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 7

11 Juin 2025

Last course (May 28)

- 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 (introduction)

This course

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

- Variational assimilation in the unstable subspace

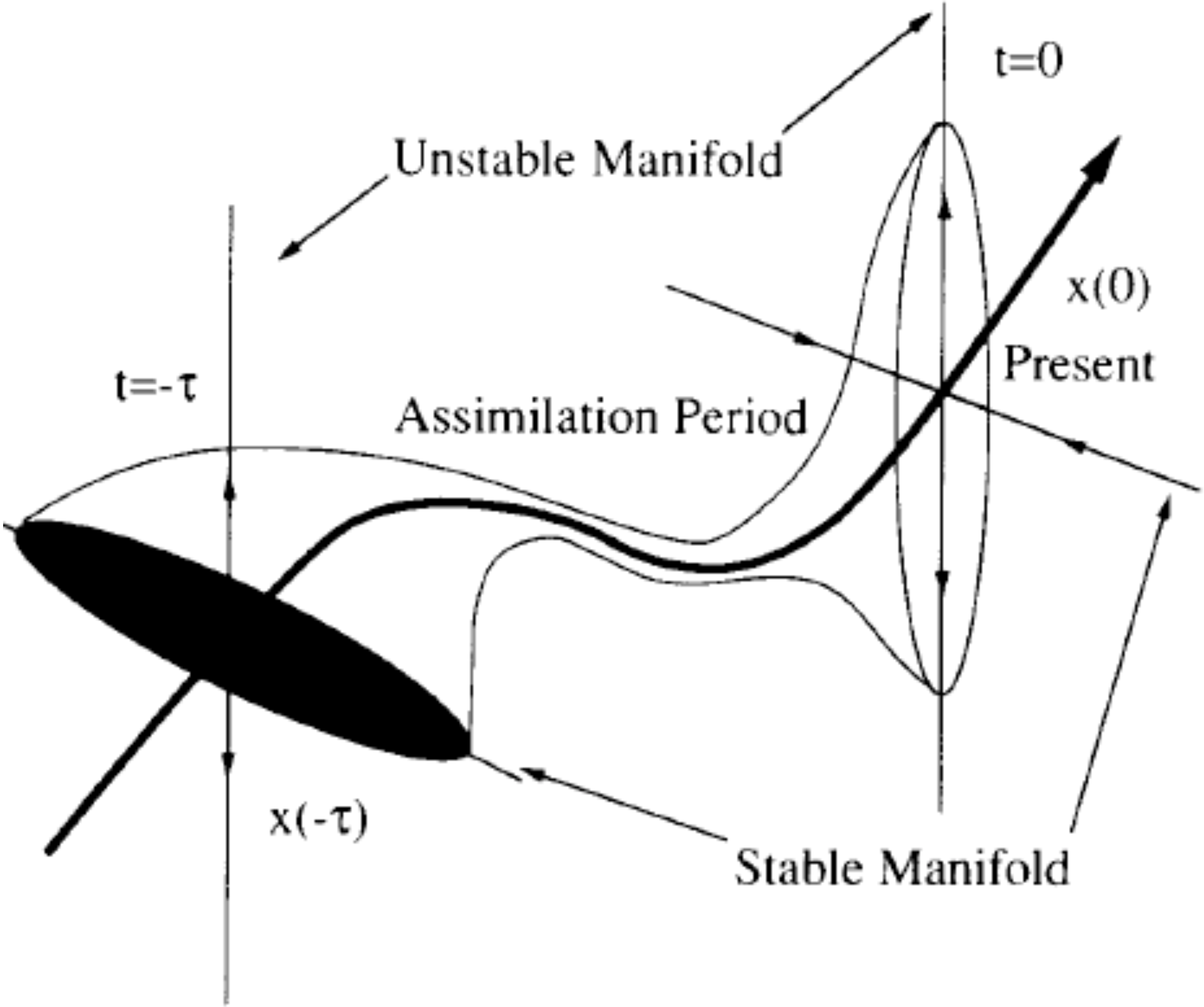
- Basics on dynamical systems

- Brief history of Numerical Weather Prediction

- 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.

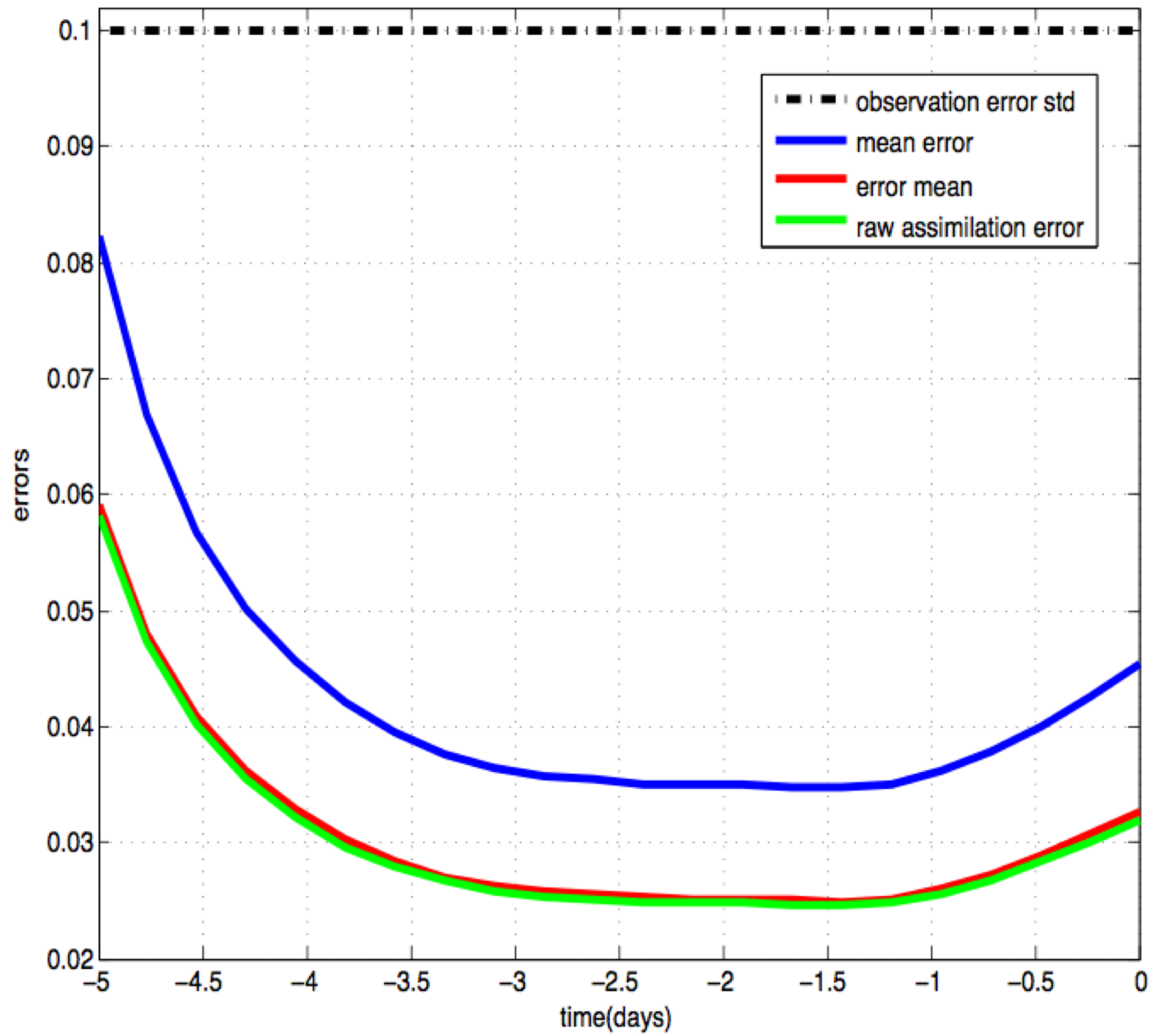


t=0
Unstable Manifold
x(0)
Present
t=-τ
Assimilation Period
x(-τ)
Stable Manifold

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

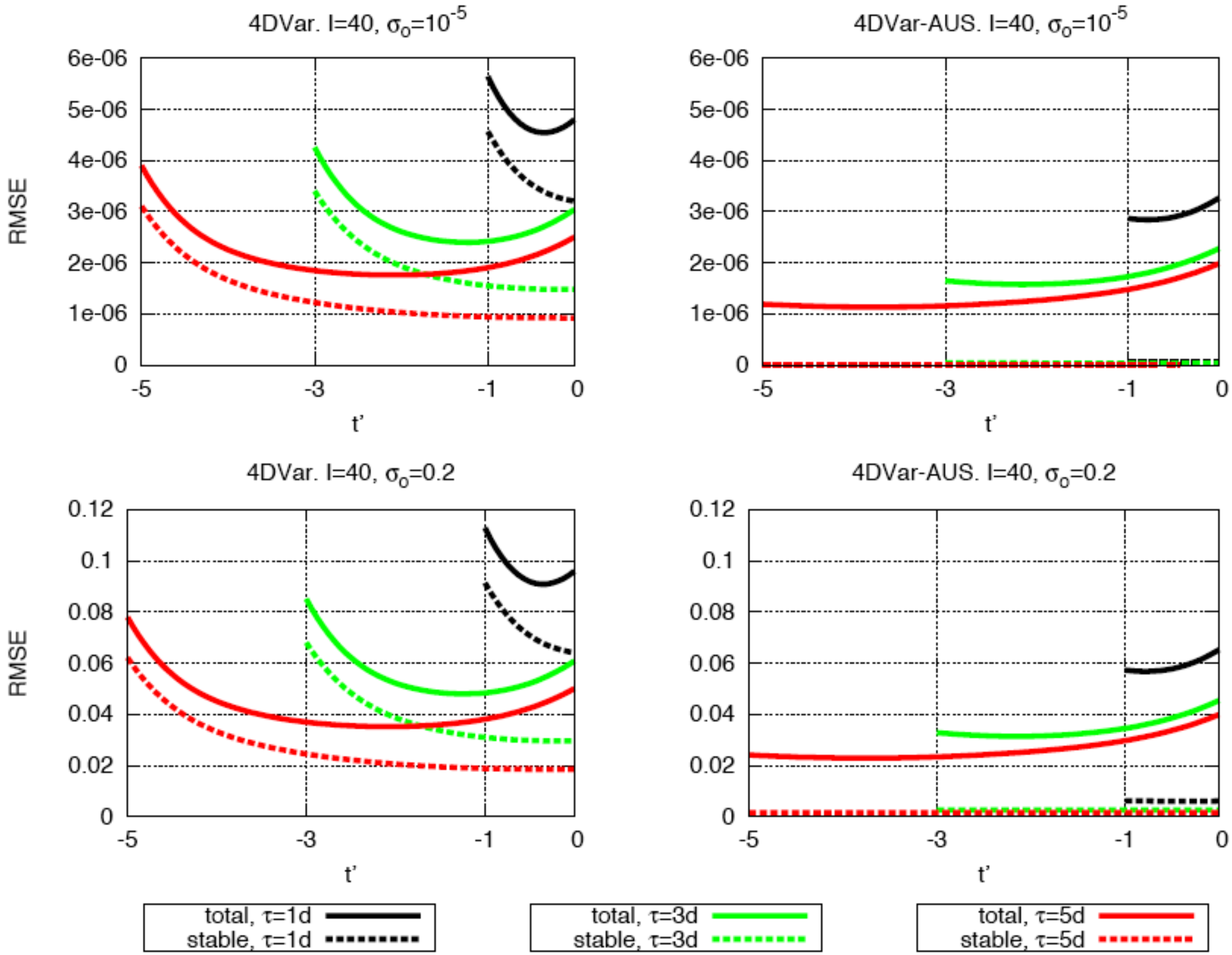


Figure 3. Time average RMS error within 1, 3, 5 days assimilation windows as a function of $t' = t - \tau$, with $\sigma_o = .2, 10^{-5}$ for the model configuration $I = 40$. Left panel: 4DVar. Right panel: 4DVar-AUS with $N = 15$. Solid lines refer to total assimilation error, dashed lines refer to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

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

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

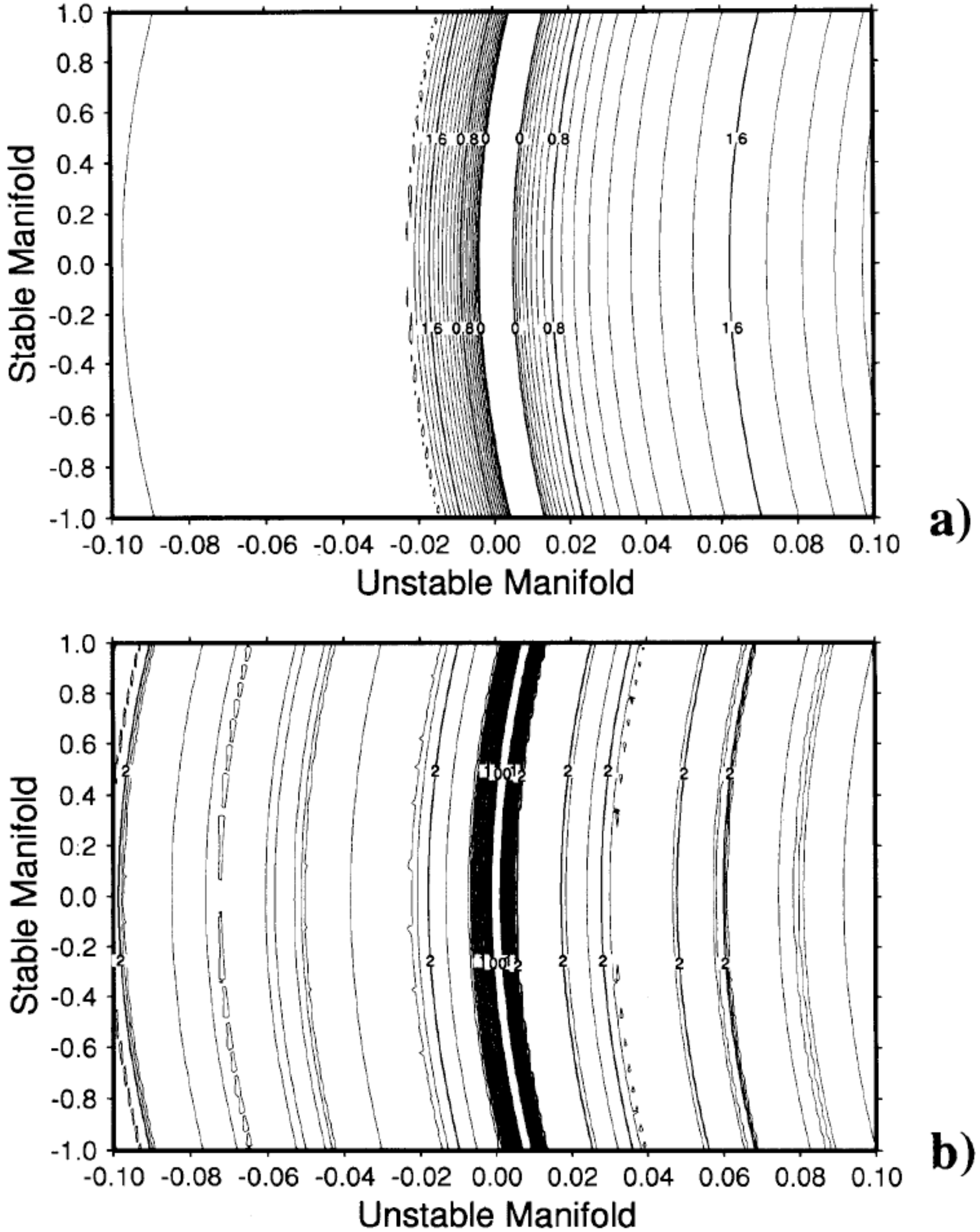
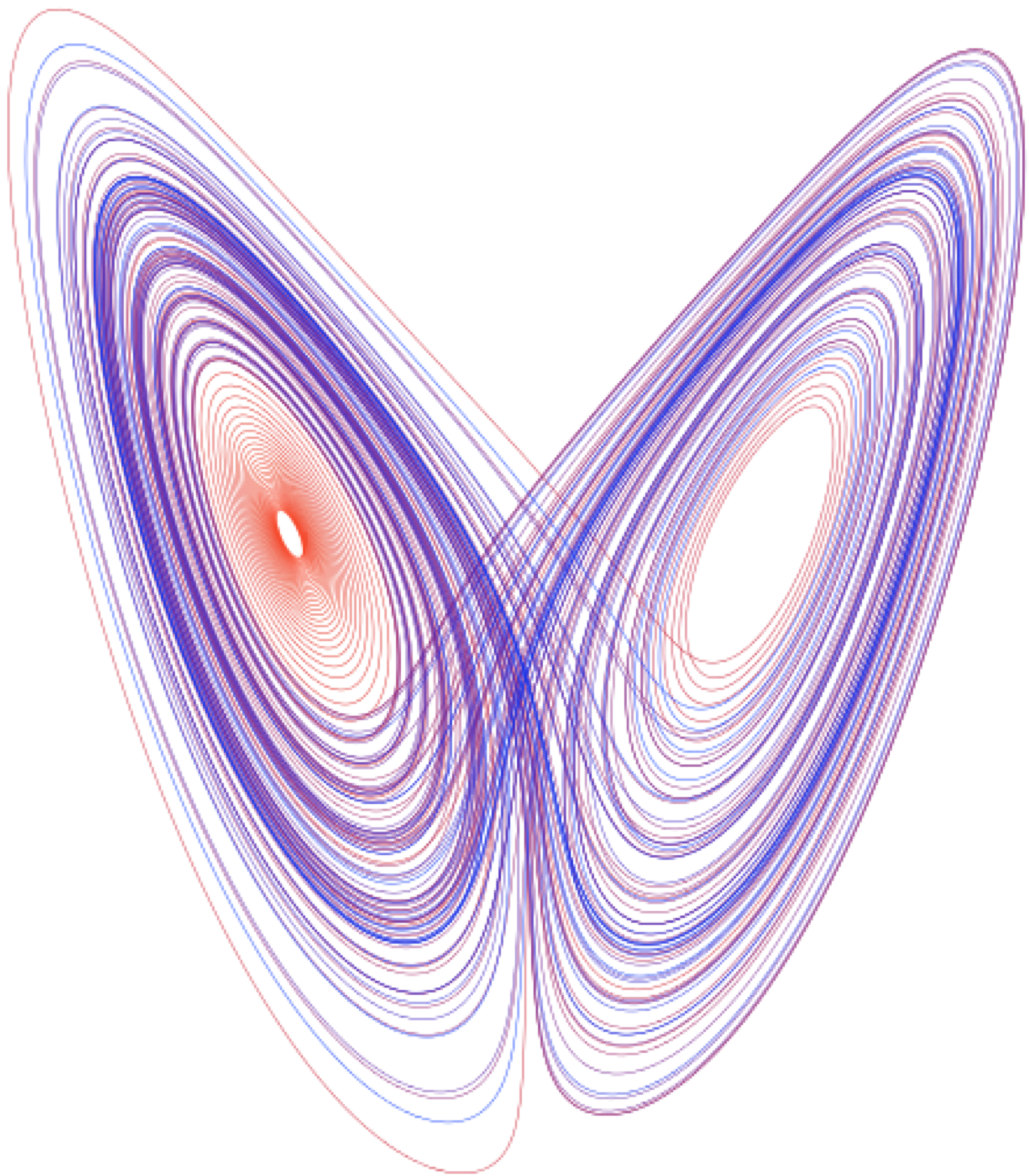


