École Doctorale des Sciences de l'Environnement d'Île-de-France
Année Universitaire 2024-2025

Modélisation Numérique de l'Écoulement Atmosphérique et Assimilation de Données

Olivier Talagrand

Cours 6

28 Mai 2025

Last course (May 12)

- Lissage de Kalman.
- Assimilation variationnelle. Principe
- Méthode adjointe. Principe.
- Assimilation variationnelle. Résultats
- La Méthode incrémentale (bref)
- Compléments sur l'Estimation Statistique (*BLUE*)

Entropy of a probability distribution

Probability distribution over domain described by coordinate ξ, with probability density $p(\xi)$. *Entropy*

$$S \equiv -\int p \ln p \, d\xi$$

Entropy of a probability distribution is a measure of the associated uncertainty. The larger the entropy, the larger the uncertainty. A uniform probability distribution over an interval of length a has entropy $\ln a$, which tends to $-\infty$ as a tends to zero. A one-dimensional Gaussian probability distribution with variance s has entropy $\ln\sqrt{(2\pi e s)}$.

For given variance s, entropy is largest for the Gaussian distribution.

Entropy of a probability distribution (continuation)

Data of the form (see slide 51, course 5)

$$z = \boldsymbol{\Gamma x} + \zeta$$

The knowledge of a probability distribution for ζ defines a conditional probability distribution $P(\boldsymbol{x}|z)$ for $\boldsymbol{x}$. Assuming that only the expectation and covariance matrix $\boldsymbol{S}$ of ζ are known, for which distribution of ζ is the entropy of $P(\boldsymbol{x}|z)$ largest ?

Response. The entropy of $P(\boldsymbol{x}|z)$ is largest when ζ is Gaussian.

If the probability distribution for ζ is unknown, assuming that it is Gaussian is in a sense the 'least committing' choice.

This course

- Weak-constraint Variational Assimilation.

- Dual Algorithm for Variational Assimilation

- Complements on Variational Assimilation.
  - How to write (and validate) an adjoint code
  - Value of objective function at minimum. χ^2 test

- Compared qualities of Sequential and Variational Assimilation

- Assimilation and (In)stability. Quasi-Static Variational Assimilation

- Weak-constraint Variational Assimilation.

How to take model error into account in variational assimilation ?

Weak constraint variational assimilation

Allows for errors in the assimilating model

Data over time interval $k = 0, \ldots, K$

- Background estimate at time 0

$$x_0^b = x_0 + \zeta_0^b \qquad E(\zeta_0^b \zeta_0^{bT}) = P_0^b$$

- Observations at times $k = 0, \ldots, K$

$$y_k = H_k x_k + \varepsilon_k \qquad E(\varepsilon_k \varepsilon_{k'}^{T}) = R_k \delta_{kk'}$$

- Model

$$x_{k+1} = M_k x_k + \eta_k \qquad E(\eta_k \eta_{k'}^{T}) = Q_k \delta_{kk'} \qquad k = 0, \ldots, K\text{-}1$$

Errors assumed to be unbiased and uncorrelated in time, H_k and M_k linear

These data are of the general form

$$\boldsymbol{z} = \boldsymbol{\Gamma x} + \boldsymbol{\zeta}$$

the unknown $\boldsymbol{x}$ being now the temporal sequence of states $\boldsymbol{x} \equiv (\boldsymbol{x}_0^{\mathrm{T}}, \boldsymbol{x}_1^{\mathrm{T}}, \ldots, \boldsymbol{x}_K^{\mathrm{T}})^{\mathrm{T}}$, and the data vector $\boldsymbol{z}$ consisting of the initial background $\boldsymbol{x}_0^b$, the observations $\boldsymbol{y}_k$ $(k = 0, \ldots, K)$,

and the model equation $0 = \boldsymbol{M}_k \boldsymbol{x}_k - \boldsymbol{x}_{k+1} + \boldsymbol{\eta}_k$ $(k = 0, \ldots, K\text{-}1)$

Minimize corresponding scalar objective function

$$\xi \;\rightarrow\; \mathcal{J}(\xi) \;\equiv\; (1/2)\,[\boldsymbol{\Gamma\xi} - \boldsymbol{z})]^{\mathrm{T}}\, \boldsymbol{S}^{-1}\, [\boldsymbol{\Gamma\xi} - \boldsymbol{z}]$$

Objective function

$(\xi_0, \xi_1, \ldots, \xi_K) \rightarrow$

$\mathcal{J}(\xi_0, \xi_1, \ldots, \xi_K) \equiv$

$$(1/2)\ (\boldsymbol{x}_0{}^b - \xi_0)^{\mathrm{T}}\, [\boldsymbol{P}_0{}^b]^{-1}\, (\boldsymbol{x}_0{}^b - \xi_0)$$

$$+\ (1/2)\ \Sigma_{k=0,\ldots,K}\, [\boldsymbol{y}_k - \boldsymbol{H}_k \xi_k]^{\mathrm{T}}\, \boldsymbol{R}_k{}^{-1}\, [\boldsymbol{y}_k - \boldsymbol{H}_k \xi_k]$$

$$+\ (1/2)\ \Sigma_{k=0,\ldots,K-1} [\xi_{k+1} - \boldsymbol{M}_k \xi_k]^{\mathrm{T}}\, \boldsymbol{Q}_k{}^{-1}\, [\xi_{k+1} - \boldsymbol{M}_k \xi_k]$$

Can include nonlinear $\boldsymbol{M}_k$ and/or $\boldsymbol{H}_k$.

Becomes singular in the strong constraint limit $\boldsymbol{Q}_k \rightarrow 0$

Dual Algorithm for Variational Assimilation (aka *Physical Space Analysis System*, *PSAS,* pronounced *'pizzazz'*; see in particular book and papers by Bennett)

$$x^a = x^b + P^b H^T [HP^bH^T + R]^{-1} (y - Hx^b)$$

$$x^a = x^b + P^b H^T \Lambda^{-1} d = x^b + P^b H^T m$$

where $\Lambda \equiv HP^bH^T + R$, $d \equiv y - Hx^b$ and $m \equiv \Lambda^{-1} d$ maximises

$$\mu \rightarrow \mathcal{K}(\mu) = - (1/2)\ \mu^T \Lambda\ \mu + d^T \mu$$

Maximisation is performed in (dual of) observation space.

