École Doctorale des Sciences de l'Environnement d'Île-de-France

Année Universitaire 2019-2020

Modélisation Numérique de l'Écoulement Atmosphérique et Assimilation de Données

Olivier Talagrand

Cours 6

28 Avril 2020

From course 5

Case of data that are distributed over time

Suppose for instance available data consist of

- Background estimate at time 0

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

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

$y_k = H_k x_k + \varepsilon_k \qquad\qquad E(\varepsilon_k \varepsilon_j^T) = R_k\, \delta_{kj}$

- Model (supposed for the time being to be exact)

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

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

Then objective function

$$\xi_0 \in \mathcal{S} \rightarrow$$

$$\mathcal{J}(\xi_0) \equiv (1/2)\,(x_0^b - \xi_0)^T\,[P_0^b]^{-1}\,(x_0^b - \xi_0) + (1/2)\,\Sigma_k[y_k - H_k\xi_k]^T\,R_k^{-1}\,[y_k - H_k\xi_k]$$

$$\equiv \qquad\qquad \mathcal{J}_b \qquad\qquad + \qquad\qquad \mathcal{J}_o$$

subject to $\xi_{k+1} = M_k \xi_k\,, \qquad\qquad k = 0, \ldots, K\text{-}1$

From course 5

$$\mathcal{J}(\xi_0) = (1/2)\,(x_0^{\,b} - \xi_0)^{\mathrm{T}}\,[P_0^{\,b}]^{-1}\,(x_0^{\,b} - \xi_0) + (1/2)\,\Sigma_k[y_k - H_k\xi_k]^{\mathrm{T}}\,R_k^{-1}\,[y_k - H_k\xi_k]$$

Background is not necessary, if observations are in sufficient number to overdetermine the problem. Nor is strict linearity.

Four-Dimensional Variational Assimilation

'*4D-Var*'

From course 5

How to minimize objective function with respect to initial state $u = \xi_0$ (u is called the *control variable* of the problem) ?

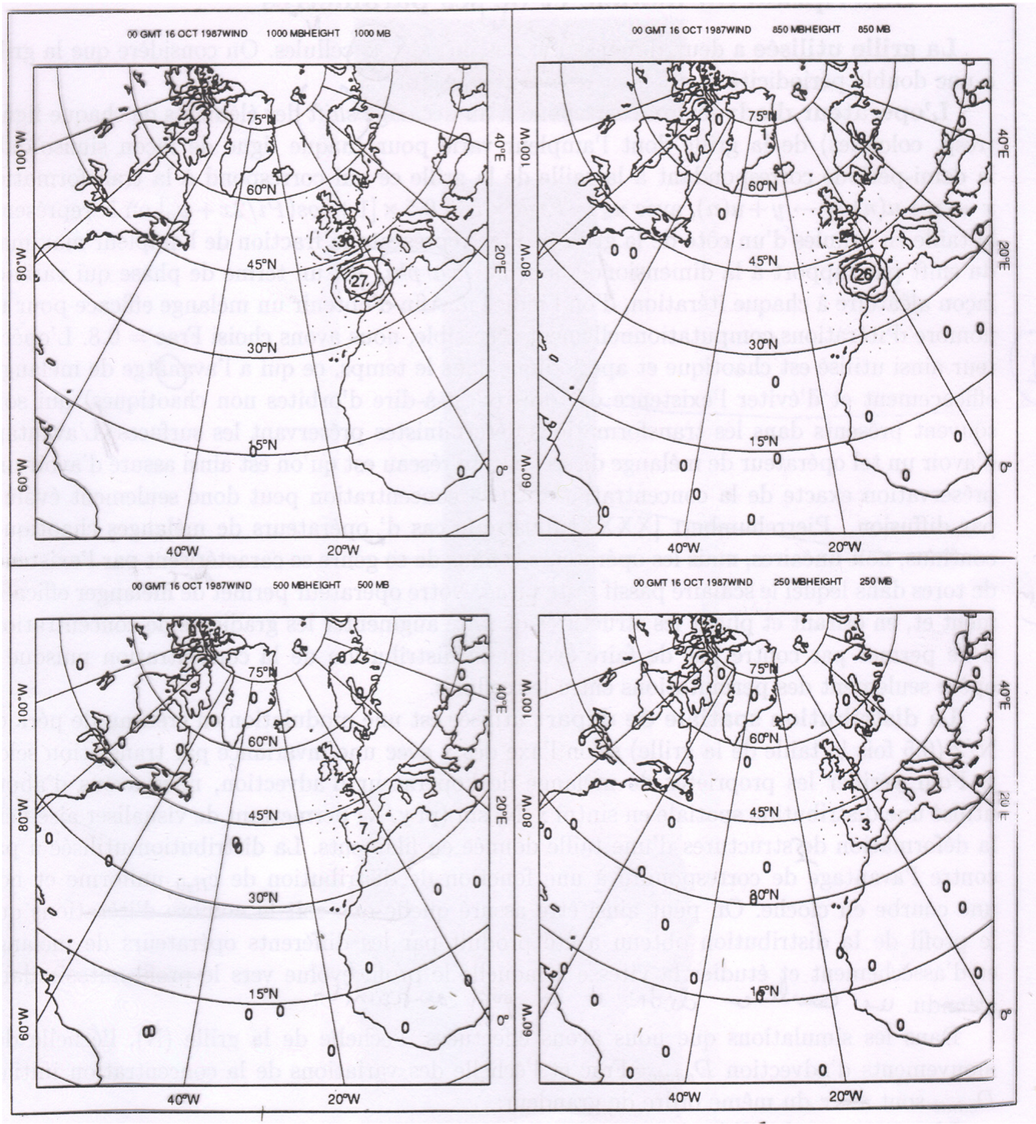
Use iterative minimization algorithm, each step of which requires the explicit knowledge of the local gradient $\nabla_u \mathcal{J} \equiv (\partial \mathcal{J}/\partial u_i)$ of $\mathcal{J}$ with respect to u.

From course 5

How to numerically compute the gradient $\nabla_u \mathcal{J}$?

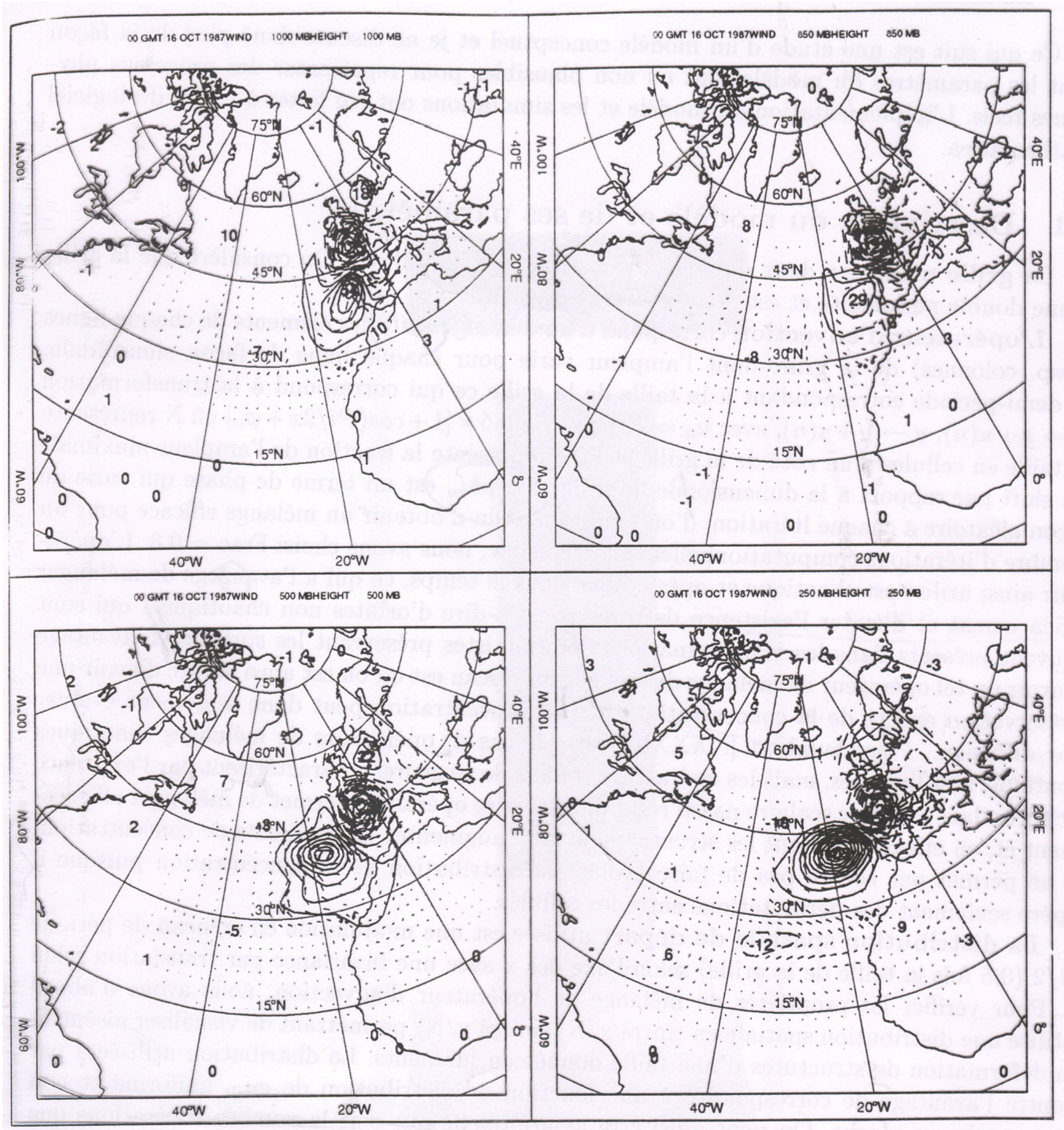
Direct perturbation, in order to obtain partial derivatives $\partial \mathcal{J}/\partial u_i$ by finite differences ? That would require as many explicit computations of the objective function $\mathcal{J}$ as there are components in u. Practically impossible.

Gradient computed by *adjoint method*.



Analysis increments in a 3D-Var corresponding to a *u*-component wind observation at the 1000-hPa pressure level (no temporal evolution of background error covariance matrix)

Thépaut *et al*., 1993, *Mon. Wea. Rev.*, **121**, 3393-3414



Same as before, but at the end of a 24-hr 4D-Var

Thépaut *et al.*, 1993, *Mon. Wea. Rev.*, **121**, 3393-3414

- Variational assimilation. Complements. A few practical tips for adjoint modeling.

- The 'Incremental approach' to variational assimilation

- Weak constraint variational assimilation. Principle. The dual algorithm for variational assimilation. Examples.

- Assimilation and (In)stabilities. *Quasi-Static Variational Assimilation* (*QSVA*).