Fig. 3. Variations of the error-free forward cost-function $J'_e(\tau, \hat{x}, x)$ (Lorenz system) in the plane spanned by the stable and unstable directions, as determined from the tangent linear system (see text), and for $\tau = 6$ (panel (a)) and $\tau = 8$ (panel (b)) respectively. The metric has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



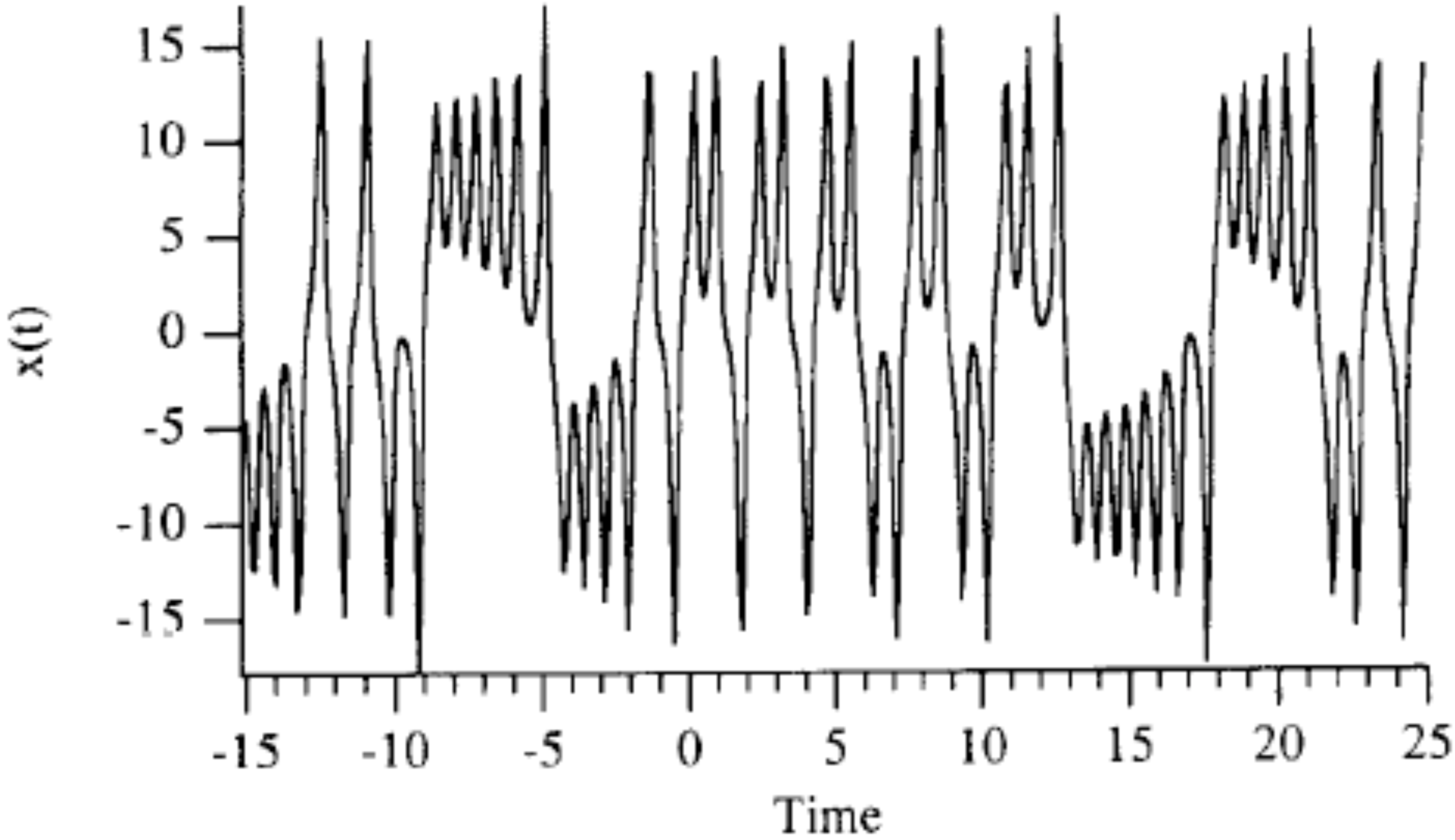


Fig. 2. Time variations, along the reference solution, of the variable $x(t)$ of the Lorenz system.

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 $k = 0$.

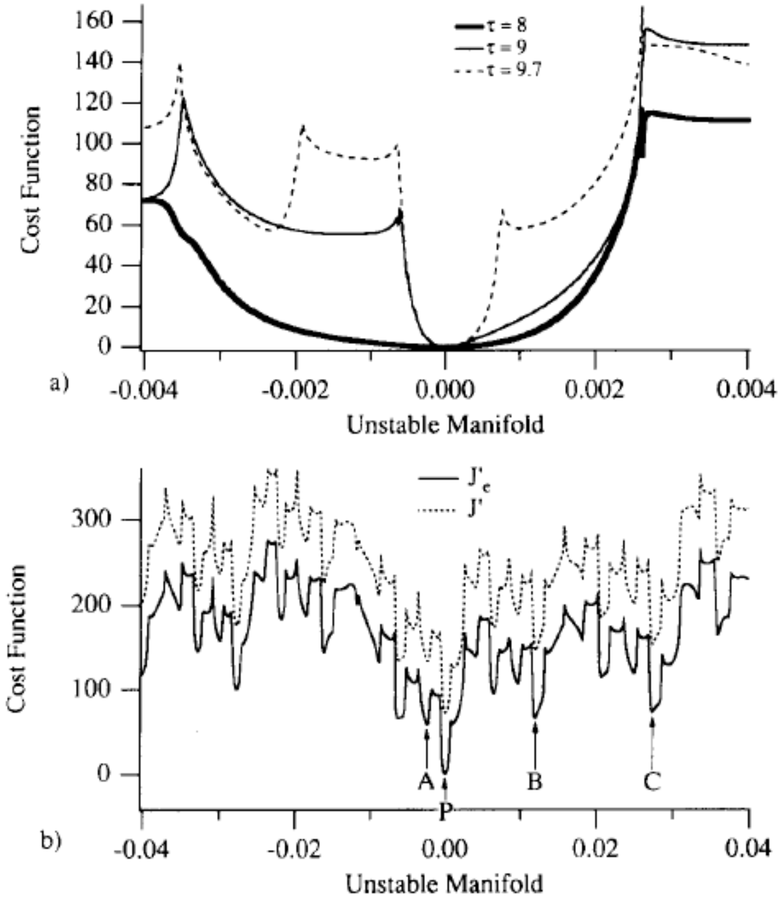


Fig. 4. Panel (a): Cross-section of the error-free forward cost-function $J'_e(\tau, \hat{x}, x)$ along the unstable manifold, for various values of τ. Panel (b). As in panel (a), for $\tau = 9.7$, and with a display interval ten times as large, respectively for the error-free forward cost-function $J'_e(\tau,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

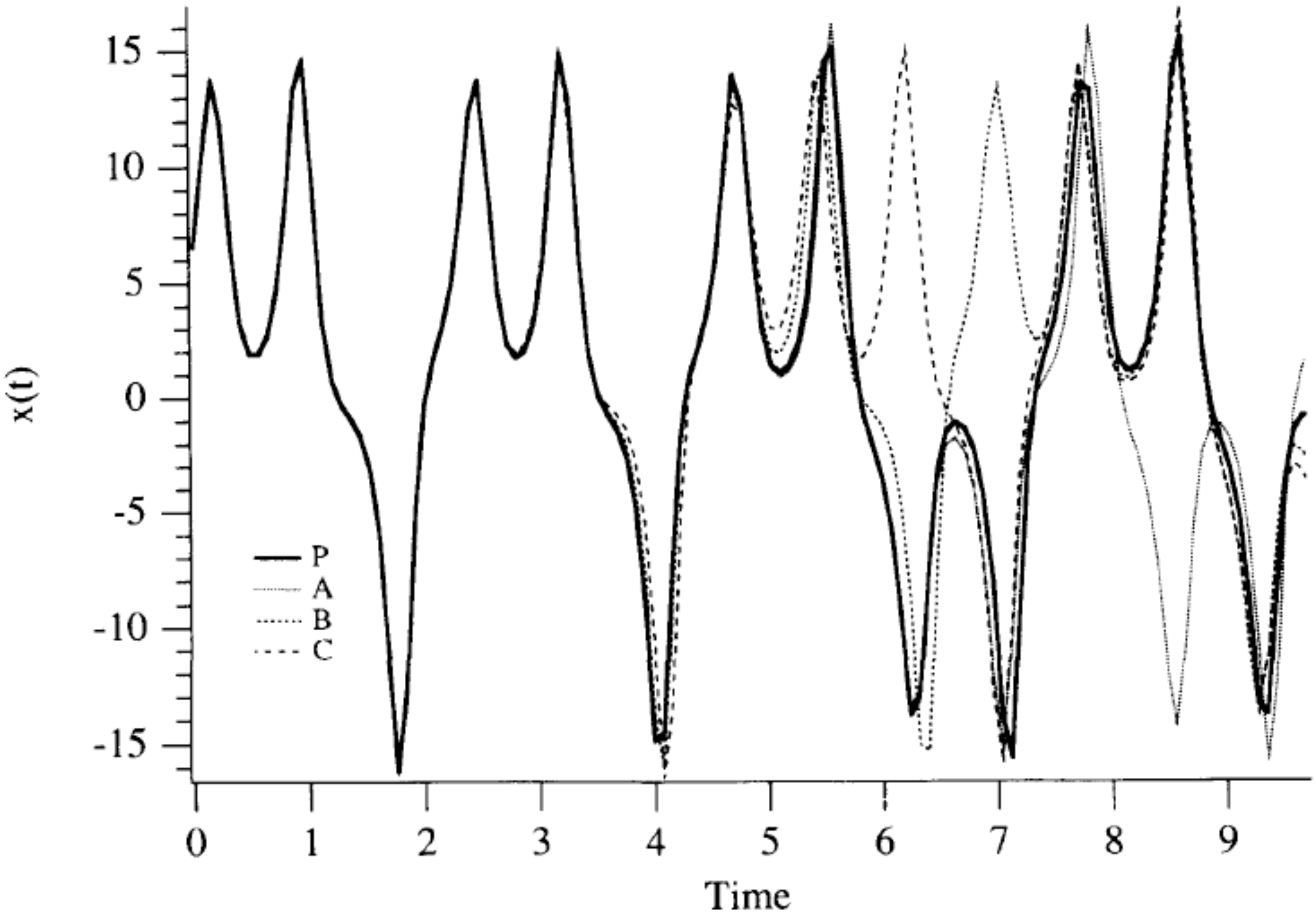


Fig. 5. Variations of the coordinate x along the orbits originating from the minima *P*, *A*, *B*, *C* (indicated in Fig. 4b) of the error-free cost-function.