Dual Algorithm for Variational Assimilation (continuation 2)

Extends to time dimension, and to weak-constraint case, by defining state vector as

$$x \equiv (x_0^{\mathrm{T}}, x_1^{\mathrm{T}}, \ldots, x_K^{\mathrm{T}})^{\mathrm{T}}$$

or, equivalently, but more conveniently, as

$$x \equiv (x_0^{\mathrm{T}}, \eta_0^{\mathrm{T}}, \ldots, \eta_{K-1}^{\mathrm{T}})^{\mathrm{T}}$$

where, as before

$$\eta_k = x_{k+1} - M_k x_k \quad , \quad k = 0, \ldots, K\text{-}1$$

The background for x_0 is x_0^b, the background for η_k is 0. Complete background is

$$x^b = (x_0^{b\mathrm{T}}, 0^{\mathrm{T}}, \ldots, 0^{\mathrm{T}})^{\mathrm{T}}$$

It is associated with error covariance matrix

$$P^b = \mathrm{diag}(P_0^b, Q_0, \ldots, Q_{K-1})$$

Dual Algorithm for Variational Assimilation (continuation 3)

Define global observation vector as

$$y \equiv (y_0^{\mathrm{T}}, y_1^{\mathrm{T}}, \ldots, y_K^{\mathrm{T}})^{\mathrm{T}}$$

and global innovation vector as

$$d \equiv (d_0^{\mathrm{T}}, d_1^{\mathrm{T}}, \ldots, d_K^{\mathrm{T}})^{\mathrm{T}}$$

where $d_k \equiv y_k - H_k\, x_k^b$, with $x_{k+1}^b \equiv M_k x_k^b$, $\quad k = 0, \ldots, K\text{-}1$

Dual Algorithm for Variational Assimilation (continuation 4)

For any state vector $\xi = (\xi_0^{\mathrm{T}}, \upsilon_0^{\mathrm{T}}, \ldots, \upsilon_{K-1}^{\mathrm{T}})^{\mathrm{T}}$, the observation operator H

$$\xi \rightarrow H\xi = (u_0^{\mathrm{T}}, \ldots, u_K^{\mathrm{T}})^{\mathrm{T}}$$

is defined by the sequence of operations

$$u_0 = H_0 \xi_0$$

then for $k = 0, \ldots, K\text{-}1$

$$\begin{aligned} \xi_{k+1} &= M_k \xi_k + \upsilon_k \\ u_{k+1} &= H_{k+1} \xi_{k+1} \end{aligned}$$

The observation error covariance matrix is equal to

$$R = \mathrm{diag}(R_0, \ldots, R_K)$$

Dual Algorithm for Variational Assimilation (continuation 5)

Maximization of dual objective function

$$\mu \rightarrow \mathcal{K}(\mu) = -(1/2)\ \mu^{\mathrm{T}} \Lambda\, \mu + d^{\mathrm{T}} \mu$$

requires explicit repeated computations of its gradient

$$\nabla_{\mu} \mathcal{K} = -\Lambda \mu + d = -(HP^b H^{\mathrm{T}} + R)\mu + d$$

Starting from $\mu = (\mu_0{}^{\mathrm{T}}, \ldots, \mu_K{}^{\mathrm{T}})^{\mathrm{T}}$ belonging to (dual) of observation space, this requires 5 successive steps

- Step 1. Multiplication by H^{T}. This is done by applying the transpose of the process defined above, *viz.*,

Set $\chi_K = 0$

Then, for $k = K$-1, …, 0

$$\nu_k = \chi_{k+1} + H_{k+1}{}^{\mathrm{T}} \mu_{k+1}$$
$$\chi_k = M_k{}^{\mathrm{T}} \nu_k$$

Finally $\lambda_0 = \chi_0 + H_0{}^{\mathrm{T}} \mu_0$

The output of this step, which includes a backward integration of the adjoint model, is the vector $(\lambda_0{}^{\mathrm{T}}, \nu_0{}^{\mathrm{T}}, \ldots, \nu_{K\text{-}1}{}^{\mathrm{T}})^{\mathrm{T}}$

Dual Algorithm for Variational Assimilation (continuation 6)

- Step 2. Multiplication by P^b. This reduces to

$$\xi_0 = P_0^b \lambda_0$$
$$\upsilon_k = Q_k \nu_k \ , \ \ k = 0, \ldots, K\text{-}1$$

- Step 3. Multiplication by H. Apply the process defined above on the vector $(\xi_0^{\mathrm{T}}, \upsilon_0^{\mathrm{T}}, \ldots, \upsilon_{K\text{-}1}^{\mathrm{T}})^{\mathrm{T}}$, thereby producing vector $(u_0^{\mathrm{T}}, \ldots, u_K^{\mathrm{T}})^{\mathrm{T}}$.

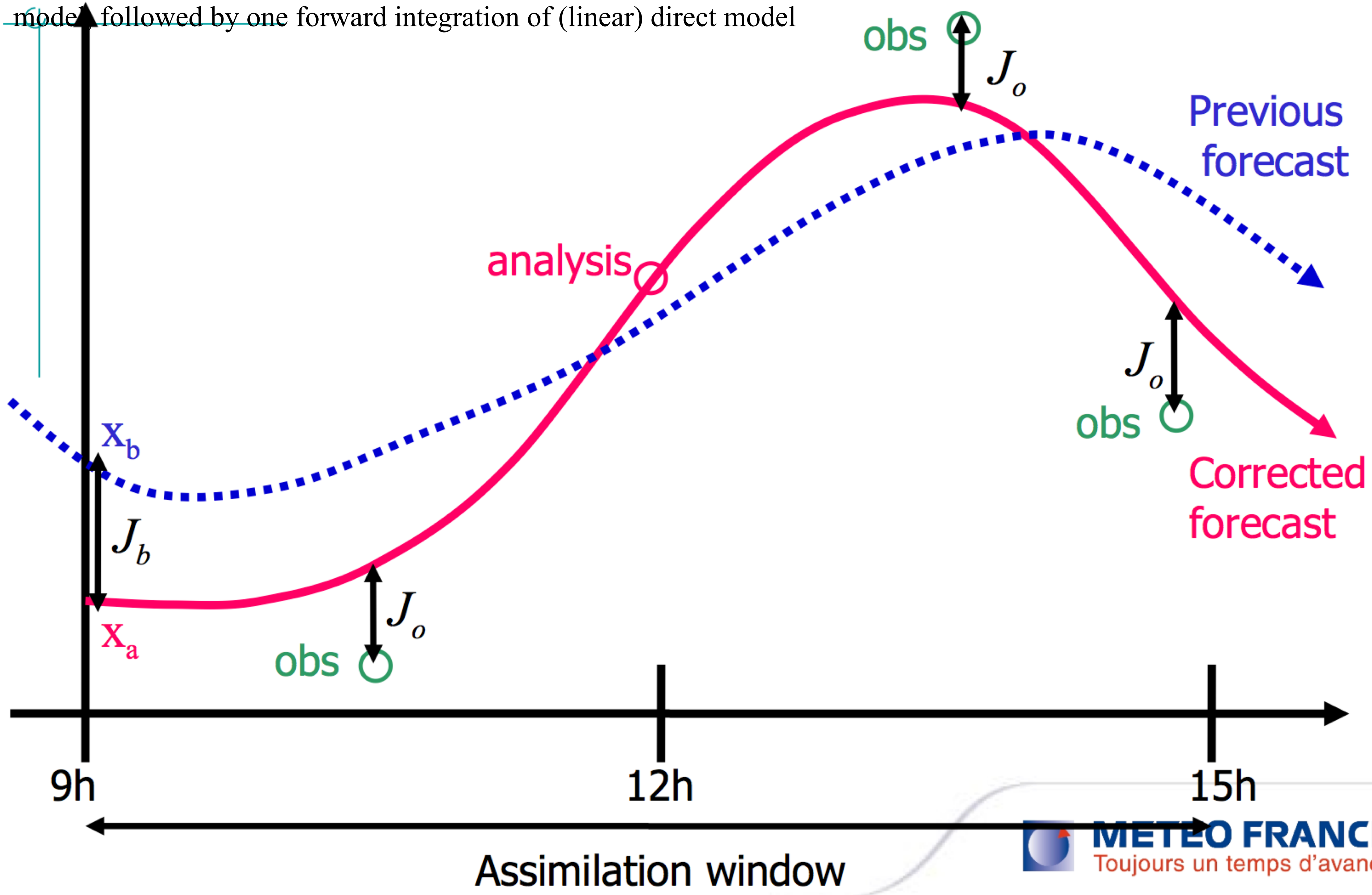
- Step 4. Add vector $R\mu$, *i. e.* compute