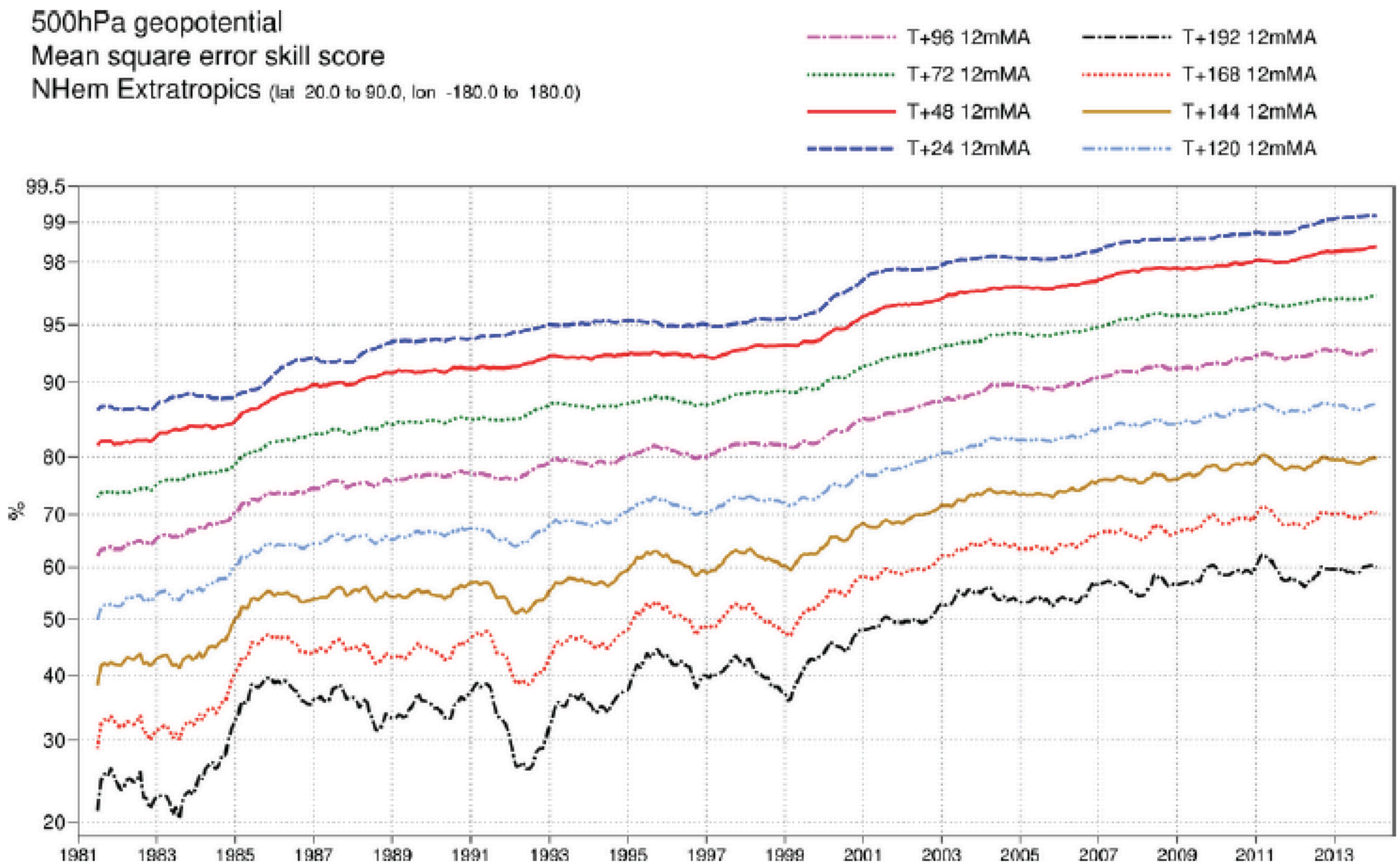
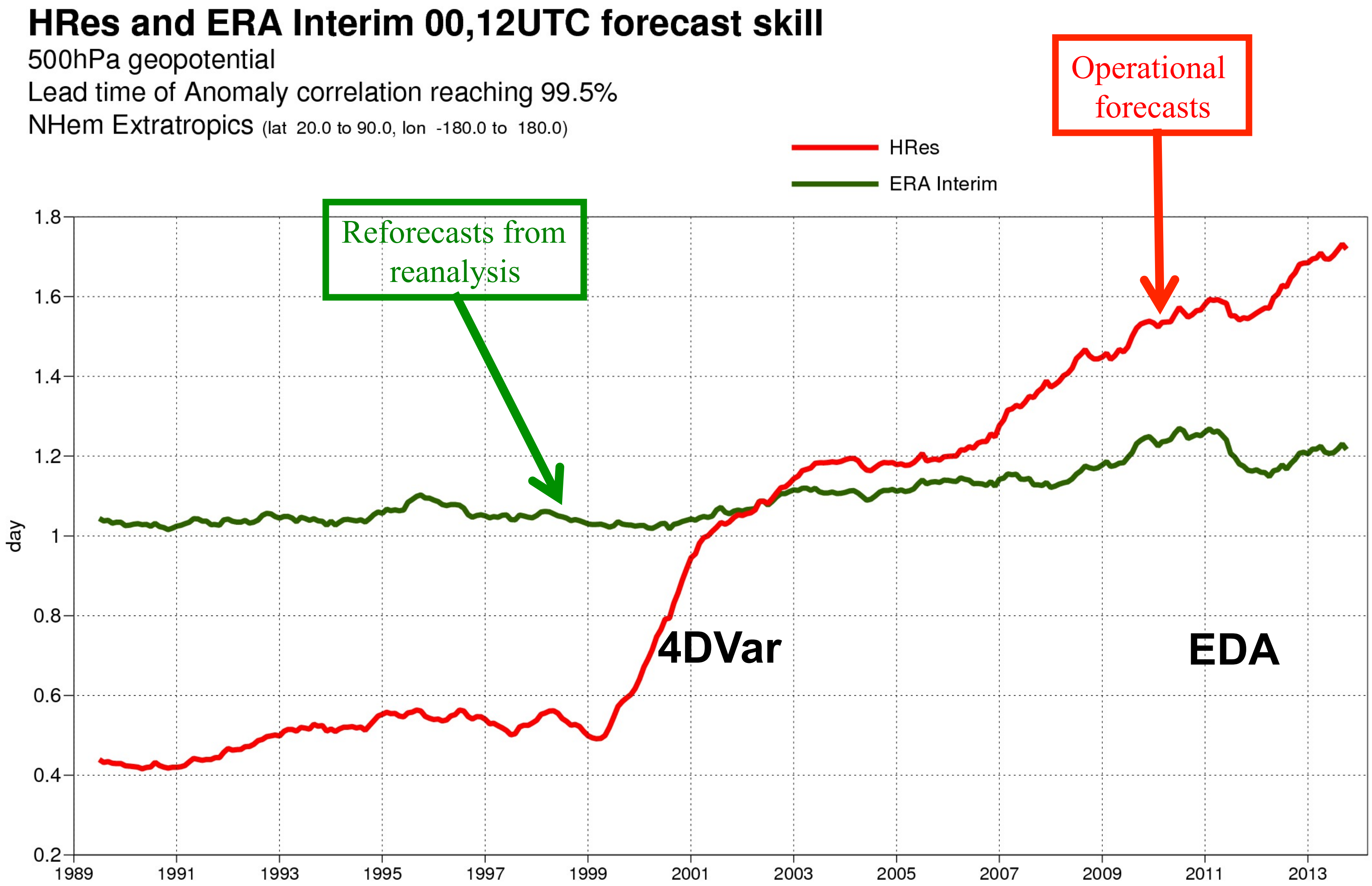


Figure 3: 500 hPa geopotential height mean square error skill score for Europe (top) and the northern hemisphere extratropics (bottom), showing 12-month moving averages for forecast ranges from 24 to 192 hours. The last point on each curve is for the 12-month period August 2013–July 2014.

Persistence = 0 ; climatology = 50 at long range

Initial state error reduction



Credit E. Källén, ECMWF

Time-correlated Errors (continuation 3)

Moral. If data errors are correlated in time, it is not possible to discard observations as they are used. In particular, if model error is correlated in time, all observations are liable to be reweighted as assimilation proceeds.

Variational assimilation can take time-correlated errors into account.

Example of time-correlated observation errors. Global covariance matrix

$$\mathcal{R} = (R_{kk'} = E(\varepsilon_k \varepsilon_{k'}^{\mathrm{T}}))$$

Objective function

$$\xi_0 \in \mathcal{S} \rightarrow$$

$$\mathcal{J}(\xi_0) = (1/2)\,(x_0^b - \xi_0)^{\mathrm{T}}\,[P_0^b]^{-1}\,(x_0^b - \xi_0) + (1/2)\,\Sigma_{kk'}[y_k - H_k\xi_k]^{\mathrm{T}}\,[\mathcal{R}^{-1}]_{kk'}\,[y_{k'} - H_{k'}\xi_{k'}]$$

where $[\mathcal{R}^{-1}]_{kk'}$ is the kk'-sub-block of global inverse matrix $\mathcal{R}^{-1}$.

Similar approach for time-correlated model error.

Time-correlated Errors (continuation 4)

Temporal correlation of observational error has been introduced by ECMWF (Järvinen *et al.*, 1999) in variational assimilation of high-frequency surface pressure observations (correlation originates in that case in representativeness error).

Identification and quantification of time correlation of errors, especially model errors ?

In the linear case, Kalman Smoother and Variational Assimilation are algorithmically equivalent. If errors are uncorrelated in time, they produce the *BLUE* of the state of the system from all available data, over the whole assimilation window (Kalman Filter produces the *BLUE* only at the end of the final time of the window). If in addition errors are globally Gaussian, both algorithms achieve Bayesian estimation.

If errors are correlated in time, only some Kalman Smoothers are equivalent with Variational Assimilation.

From course 5

Adjoint Method (continued 3)

Nonlinearities ?

$$\mathcal{J}(\xi_0) = (1/2)\,(x_0^b - \xi_0)^T [P_0^b]^{-1} (x_0^b - \xi_0) + (1/2)\,\Sigma_k [y_k - H_k(\xi_k)]^T R_k^{-1} [y_k - H_k(\xi_k)]$$

subject to $\xi_{k+1} = M_k(\xi_k)$, $\quad k = 0, \dots, K\text{-}1$

Control variable $\quad \xi_0 = \boldsymbol{u}$

Adjoint equation

$$\lambda_K = \quad H_K{}'^T R_K^{-1} [H_K(\xi_K) - y_K]$$

....

$$\lambda_k = M_k{}'^T \lambda_{k+1} + H_k{}'^T R_k^{-1} [H_k(\xi_k) - y_k] \qquad k = K\text{-}1, \dots, 1$$

....

$$\lambda_0 = M_0{}'^T \lambda_1 + H_0{}'^T R_0^{-1} [H_0(\xi_0) - y_0] + [P_0^b]^{-1} (\xi_0 - x_0^b)$$

$$\nabla_u \mathcal{J} = \lambda_0$$

Not approximate (it gives the exact gradient $\nabla_u \mathcal{J}$), and really used as described here.