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

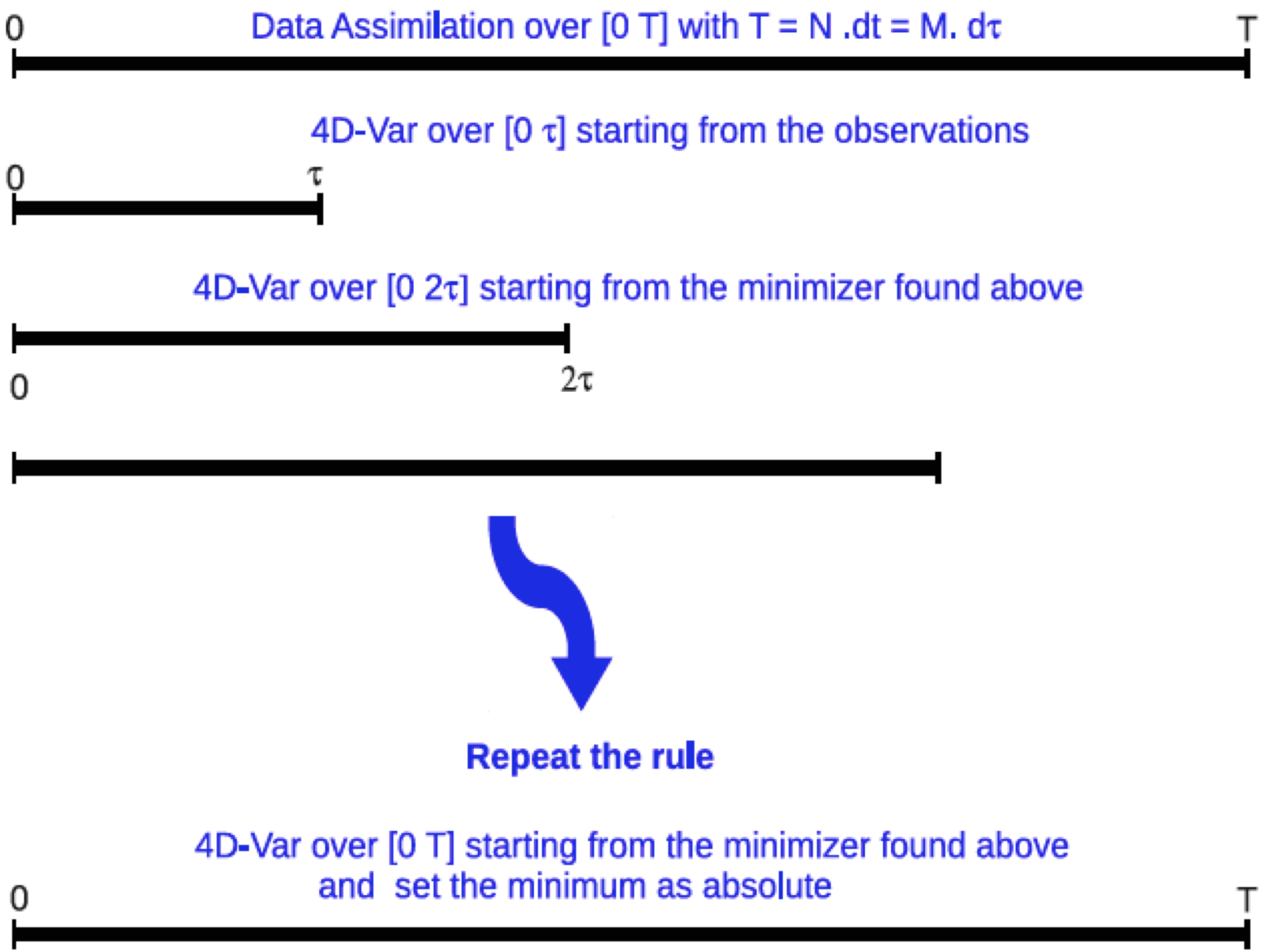
$\mu(C(\tau, x))$	Cloud of points QSVA	Cloud of points raw assimilation	Linear tangent system	Upper bound
$\tau = 0$	1	1	1	1
$\tau = 1$	0.36	0.37	0.39	0.46
$\tau = 2$	5.9×10^{-2}	5.74	4.5×10^{-2}	0.401
$\tau = 3$	3.3×10^{-2}	29.4	2.9×10^{-2}	0.397
$\tau = 8$	1.4×10^{-2}	59.9	*	0.396

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).

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)



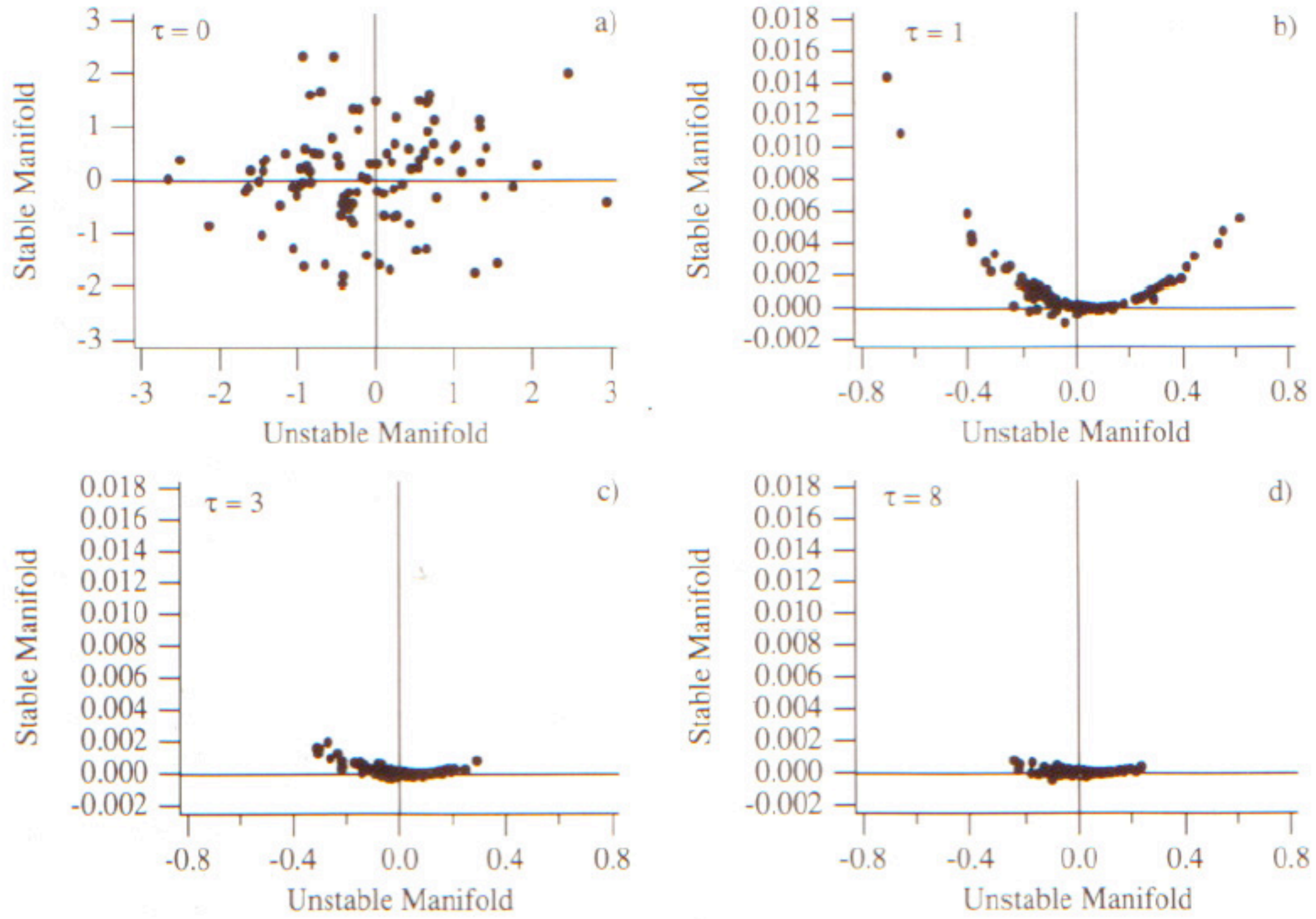


Fig. 7. Projection of the 100 minimizing solutions, at the end of the assimilation period, onto the plane spanned by the stable and unstable directions, defined as in Fig. 3. Values of τ are indicated on the panels. The projection is not an orthogonal projection, but a projection parallel to the local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

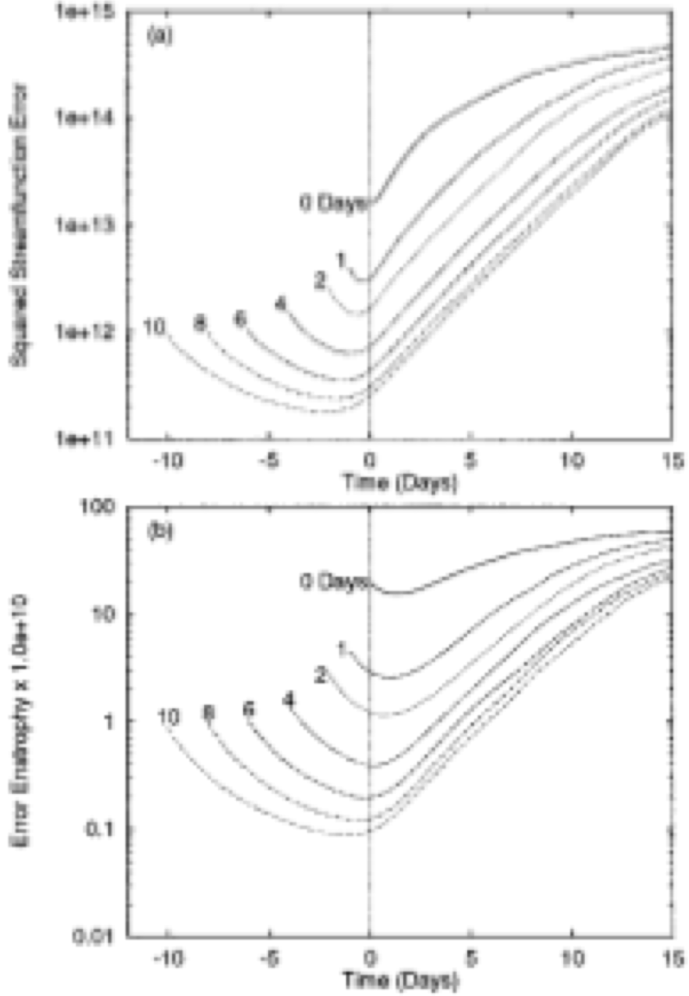


Fig. 5. Median values of the (a) streamfunction squared error, and (b) enstrophy error for the 200 forecast set as a function of forecast time and of the assimilation time T_a.

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

Since, after an assimilation has been performed over a period of time, uncertainty is likely to be concentrated in modes that have been unstable, it might be useful for the next assimilation, and at least in terms of cost efficiency, to concentrate corrections on the background in those modes.

Actually, presence of residual noise in stable modes can be damageable for analysis and subsequent forecast.

Assimilation in the Unstable Subspace (AUS) (Carrassi *et al*., 2007, 2008, for the case of 3D-Var)

Four-dimensional variational assimilation in the unstable subspace (*4DVar-AUS*)

Trevisan *et al*., 2010, Four-dimensional variational assimilation in the unstable subspace and the optimal subspace dimension, *Q. J. R. Meteorol. Soc.*, **136,** 487-496.

Experiments performed on the Lorenz (1996) model

$$\frac{d}{dt}x_j = (x_{j+1} - x_{j-2})x_{j-1} - x_j + F$$

with $j = 1, ..., I.$

with periodic conditions in j, and value $F = 8$, which gives rise to chaos.

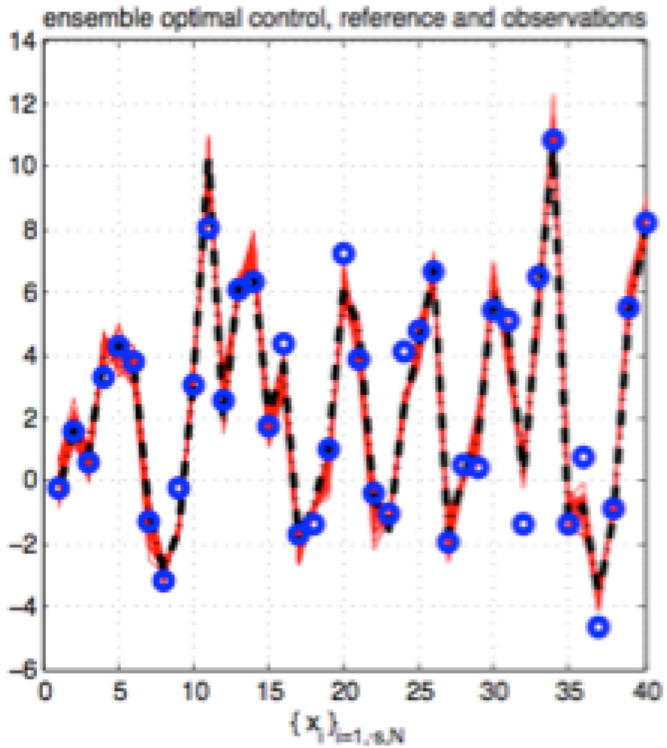
Three values of I have been used, namely $I = 40, 60, 80$, which correspond to respectively $N^+ = 13, 19$ and 26 positive Lyapunov exponents.

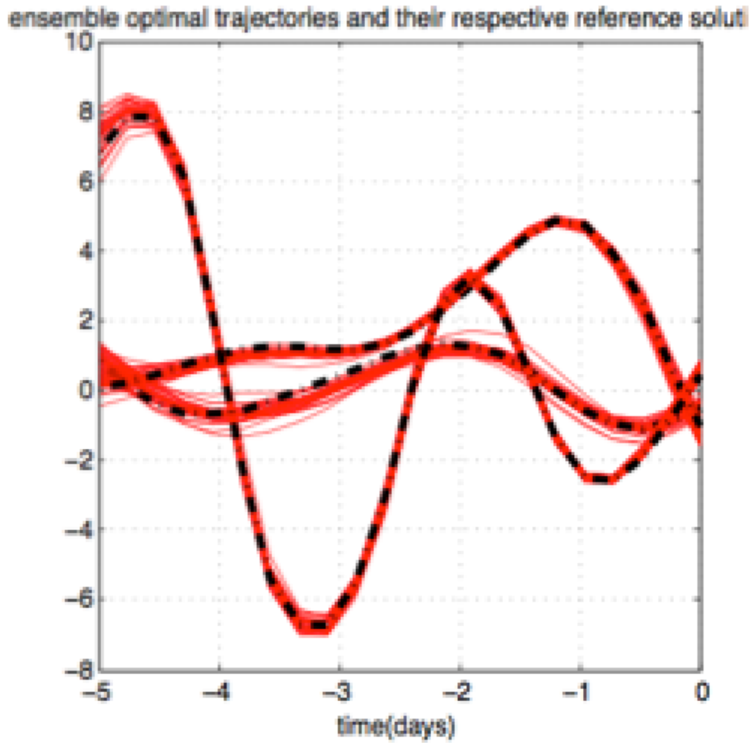
In all three cases, the largest Lyapunov exponent corresponds to a doubling time of about 2 days (with 1 'day' = 1/5 model time unit).

Identical twin experiments (perfect model)

System produces wavelike chaotic motions, with properties similar to those of midlatitude atmospheric waves

- generally westward phase velocity
- typical predictability time : 5 ‘days’
- in addition, quadratic terms conserve ‘energy’





4D-Var-AUS

Algorithmic implementation

Define N perturbations to the current state, and evolve them according to the tangent linear model, with periodic reorthonormalization in order to avoid collapse onto the dominant Lyapunov vector (same algorithm as for computation of Lyapunov exponents).

Cycle successive 4D-Var's, restricting at each cycle the modification to be made on the current state to the space spanned by the N perturbations emanating from the previous cycle (if N is the dimension of state space, that is identical with standard 4D-Var).

Observing system' defined as in Fertig *et al*. (*Tellus*, 2007):

At each observation time, one observation every four grid points (observation points shifted by one grid point at each observation time).

Observation frequency : 1.5 hour

Random gaussian observation errors with expectation 0 and standard deviation $\sigma_0 = 0.2$ ('climatological' standard deviation 5.1).

Sequences of variational assimilations have been cycled over windows with length $\tau = 1, \ldots , 5$ days. Results are averaged over 5000 successive windows.