$$\varphi_k = u_k + R_k \mu_k \qquad , \ \ k = 0, \ldots, K$$

- Step 5. Change sign of vector $\varphi = (\varphi_0^{\mathrm{T}}, \ldots, \varphi_K^{\mathrm{T}})^{\mathrm{T}}$, and add vector $d = y - Hx^b$. It is through the addition of d that the observation y enters the algorithm.

Principle of 4D-VAR assimilation

One iteration of dual algorithm : one backward integration of adjoint model, followed by one forward integration of (linear) direct model



Dual Algorithm for Variational Assimilation (continuation 7)

$$x^a = x^b + P^b H^T [HP^b H^T + R]^{-1} (y - Hx^b)$$

$$x^a = x^b + P^b H^T \Lambda^{-1} d = x^b + P^b H^T m$$

Finally, multiplication by $P^b H^T$ and addition to background

Dual Algorithm for Variational Assimilation (continuation 8)

Temporal correlations can be introduced, by taking non-block diagonal observation and model errors covariance matrices.

So can also correlations between model and observation errors, by introducing additional terms in matrix Λ.

Dual algorithm remains regular in the limit of vanishing model error. Can be used for both strong- and weak-constraint assimilation.

No significant increase of computing cost in comparison with standard strong constraint variational assimilation. The cost depends mainly on the number of model integrations (Courtier, Louvel)

S. Louvel (Doctoral Dissertation, Université Paul-Sabatier, Toulouse, 1999, *J. Geophys. Res.,* 2001)

Assimilation of altimetric observations performed by satellites Topex/Poseidon and ERS-1.

Assimilation performed with primitive equation ocean *Miami Isopycnic Coordinate Ocean Model* (*MICOM*)

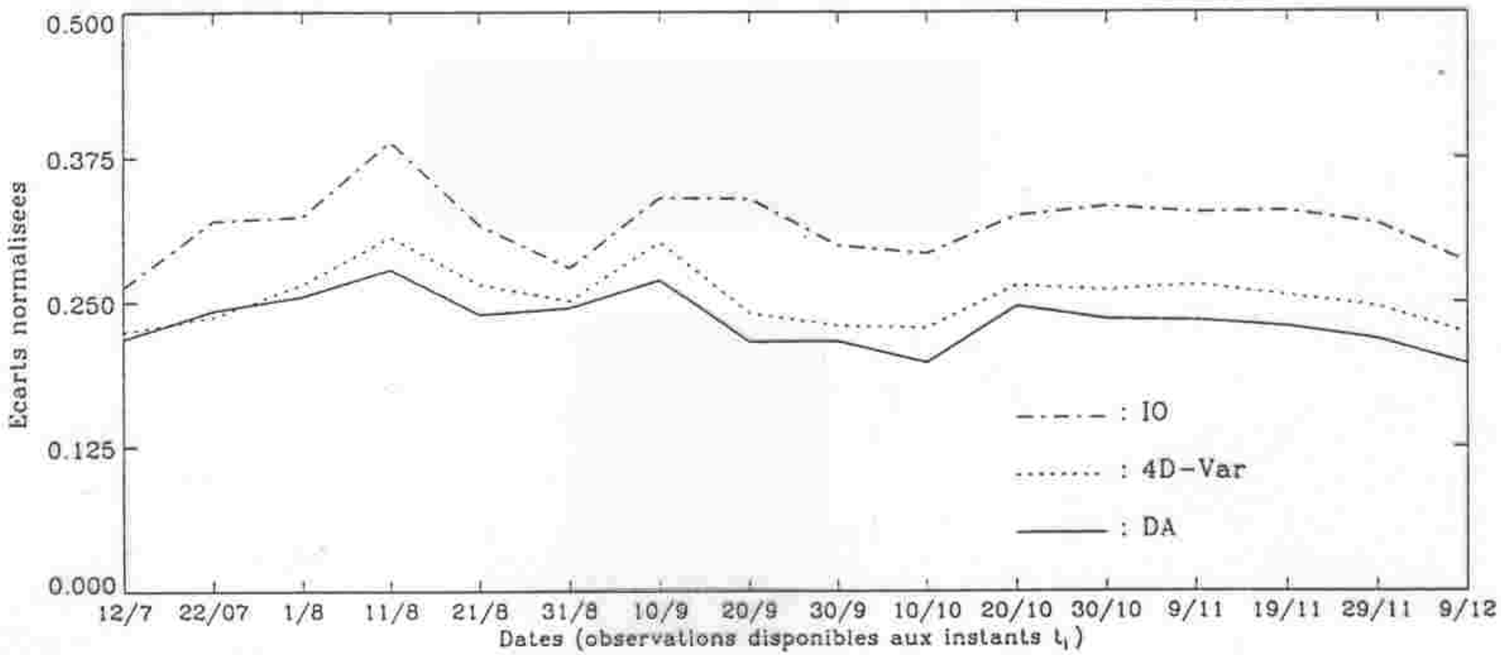


FIG. 9.11 – *Ecarts normalisés prévision/observations sur l'ensemble de la période étudiée*

Louvel, Doctoral Dissertation, Université Paul-Sabatier, Toulouse, 1999 nous permet d'observer les performances des différentes techniques d'assimilation.

