École Doctorale des Sciences de l'Environnement d'Île-de-France
Année Universitaire 2022-2023

Modélisation Numérique de l'Écoulement Atmosphérique et Assimilation de Données

Olivier Talagrand
Cours 6

9 Mai 2023

Last course (May 2)

- Assimilation variationnelle. Principe

- Méthode adjointe. Principe.

- Assimilation variationnelle. Résultats

- La Méthode incrémentale

- Compléments sur l'Estimation Statistique (*BLUE*)

- Weak-constraint Variational Assimilation.
  Dual Algorithm for Variational Assimilation
- Complements on Variational Assimilation.
  - Mahalanobis Norm
  - 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 errors $\boldsymbol{M}_k\boldsymbol{x}_k - \boldsymbol{x}_{k+1} + \boldsymbol{\eta}_k = 0$ $(k = 0, \ldots, K\text{-}1)$

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)

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

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

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

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

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

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

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

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

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

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 $(\xi_0^T, \upsilon_0^T, \ldots, \upsilon_{K-1}^T)^T$, thereby producing vector $(u_0^T, \ldots, u_K^T)^T$.

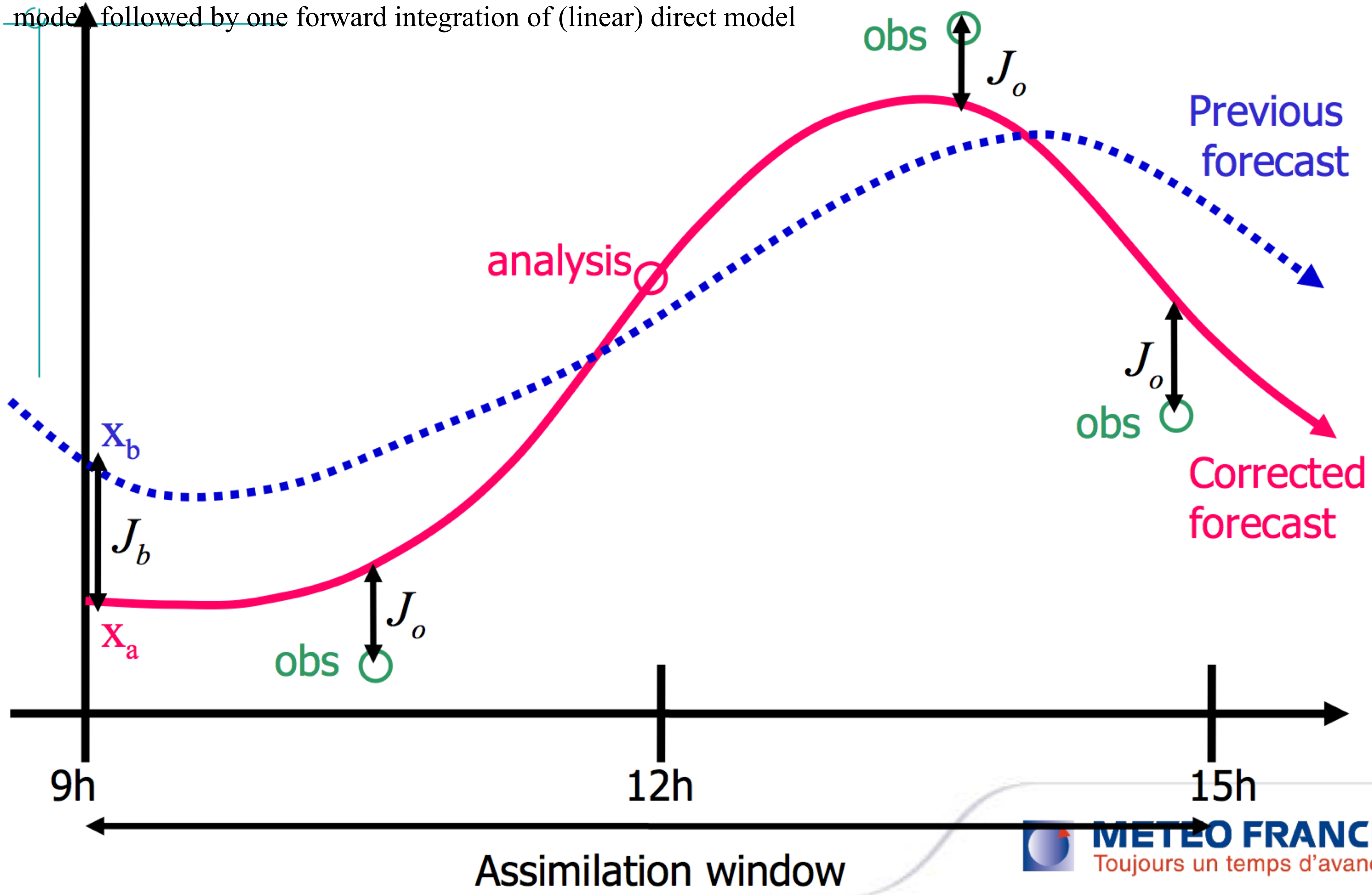
- Step 4. Add vector $R\mu$, *i. e.* compute

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

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

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. The cost depends mainly on the number of model integrations (Courtier, Louvel)

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

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

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

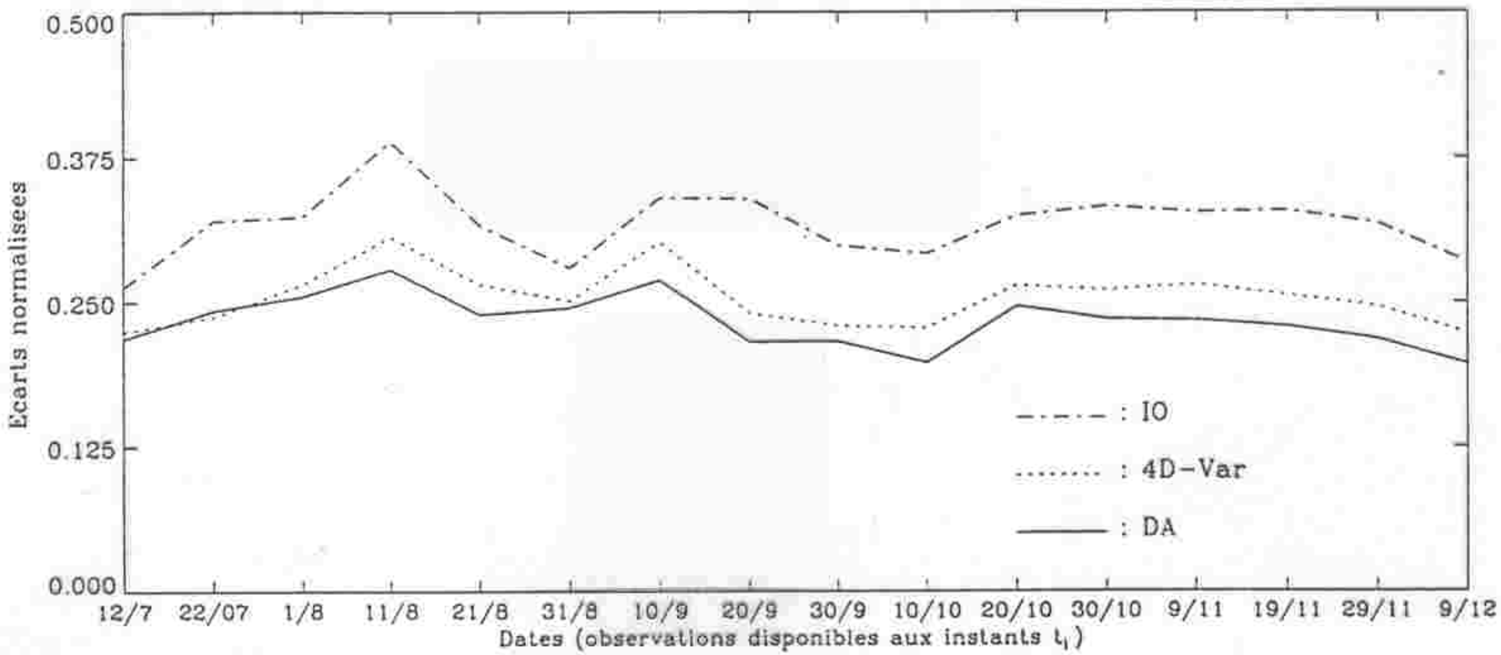


FIG. 9.11 – *Ecarts normalisés prévision/observations sur l'ensemble de la période étudiée*