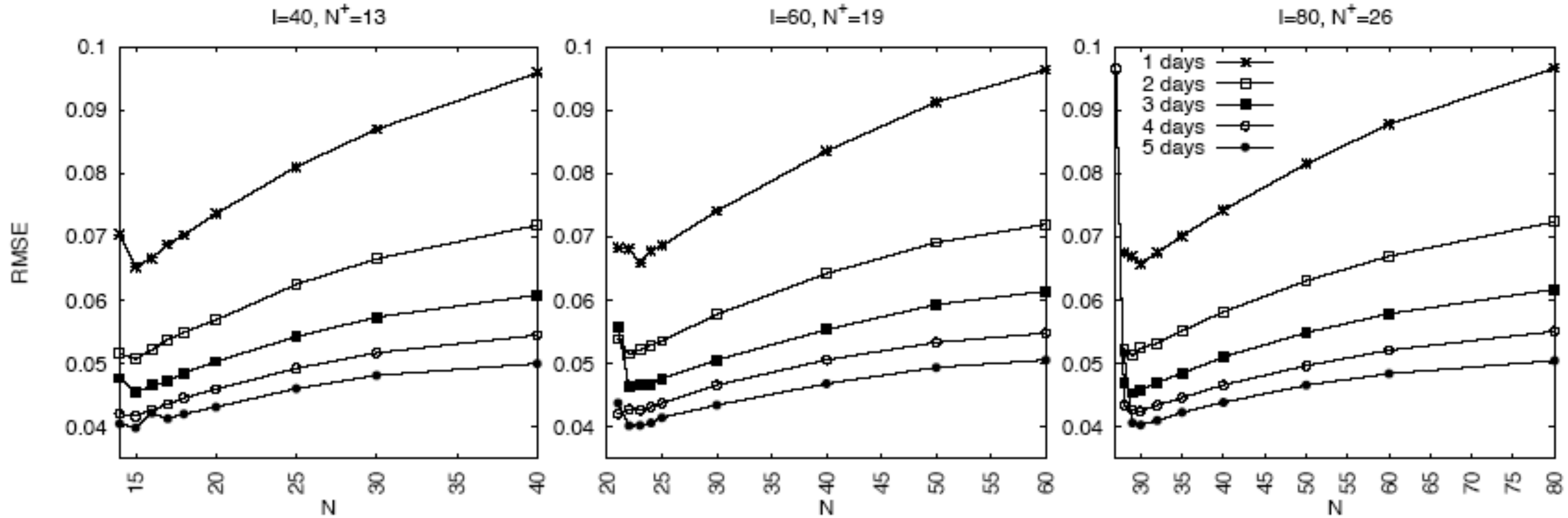


Figure 1. Time average RMS analysis error at $t = \tau$ as a function of the subspace dimension N for three model configurations: I=40, 60, 80. Different curves in the same panel refer to different assimilation windows from 1 to 5 days. The observation error standard deviation is $\sigma_o = 0.2$.

No explicit background term *(i. e.,* with error covariance matrix) in objective function : information from past lies in the background to be updated, and in the N perturbations which define the subspace in which updating is to be made.

Best performance for N slightly above number N^+ of positive Lyapunov exponents.

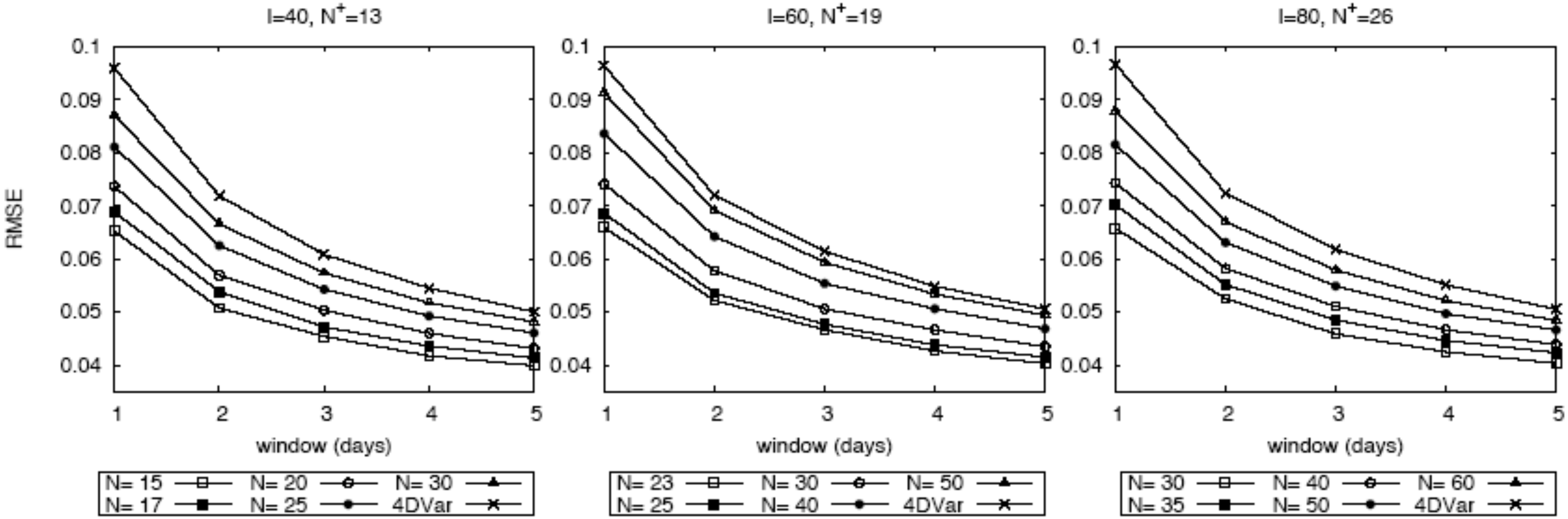


Figure 2. Time average RMS analysis error at $t = \tau$ as a function of the length of the assimilation window for three model configurations: I=$40, 60, 80$. Different curves in the same panel refer to a different subspace dimension N of 4DVar-AUS and to standard 4DVar. $\sigma_o = 0.2$.

Different curves are almost identical on all three panels. Relative improvement obtained by decreasing subspace dimension N to its optimal value is largest for smaller window length τ (for small window lengths, the system does not know which modes are stable or unstable).

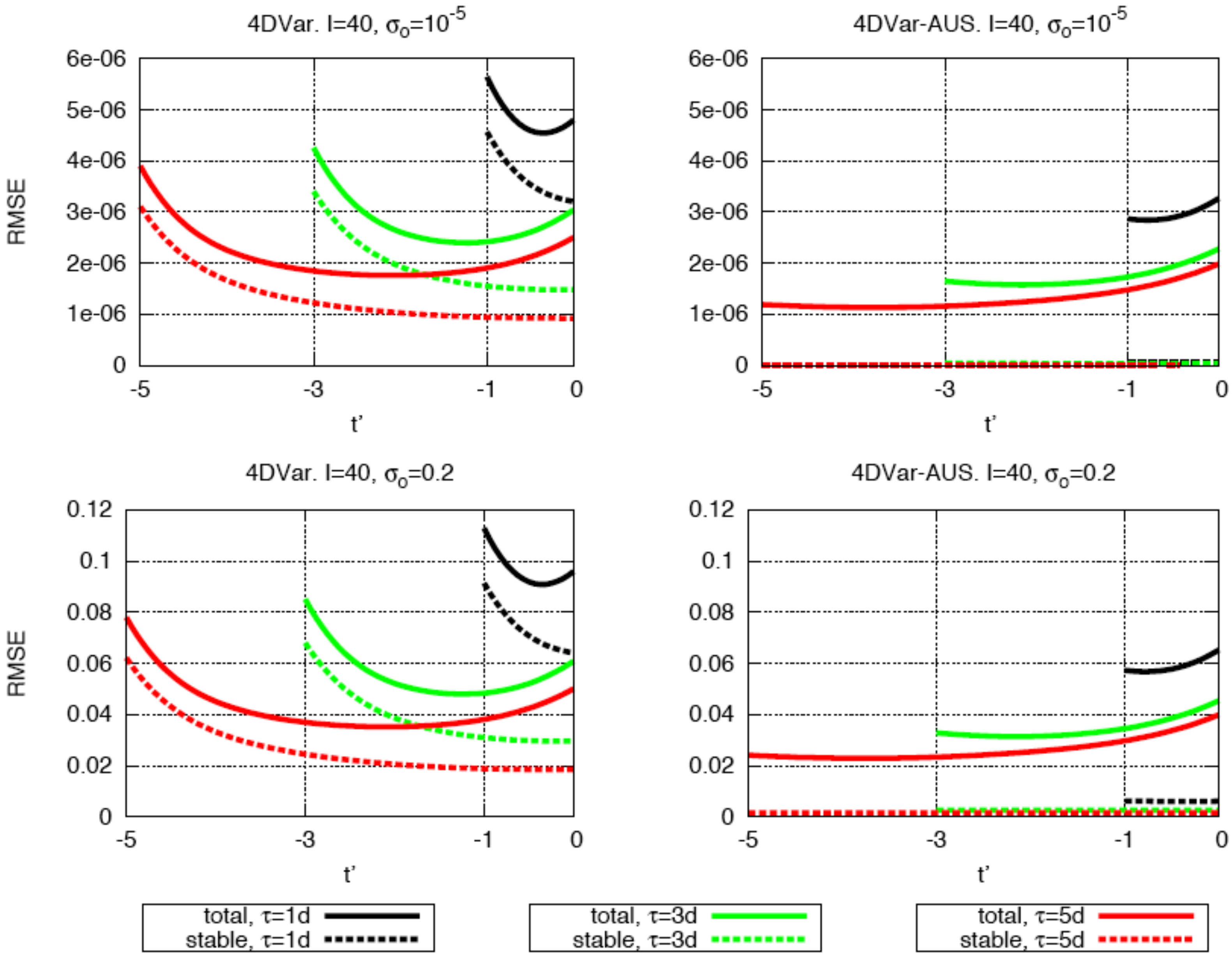


Figure 3. Time average RMS error within 1, 3, 5 days assimilation windows as a function of $t' = t - \tau$, with $\sigma_o = .2, 10^{-5}$ for the model configuration $I = 40$. Left panel: 4DVar. Right panel: 4DVar-AUS with $N = 15$. Solid lines refer to total assimilation error, dashed lines refer to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

Experiments have been performed in which an explicit background term was present, the associated error covariance matrix having been obtained as the average of a sequence of full 4D-Var's.

The estimates are systematically improved, and more for full 4D-Var than for 4D-Var-AUS. But they remain qualitatively similar, with best performance for 4D-Var-AUS with N slightly above N^+.

Minimum of objective function cannot be made smaller by reducing control space. Numerical tests show that minimum of objective function is smaller (by a few percent) for full 4D-Var than for 4D-Var-AUS. Full 4D-Var is closer to the noisy observations, but farther away from the truth. And tests also show that full 4D-Var performs best when observations are perfect (no noise).

Results show that, if all degrees of freedom that are available to the model are used, the minimization process introduces components along the stable modes of the system, in which no error is present, in order to ensure a closer fit to the observations. This degrades the closeness of the fit to reality. The optimal choice is to restrict the assimilation to the unstable modes.

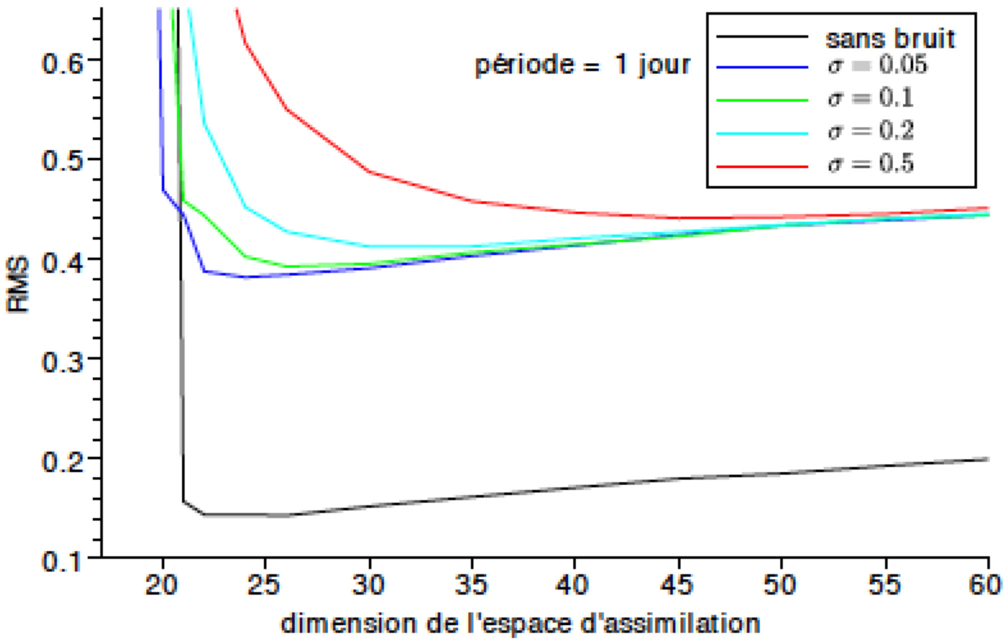
Results are improved when an explicit background is available at the initial time of the assimilation window. One can expect that a proper background (obtained for instance from a properly implemented Kalman Filter, or from an Ensemble Variational Assimilation) would not only say that the uncertainty is restricted to the unstable space, but how it is distributed in that subspace. The 'restriction' to the unstable subspace would be automatically made.

Can have major practical algorithmic implications.

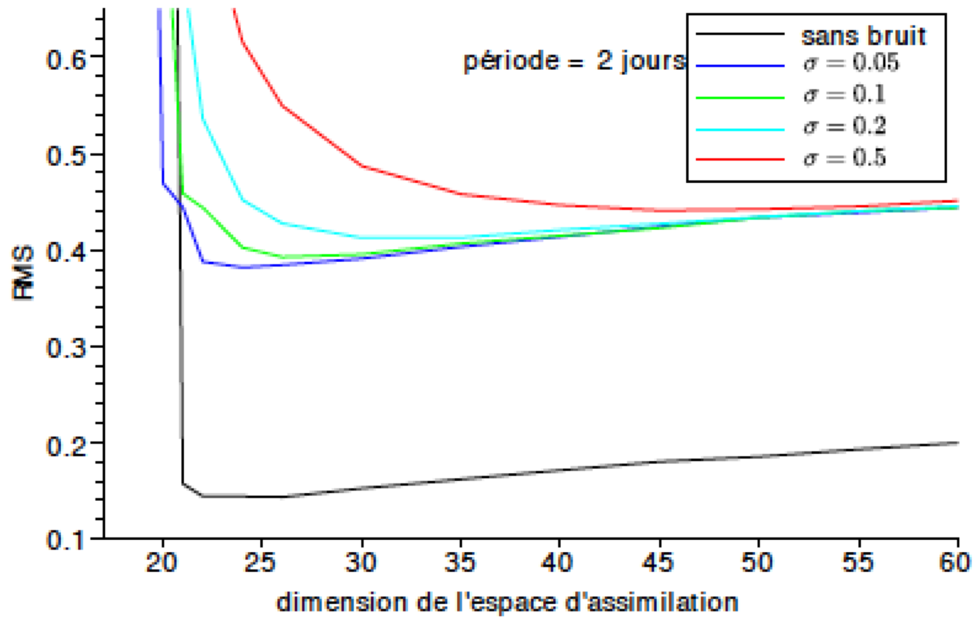
Questions.

- Degree of generality of results ?

- Impact of model errors ?



$\tau = 1$ day



$\tau = 2$ days

Time averaged rms analysis error at the end of the assimilation window (with length τ) as a function of increment subspace dimension ($I = 60$, $N^{+}=19$), for different amplitudes of white model noise.

(W. Ohayon and O. Pannekoucke, 2011).

Conclusions

Error concentrates in unstable modes at the end of assimilation window. It must therefore be sufficient, at the beginning of new assimilation cycle, to introduce increments only in the subspace spanned by those unstable modes.

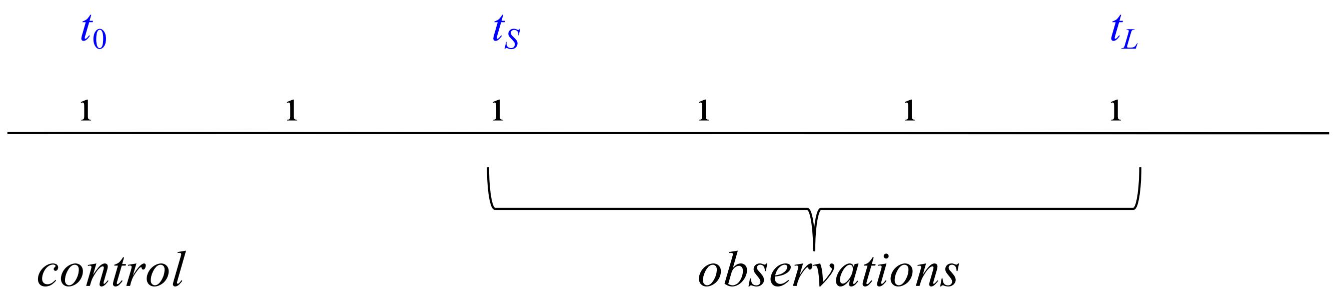
In the perfect model case, assimilation is most efficient when increments are introduced in a space with dimension slightly above the number of non-negative Lyapunov exponents.