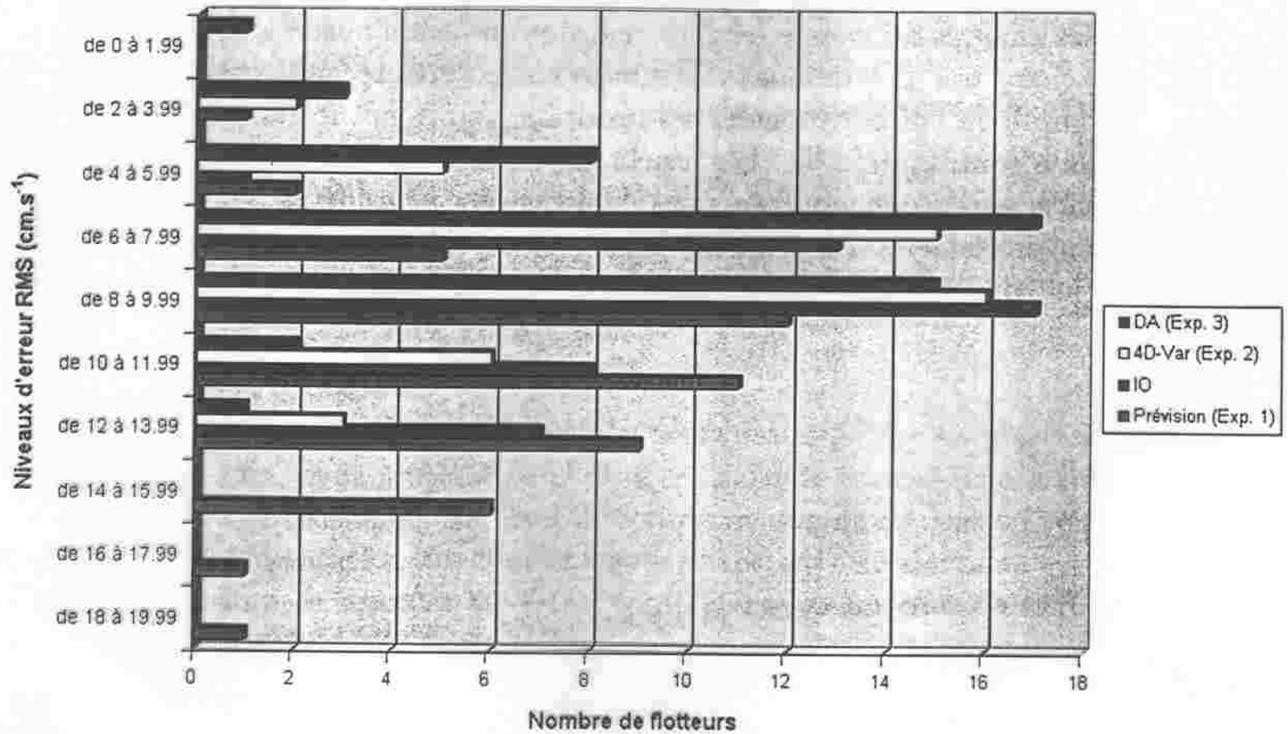


FIG. 9.15 – *Description des écarts flotteurs/modèle en terme de vitesse (à 150 m de profondeur) pour les différents algorithmes d'assimilation*

Louvel, Doctoral Dissertation, Université Paul-Sabatier, Toulouse, 1999

Dual Algorithm for Variational Assimilation (continuation)

Requires

- Explicit background (not much of a problem)
- Exact linearity (much more of a problem). Definition of iterative nonlinear procedures is being studied (Auroux, …)

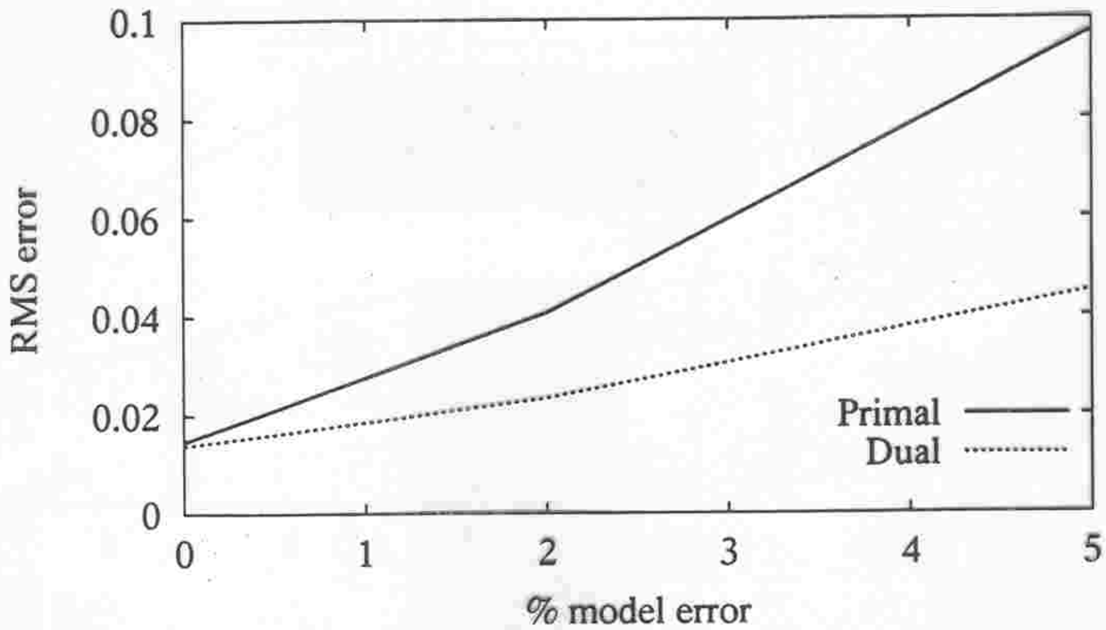


FIG. 6.13 – Normes RMS des erreurs d'assimilation obtenues pour les deux méthodes en fonction de l'erreur introduite dans le modèle au cours de la période d'assimilation.