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

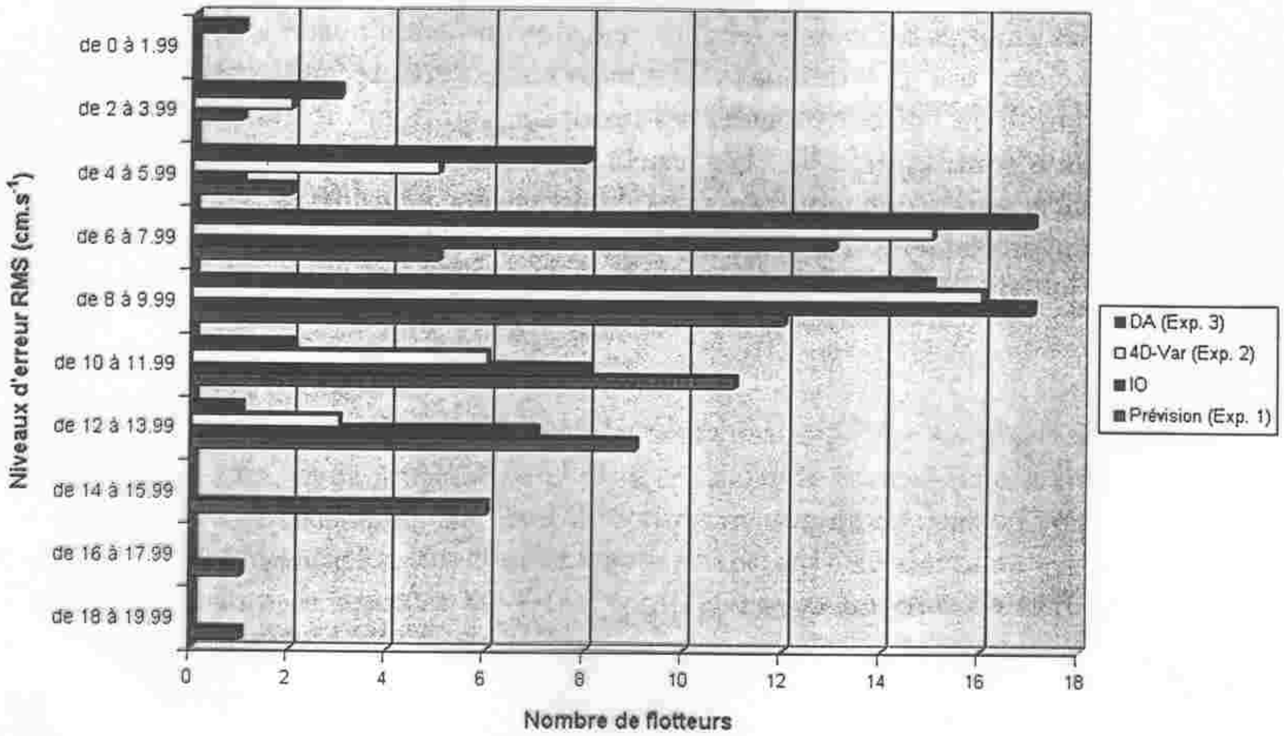


FIG. 9.15 – *Description des écarts flotteurs/modèle en terme de vitesse (à 150 m de profondeur) pour les différents algorithmes d'assimilation*

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, …)

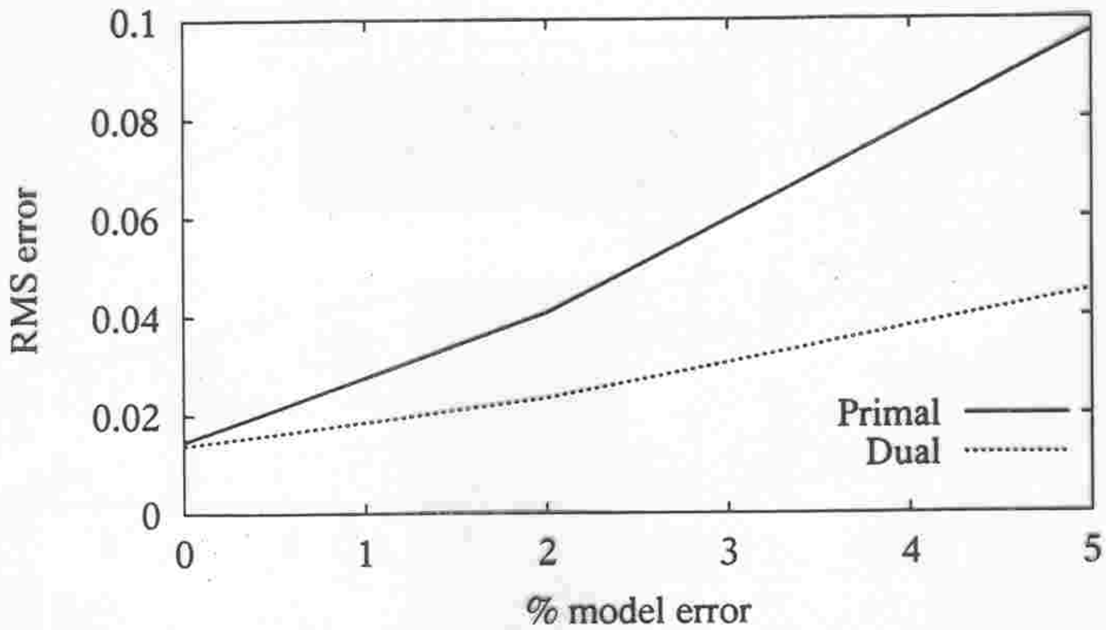


FIG. 6.13 – Normes RMS des erreurs d'assimilation obtenues pour les deux méthodes en fonction de l'erreur introduite dans le modèle au cours de la période d'assimilation.

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

Dual 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.

Data

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

Analysis minimizes scalar objective function

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

where $\boldsymbol{S} = \mathrm{E}(\boldsymbol{\zeta\zeta}^{\mathrm{T}})$ is covariance matrix of data error $\boldsymbol{\zeta}$

Consider quantity $D = z_1^{\mathrm{T}} S^{-1} z_2 = z_1^{\mathrm{T}} [\mathrm{E}(\zeta\zeta^{\mathrm{T}})]^{-1} z_2$

where z_1 and z_2 are any two vectors in data space

Change of coordinates $z \equiv Tw$

$$\zeta = T\chi \quad \Rightarrow S = \mathrm{E}(\zeta\zeta^{\mathrm{T}}) = \mathrm{E}[T\chi(T\chi)^{\mathrm{T}}] = T\,\mathrm{E}(\chi\chi^{\mathrm{T}})T^{\mathrm{T}}$$

$$D = w_1^{\mathrm{T}}\, T^{\mathrm{T}}\, [T\,\mathrm{E}(\chi\chi^{\mathrm{T}})T^{\mathrm{T}}]^{-1} T w_2$$

$$D = w_1^{\mathrm{T}}\, [\mathrm{E}(\chi\chi^{\mathrm{T}})]^{-1}\, w_2$$

Expression $D = z_1^{\mathrm{T}} S^{-1} z_2$

defines proper scalar product, and associated norm, on data space

Called ***Mahalanobis norm***

Prasanta Chandra Mahalanobis (1893 -1972)

Gaussian variables

Unidimensional

$$\mathcal{N}[m, a] \sim (2\pi a)^{-1/2} \exp[-(1/2a)(\xi - m)^2]$$

Dimension n

$$\mathcal{N}[\boldsymbol{m}, \boldsymbol{A}] \sim [(2\pi)^n \det\boldsymbol{A}]^{-1/2} \exp[-(1/2)\underbrace{(\boldsymbol{\xi}-\boldsymbol{m})^{\mathrm{T}}\boldsymbol{A}^{-1}(\boldsymbol{\xi}-\boldsymbol{m})}_{\textit{Mahalanobis norm}}]$$

Minimum of objective function

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

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

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

where $\boldsymbol{d}$ is innovation ($\boldsymbol{d} = \boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}^b = \boldsymbol{\varepsilon} - \boldsymbol{H}\zeta^b$, or more generally what is obtained by eliminating the unknown $\boldsymbol{x}$ from the data 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)*