In the case of imperfect model (and of strong constraint assimilation), preliminary results lead to similar conclusions, with larger optimal subspace dimension, and less well marked optimality. Further work necessary.

In agreement with theoretical and experimental results obtained for Kalman Filter assimilation (Trevisan and Palatella, McLaughlin).

Iterative Ensemble Kalman Smoother (*IEnKS*, Bocquet and Sakov, 2014)

Minimization performed at time t_0, in an appropriately chosen reduced subspace, assimilating observations performed between times t_S and t_L, with $t_0 \leq t_S \leq t_L$



If the dimension of the reduced subspace is small enough, gradient of objective function can be computed by finite differences, and approximate Hessian can be determined. Once the minimization has been achieved, a new ensemble of perturbations can be obtained by transport of the approximate inverse Hessian.

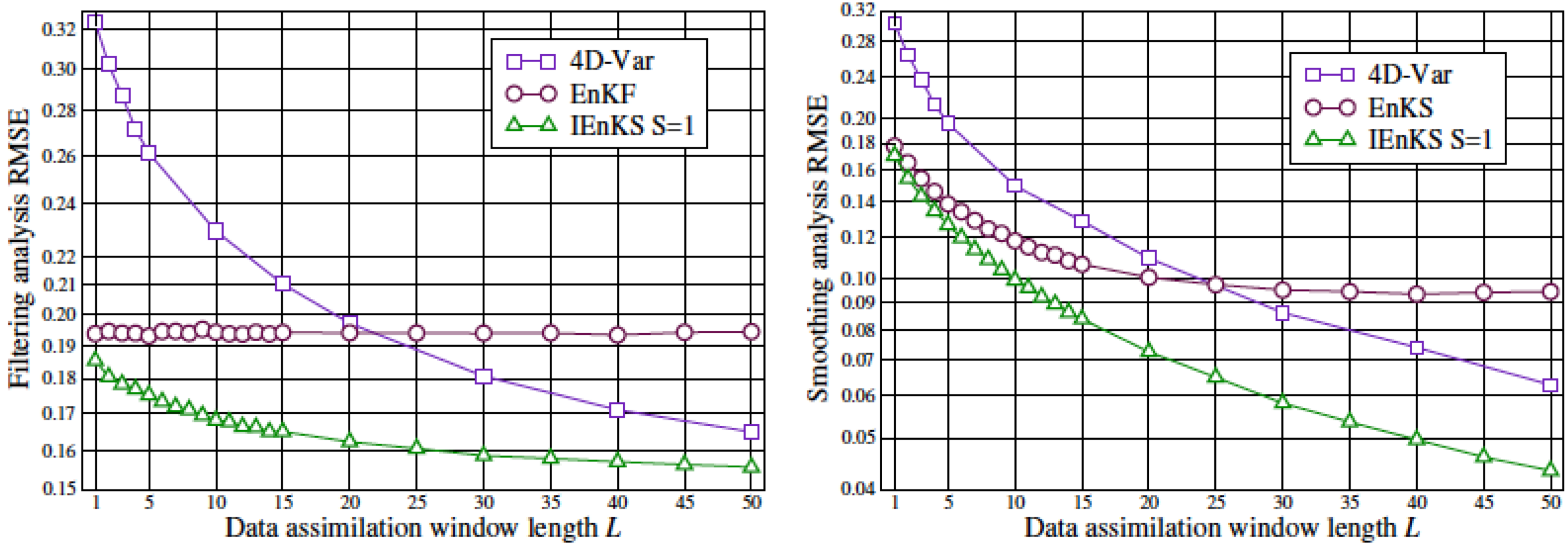


Figure 5: Average root mean square error of several DA methods computed from synthetic experiments with the Lorenz-96 model. The left panel shows the filtering analysis root mean square error of optimally tuned EnKF, 4DVar, IEnKS assimilation experiments, as a function of the length of the DAW. The right panel shows the smoothing analysis root mean square error of optimally tuned EnKS, 4DVar and IEnKS as a function of the length of their data assimilation window. The optimal RMSE is chosen within the window for 4DVar and it is taken at the beginning of the window for the IEnKS. The EnKF, EnKS and IEnKS use an ensemble of $N = 20$, which avoids the need for localization but requires inflation. The length of the DAW is $L \times \Delta t$, where $\Delta t = 0.05$.

Carrassi *et al.*, 2018

Dynamical system

State vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$. Evolves in time according to equation

$$d\boldsymbol{x}/dt = \boldsymbol{F}(\boldsymbol{x}) \qquad (1)$$

or, componentwise

$$dx_i/dt = F_i(\boldsymbol{x}), \quad i = 1, \ldots, n$$

Purely deterministic (no stochastic component)

$$d\boldsymbol{x}/dt = \boldsymbol{F}(\boldsymbol{x}) \qquad (1)$$

Initial condition $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$ defines unique solution (or *orbit*)

$$\boldsymbol{x}(t) = \boldsymbol{R}(t, t_0)\,(\boldsymbol{x}_0)$$

$\boldsymbol{R}(t, t_0)$ is the *resolvent* of Eq. (1) between times t_0 and t.

System can be discretized in time

$$\boldsymbol{x}_{k+1} = \boldsymbol{M}_k\,(\boldsymbol{x}_k)$$

Typical questions about dynamical systems

- *Stationary points* ($\boldsymbol{F}(\boldsymbol{x}) = 0$) and *associated stability ?*

- *Stability of orbits ?*

- *Long term behaviour of orbits* (convergence to fixed points, periodicity, convergence to limit cycle or torus, divergence to infinity, non-periodic oscillations, …) *?*

- *Uncertainty in initial conditions. How does it evolve ?*

$$d\boldsymbol{x}/dt = \boldsymbol{F}(\boldsymbol{x}) \qquad (1)$$

Solution $\boldsymbol{x}(t)$. Perturbation $\delta\mathbf{x}(t)$. Evolves according to

$$d\delta\mathbf{x}/dt = \boldsymbol{F}[\mathbf{x}(t) + \delta\mathbf{x}] - \boldsymbol{F}[\boldsymbol{x}(t)] \approx \mathbf{F}'(t)\ \delta\boldsymbol{x}$$

where $\mathbf{F}'(t)$ is *Jacobian* (matrix of partial derivatives) of operator $\boldsymbol{F}$ at point $\boldsymbol{x}(t)$

$$d\delta\mathbf{x}/dt = \mathbf{F}'(t)\ \delta\boldsymbol{x} \qquad \text{(TLM)}$$

is *tangent linear system* of system (1) along solution $\boldsymbol{x}(t)$. Describes evolution of perturbation $\delta\boldsymbol{x}$ on $\boldsymbol{x}(t)$ to first order wrt initial value of perturbation.

$\delta\mathbf{x}(t) = \boldsymbol{F}[\boldsymbol{x}(t)]$ is solution of (TLM)

$$d\delta\boldsymbol{x}/dt = \mathbf{F}'(t)\ \delta\boldsymbol{x} \qquad \text{(TLM)}$$

Adjoint equation

$$d\lambda/dt = -\ [\mathbf{F}'(t)]^{\mathrm{T}}\ \lambda \qquad \text{(ADJ)}$$

For system discretized in time

$$\boldsymbol{x}_{k+1} = \boldsymbol{M}_k(\boldsymbol{x}_k)$$

$$\delta\boldsymbol{x}_{k+1} = \mathbf{M}_k'\ \delta\boldsymbol{x}_k \qquad \text{(TLM)}$$

Adjoint

$$\lambda_k = [\mathbf{M}_k']^{\mathrm{T}}\ \lambda_{k+1} \qquad \text{(ADJ)}$$

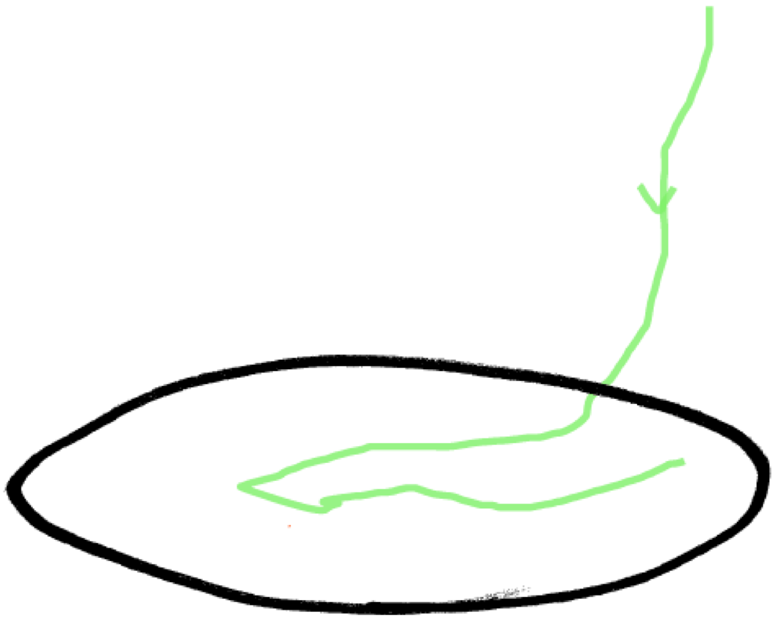
Lorenz (1963)

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

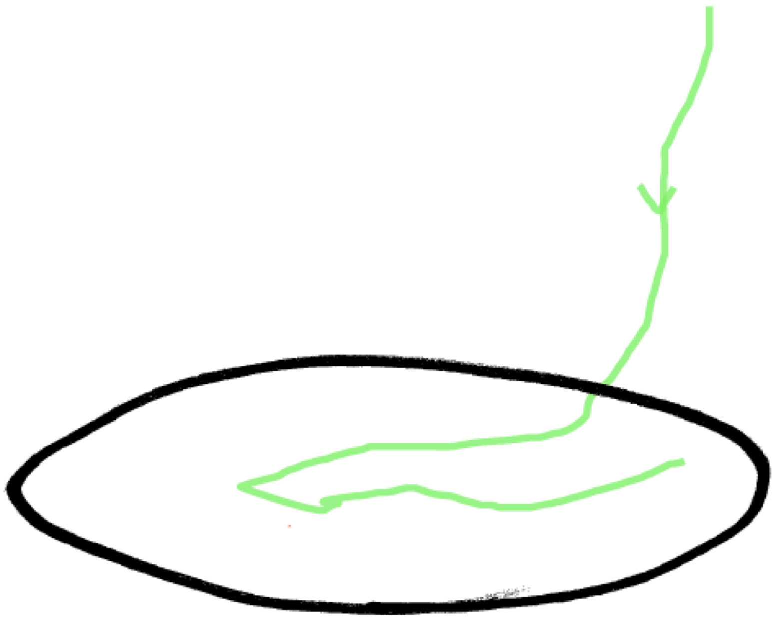
$$dy/dt = \rho x - y - xz$$

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

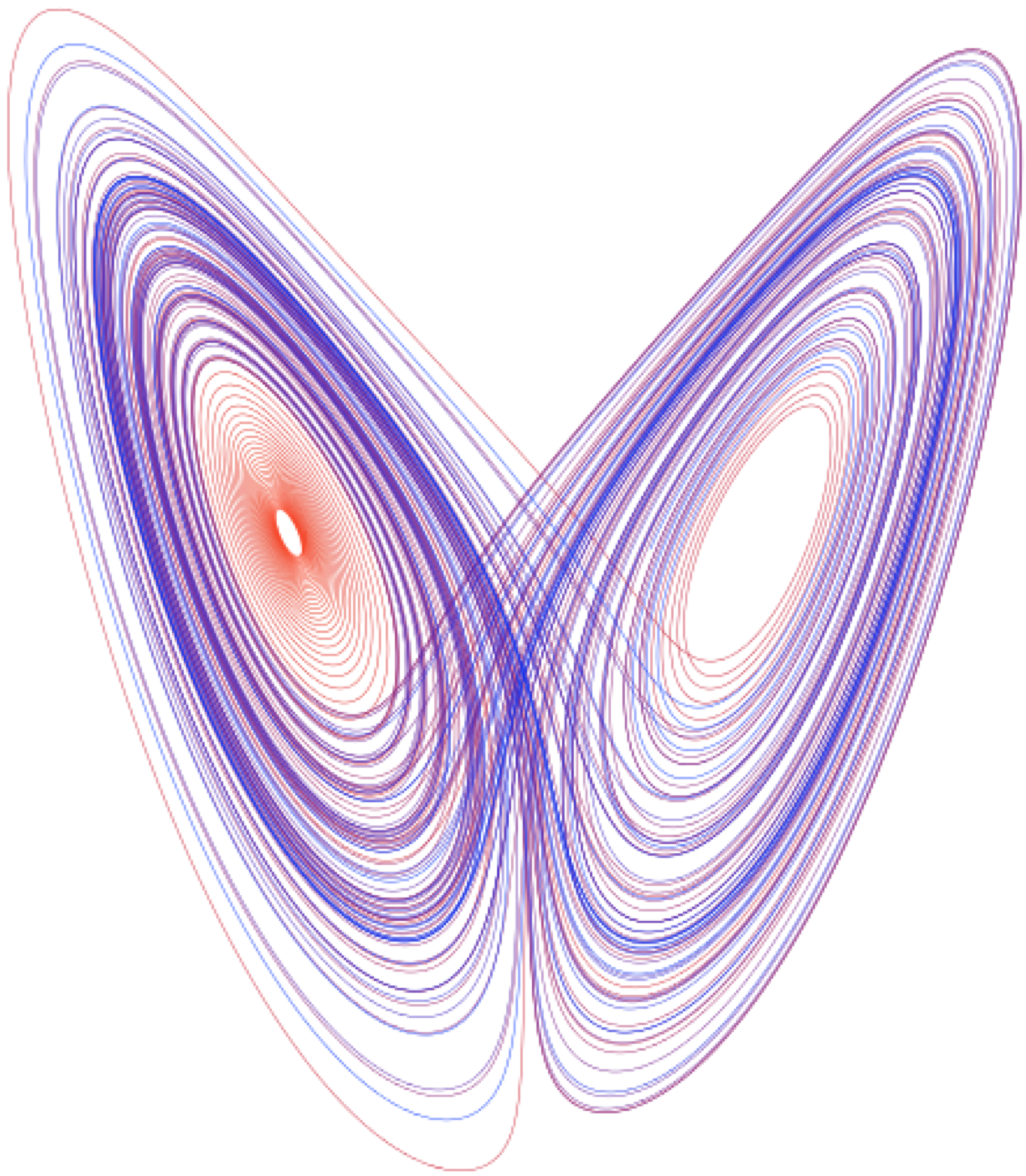
with parameter values $\sigma = 10$, $\rho = 28$, $\beta = 8/3$ $\Rightarrow$ chaos



All orbits end up trapped in the same neighbourhood, within which they have accumulation points



By continuity, the set of accumulation points of any orbit consists of full orbits. The accumulation points of almost all orbits are the same, and make up the *attractor* of the system.