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

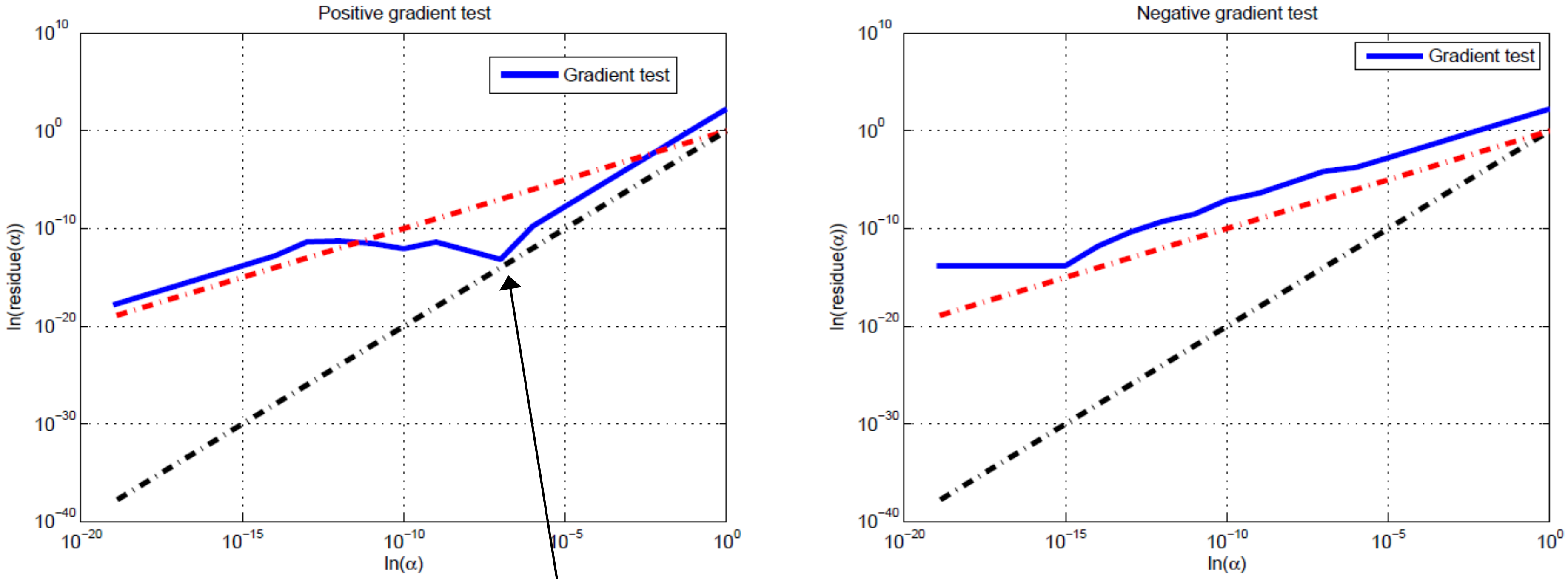
$$\partial J/\partial b = (\partial J/\partial a)\underbrace{(\partial a/\partial b)}_{c} + (\partial J/\partial b')\underbrace{(\partial b'/\partial b)}_{1} + (\partial J/\partial c')\underbrace{(\partial c'/\partial b)}_{0}$$

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

Similarly

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

Gradient test



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

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

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

Incremental Method for Variational Assimilation

Variational assimilation, as it has been described, requires the use of the adjoint of the full model.

Simplifying the adjoint as such can be very dangerous. The computed gradient would not be exact, and experience shows that optimization algorithms (and especially efficient ones) are very sensitive to even slight misspecification of the gradient.

Principle of *Incremental Method* (Courtier *et al.*, 1994, *Q. J. R. Meteorol. Soc.*) : simplify simultaneously the (local tangent linear) dynamics and the corresponding adjoint.

Incremental Method (continuation 1)

- Basic (nonlinear) model

$\xi_{k+1} = M_k(\xi_k)$

- Tangent linear model

$\delta\xi_{k+1} = M_k' \, \delta\xi_k$

where M_k' is jacobian of M_k at point ξ_k.

- Adjoint model

$\lambda_k = M_k'^{\mathrm{T}} \lambda_{k+1} + \ldots$

Incremental Method. Simplify both M_k' and $M_k'^{\mathrm{T}}$ consistently.

Incremental Method (continuation 2)

More precisely, for given solution $\xi_k^{(0)}$ of nonlinear model, replace tangent linear and adjoint models respectively by

$$\delta\xi_{k+1} = L_k\, \delta\xi_k \qquad (2)$$

and

$$\lambda_k = L_k^{\mathrm{T}}\, \lambda_{k+1} + \ldots$$

where L_k is an appropriate simplification of jacobian M_k'.

It is then necessary, in order to ensure that the result of the adjoint integration is the exact gradient of the objective function, to modify the basic model in such a way that the solution emanating from $\xi_0^{(0)} + \delta\xi_0$ is equal to $\xi_k^{(0)} + \delta\xi_k$, where $\delta\xi_k$ evolves according to (2). This makes the basic dynamics exactly linear.

Incremental Method (continuation 3)

As concerns the observation operators in the objective function, a similar procedure can be implemented if those operators are nonlinear. This leads to replacing $H_k(\xi_k)$ by $H_k(\xi_k^{(0)}) + N_k \delta\xi_k$, where N_k is an appropriate 'simple' linear operator (possibly, but not necessarily, the jacobian of H_k at point $\xi_k^{(0)}$). The objective function depends only on the initial $\delta\xi_0$ deviation from $\xi_0^{(0)}$, and reads

$$\mathcal{J}_\mathrm{I}(\delta\xi_0) = (1/2)\,(x_0^b - \xi_0^{(0)} - \delta\xi_0)^\mathrm{T}\,[P_0^b]^{-1}\,(x_0^b - \xi_0^{(0)} - \delta\xi_0) + (1/2)\,\Sigma_k[d_k - N_k\delta\xi_k]^\mathrm{T}\,R_k^{-1}\,[d_k - N_k\delta\xi_k]$$

where $d_k \equiv y_k - H_k(\xi_k^{(0)})$ is the innovation at time k, and the $\delta\xi_k$ evolve according to

$$\delta\xi_{k+1} = L_k\,\delta\xi_k \qquad (2)$$

With the choices made here, $\mathcal{J}_\mathrm{I}(\delta\xi_0)$ is an exactly quadratic function of $\delta\xi_0$. The minimizing perturbation $\delta\xi_{0,m}$ defines a new initial state $\xi_0^{(1)} \equiv \xi_0^{(0)} + \delta\xi_{0,m}$, from which a new solution $\xi_k^{(1)}$ of the basic nonlinear equation is determined. The process is restarted in the vicinity of that new solution.

Incremental Method (continuation 4)

This defines a system of two-level nested loops for minimization. Advantage is that many degrees of freedom are available for defining the simplified operators L_k and N_k, and for defining an appropriate trade-off between practical implementability and physical usefulness and accuracy. It is the incremental method which, together with the adjoint method, makes variational assimilation possible.

First-Guess-At-the-right-Time 3D-Var (*FGAT 3D-Var*). Corresponds to $L_k = I_n$. Assimilation is four-dimensional in that observations are compared to a first-guess which evolves in time, but is three-dimensional in that no dynamics other than the trivial dynamics expressed by the unit operator is present in the minimization.

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.

How to take model error into account in variational assimilation ?

Weak constraint variational assimilation

Allows for errors in the assimilating model

- Data

  - Background estimate at time 0

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

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

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

  - Model

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

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

Then objective function

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

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

$$= (1/2)\, (x_0^b - \xi_0)^{\mathrm{T}}\, [P_0^b]^{-1}\, (x_0^b - \xi_0)$$

$$+ (1/2)\, \Sigma_{k=0,\ldots,K} [y_k - H_k \xi_k]^{\mathrm{T}}\, R_k^{-1}\, [y_k - H_k \xi_k]$$

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

Can include nonlinear M_k and/or H_k.

Implemented operationally at ECMWF for the assimilation in the stratosphere.

Becomes singular in the strong constraint limit $Q_k \rightarrow 0$

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

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

or, equivalently, but more conveniently, as

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

where, as before

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

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

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

It is associated with error covariance matrix

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

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 $\xi = (\xi_0{}^{\mathrm{T}}, \upsilon_0{}^{\mathrm{T}}, \ldots, \upsilon_{K\text{-}1}{}^{\mathrm{T}})^{\mathrm{T}}$, 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 $k = 0, \ldots, K\text{-}1$

$$\xi_{k+1} = M_k \xi_k + \upsilon_k$$
$$u_{k+1} = H_{k+1} \xi_{k+1}$$

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

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

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

$$\varphi_0 = \xi_0 + R_0 \mu_0$$
$$\varphi_k = \upsilon_{k\text{-}1} + R_k \mu_k \qquad , \ k = 1, \ldots,$$

- Step 5. Change sign of vector $\varphi = (\varphi_0^{\ \mathrm{T}}, \ldots, \varphi_K^{\ \mathrm{T}})^{\mathrm{T}}$, and add vector $d = y - Hx^b$,

Dual Algorithm for Variational Assimilation (continuation 7)

Temporal correlations can be introduced.

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 (Courtier, Louvel)

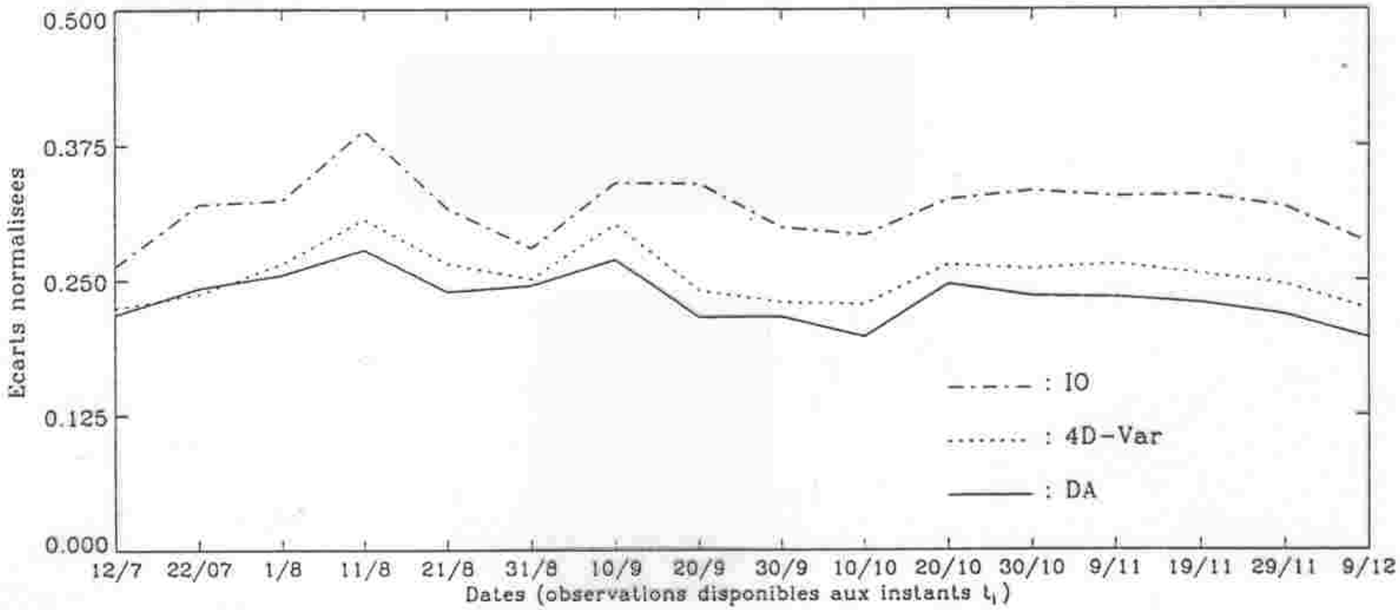


FIG. 9.11 – *Ecarts normalisés prévision/observations sur l'ensemble de la période étudiée*