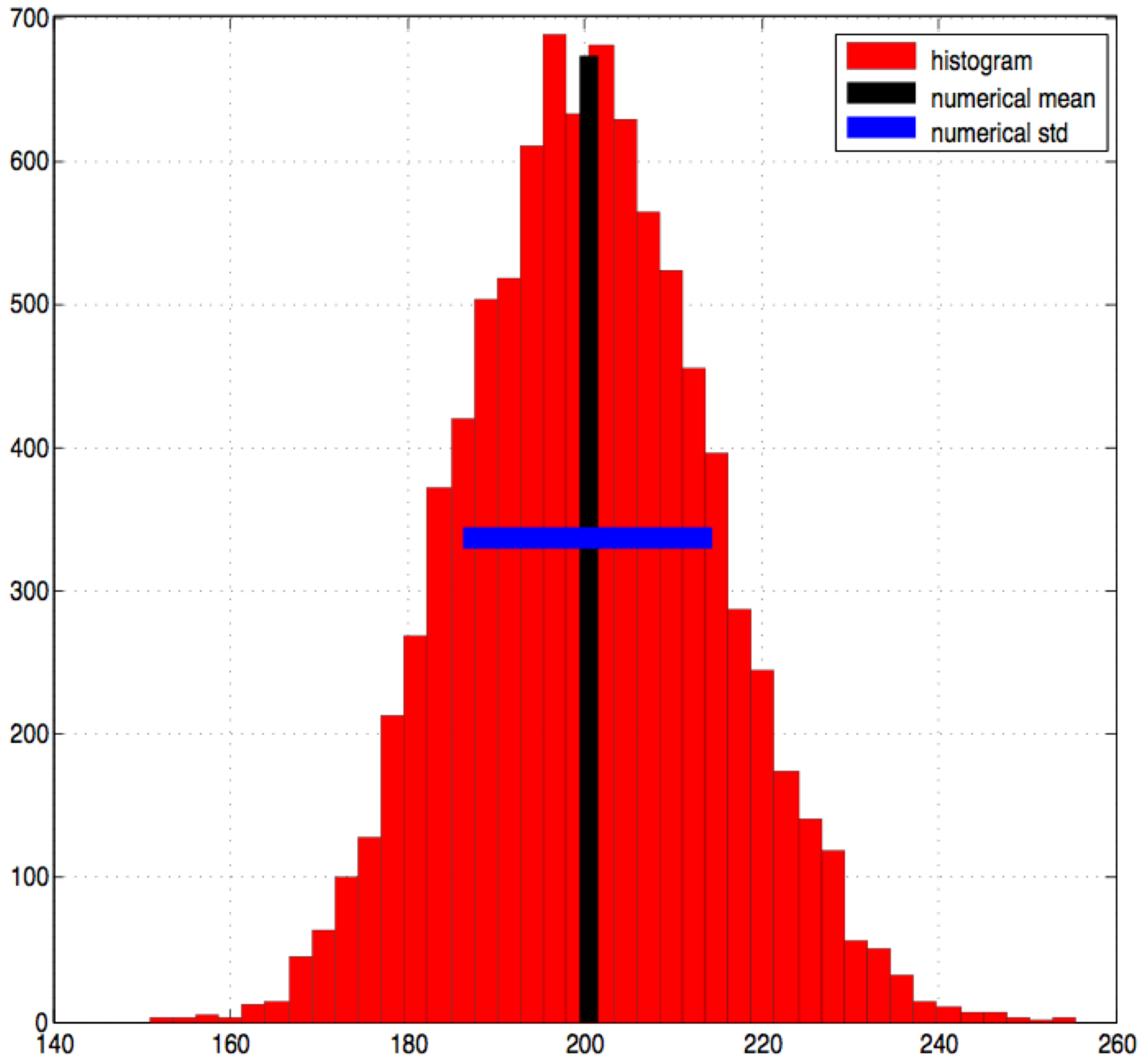
$$\mathcal{J}_{min} = (1/2)\ \mathrm{Tr}\left\{[E(\boldsymbol{d}\boldsymbol{d}^{\mathrm{T}})]^{-1}\ \boldsymbol{d}\boldsymbol{d}^{\mathrm{T}}\right\}$$

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

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

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

Remark. 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.



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).

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 $a = b \times c$

Differentiate $\delta a = b \times \delta c + c \times \delta b$

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

$\boldsymbol{bad} = \boldsymbol{bad} + \boldsymbol{c} \times \boldsymbol{aad}$

$\boldsymbol{cad} = \boldsymbol{cad} + \boldsymbol{b} \times \boldsymbol{aad}$

$\boldsymbol{aad} = 0$

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

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

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

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

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

$\Longrightarrow$ 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)*

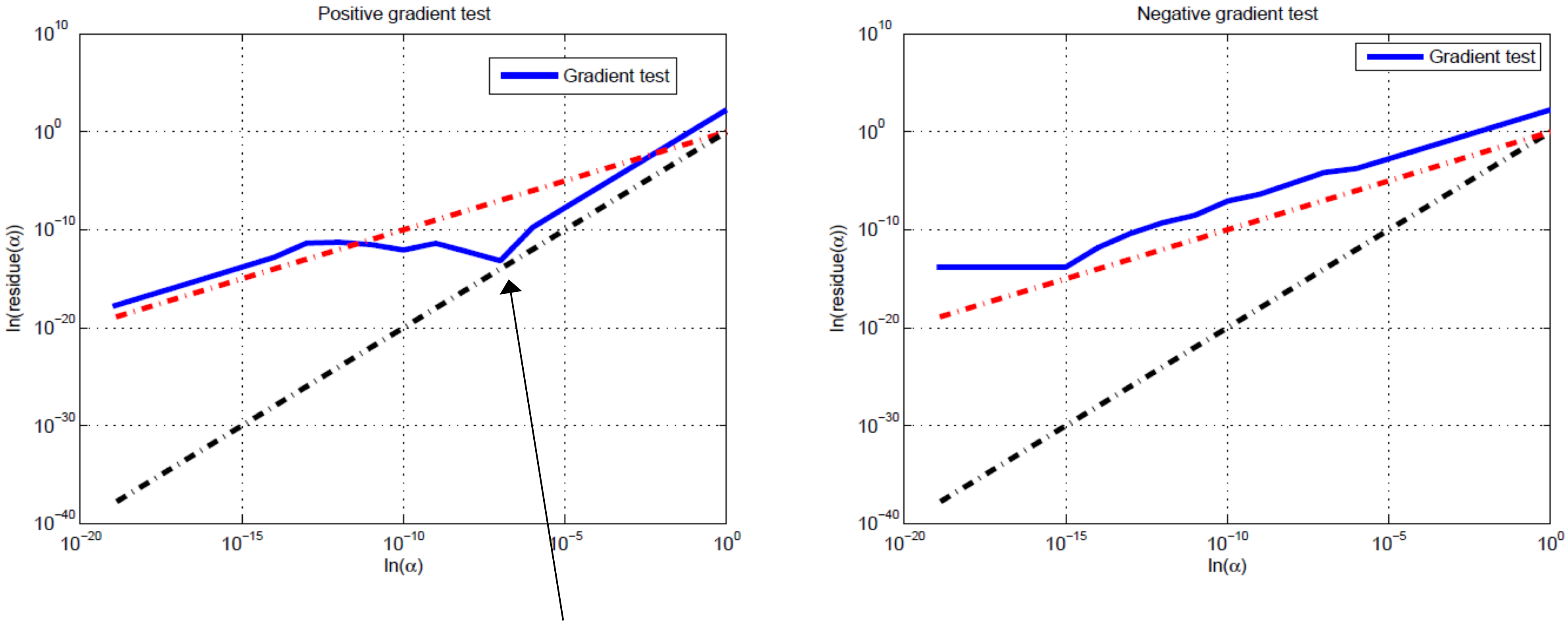
Adjoint compilers

TAPENADE (Laurent Hascoet, 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

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).