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

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

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

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

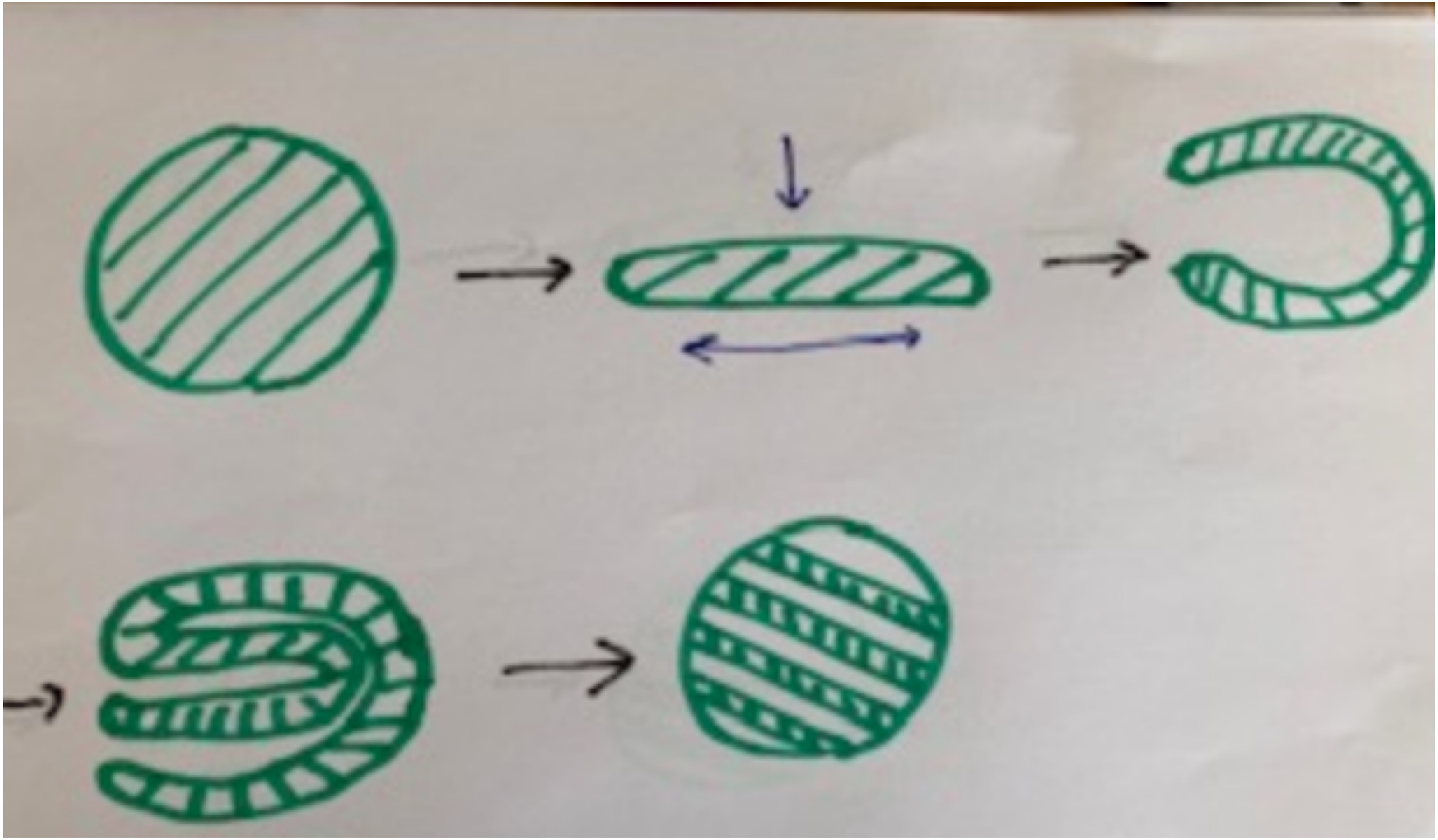
$$\mathrm{div}F = \partial(dx/dt)/\partial x + \partial(dy/dt)/\partial y + \partial(dz/dt)/\partial z$$
$$= -(\sigma + 1 + \beta) = -13.666\ldots < 0$$

Volume element $V(t) = V(0) \exp[-13.67\, t]$ decreases exponentially with time

Probability Density Function (*PDF*) $p(\boldsymbol{x}, t)$ for state vector. Evolves in time according to equation

$$dp/dt + p \, div \, \boldsymbol{F} = 0$$

which expresses conservation of probability in the flow $\boldsymbol{F}$. It is fundamentally the same equation as the 'continuity' equation, which expresses conservation of mass in physical motion. It is called in the present context the *Liouville equation*.



Loss of predictability in dissipative chaos

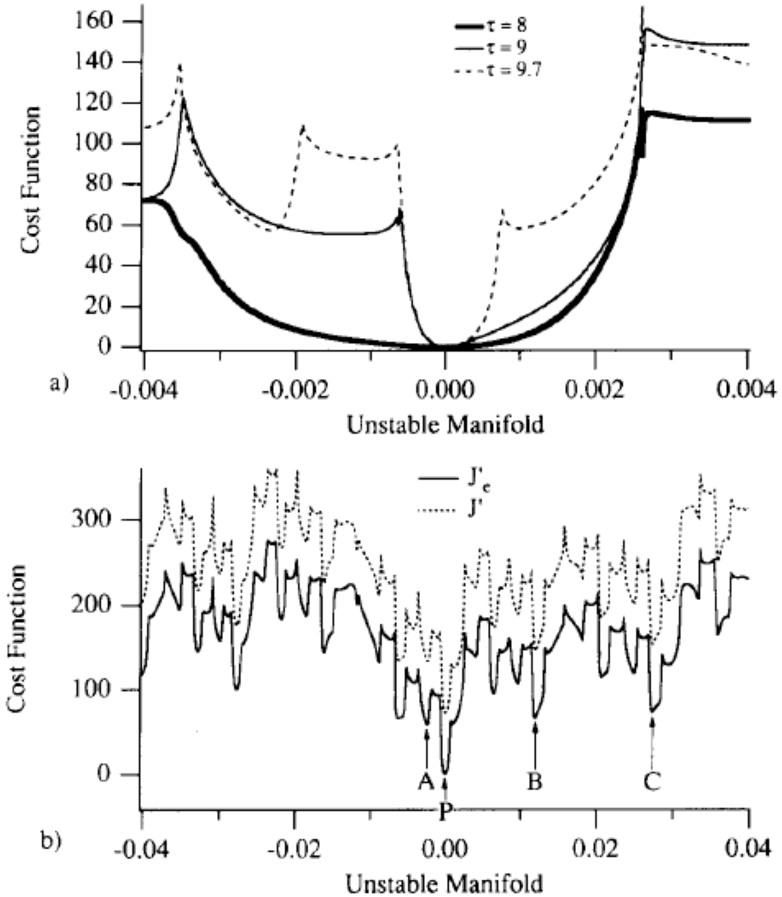
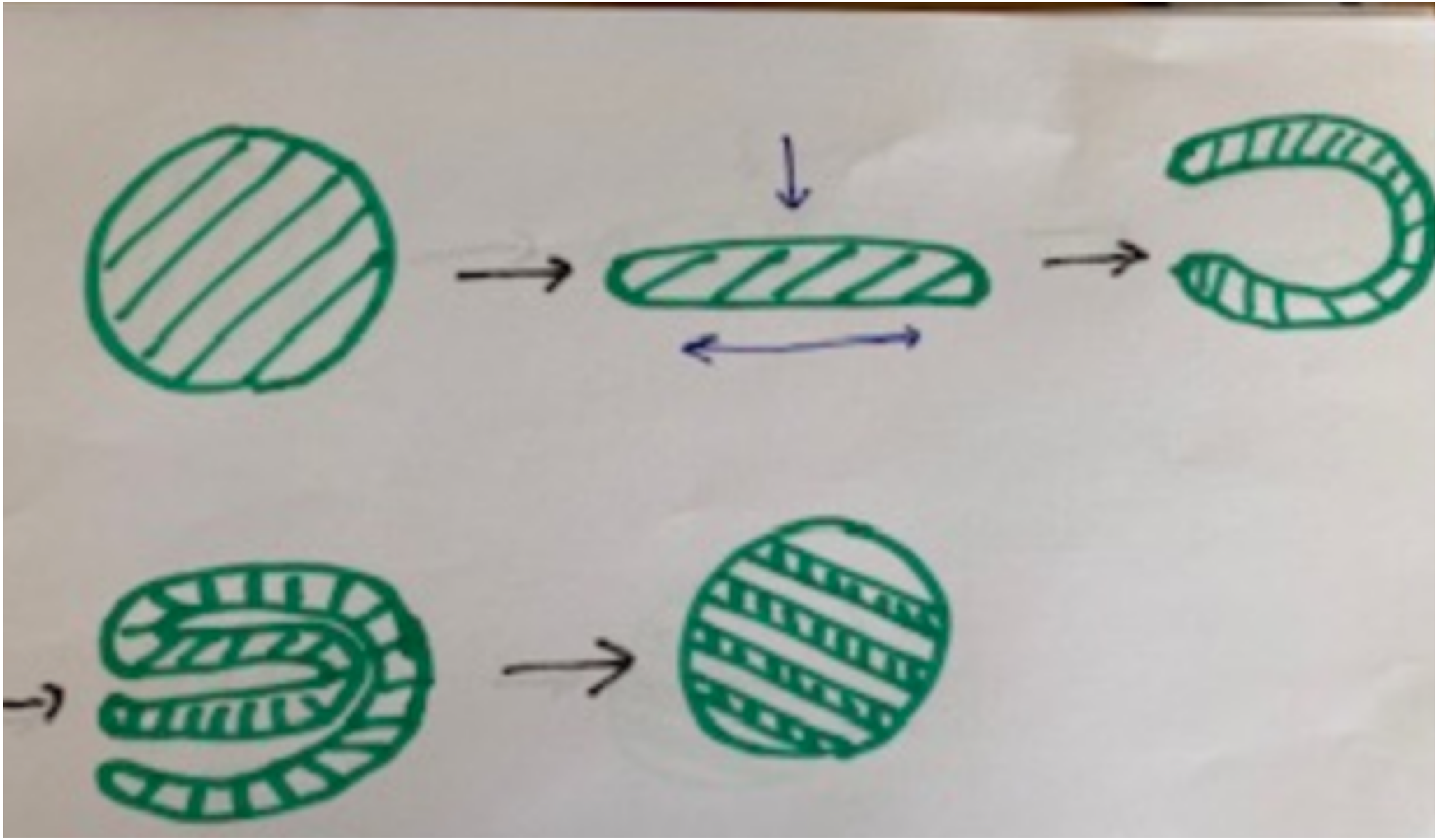


Fig. 4. Panel (a): Cross-section of the error-free forward cost-function $J'_e(\tau, \hat{x}, x)$ along the unstable manifold, for various values of τ. Panel (b). As in panel (a), for $\tau = 9.7$, and with a display interval ten times as large, respectively for the error-free forward cost-function $J'_e(\tau,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



The attractor, which consists of infinitely many foliations, is called *strange*

Linear constant coefficient system with dimension n

$$d\boldsymbol{x}/dt = \boldsymbol{A}\boldsymbol{x}$$

Eigenvectors $\boldsymbol{e}_j$, $j = 1, \ldots, n$

Eigenvalues $\mu_j = \lambda_j + i\,\nu_j$, $\lambda_1 > \ldots > \lambda_n$

$$\boldsymbol{x}(t_0) = \Sigma_j\, x_j(t_0)\, \boldsymbol{e}_j \qquad\qquad x_1(t_0) \neq 0$$

$$\begin{aligned}\boldsymbol{x}(t_0+\tau) &= \Sigma_j \exp(\mu_j\,\tau)\, x_j(t_0)\, \boldsymbol{e}_j \\ &= \exp(\mu_1\,\tau)\, x_1(t_0) \\ &\quad \text{x } \{\boldsymbol{e}_1 + \Sigma_{j>1} \exp[(\mu_j - \mu_1)\,\tau]\, x_j(t_0)/x_1(t_0)\, \boldsymbol{e}_j\}\end{aligned}$$

$$\boldsymbol{x}(t_0+\tau) \;=\; \exp(\mu_1 \tau)\, x_1(t_0)\, [\boldsymbol{e}_1 + o(1)]$$

$$\|\boldsymbol{x}(t_0+\tau)\| = \exp(\lambda_1 \tau)\, |x_1(t_0)|\, [\, \|\boldsymbol{e}_1\| + o(1)\,]$$

$$\lim_{t\to\infty} [(1/t) \ln \|\boldsymbol{x}(t)\|\,] = \lambda_1$$

If $x_1(t_0) = 0$, $x_2(t_0) \neq 0$, the limit is λ_2, and so on …

There exist a sequence of real numbers (real parts of eigenvalues of matrix $\boldsymbol{A}$)

$$\lambda_1 > \ldots > \lambda_m \qquad m \leq n$$

and a sequence of (sub)spaces of $\mathcal{R}^n$

$$\mathcal{E}_{m+1} = \varnothing \subset \mathcal{E}_m \subset \ldots \subset \mathcal{E}_j \subset \ldots \subset \mathcal{E}_1 = \mathcal{R}^n$$

such that $\lim_{t \to \infty} [(1/t) \ln \| \boldsymbol{x}(t) \|] = \lambda_j$ when $\boldsymbol{x}(t_0) \in \mathcal{E}_j / \mathcal{E}_{j+1}$

The same is fundamentally true for dynamical systems with attractors (solutions constantly return to the vicinity of same points → *ergodicity*)

$$d\boldsymbol{x}/dt = \boldsymbol{F}(\boldsymbol{x})$$

Solution $\boldsymbol{x}(t)$. Associated TLM

$$d\delta\mathbf{x}/dt = \mathbf{F}'(t)\ \delta\mathbf{x} \qquad \text{(TLM)}$$

Oseledets theorem. There exist a sequence of real numbers (*Lyapunov exponents*)

$$\lambda_1 > \ldots > \lambda_m \qquad\qquad m \leq n$$

and a sequence of (sub)spaces of $\mathcal{R}^n$

$$\mathcal{E}_{m+1} = \varnothing \subset \mathcal{E}_m \subset \ldots \subset \mathcal{E}_j \subset \ldots \subset \mathcal{E}_1 = \mathcal{R}^n$$

such that $\lim_{t\to\infty} [(1/t) \ln \| \delta\mathbf{x}(t) \|] = \lambda_j$ when $\delta\mathbf{x}(t_0) \in \mathcal{E}_j / \mathcal{E}_{j+1}$

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)

Lyapunov exponents

$$0.9056,\ 0,\ -14.5723\ (\text{sum} = -13.6667)$$

$$[\mathrm{div}F = -(\sigma + 1 + \beta) = -41/3 = -13.6666.. < 0]$$

Lyapunov exponents

$$0.9056,\ \ 0, -14.5723$$

Lyapunov exponents measure rate of growth of perturbations, averaged over the whole attractor

Presence of at least one positive Lyapunov exponent is signature of chaos.

In an ergodic system, one exponent is equal to 0. It corresponds to perturbations in the direction of the motion, which will be neither amplified nor damped over long periods.

Experiments performed on the Lorenz (1996) model

$$\frac{d}{dt}x_j = (x_{j+1} - x_{j-2})x_{j-1} - x_j + F$$

with $j = 1, ..., I$.

with periodic co

Divergence. $\mathrm{div}F = \Sigma_j\, \partial(dx_j/dt)\,/\partial x_j = -\,I\ < 0$

Three values of I have been used, namely $I = 40, 60, 80$, which correspond to respectively $N^+ = 13, 19$ and 26 positive Lyapunov exponents.

In all three cases, the largest Lyapunov exponent corresponds to a doubling time of about 2 days (with 1 ‘day’ = 1/5 model time unit).

Identical twin experiments (perfect model)

Lyapunov exponents

$$\lambda_1 > \ldots > \lambda_m \qquad\qquad m \le n$$

associated (sub)spaces of $\mathcal{R}^n$

$$\mathcal{E}_{m+1} = \varnothing \subset \mathcal{E}_m \subset \ldots \subset \mathcal{E}_j \subset \ldots \subset \mathcal{E}_1 = \mathcal{R}^n$$

$$\lim_{t\to\infty} [(1/t) \ln \| \delta\mathbf{x}(t) \|] = \lambda_j \text{ when } \delta\mathbf{x}(t_0) \in \mathcal{E}_j / \mathcal{E}_{j+1}$$

Modulus $\| \delta\mathbf{x}(t) \|$ depends on choice of norm, but asymptotic exponential rate of growth or decay does not. Lyapunov exponents do not depend on position on orbit, and are the same for all orbits with the same attractor. Subspaces $\mathcal{E}_j$ depend on position on orbit, but evolve with the motion.