Auroux, Doctoral Dissertation, Université de Nice-Sophia Antipolis, Nice, 2003

Dual : weak constraint
Primal : strong constraint

Dual Algorithm for Variational Assimilation is now used, in the weak-constraint form, at *Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique* (*CERFACS*) in Toulouse (A. Weaver, S. Gürol) for assimilation of oceanographical observations.

Weak-constraint assimilation used (in primal form) at ECMWF for assimilation of stratospheric observations.

From course 5

Data of the form

$$\boldsymbol{z} = \boldsymbol{\Gamma x} + \boldsymbol{\zeta},$$

Known data vector $\boldsymbol{z}$ belongs to *data space* D, $dimD = m$,
Unknown state vector $\boldsymbol{x}$ belongs to *state space* X, $dimX = n$
$\boldsymbol{\Gamma}$ known (mxn)-matrix, $\boldsymbol{\zeta}$ unknown 'error'

Determinacy condition : rank$\boldsymbol{\Gamma} = n$. Data contain information, directly or indirectly, on every component of state vector $\boldsymbol{x}$. Requires $m \geq n$.
Set $m = n + p$

p is the *degree of overdeterminacy* of the system. In the background + observations format, p is the dimension of the innovation

Minimum of objective function

$$\mathcal{J}(\boldsymbol{\xi}) \equiv (1/2)\,[\Gamma\boldsymbol{\xi} - \boldsymbol{z}]^{\mathrm{T}}\, S^{-1}\, [\Gamma\boldsymbol{\xi} - \boldsymbol{z}]$$

$$\mathcal{J}_{min} \equiv \mathcal{J}(\boldsymbol{x}^{a}) = (1/2)\,[\Gamma\boldsymbol{x}^{a} - \boldsymbol{z}]^{\mathrm{T}}\, S^{-1}\, [\Gamma\boldsymbol{x}^{a} - \boldsymbol{z}]$$

$$= (1/2)\,\boldsymbol{d}^{\mathrm{T}}\,[E(\boldsymbol{d}\boldsymbol{d}^{\mathrm{T}})]^{-1}\,\boldsymbol{d}$$

where $\boldsymbol{d}$ is innovation ($\boldsymbol{d} = \boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}^{b} = \boldsymbol{\varepsilon} - \boldsymbol{H}\boldsymbol{\zeta}^{b}$, or more generally what is obtained by eliminating the unknown $\boldsymbol{x}$ from the data $\boldsymbol{z}$). Innovation is only objective case-to-case measure of the errors affecting the data

$$\mathcal{J}_{min} = (1/2)\,\boldsymbol{d}^{\mathrm{T}}\,[E(\boldsymbol{d}\boldsymbol{d}^{\mathrm{T}})]^{-1}\,\boldsymbol{d} = (1/2)\,\mathrm{Tr}\left\{\boldsymbol{d}^{\mathrm{T}}\,[E(\boldsymbol{d}\boldsymbol{d}^{\mathrm{T}})]^{-1}\,\boldsymbol{d}\right\}$$

$$= (1/2)\,\mathrm{Tr}\left\{[E(\boldsymbol{d}\boldsymbol{d}^{\mathrm{T}})]^{-1}\,\boldsymbol{d}\boldsymbol{d}^{\mathrm{T}}\right\}$$

Minimum of objective function *(continuation 1)*

$$\mathcal{J}_{min} = (1/2)\ \mathrm{Tr}\left\{[E(\boldsymbol{d}\boldsymbol{d}^{\mathrm{T}})]^{-1}\ \boldsymbol{d}\boldsymbol{d}^{\mathrm{T}}\right\}$$

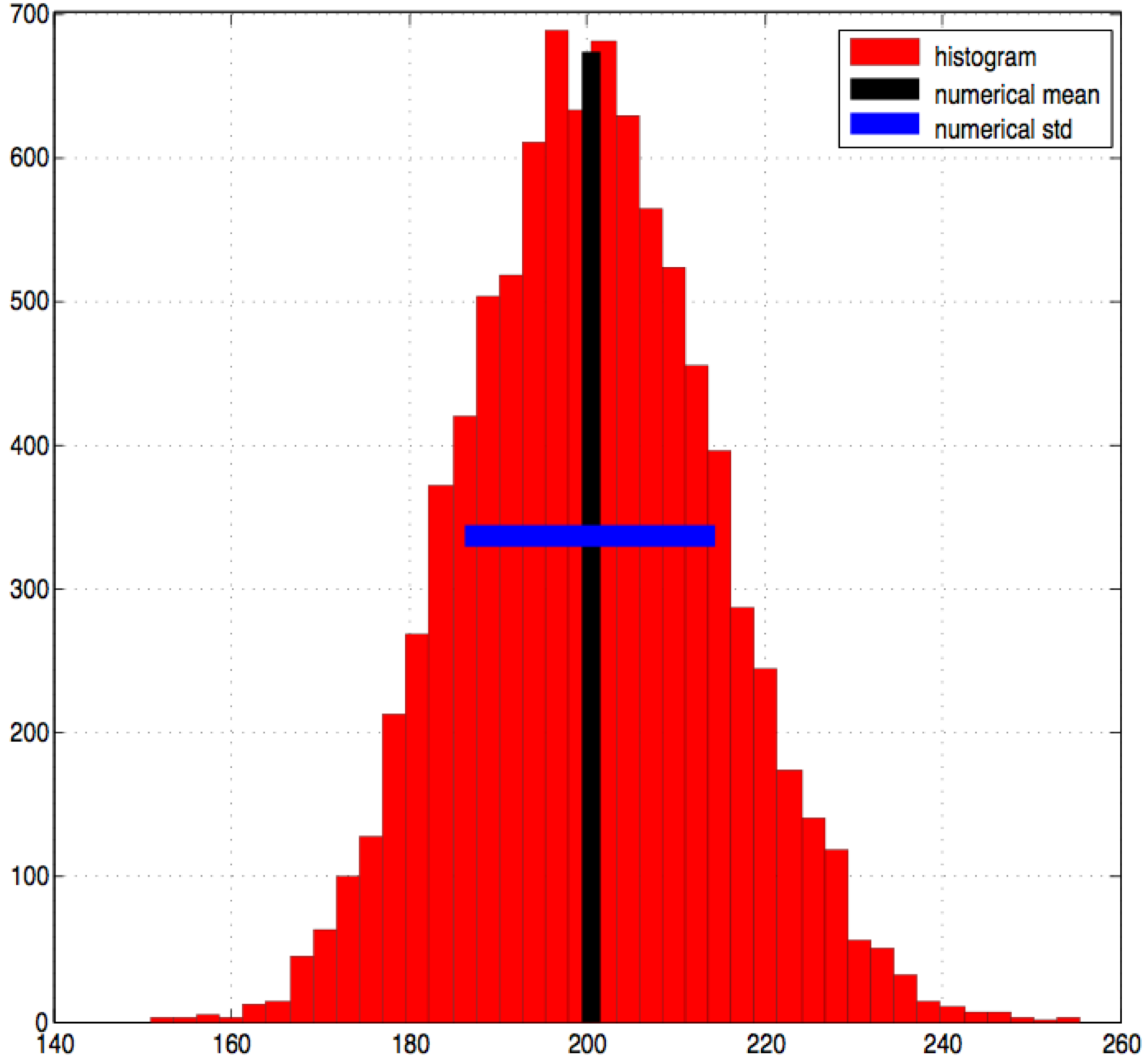
$$E(\mathcal{J}_{min}) = (1/2)\ \mathrm{Tr}\{[E(\boldsymbol{d}\boldsymbol{d}^{\mathrm{T}})]^{-1}\ E(\boldsymbol{d}\boldsymbol{d}^{\mathrm{T}})\} = (1/2)\ \mathrm{Tr}\ \boldsymbol{I}_p = p/2$$