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

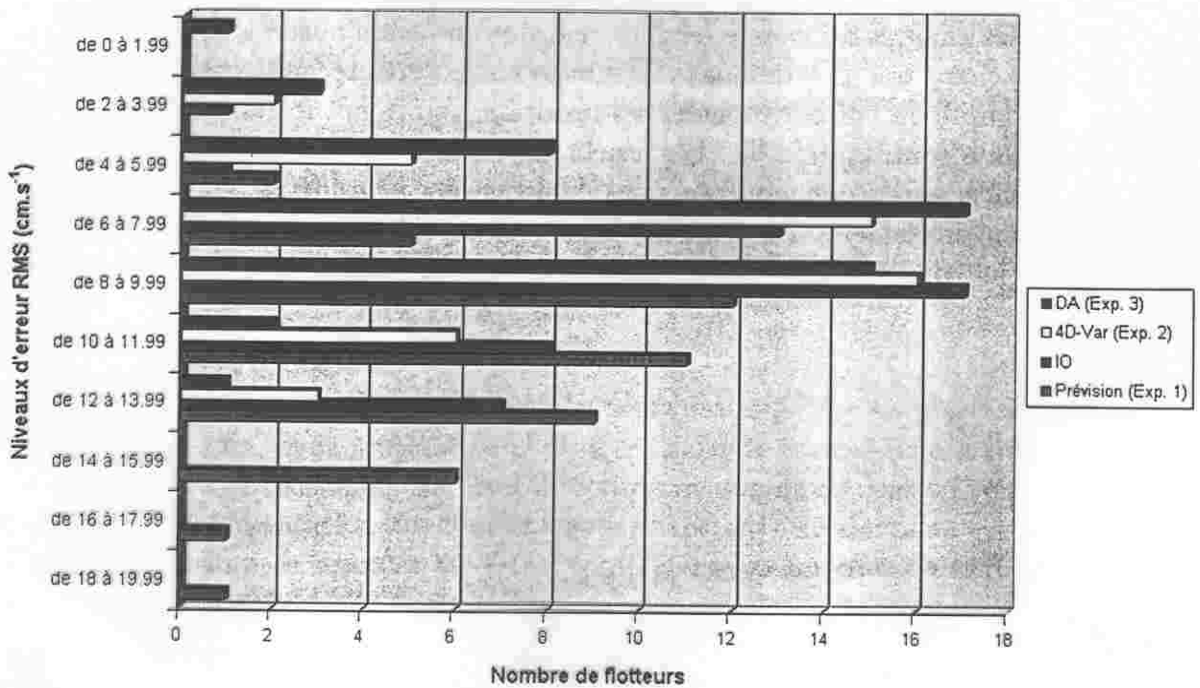


FIG. 9.15 – *Description des écarts flotteurs/modèle en terme de vitesse (à 150 m de profondeur) pour les différents algorithmes d'assimilation*

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, …)

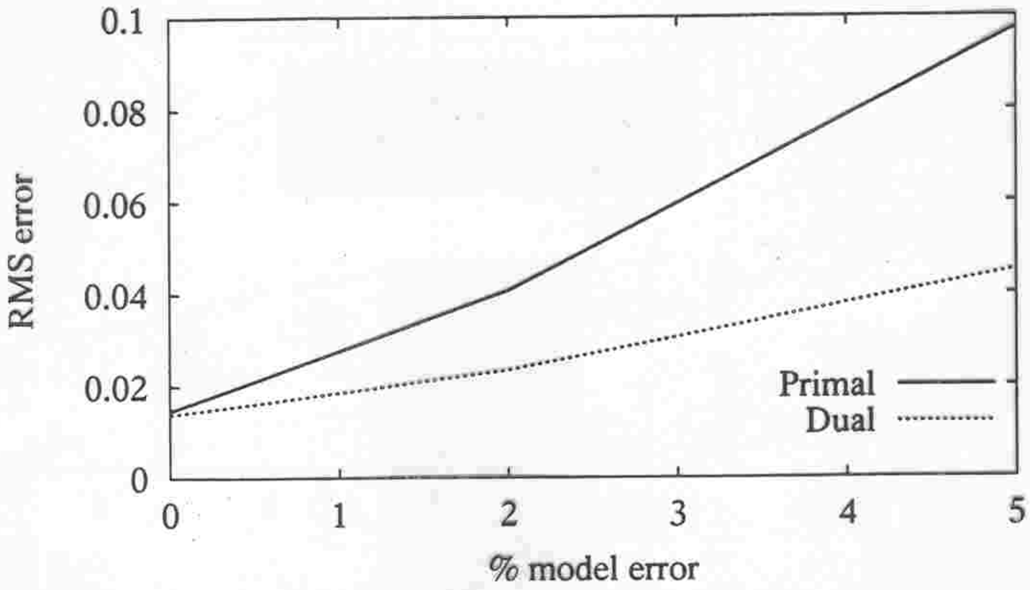


FIG. 6.13 – Normes RMS des erreurs d'assimilation obtenues pour les deux méthodes en fonction de l'erreur introduite dans le modèle au cours de la période d'assimilation.

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

Dual 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.

Conclusion on Sequential Assimilation

Pros

'Natural', and well adapted to many practical situations

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.

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

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 still needs further development for application in non exactly linear cases.

- Is ensemble variational assimilation possible ? Probably yes. But also needs development.

Variational assimilation has been extended to non Gaussian probability distributions (lognormal distributions), the unknown being the mode of the conditional distribution (M. Zupanski, Fletcher).

Bayesian character of variational assimilation ?

- If everything is linear and gaussian, ready recipe for obtaining bayesian sample

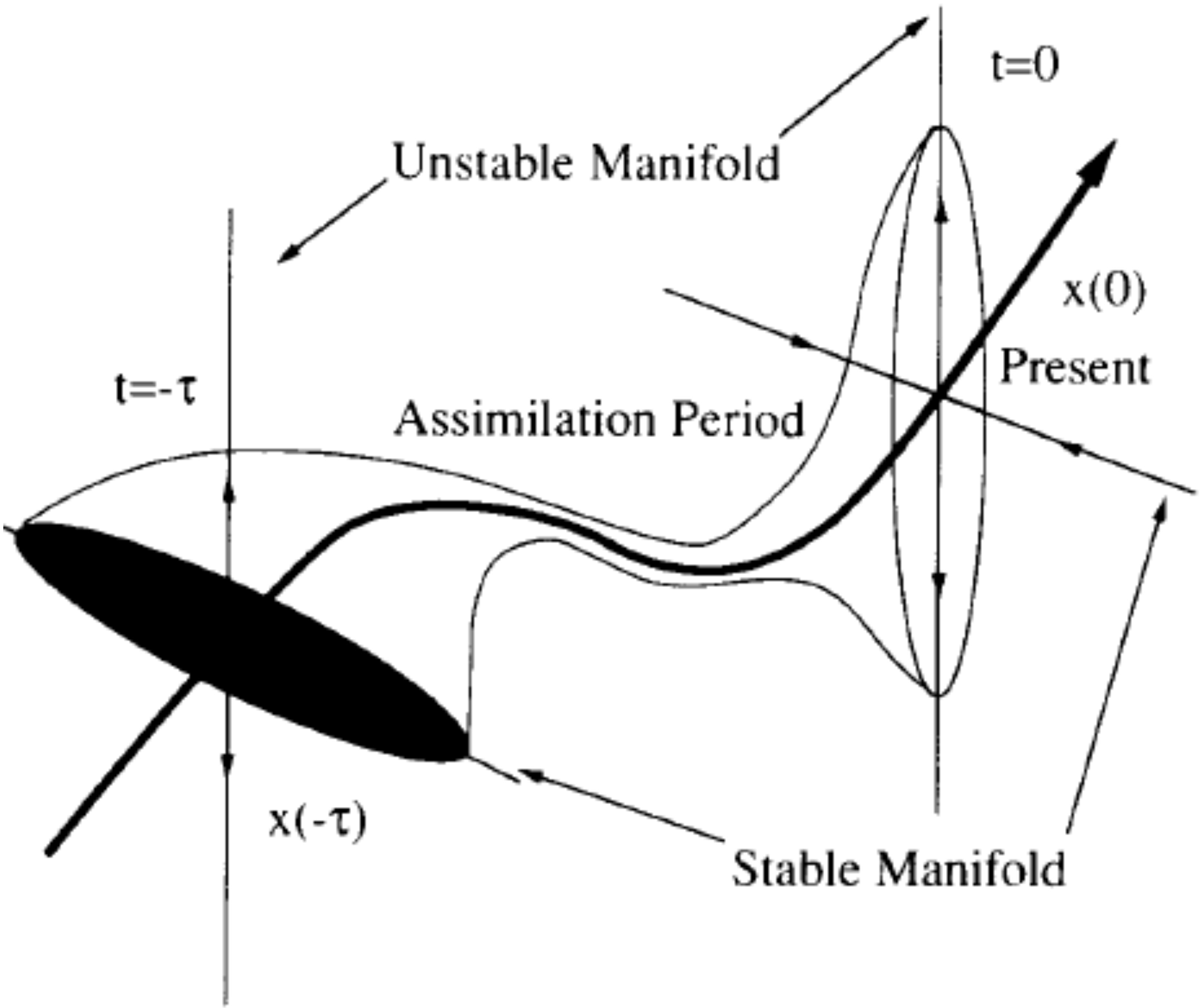
Perturb data (background, observations and model) according to their error probability distributions, do variational assimilation, and repeat process

Sample of system orbits thus obtained is bayesian

- If not, very little can be said at present

If there is uncertainty on the state of the system, and dynamics of the system is perfectly known, uncertainty on the state along stable modes decreases over time, while uncertainty along unstable modes increases.

