an optimal scheme
  - Can even degrade the analysis in a sub-optimal scheme
  - An optimal sampling can extract the major part of the independent information present in the complete observation network