$$E(\mathcal{J}_{min}) = p/2 \qquad\qquad (p = \dim\boldsymbol{y} = \dim\boldsymbol{d})$$

If p is large, a few realizations are sufficient for determining $E(\mathcal{J}_{min})$

Remark 1. If in addition errors are gaussian, the quantity $2E(\mathcal{J}_{min})$ follows a χ^2-probability distribution of order p. For that reason the criterion $E(\mathcal{J}_{min}) = p/2$ is often called the χ^2 criterion. Also $Var(\mathcal{J}_{min}) = p/2$ in the gaussian case.

Remark 2. The result is the same for the maximum of the objective function of the dual approach.



Linearized Lorenz'96. 5 days. Histogram of $\mathcal{J}_{min}$

$E(\mathcal{J}_{min}) = p/2$ (=200) ; $\sigma(\mathcal{J}_{min}) = \sqrt{(p/2)}$ (≈14.14)

Observed values 199.39 and 14.27

Credit M. Jardak

Minimum of objective function *(continuation 2)*

$$\mathcal{J}_{min} = (1/2)\ \mathrm{Tr}\left\{\underbrace{[E(\boldsymbol{dd}^{\mathrm{T}})]^{-1}}_{assumed}\ \underbrace{\boldsymbol{dd}^{\mathrm{T}}}_{real}\right\}$$

$$E_{\mathrm{R}}(\mathcal{J}_{min}) = (1/2)\ \mathrm{Tr}\{[E(\boldsymbol{dd}^{\mathrm{T}})]^{-1}\ E_{\mathrm{R}}(\boldsymbol{dd}^{\mathrm{T}})\} = p/2$$

If $E(\mathcal{J}_{min}) > p/2$ ($< p/2$), $E(\boldsymbol{dd}^{\mathrm{T}})$, as resulting from the *a priori* specification of data errors, is too small (too large).