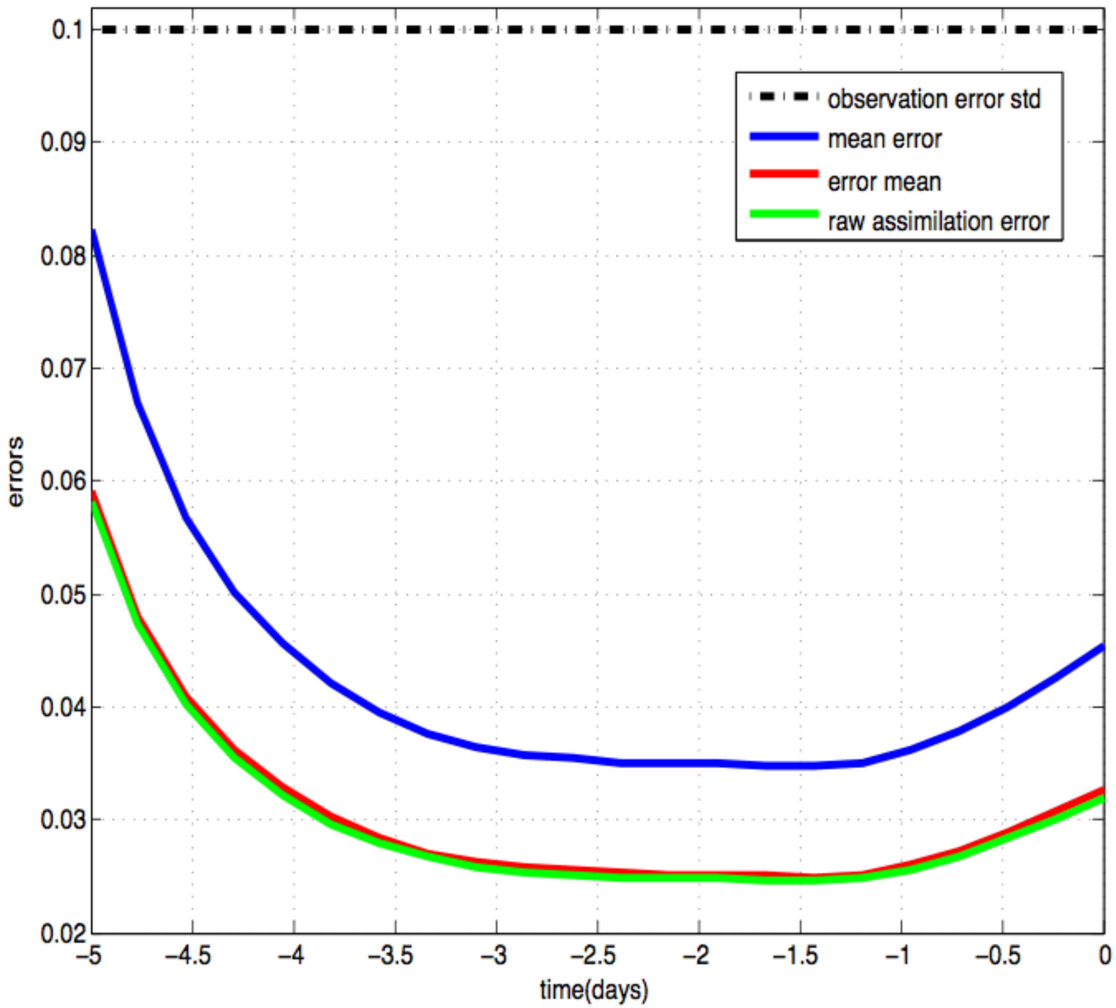
Stable (unstable) modes : perturbations to the basic state that decrease (increase) over time.



t=0
Unstable Manifold
x(0)
Present
t=-τ
Assimilation Period
x(-τ)
Stable Manifold

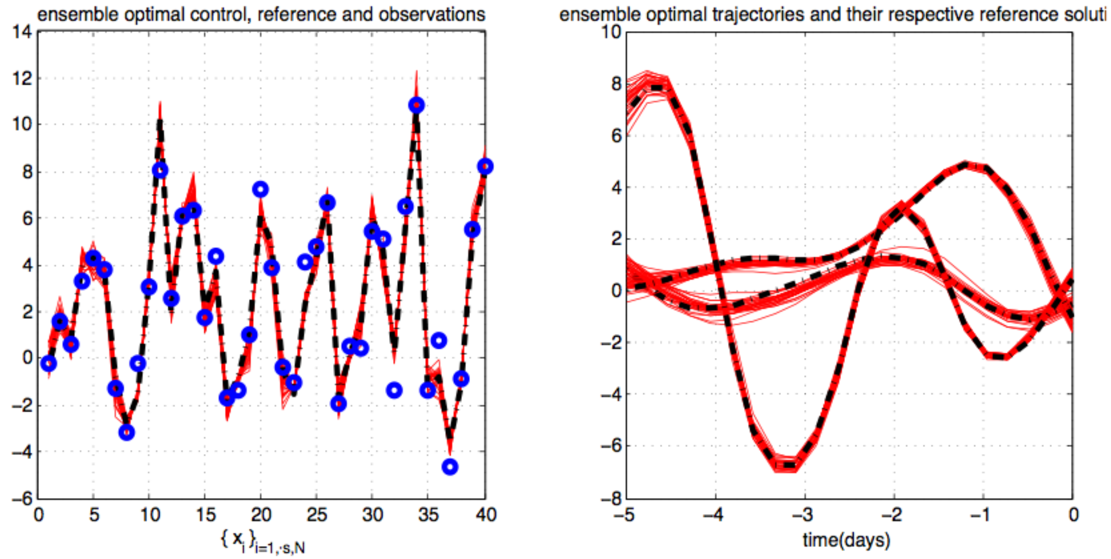
Consequence : Consider 4D-Var assimilation, or any form of smoother, which carries information both forward and backward in time, performed over time interval $[t_0, t_1]$ over uniformly distributed noisy data. If assimilating model is perfect, estimation error is concentrated in stable modes at time t_0, and in unstable modes at time t_1. Error is smallest somewhere within interval $[t_0, t_1]$.

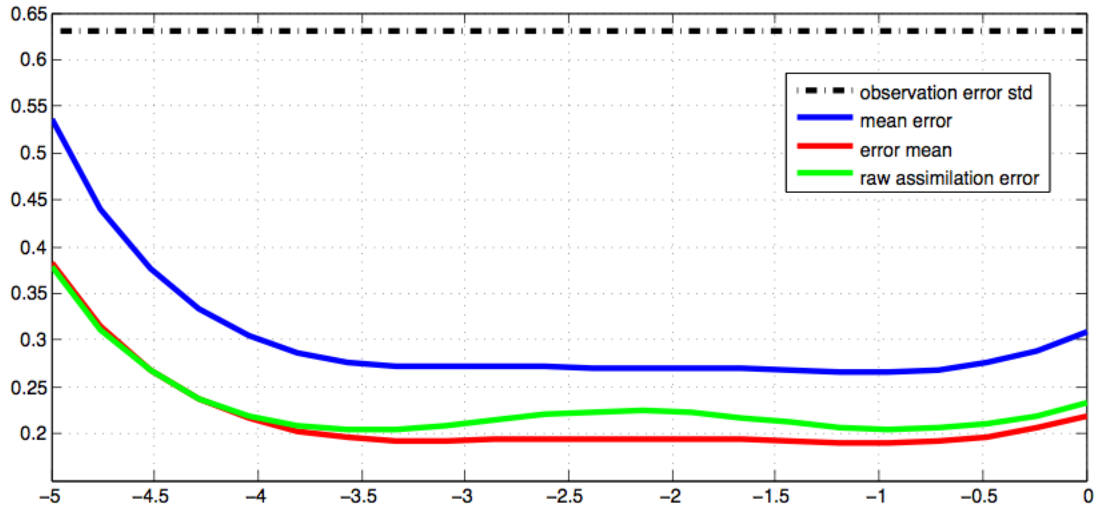
Similar result holds true for Kalman filter (or more generally any form of sequential assimilation), in which estimation error is concentrated in unstable modes at any time.



Linearized Lorenz'96. 5 days

Jardak and Talagrand





Nonlinear Lorenz'96. 5 days

Jardak and Talagrand

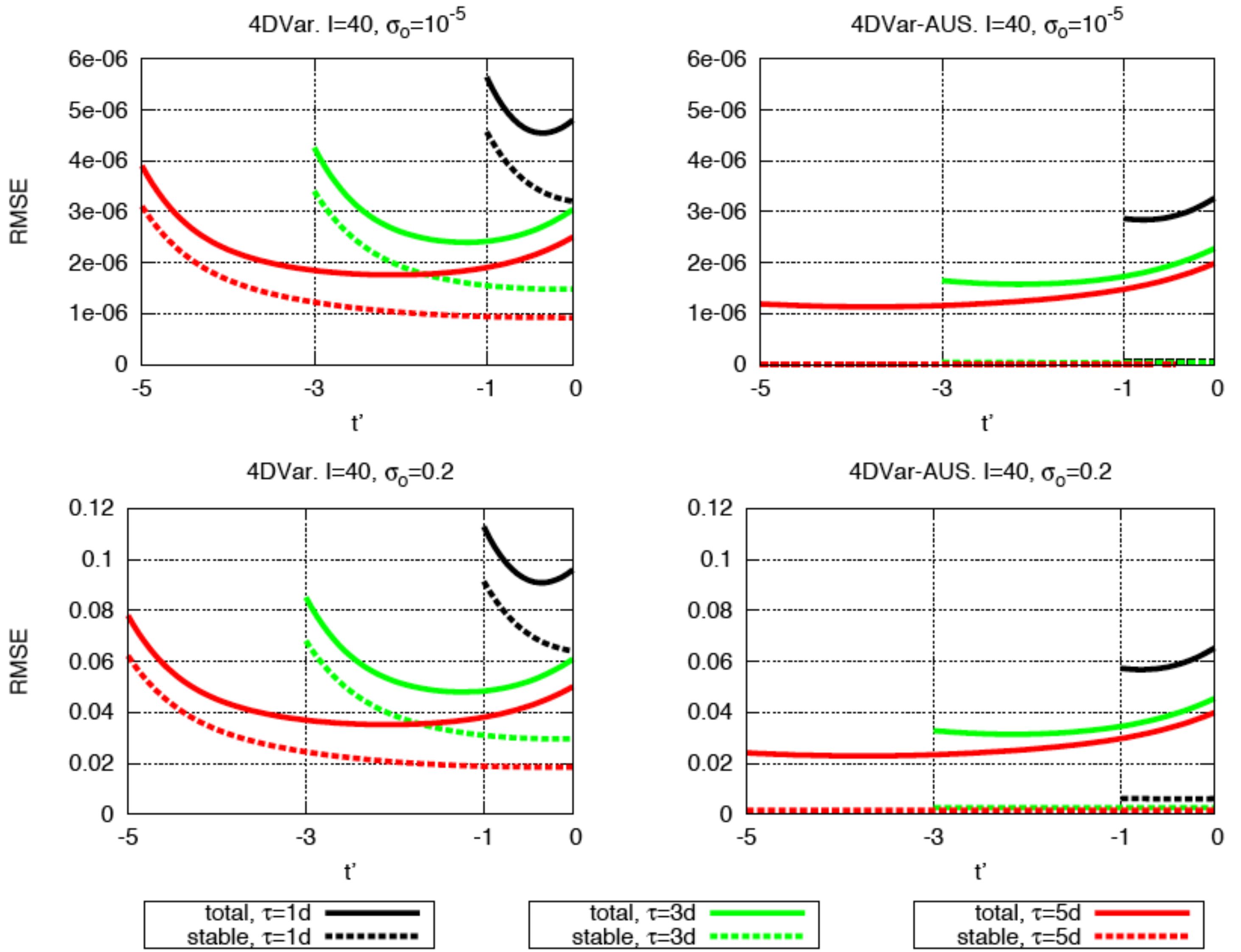


Figure 3. Time average RMS error within 1, 3, 5 days assimilation windows as a function of $t' = t - \tau$, with $\sigma_o = .2, 10^{-5}$ for the model configuration $I = 40$. Left panel: 4DVar. Right panel: 4DVar-AUS with $N = 15$. Solid lines refer to total assimilation error, dashed lines refer to the error component in the stable subspace $\mathbf{e}_{16}, ..., \mathbf{e}_{40}$.