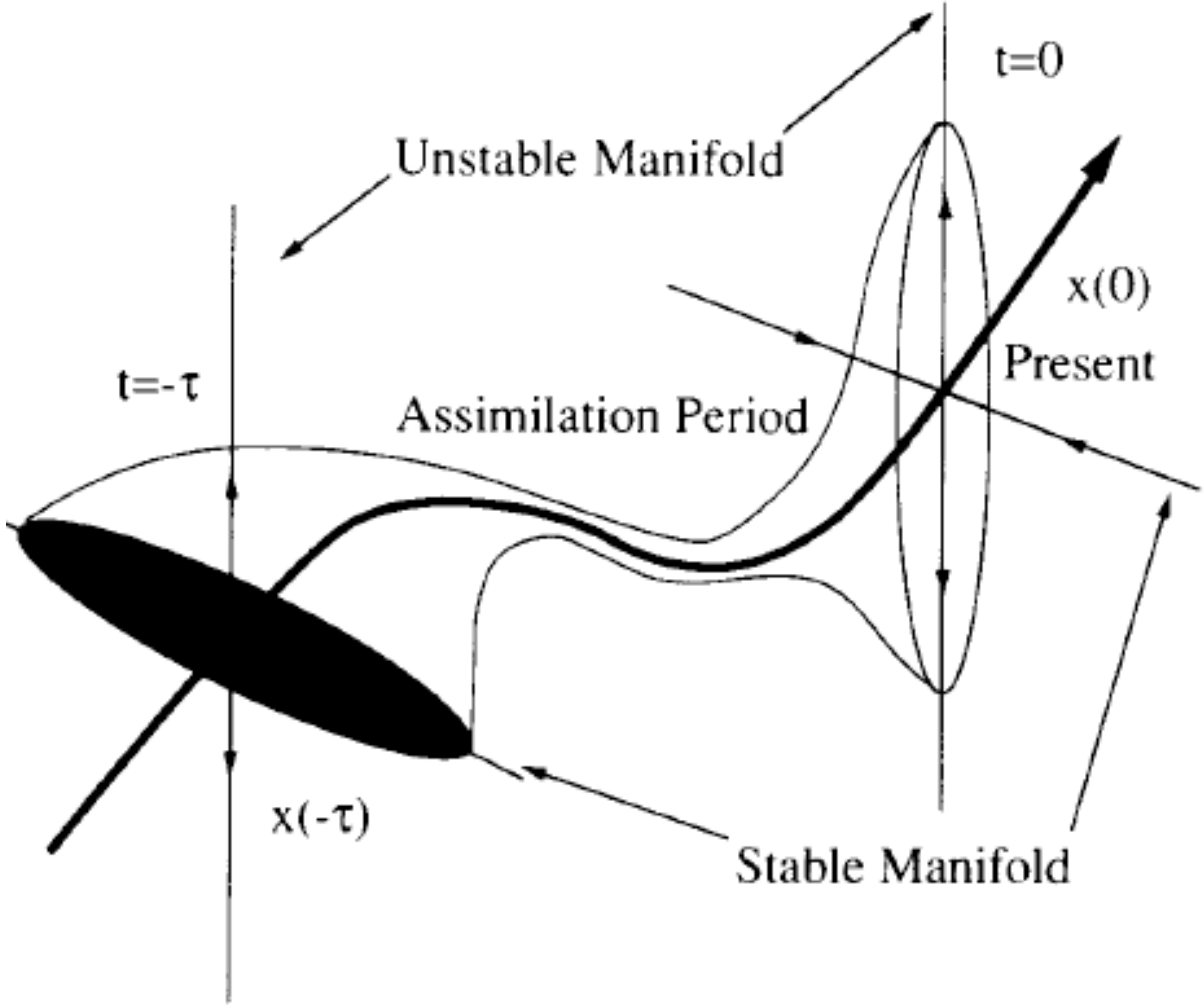
Buehner *et al*. (*Mon. Wea. Rev*., 2010)

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

- Assimilation and (In)stability

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.

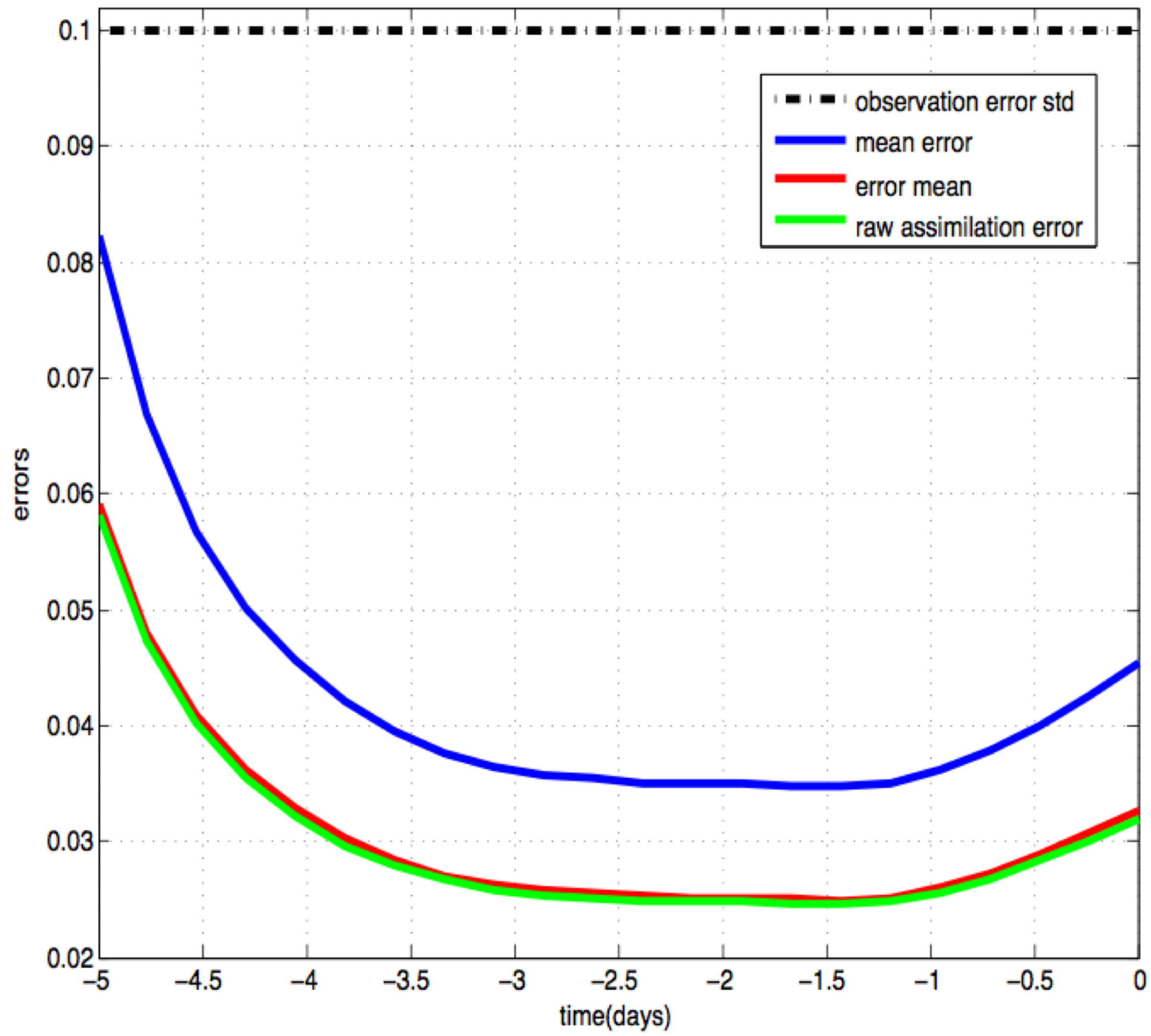
Stable (unstable) modes : perturbations to the basic state that decrease (increase) over time.


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.

Gurumoorthy *et al*. (2017*a*, 2017*b*) have shown that in the linear perfect model case, the error covariance matrix of the Kalman filter converges to the neutral-unstable subspace of the system (space spanned by the non-negative Lyapunov exponents of the system)