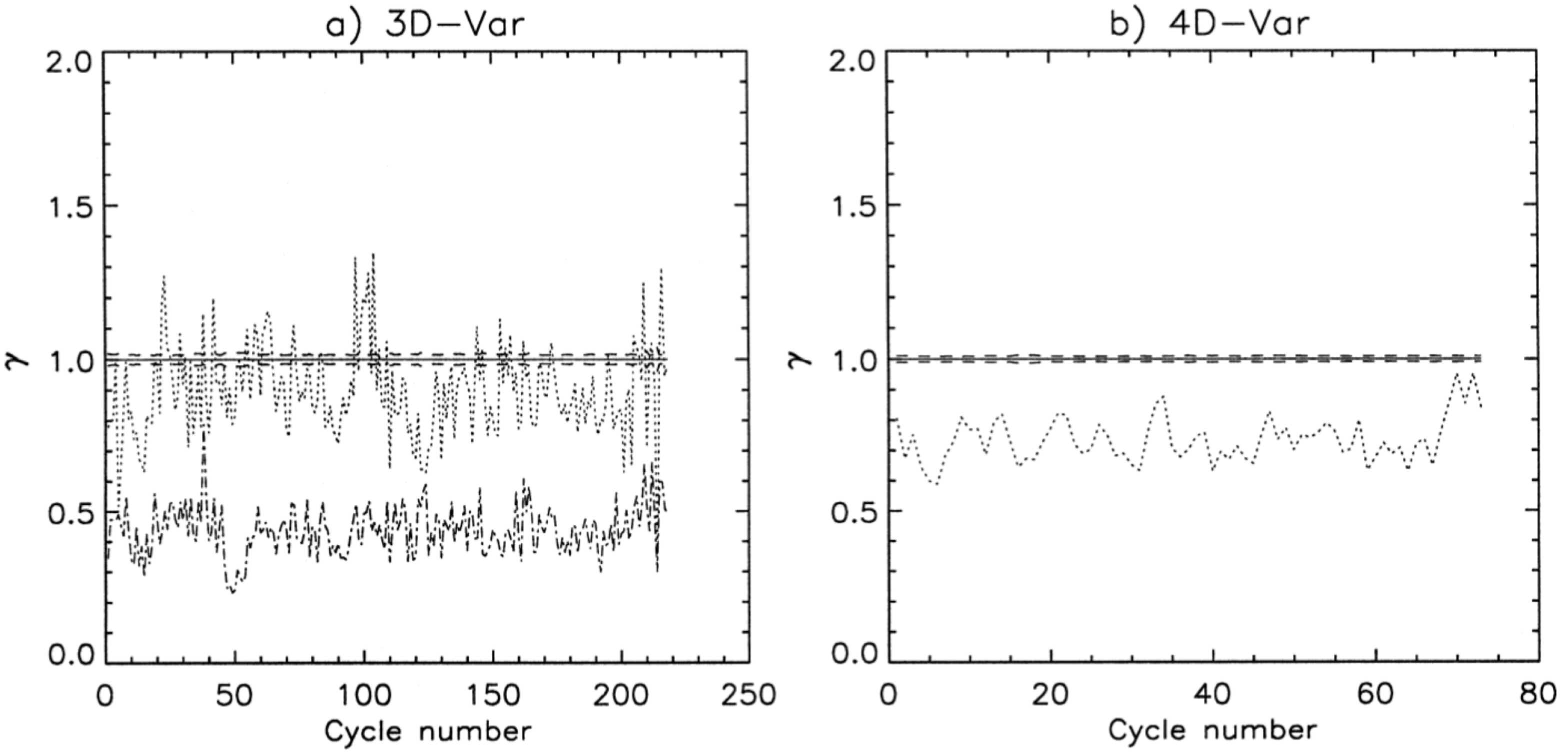


FIG. 7. The value of $\gamma = 2J_{\min}/p$ plotted as a function of the assimilation cycle during the period 1993–98 in (a) 3DVAR and (b) 4DVAR. On each cycle, $J_{\min}$ represents the value of the cost function at the end of the minimization and p the total number of observations assimilated through the J_o term. A total of 219 (73) 10-day (30 day) cycles were performed in 3DVAR (4DVAR). The *expected* value of $\gamma = 1$ (solid line) is plotted together with error bars (dashed lines) at $1 \pm \sigma_\gamma$, where $\sigma_\gamma = \sqrt{2/p}$ is the *expected* standard deviation of γ. The dotted curves in (a) and (b) correspond to the *actual* values of γ computed on each cycle in the reference experiments EX3D and EX4D. The dashed–dotted curve in (a) corresponds to the *actual* values of γ from a 3DVAR experiment in which the observation-error variance for the TAO data is set to twice the value used in EX3D.

How to write the adjoint of a code ?

Operation $a = b \times c$

Input b, c Output a but also b, c

For clarity, we write

$$a = b \times c$$
$$b' = b$$
$$c' = c$$

$\partial J/\partial a$, $\partial J/\partial b'$, $\partial J/\partial c'$ available. We want to determine $\partial J/\partial b$, $\partial J/\partial c$

Chain rule

$$\partial J/\partial b = (\partial J/\partial a)\underbrace{(\partial a/\partial b)}_{c} + (\partial J/\partial b')\underbrace{(\partial b'/\partial b)}_{1} + (\partial J/\partial c')\underbrace{(\partial c'/\partial b)}_{0}$$

$$\partial J/\partial b = (\partial J/\partial a)\, c + \partial J/\partial b'$$

Similarly

$$\partial J/\partial c = (\partial J/\partial a)\, b + \partial J/\partial c'$$

How to write the adjoint of a code ? *(continuation 1)*

Operation $a = b \text{ x } c$

Differentiate $\delta a = b \text{ x } \delta c + c \text{ x } \delta b$

$\partial J/\partial b = (\partial J/\partial a)\ c + \partial J/\partial b'$

$\boldsymbol{bad} = \boldsymbol{bad} + \boldsymbol{c} \text{ x } \boldsymbol{aad}$

$\boldsymbol{cad} = \boldsymbol{cad} + \boldsymbol{b} \text{ x } \boldsymbol{aad}$

$\boldsymbol{aad} = 0$

Start adjoint computations by setting all adjoint variables to 0 (except whatever is necessary to start the whole computation, *e.g.*, $\boldsymbol{Jad} = 1$)

How to write the adjoint of a code ? *(continuation 2)*

There are shortcuts

General expression for transpose (adjoint) operator

$$< \boldsymbol{x} , \boldsymbol{Ay} > = < \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x} , \boldsymbol{y} >$$

If direct computation is multiplication by matrix $\boldsymbol{A}$, corresponding adjoint computation is multiplication by transpose matrix $\boldsymbol{A}^{\mathrm{T}}$. In particular, if $\boldsymbol{A}$ is symmetric (skew symmetric), adjoint computation is identical with (is minus) the direct computation.

Example 1. Poisson solver. $\Delta\varphi = f$, where Δ is Laplacian

$\Delta\varphi = f \Rightarrow \varphi = \Delta^{-1} f$.

Laplacian Δ, and inverse Δ^{-1}, are symmetric

$\Rightarrow$ adjoint of Poisson solver is Poisson solver

How to write the adjoint of a code ? *(continuation 3)*

Example 2. Fourier transform $\boldsymbol{F}$

$< \boldsymbol{Fx} , \boldsymbol{Fy} > \; = \; < \boldsymbol{x} , \boldsymbol{y} >$ (Parseval)

$< \boldsymbol{x} , \boldsymbol{F}^{\mathrm{T}} \boldsymbol{Fy} > \; = \; < \boldsymbol{x} , \boldsymbol{y} >. \quad \Longrightarrow \quad \boldsymbol{F}^{\mathrm{T}} = \boldsymbol{F}^{-1}$