Lyapunov vectors

At a given point along an orbit, *forward Lyapunov vectors* are vectors that will concentrate most rapidly on the Lyapunov rate of growth or decay.

Similarly *backward Lyapunov vectors* are vectors that have concentrated most rapidly in the past on the Lyapunov rate of growth or decay. In assimilation, they tend to dominate the background error.

These vectors depend on the choice of a norm and are orthogonal with respect to the chosen norm. They do not follow the evolution of the flow.

One also defines *covariant Lyapunov vectors*, which are exactly amplified or damped according to the Lyapunov exponents, and evolve with the motion. They do not depend on the choice of a norm, and are not orthogonal wrt to a time-independent norm.

The notions of Lyapunov exponents and vectors have turned out to be very useful for the study of the dynamics of the atmosphere and the ocean, They are relatively easy to determine (identifying them does not require long numerical integrations, which means that the atmosphere and the ocean have in a sense ‘good ergodicity’). They more or less explicitly underlie the approach of Assimilation in the Unstable Subspace

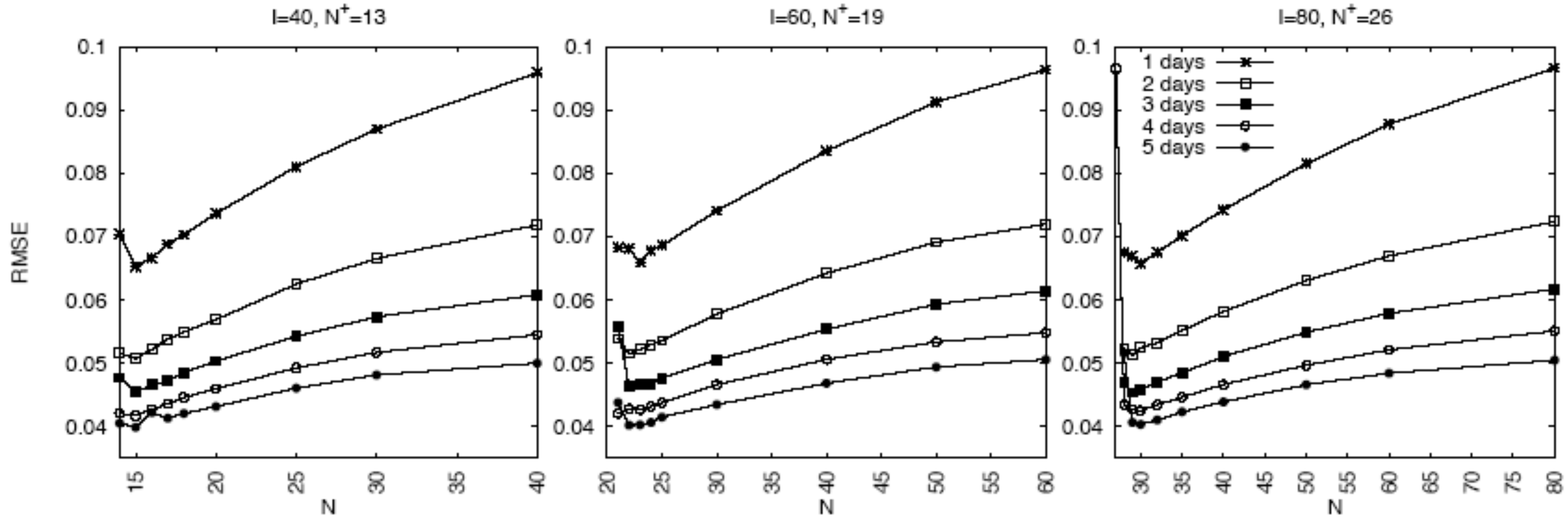


Figure 1. Time average RMS analysis error at $t = \tau$ as a function of the subspace dimension N for three model configurations: I=40, 60, 80. Different curves in the same panel refer to different assimilation windows from 1 to 5 days. The observation error standard deviation is $\sigma_o = 0.2$.

No explicit background term *(i. e.,* with error covariance matrix) in objective function : information from past lies in the background to be updated, and in the N perturbations which define the subspace in which updating is to be made.

Best performance for N slightly above number N^{+} of positive Lyapunov exponents.

Trevisan *et al*., 2010

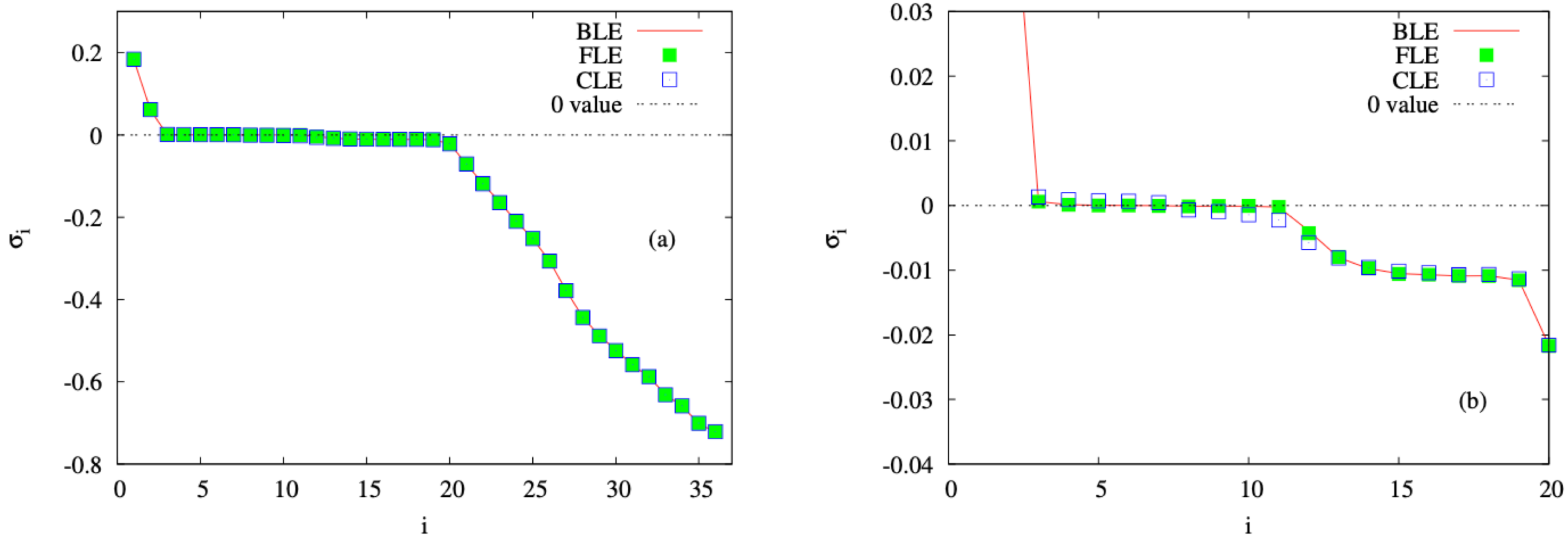


Figure 1: Lyapunov spectra based on the forward integration (FLEs, red line), backward integration (BLEs, full green squares), and on the CLVs (CLEs, empty blue squares). Panel (a): the full spectrum of LEs. Panel (b) a zoom around 0. The parameter values are $C_o = 350$ W m^{-2} and $d = 1 \times 10^{-8} s^{-1}$.

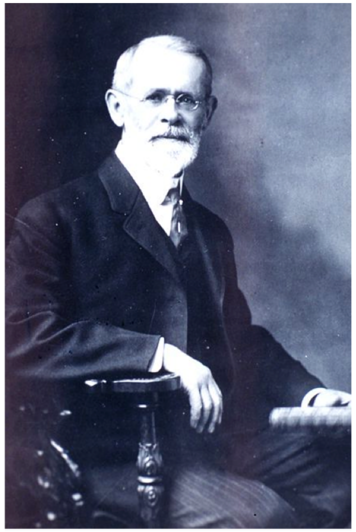
Lyapunov exponents of a low-order coupled atmosphere-ocean model (Vannitsem and Lucarini, *J. Phys. A*, 2016)

History of Numerical Weather Prediction

Cleveland Abbe

The Physical Basis of Long Range Weather Forecasts, 1901,

Monthly Weather Review



Wilhelm Bjerknes

Das Problem der Wettervorhersage, betrachtet von Standpunkt der Mechanik und Physik, 1904, *Meteorologische Zeitschrift*

V. Bjerknes at the origin of the ‘Bergen School of Meteorology'

From course 2

Physical laws governing the flow

- Conservation of mass

  $D\rho/Dt + \rho\,\mathrm{div}\underline{U} = 0$

- Conservation of energy

  $De/Dt - (p/\rho^2)\,D\rho/Dt = Q$

- Conservation of momentum

  $D\underline{U}/Dt + (1/\rho)\,\underline{\mathrm{grad}}p - \underline{g} + 2\,\underline{\Omega} \wedge \underline{U} = \underline{F}$

- Equation of state

  $f(p, \rho, e) = 0$ $\qquad$ $(p/\rho = rT,\ e = C_v T)$

- Conservation of mass of secondary components (water in the atmosphere, salt in the ocean, chemical species, …)

  $Dq/Dt + q\,\mathrm{div}\underline{U} = S$

These physical laws must be expressed in practice in discretized (and necessarily imperfect) form, both in space and time

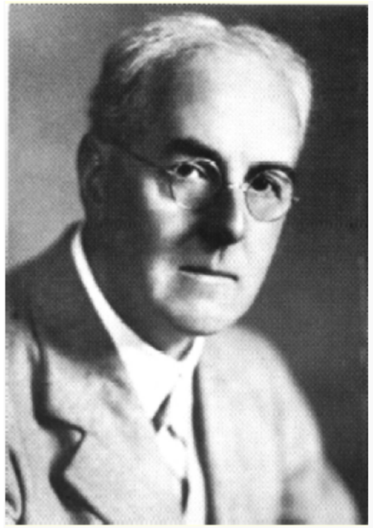
History of Numerical Weather Prediction (continuation)

Lewis Fry Richardson

Weather Prediction by Numerical Process, 1922

Cambridge University Press *

Forecast Factory



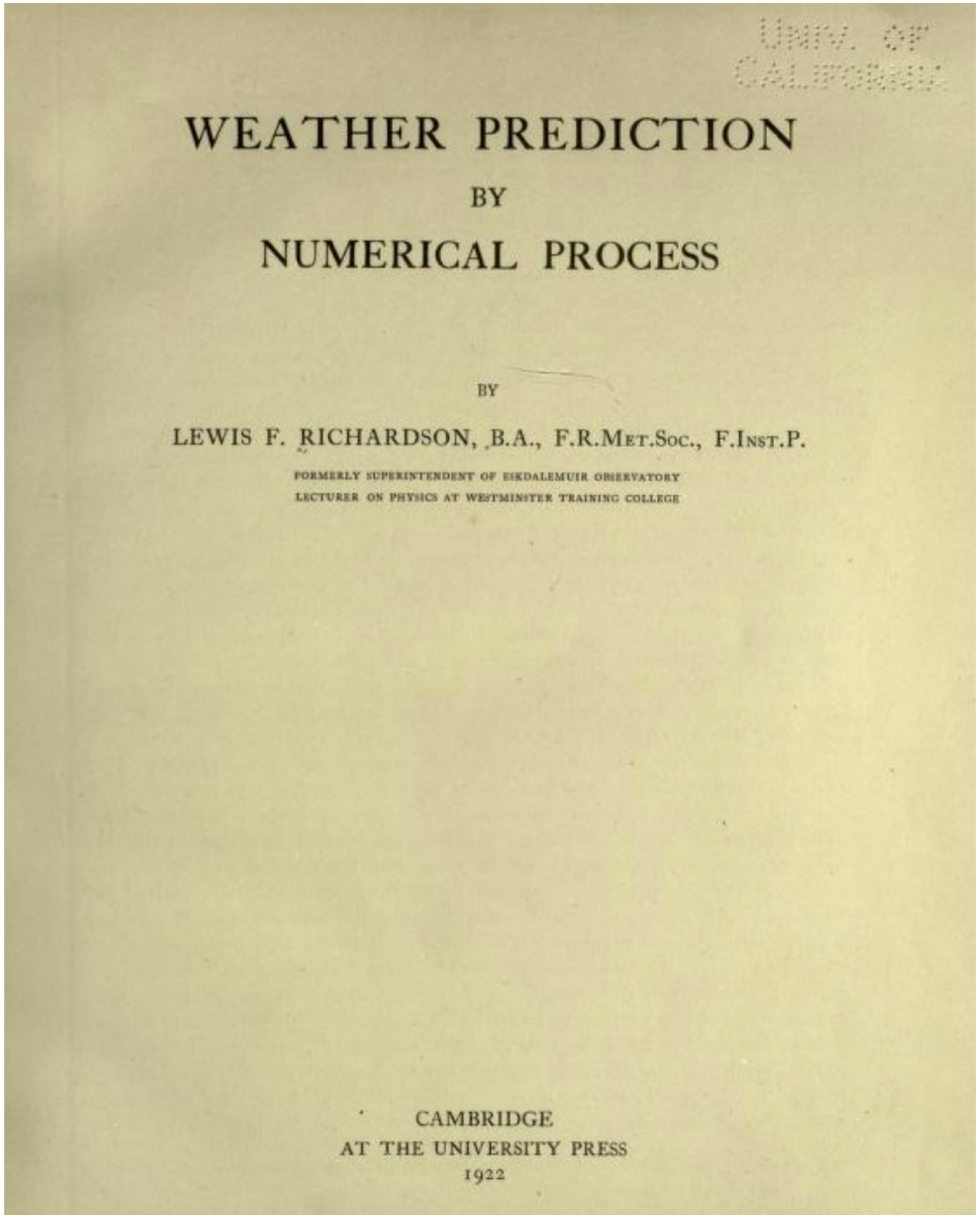
Richardson number, fractals, pacifism

* Accessible at URL

https://energy4climate.pages.in2p3.fr/public/education/ensemble_data_assimilation_tutorial/notebooks/T1%20-%20Introduction%20to%20Ensemble%20Data%20Assimilation%20for%20Numerical%20Weather%20Prediction.html

History of Numerical Weather Prediction (continuation 2)

https://energy4climate.pages.in2p3.fr/public/education/ensemble_data_assimilation_tutorial/notebooks/T1%20-%20Introduction%20to%20Ensemble%20Data%20Assimilation%20for%20Numerical%20Weather%20Prediction.html



WEATHER PREDICTION

BY

NUMERICAL PROCESS

BY

LEWIS F. RICHARDSON, B.A., F.R.MET.SOC., F.INST.P.

FORMERLY SUPERINTENDENT OF ESKDALEMUIR OBSERVATORY

LECTURER ON PHYSICS AT WESTMINSTER TRAINING COLLEGE

CAMBRIDGE

AT THE UNIVERSITY PRESS

1922

History of Numerical Weather Prediction (continuation 3)

John von Neumann

Institute for Advanced Studies, Princeton, 1946-1950

First electronic computers (ENIAC)

(J. Charney, N. A. Phillips, R. Fjørtoft, C. G. Rossby,

J. Smagorinsky, …)

Institute for Advanced Study, about 1948-50. J. von Neumann is second from left. And from right, J. Charney, C. G. Rossby (?), R. Fjørtoft (?)

History of Numerical Weather Prediction (continuation 5)

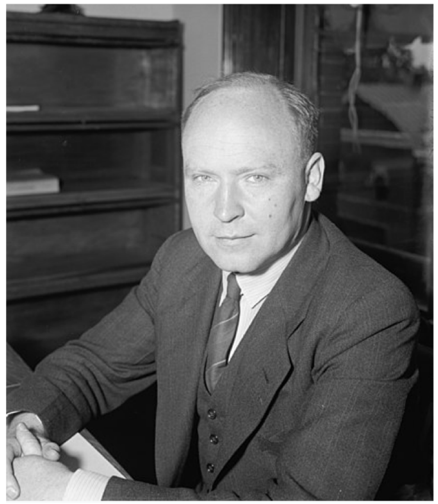
Charney developed vorticity barotropic model

First simulation of real atmospheric situation in 1950

Jule Gregory Charney en 1978.