Linearized Lorenz’96. 5 days

Jardak and Talagrand

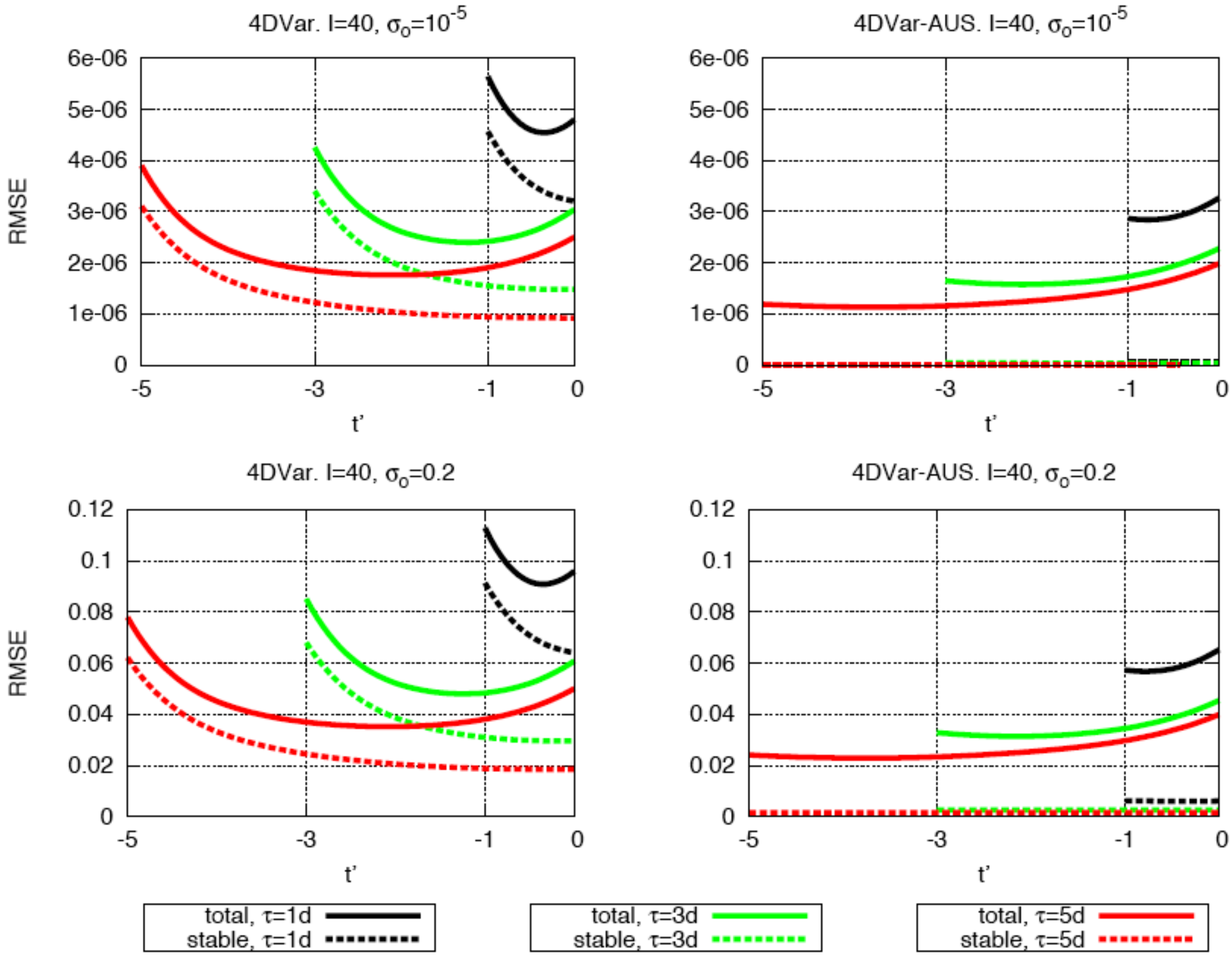


Figure 3. Time average RMS error within 1, 3, 5 days assimilation windows as a function of $t' = t - \tau$, with $\sigma_o = .2, 10^{-5}$ for the model configuration $I = 40$. Left panel: 4DVar. Right panel: 4DVar-AUS with $N = 15$. Solid lines refer to total assimilation error, dashed lines refer to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

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

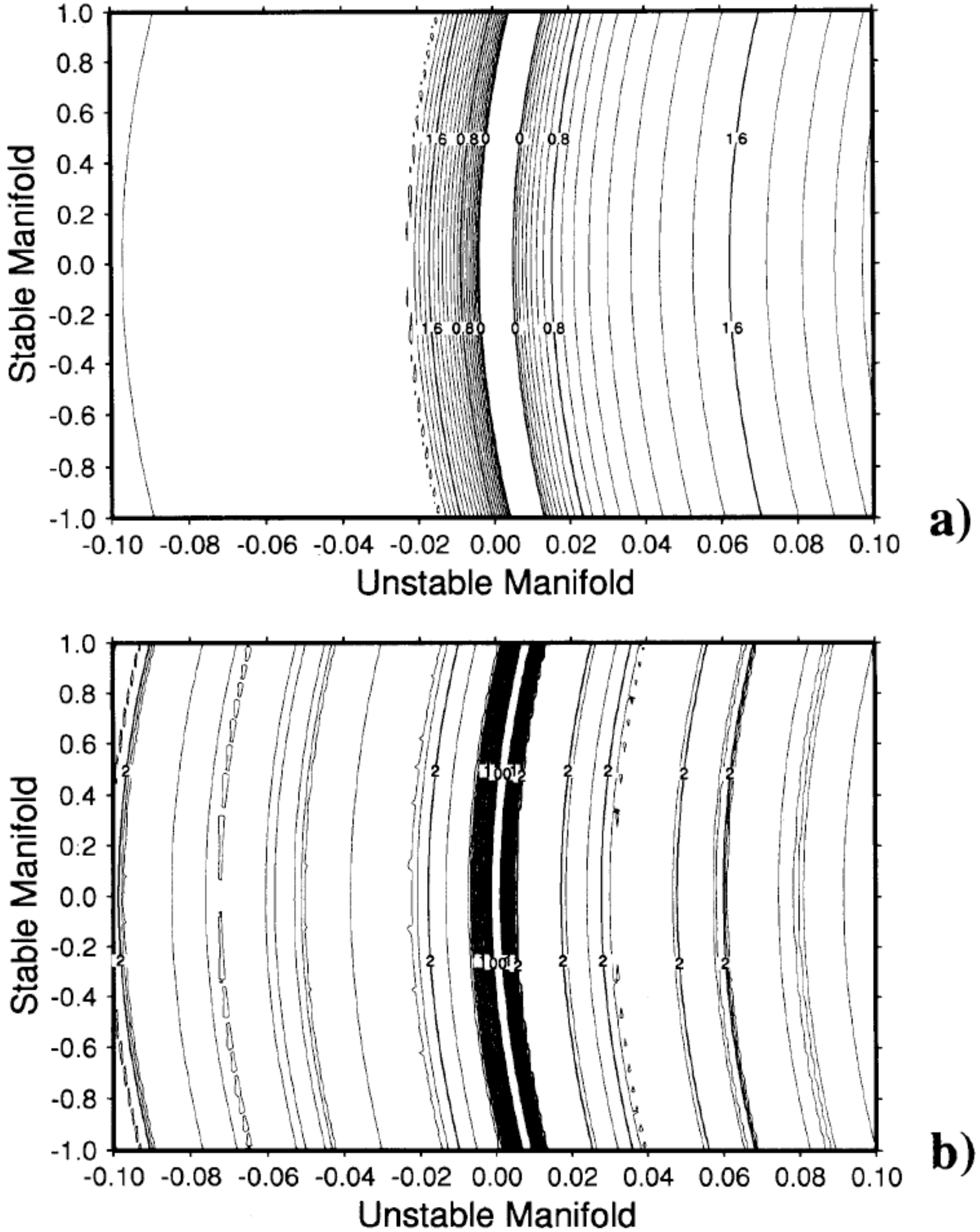


Fig. 3. Variations of the error-free forward cost-function $J'_e(\tau, \hat{x}, x)$ (Lorenz system) in the plane spanned by the stable and unstable directions, as determined from the tangent linear system (see text), and for $\tau = 6$ (panel (a)) and $\tau = 8$ (panel (b)) respectively. The metric has 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.