Adjoint of Fourier transform is inverse Fourier transform

How to write the adjoint of a code ? *(continuation 4)*

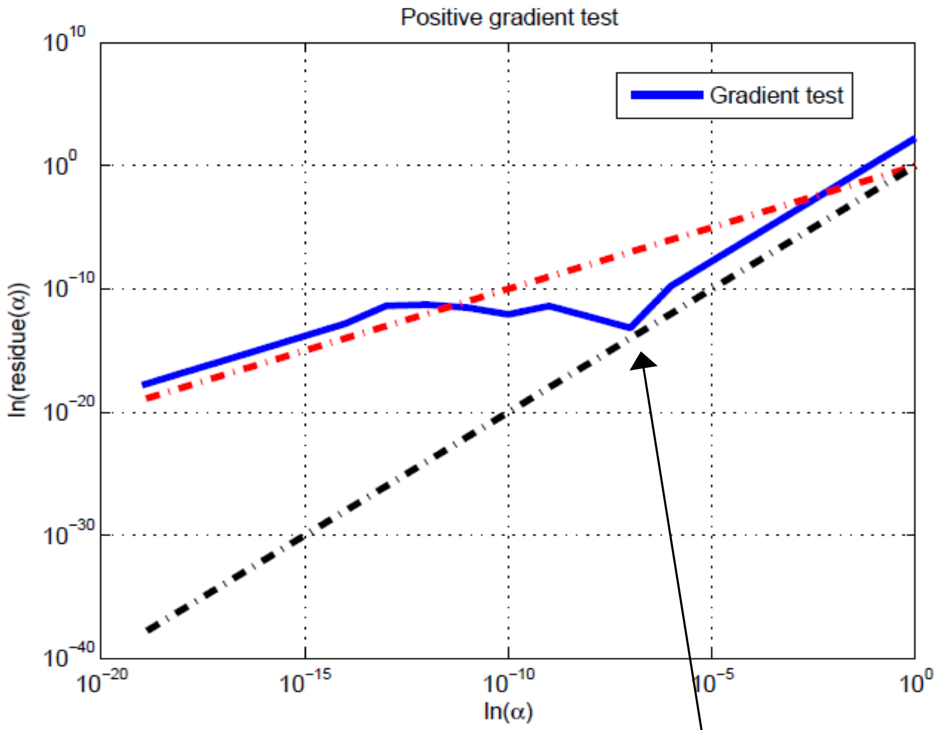
Adjoint compilers

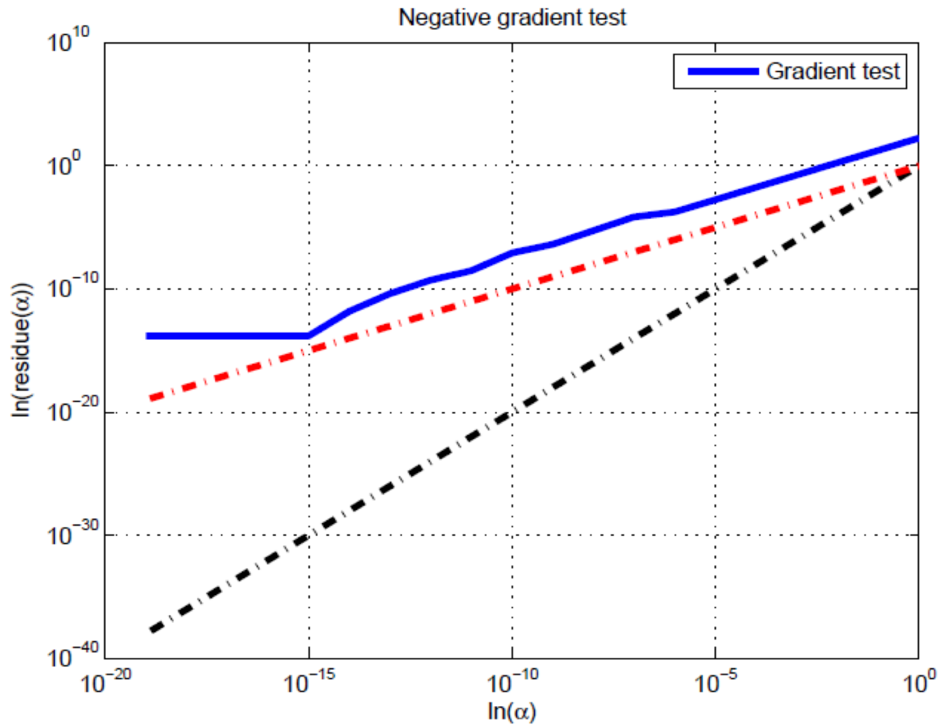
TAPENADE (Laurent Hascoët, Institut national de recherche en informatique et en automatique)

FastOpt AD-Tool (Ralf Giering and Thomas Kaminski)

- ….

Gradient test





$\epsilon \cdot \mathfrak{J}(\text{optimal control variable})$

$\epsilon = 2^{-53}$ zero machine

$residue(\alpha) = (\mathfrak{J}(x + \alpha dx) - \mathfrak{J}(x)) - \alpha \nabla \mathfrak{J}(x) dx$

Conclusion on Sequential Assimilation

Pros

‘Natural’, and well adapted to many practical situations (transition to forecast is immediate)

Provides, at least relatively easily, explicit estimate of estimation error

Cons

Carries information only forward in time (of no importance if one is interested only in doing forecast)

In a strictly sequential assimilation (*i.e.,* any individual piece of information is discarded once it has been used), optimality is possible only if errors are uncorrelated in time.

Ensemble Kalman Filter requires empirical inflation and localisation (but …)

Conclusion on Variational Assimilation

Pros

Carries information both forward and backward in time (important for reassimilation of past data).

Can easily take into account temporal statistical dependence (Järvinen *et al*.)

Does not require explicit computation of temporal evolution of estimation error

Very well adapted to some specific problems (*e. g*., identification of tracer sources)

Cons

Transition to forecast not immediate (necessary to come back in time)

Does not readily provide estimate of estimation error

Requires development and maintenance of adjoint codes. But the latter can have other uses (sensitivity studies).

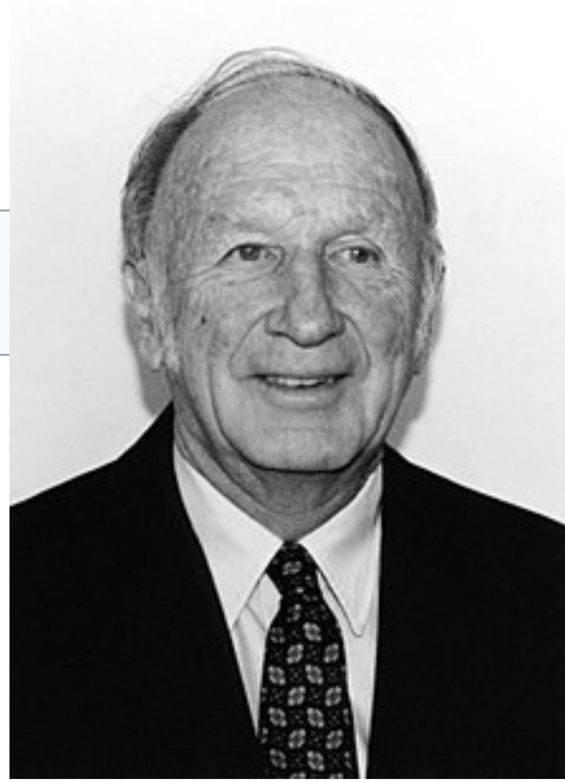
- Dual approach seems most promising. But little used.

- Can be implemented in ensemble form (see course 7).

Buehner *et al*. (*Mon. Wea. Rev*., 2010)

For the same numerical cost, and in meteorologically realistic situations, Ensemble Kalman Filter and Variational Assimilation produce results of similar quality.

Edward N. Lorenz (1917 – 2008)



Studied mathematics. Interest in theory of dynamical systems.

1963. Observation of sensitivity to initial conditions on small dimension deterministic system (*deterministic chaos*)

Notion of *available potential energy*

Introduced a number of small dimension chaotic systems, with properties somewhat similar to properties of atmospheric flow

Cours à venir

~~Mercredi 2 avril~~

~~Vendredi 11 avril~~

~~Vendredi 18 avril~~

~~Mercredi 23 avril~~

~~Lundi 12 mai~~

~~Mercredi 28 mai~~

Mercredi 11 juin

Mercredi 18 juin