Trevisan *et al*., 2010, *Q. J. R. Meteorol. Soc.*

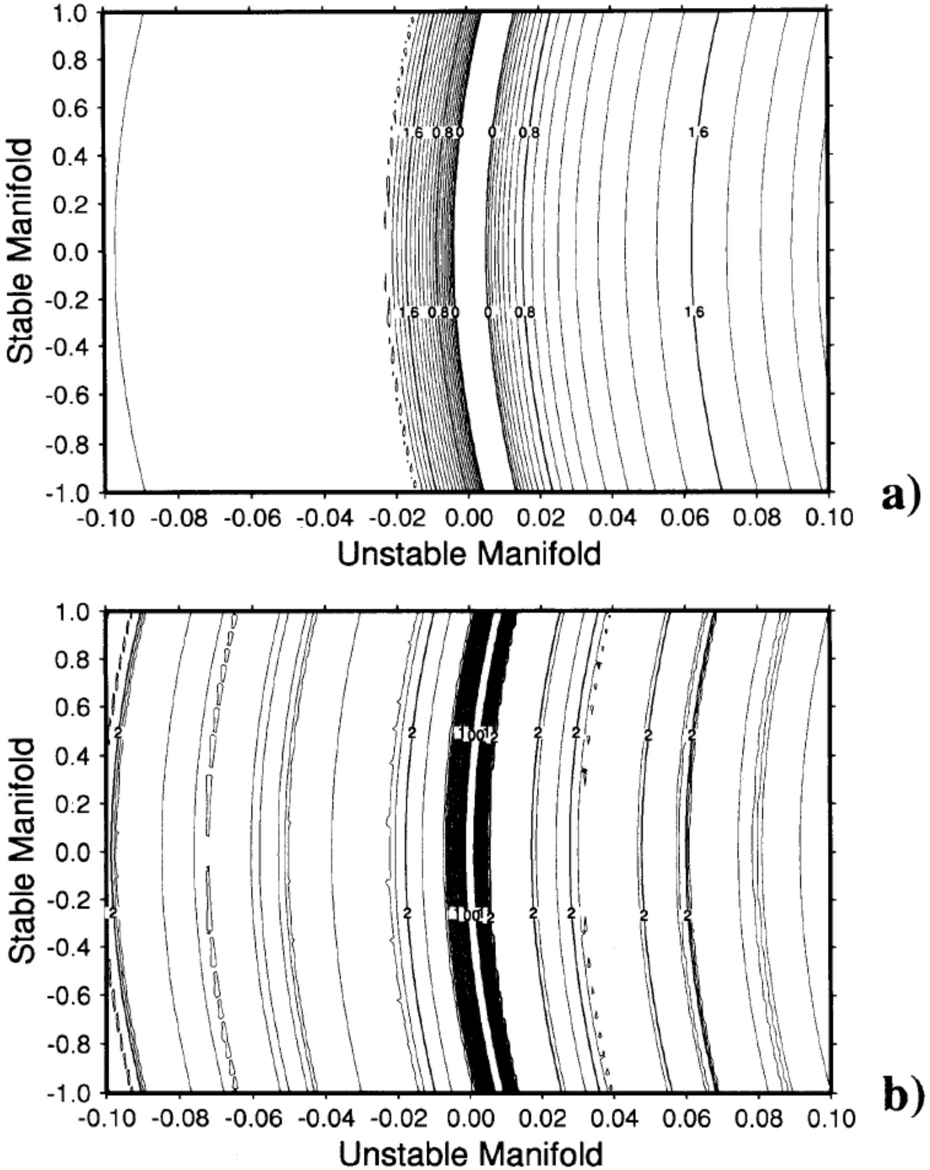


Fig. 3. Variations of the error-free forward cost-function $J'_e(\tau, \hat{x}, x)$ (Lorenz system) in the plane spanned by the stable and unstable directions, as determined from the tangent linear system (see text), and for $\tau = 6$ (panel (a)) and $\tau = 8$ (panel (b)) respectively. The metric has been distorted in order to make the stable and unstable manifolds orthogonal to each other in the figure. The scale on the contour lines is logarithmic (decimal logarithm). Contour interval: 0.1. For clarity, negative contours, which would be present only in the central "valley" directed along the stable manifold, have not been drawn.

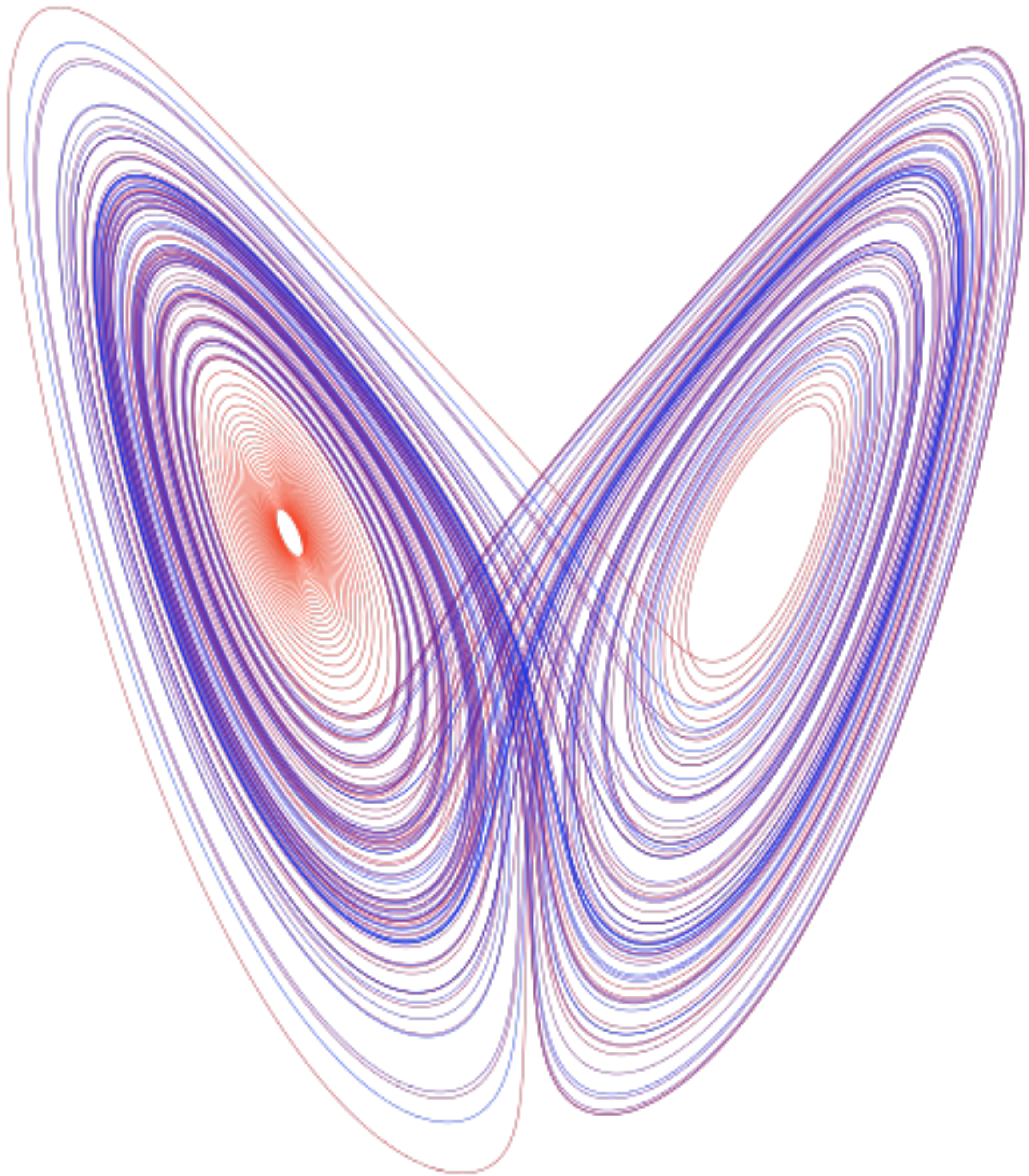
Lorenz (1963)

$$dx/dt = \sigma(y - x)$$

$$dy/dt = \rho x - y - xz$$

$$dz/dt = -\beta z + xy$$

with parameter values $\sigma = 10$, $\rho = 28$, $\beta = 8/3$ $\Rightarrow$ chaos