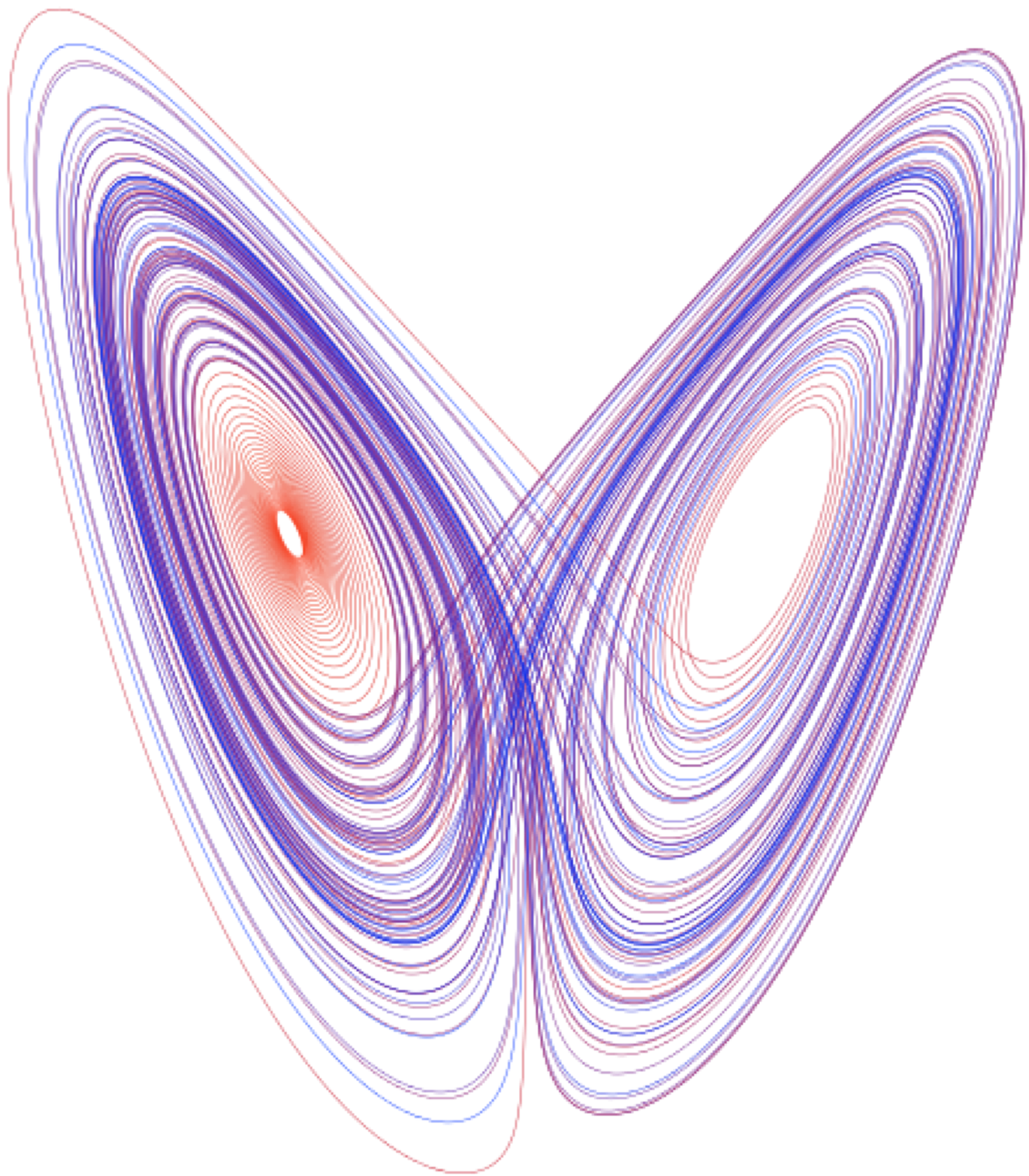
Lorenz (1963)

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

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

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

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



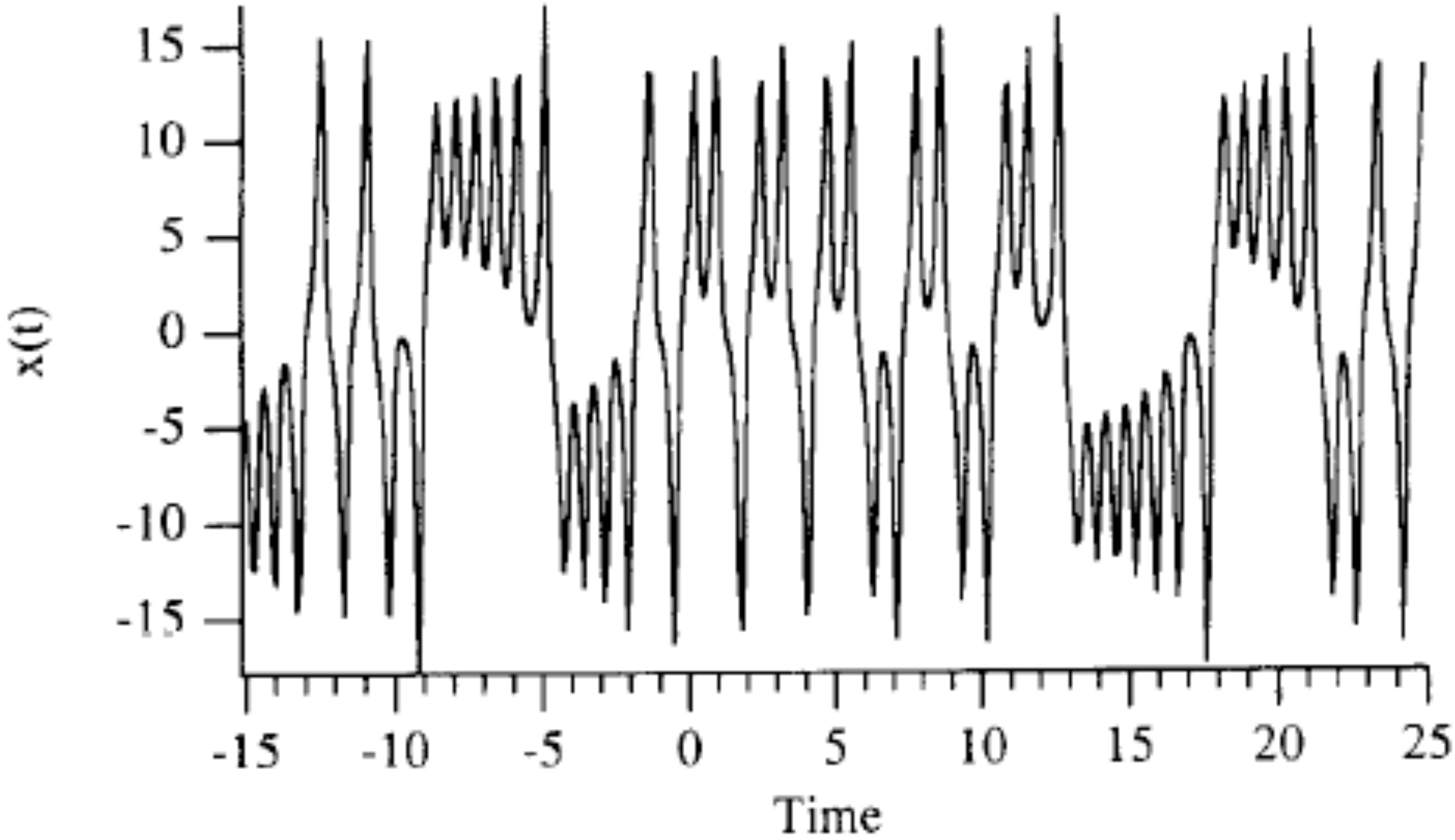


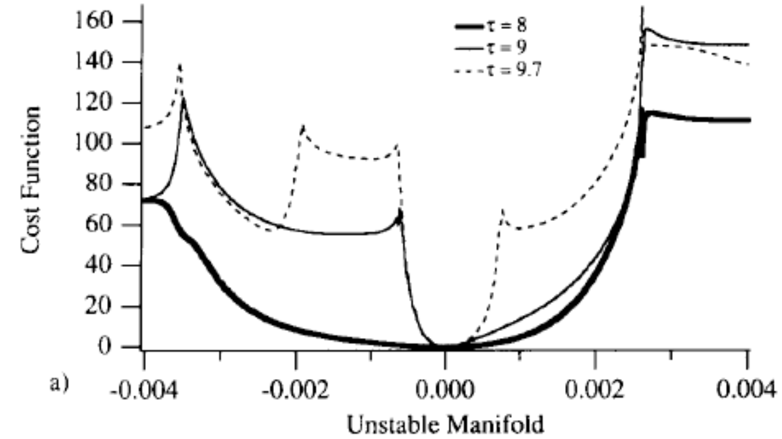
Fig. 2. Time variations, along the reference solution, of the variable $x(t)$ of the Lorenz system.

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 $k = 0$.