First operational numerical forecast performed in 1954 in Sweden

(C. G. Rossby)



History of Numerical Weather Prediction (continuation 6)

Numerical prediction has gradually been implemented in more and more meteorological services around the world.

European Centre for Medium-Range Weather Forecasts (ECMWF, 1975)

Ensemble prediction

History of Numerical Weather Prediction (continuation 7)

Extension to simulation of oceanic circulation and climate (early 1970's, S. Manabe and K. Bryan, GFDL).

Climate simulations (S. Manabe, R. Wetherald)

S. Manabe awarded Nobel Prize in Physics in 2021

History of Numerical Weather Prediction (continuation 8)

A large variety of models covering different spatial and temporal scales and phenomena (small-scale convection, monthly and seasonal prediction, atmospheric chemistry, …) have been developed over the years and are used for research and operational applications.

Intergovernmental Panel on Climate Change (IPCC, 1988)

Publishes reports that describe the state of climate science and presents ‘projections’ largely based on numerical simulations

First report in 1990

…

Fifth report in 2014

Sixth report in 2023

History of Numerical Weather Prediction (continuation 9)

Another application has been to *reanalyses of past data*, with present models and assimilation algorithms (as well as all observations which are now available, and may not have been when they were performed).

ECMWF : *ERA-15*, *ERA-40*, *ERA-Interim*, *ERA5* (1940 - present, with ~ 31-km horizontal resolution),
ERA6 is in development with 14-km horizontal resolution

NCEP/NCAR (National Centers for Environmental Prediction / National Center for Atmospheric Research) :1948 - present

History of Numerical Weather Prediction (continuation 10)

More recently, as concerns short and medium-range prediction, a major change has been the development of algorithms based on machine learning, trained on long series of past analyses. These algorithms produce forecats or quality similar to those of physical forecasts, but at a much lower numerical cost.

Artificial Intelligence

(aka ***Machine Learning*** or ***Deep Learning***)

Numerical modelling of the atmospheric and oceanic flow, as presented in the course, is fundamentally built on known physical laws (conservation of mass, momentum and energy).

Why not directly use observations (for instance, in the case of a weather forecast, why not look for analogues in the past, and make the forecast from those analogues) ?

E. N. Lorenz (1960s). Sample of past observations will never be large enough for competing with physically-based models.

But :

- there is no incompatibility between the two approaches

- there remain many processes in numerical models which we do not know how to describe on the basis of well-established physical laws (interactions between atmosphere and underlying medium, such as *e.g.* vegetation, all kinds of subgrid scale processes, …)

- amount of data of all kinds, as well as computing power, are increasing very rapidly.

Artificial Intelligence (aka ***Machine Learning***) (continuation)

Powerful numerical tools have been developed for the exploitation of very large sets of data (*big data*)

Neural networks. Define an explicit numerical link between an *input set* and an *output set.* Define function F such that, to some useful degree of approximation

$$\boldsymbol{y} = \boldsymbol{F}(\boldsymbol{x})$$

where x and y belong to the input and output set respectively.

The function $\boldsymbol{F}$ is typically built as a composition of *activation functions*

Artificial Intelligence (aka ***Machine Learning***) (continuation 2)

Activation functions

Sigmoid functions, relu function

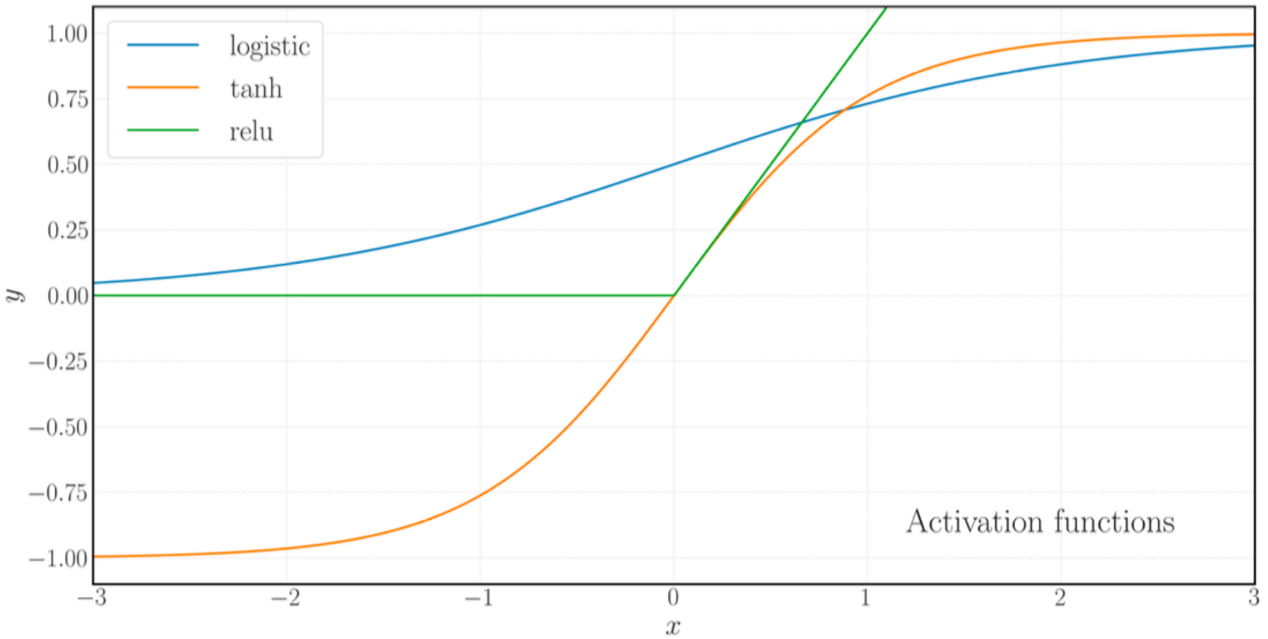


Figure 6.2: Classical activation functions.

Bocquet and Farchi, 2025

Artificial Intelligence (aka ***Machine Learning***) (continuation 2)

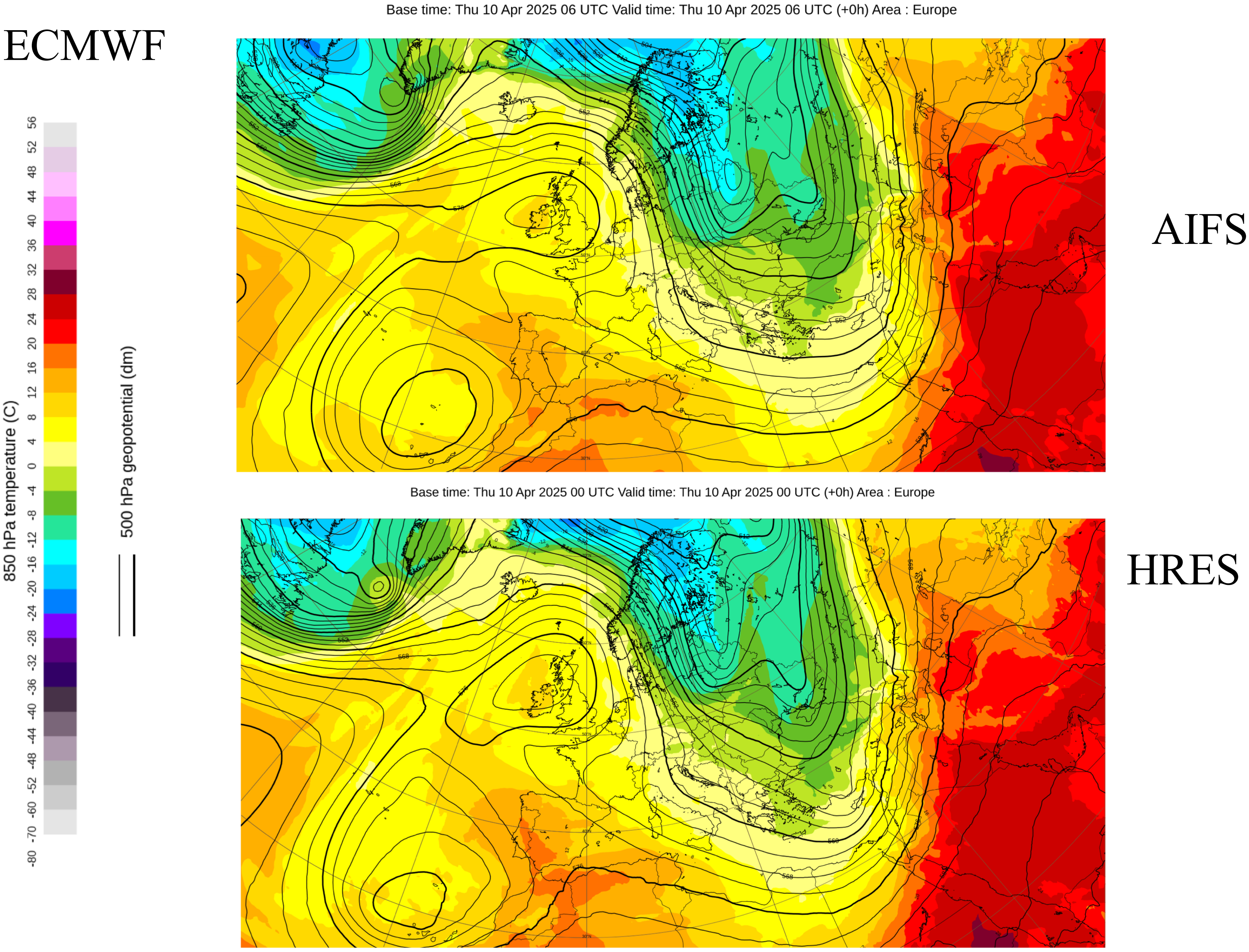
Neural networks have turned out to be extremely efficient in many applications. In the context of assimilation of observations, they have been used for defining for instance the observation operators (***H***) corresponding to satellite observations. But they have been used more recently, in evaluation studies and on idealized situations, but with some success, for determining 'dynamical laws'.

Machine Learning (continuation 3)

And, more importantly, they have been used for developing softwares for meteorological predictions at a range of a few days, using as training ensembles reanalyses produced by meteorological centres.

- *GraphCast*
- *Pangu-Weather*
- ...

Forecasts obtained are of similar quality to those of best physical models, but at a much lower numerical cost (a few minutes, instead of a few hours, for a 10-day forecast). ECMWF has for instance developed the *AIFS* software, with its own ERA5 reanalysis (1979-present) as training ensemble.


ECMWF
850 hPa temperature (C)
500 hPa geopotential (dm)
56
52
48
44
40
36
32
28
24
20
16
12
8
4
0
-4
-8
-12
-16
-20
-24
-28
-32
-36
-40
-44
-48
-52
-60
-70
-80
Base time: Thu 10 Apr 2025 06 UTC Valid time: Thu 10 Apr 2025 06 UTC (+0h) Area : Europe
AIFS
Base time: Thu 10 Apr 2025 00 UTC Valid time: Thu 10 Apr 2025 00 UTC (+0h) Area : Europe
HRES

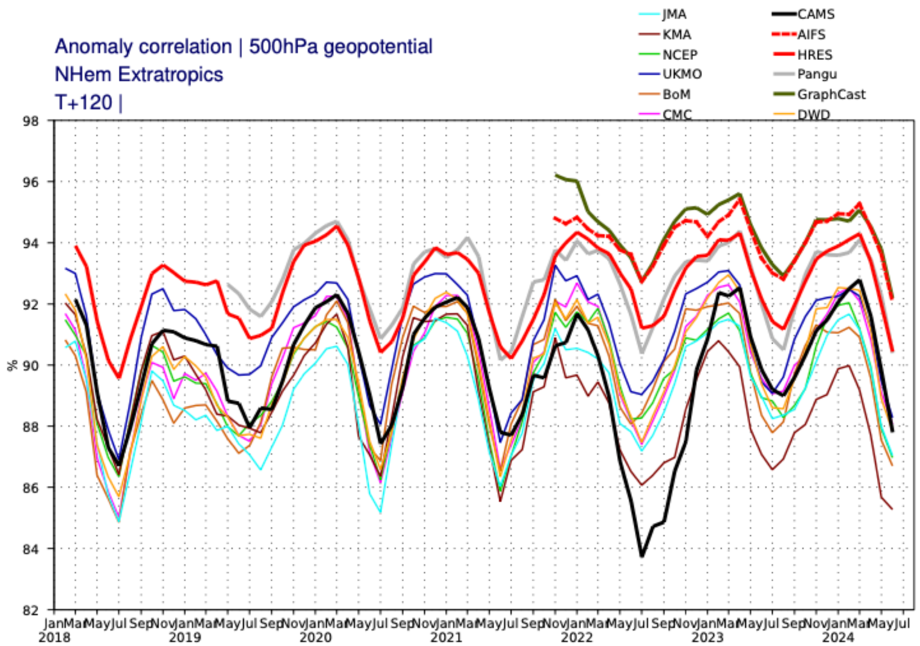


Figure 17: Anomaly correlation of 500 hPa geopotential in the northern hemisphere extratropics at day 5. CAMS forecast (black) shown in comparison to the HRES (red) and forecasts from other global centres (thin lines). Also shown are forecasts from machine learning (ML) models: GraphCast (olive), Pangu (grey), and AIFS (red dashes).

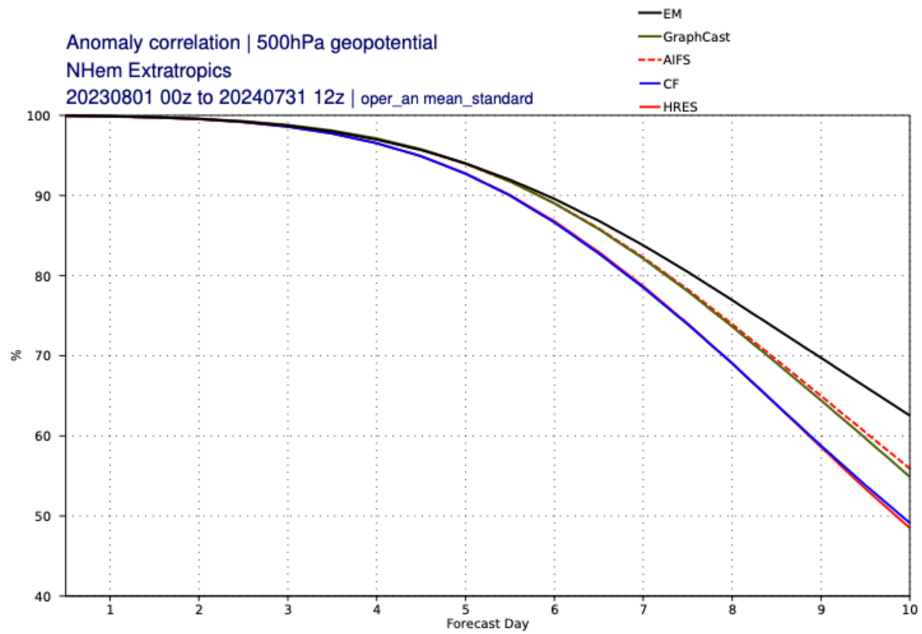


Figure 18: Anomaly correlation of 500 hPa geopotential in the northern extratropics for the 12-month period Aug 2023 to July 2024. Black: ENS mean, olive: GraphCast ML forecast, red dashes: AIFS, blue: ENS control, red: HRES.

Cours à venir

~~Mercredi 2 avril~~

~~Vendredi 11 avril~~

~~Vendredi 18 avril~~

~~Mercredi 23 avril~~

~~Lundi 12 mai~~

~~Mercredi 28 mai~~

~~Mercredi 11juin~~

Mercredi 18 juin