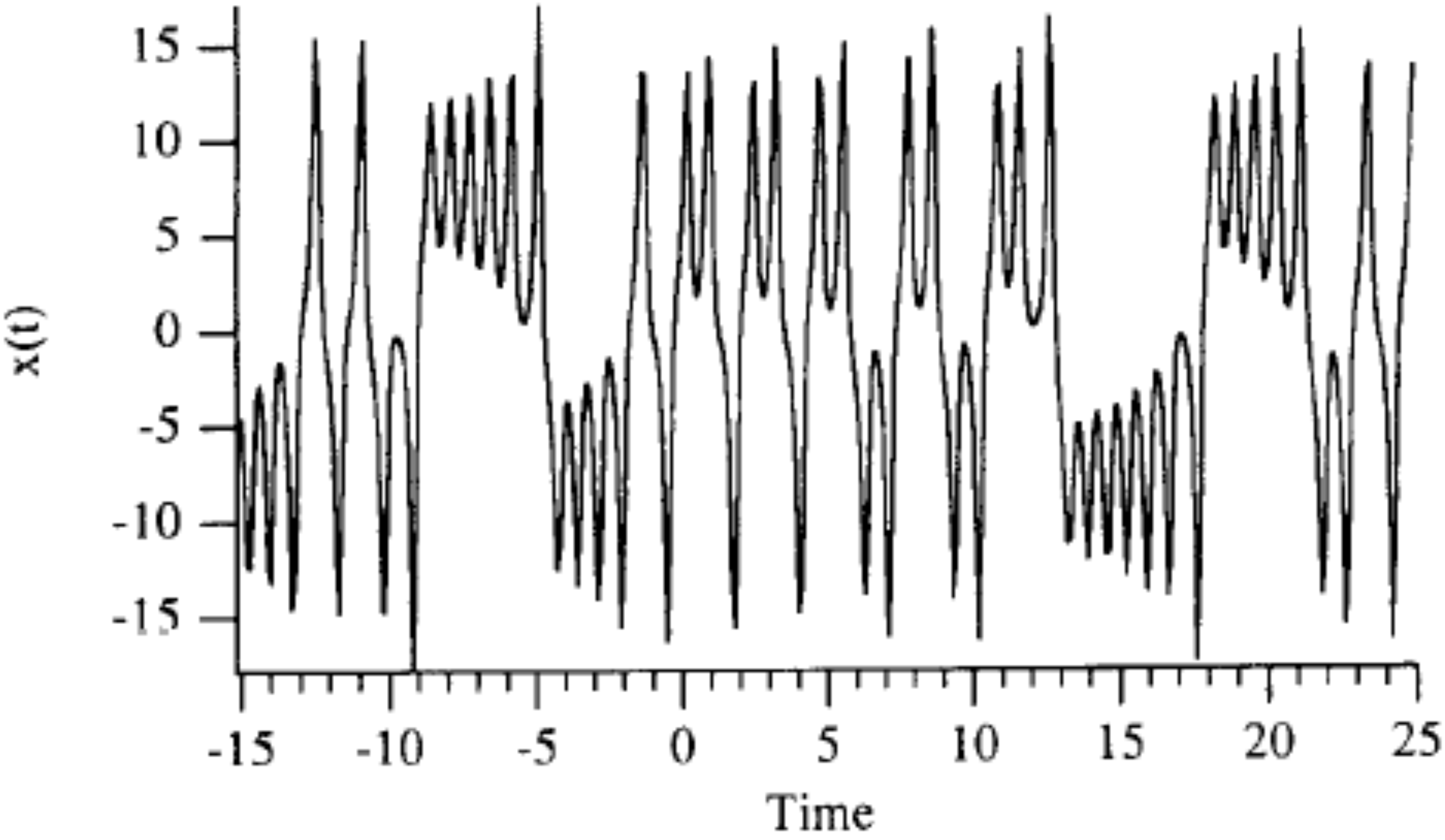


Fig. 2. Time variations, along the reference solution, of the variable $x(t)$ of the Lorenz system.

Twin (strong constraint) experiment. Observations $y_k = H_k x_k + \varepsilon_k$ at successive times k, and objective function of form

$$\mathcal{J}(\xi_0) = (1/2)\, \Sigma_k [y_k - H_k \xi_k]^T R_k^{-1} [y_k - H_k \xi_k]$$

x_k denotes here the complete state vector, and H_k is the unit operator (all three components of x_k are observed)

No 'background' term from the past, but observation y_0 at time $k = 0$.

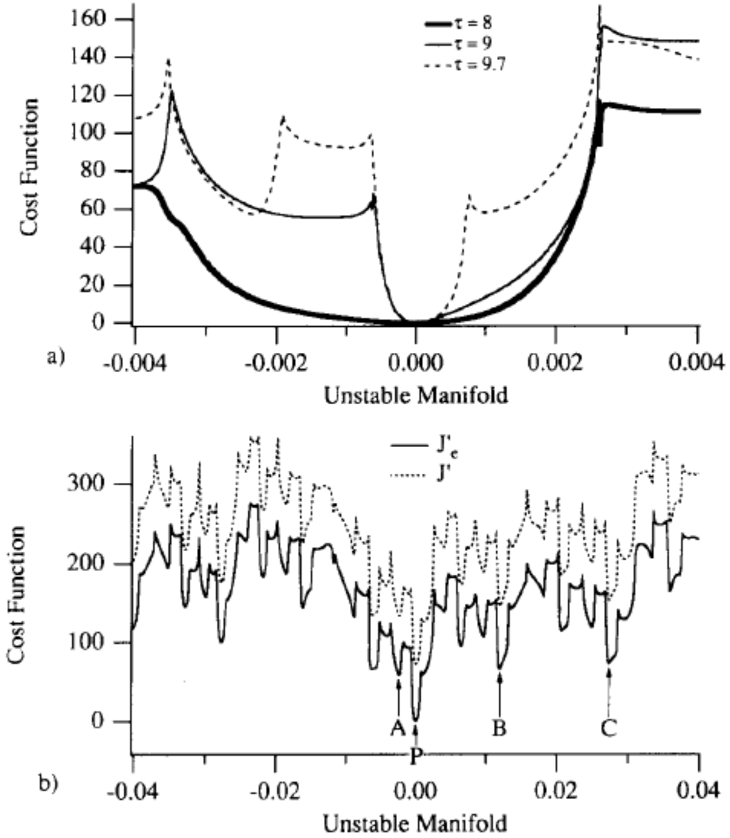


Fig. 4. Panel (a): Cross-section of the error-free forward cost-function $J'_e(\tau, \hat{x}, x)$ along the unstable manifold, for various values of τ. Panel (b). As in panel (a), for $\tau = 9.7$, and with a display interval ten times as large, respectively for the error-free forward cost-function $J'_e(\tau, \hat{x}, x)$ (solid curve) and for the error-contaminated cost-function $J_e(\tau, \hat{x}, x)$ (dashed curve). In the latter case, the total variance of the observational noise is $E^2 = 75$.

Pires *et al., Tellus*, 1996 ; Lorenz system (1963)