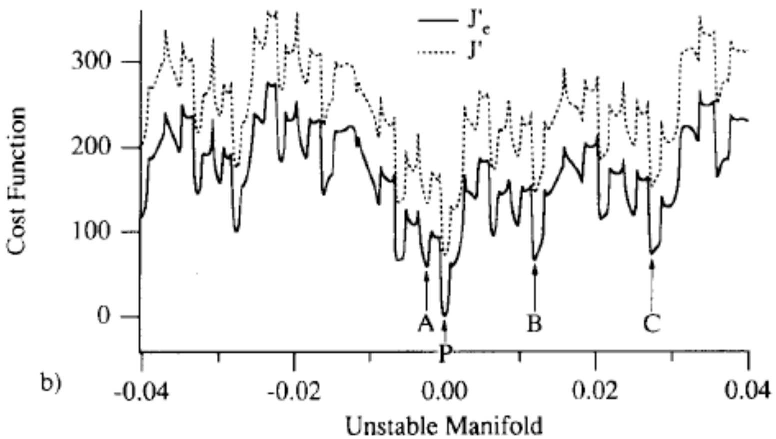


Fig. 4. Panel (a): Cross-section of the error-free forward cost-function $J_e'(\tau, \hat{x}, x)$ along the unstable manifold, for various values of τ. Panel (b). As in panel (a), for $\tau = 9.7$, and with a display interval ten times as large, respectively for the error-free forward cost-function $J_e'(\tau, \hat{x}, x)$ (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)

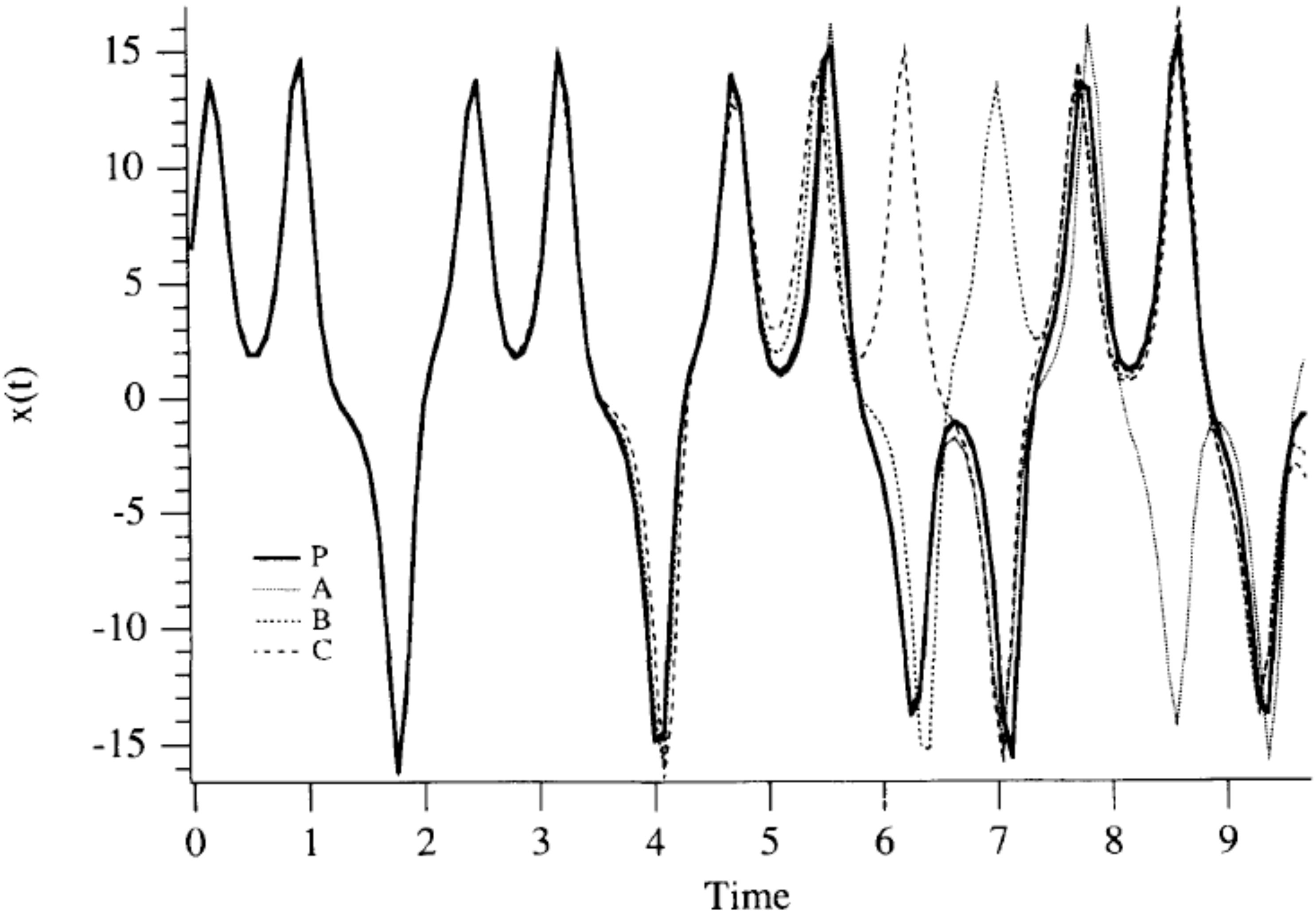


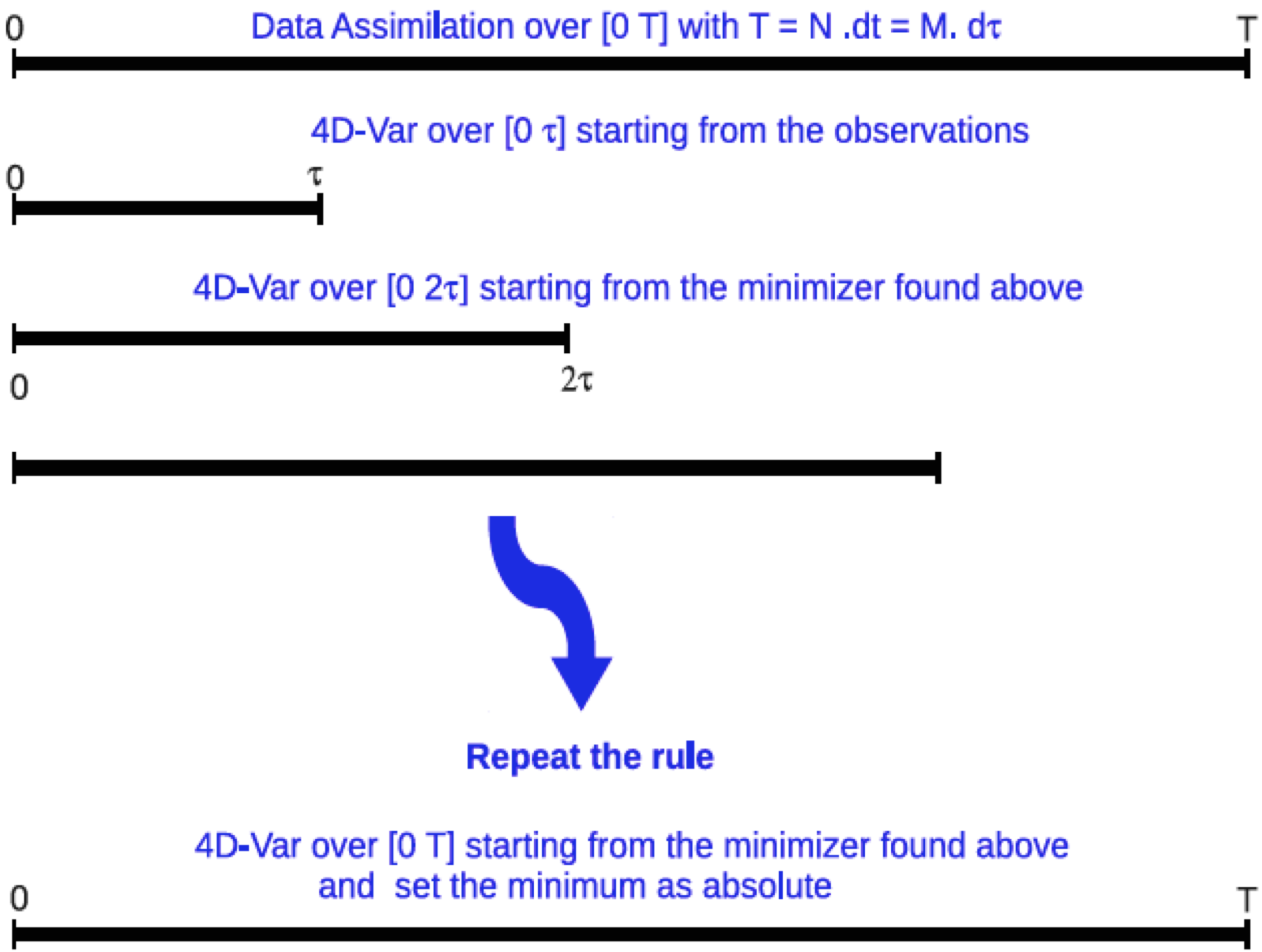
Fig. 5. Variations of the coordinate x along the orbits originating from the minima *P*, *A*, *B*, *C* (indicated in Fig. 4b) of the error-free cost-function.

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)



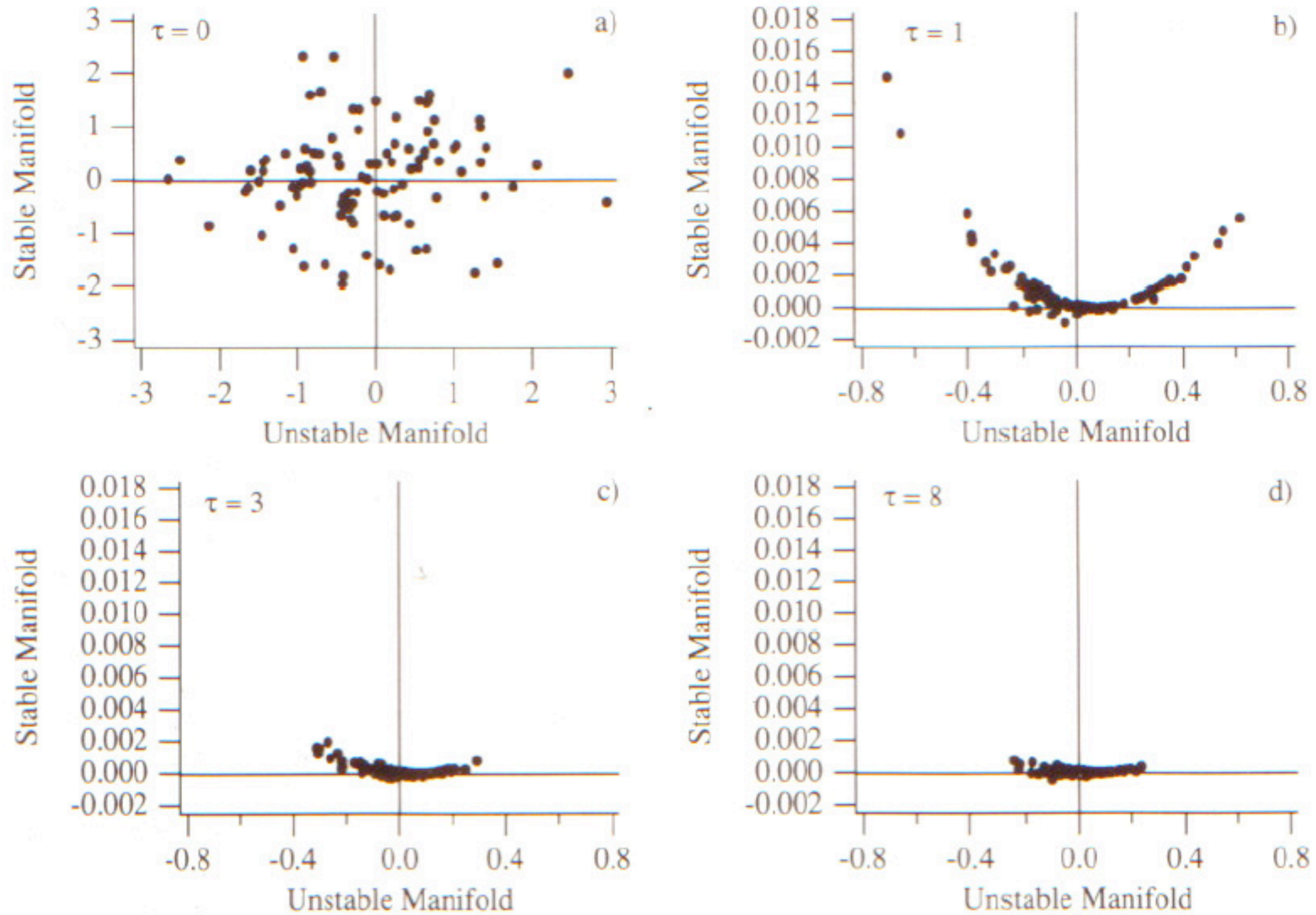


Fig. 7. Projection of the 100 minimizing solutions, at the end of the assimilation period, onto the plane spanned by the stable and unstable directions, defined as in Fig. 3. Values of τ are indicated on the panels. The projection is not an orthogonal projection, but a projection parallel to the local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

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

$\mu(C(\tau, x))$	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).

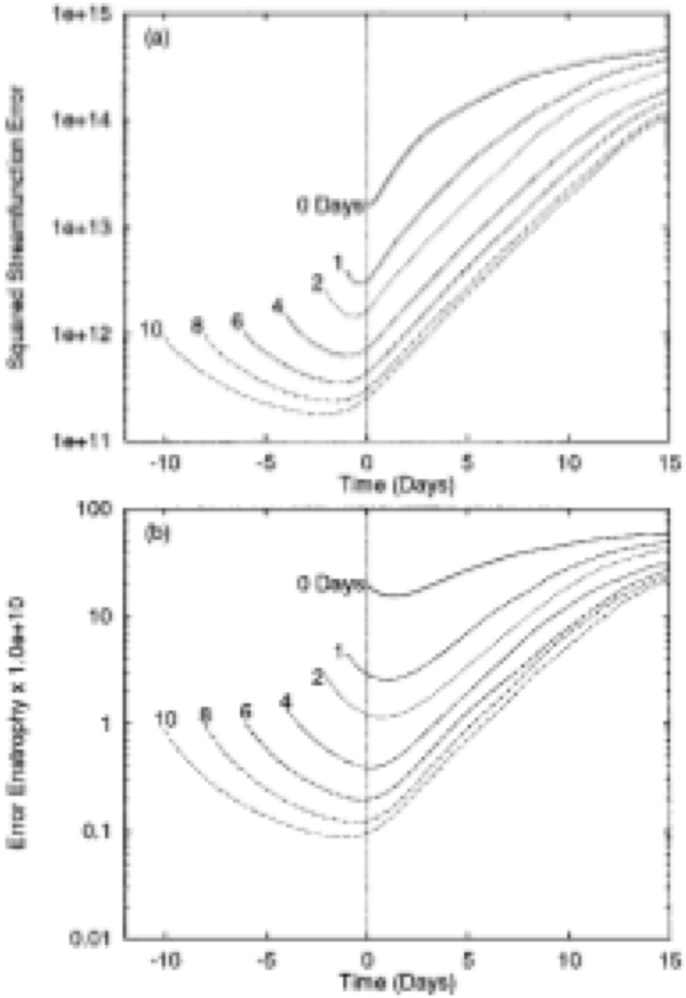
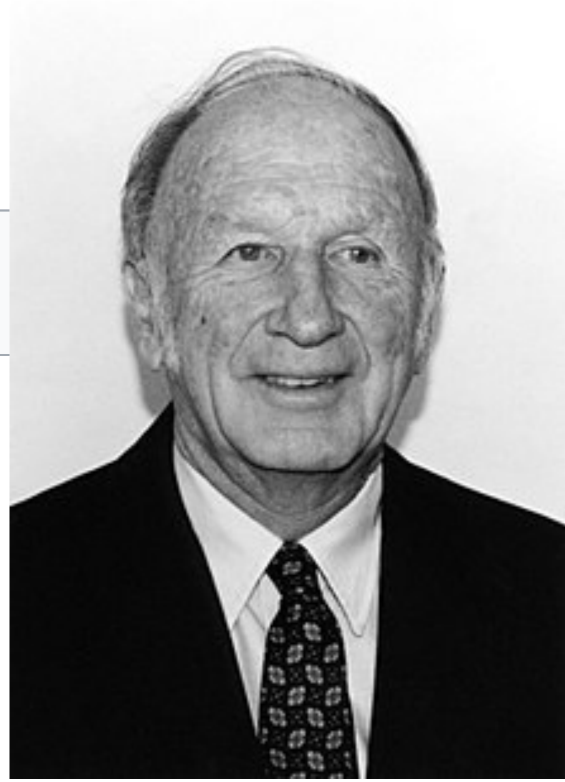


Fig. 5. Median values of the (a) streamfunction squared error, and (b) enstrophy error for the 200 forecast set as a function of forecast time and of the assimilation time T_a.

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

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

~~Mardi 21 mars~~
~~Mardi 28 mars~~
~~Mardi 4 avril~~
~~Mardi 11 avril~~
~~Mardi 2 mai~~
~~Mardi 9 mai~~
Mardi 23 mai
Mardi 30 mai