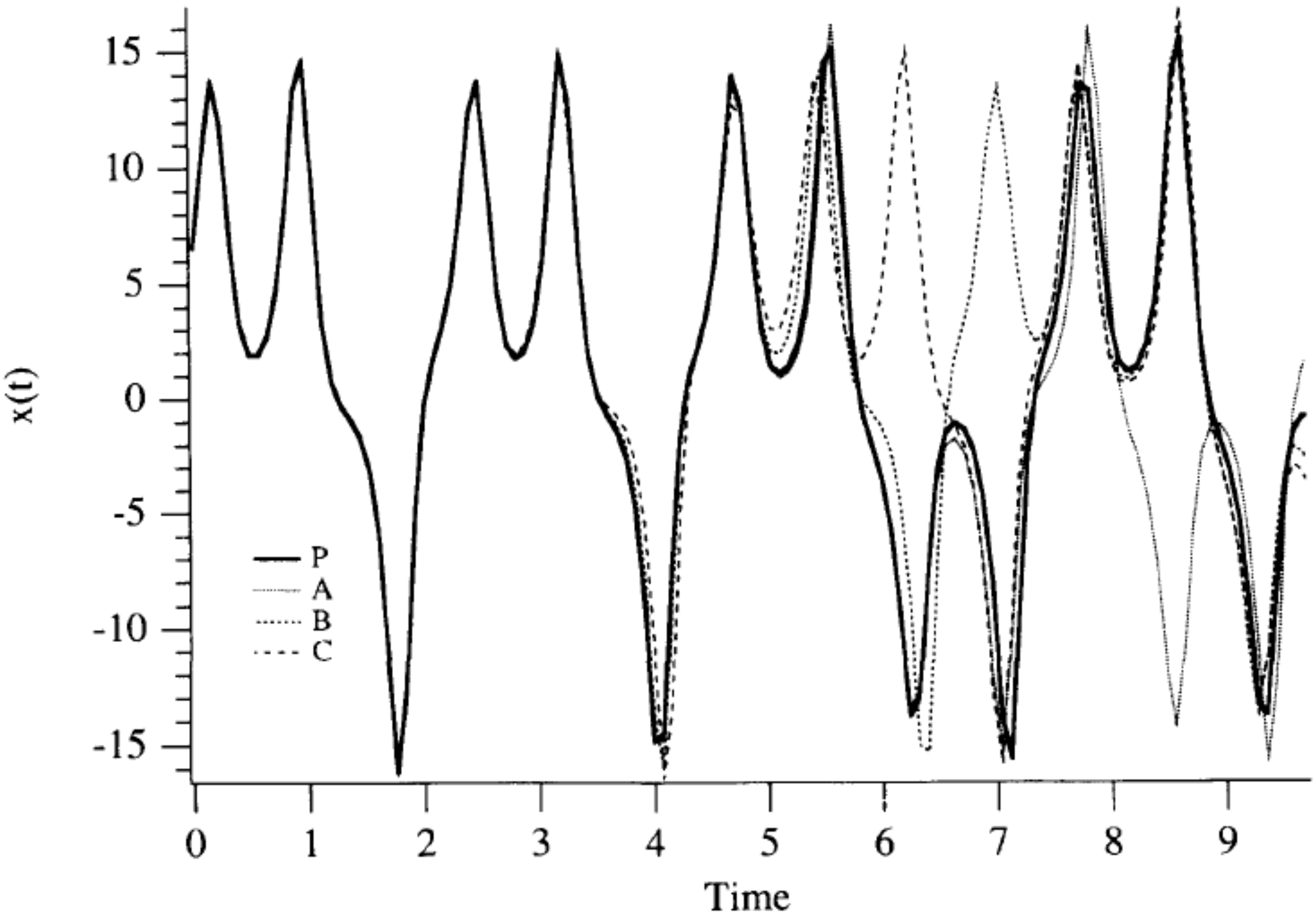


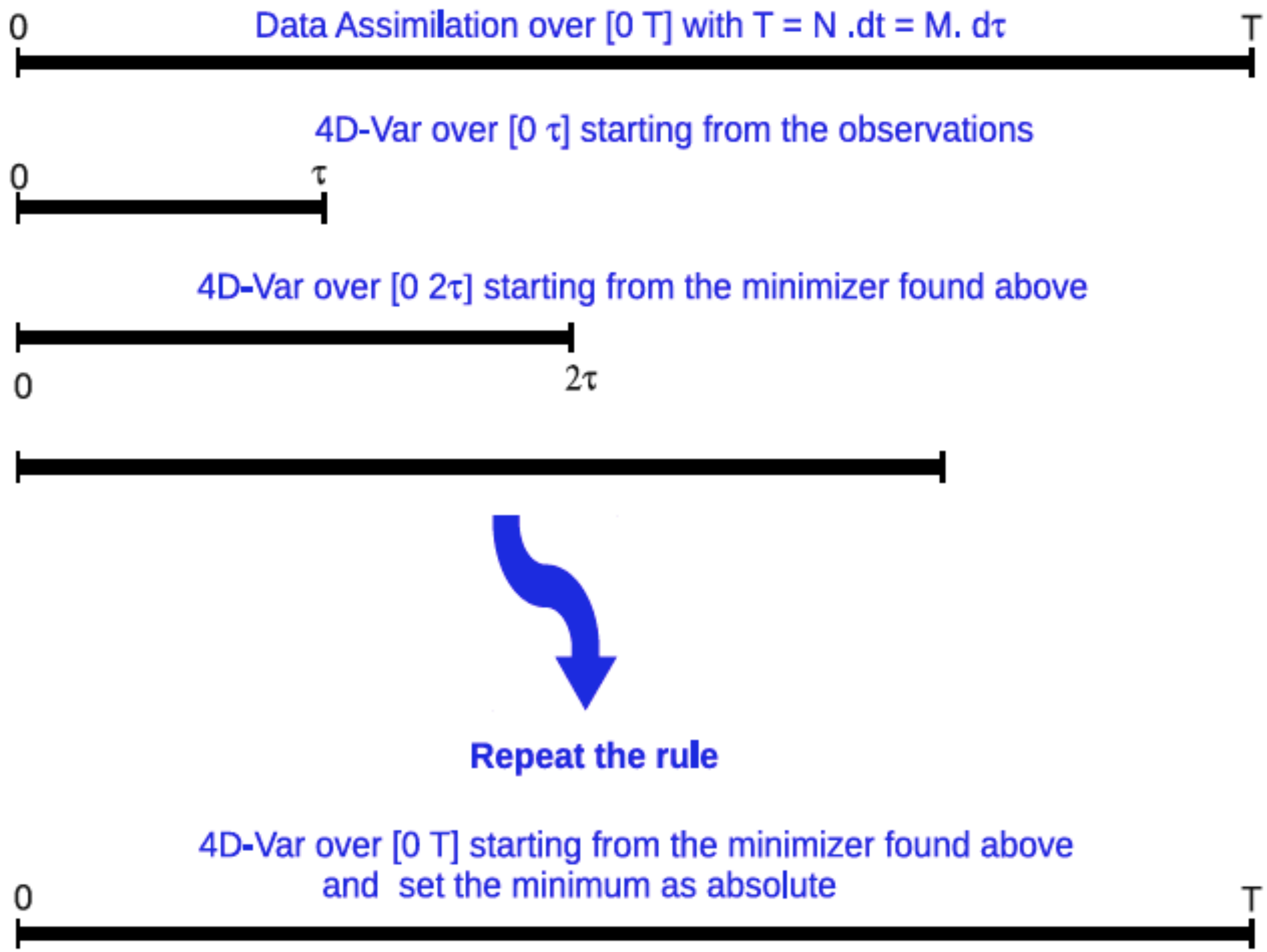
Fig. 5. Variations of the coordinate x along the orbits originating from the minima *P*, *A*, *B*, *C* (indicated in Fig. 4b) of the error-free cost-function.

Minima in the variations of objective function correspond to solutions that have bifurcated from the observed solution, and to different folds in state space.

Quasi-Static Variational Assimilation (*QSVA*). Increase progressively length of the assimilation window, starting each new assimilation from the result of the previous one. This should ensure, at least if observations are in a sense sufficiently dense in time, that current estimation of the system always lies in the attractive basin of the absolute minimum of objective function (Pires *et al.*, Swanson *et al.*, Luong, Järvinen *et al.*)

.

Quasi-Static Variational Assimilation (QSVA)



$\mu(C(\tau, x))$	Cloud of points QSVA	Cloud of points raw assimilation	Linear tangent system	Upper bound
$\tau = 0$	1	1	1	1
$\tau = 1$	0.36	0.37	0.39	0.46
$\tau = 2$	5.9×10^{-2}	5.74	4.5×10^{-2}	0.401
$\tau = 3$	3.3×10^{-2}	29.4	2.9×10^{-2}	0.397
$\tau = 8$	1.4×10^{-2}	59.9	*	0.396

In the left column, the estimates are calculated from the ensemble of 100 assimilations (see also Fig. 7). The 2nd column contains the values obtained from the raw assimilation. In the 3rd column, the estimates are obtained from the tangent linear system and eqs. (3.5–3.9) (the star indicates a computational overflow). The estimates in the right-hand column are the upper bounds defined by eq. (3.13).

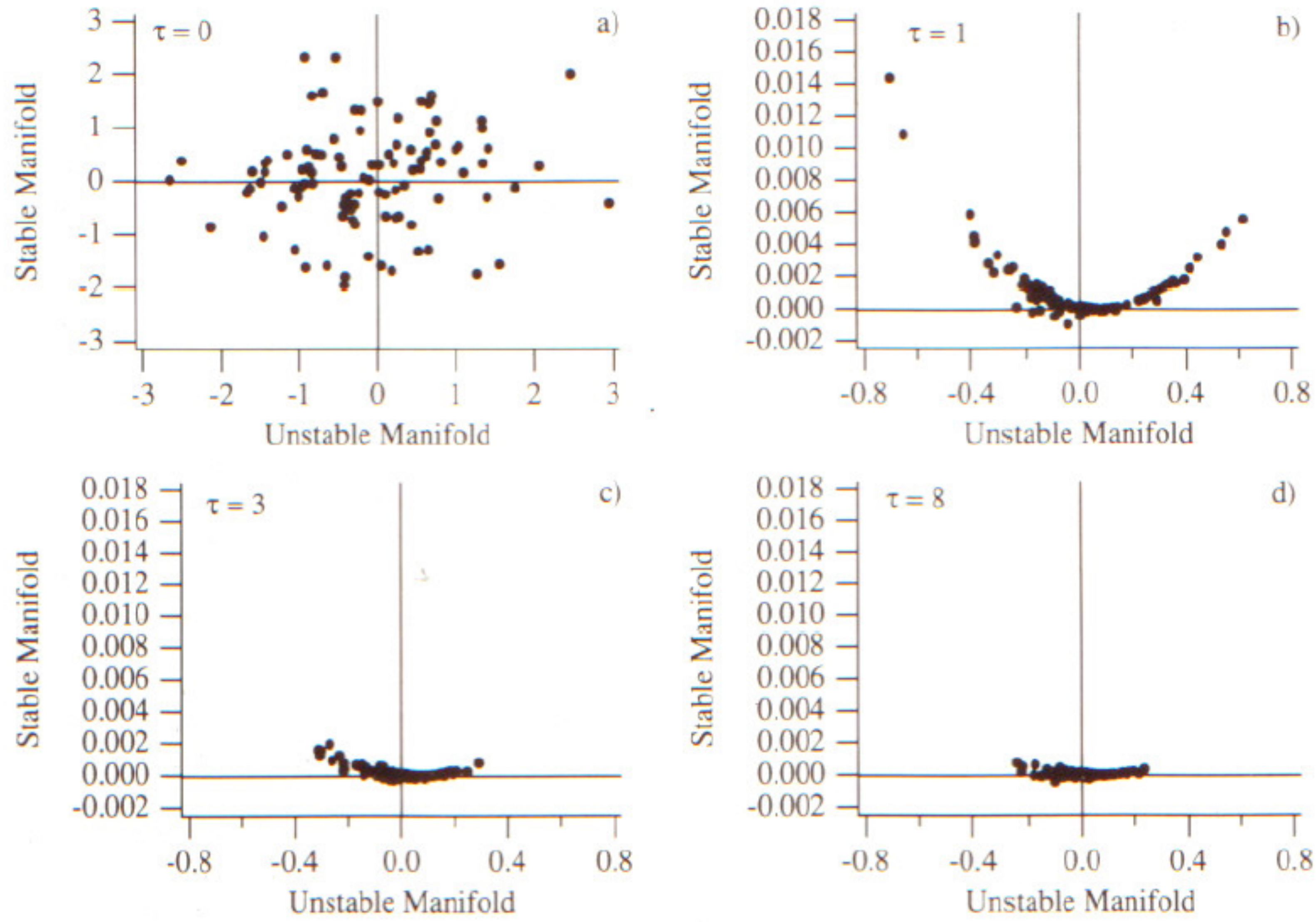


Fig. 7. Projection of the 100 minimizing solutions, at the end of the assimilation period, onto the plane spanned by the stable and unstable directions, defined as in Fig. 3. Values of τ are indicated on the panels. The projection is not an orthogonal projection, but a projection parallel to the local velocity vector (dx/dt, dy/dt, dz/dt) (central manifold).

Pires *et al., Tellus*, 1996 ; Lorenz system (1963)

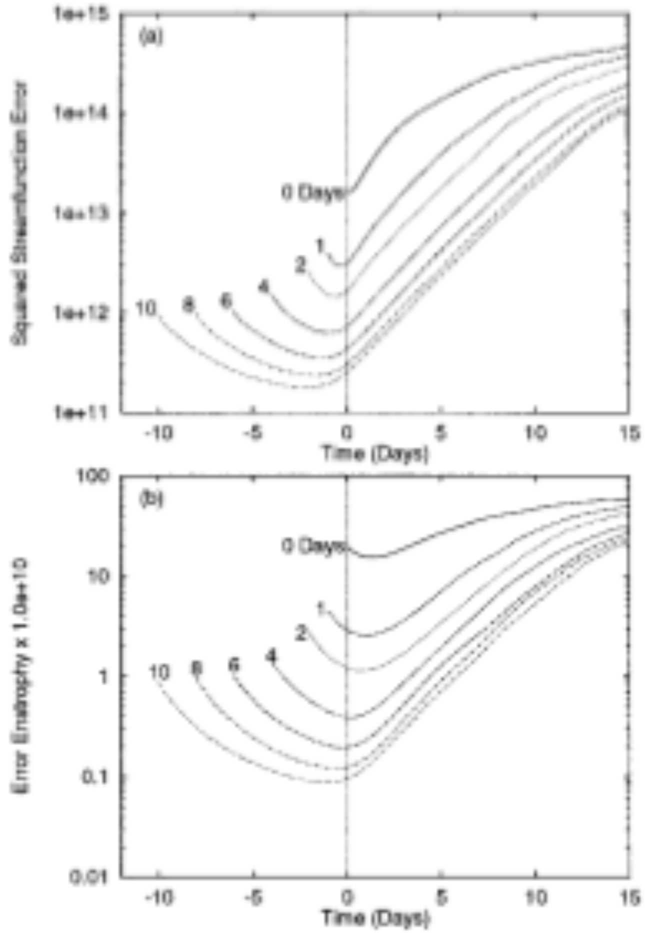


Fig. 5. Median values of the (a) streamfunction squared error, and (b) enstrophy error for the 200 forecast set as a function of forecast time and of the assimilation time T_a.

Swanson, Vautard and Pires, 1998, *Tellus*, **50A**, 369-390

Cours à venir

~~Jeudi 19 Mars~~

~~Jeudi 26 mars~~

~~Jeudi 02 avril~~

~~Jeudi 09 avril~~

~~Mardi 21 avril, 14h00~~

~~Mardi 28 avril, 14h00~~

Mardi 12 mai, 14h00

Mardi 26 mai, 14h00