Data Assimilation and (In)stability

Olivier Talagrand

Laboratoire de Météorologie Dynamique, École Normale Supérieure, Paris, France

Numerical Modeling, Predictability and Data Assimilation in Weather, Ocean and Climate

A Symposium Honoring the Legacy of Anna Trevisan

Bologna, Italy

19 October 2017

- Brief history of assimilation : ideas and methods

- Two powerful classes of algorithms : *Variational Assimilation* (*4D-Var*), (*Ensemble*) *Kalman Filter*

- Links between assimilation and stability – instability of the flow (but not really up to date !)

- Complements

  - Observability

Assimilation of observations, as it is known in meteorology and oceanography, originated from the need of defining initial conditions (ICs) for numerical weather prediction. Difficulties gradually arose

- Need for defining ICs with appropriate spatial scales ⇒ '*structure functions*' (now incorporated in background error covariance matrices)

- Need for defining ICs in approximate geostrophic balance ⇒ '*initialization*' (now also incorporated in background error covariance matrices and/or specific penalty terms)

- …

- Need for defining ICs with appropriate spatial scales ⇒ ‘*structure functions*‘ (now incorporated in background error covariance matrices)

- Need for defining ICs in approximate geostrophic balance ⇒ ‘*initialization’* (now also incorporated in background error covariance matrices and/or specific penalty terms)

- Realization that forecast was very sensitive to ICs ⇒ stressed still more need for accurate ICs

- Use of satellite observations, which are

  - distributed continuously in time ⇒ need for using a dynamical model of the flow before even the forecast is started (word *assimilation* was coined in 1967-68)

  - indirect ⇒ need for some form of ‘inversion’

Relative cost of the various components of the operational forecast suite at ECMWF (september 2015, J.-N. Thépaut) :

4DVAR: 9.5%
HRES FC: 4.5%
EDA: 30%
ENS: 22%
ENS: hindcasts 14%

Other: 20% of which BC AN: 3.5% BC FC: 4% BC ENS: 9.5%

EDA (Ensemble Data Assimilation) provides both the variances of background errors for the 4D-Var, and initial perturbations (in addition of singular vectors) for the EPS.

Ratio 1 day of assimilation / 1 day of forecast $\approx$ 15

One early attempt at assimilation

Charney, Halem and Jastrow, *JAS*, 1969

Twin experiment with 2-level primitive equation model, $7° \times 9°$ lat-long grid

Question : is it possible to reconstruct the entire flow from history of temperature field ?

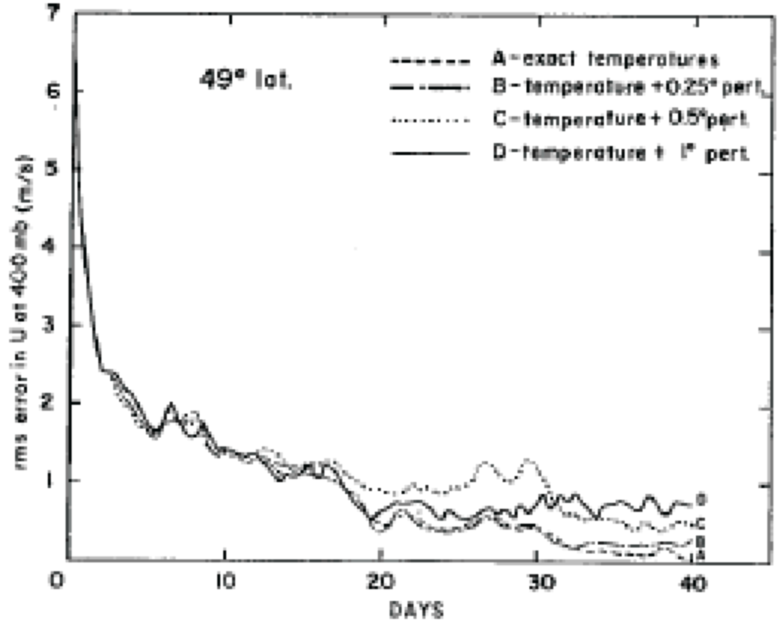


Fig. 1. The rms error in zonal wind (m sec^{-1}) at 400 mb at 49° latitude, in cases where temperatures with random error perturbations of 0, 0.25, 0.5 and 1C are inserted every 12 hr at all grid points.

Charney *et al*., *JAS*,1969

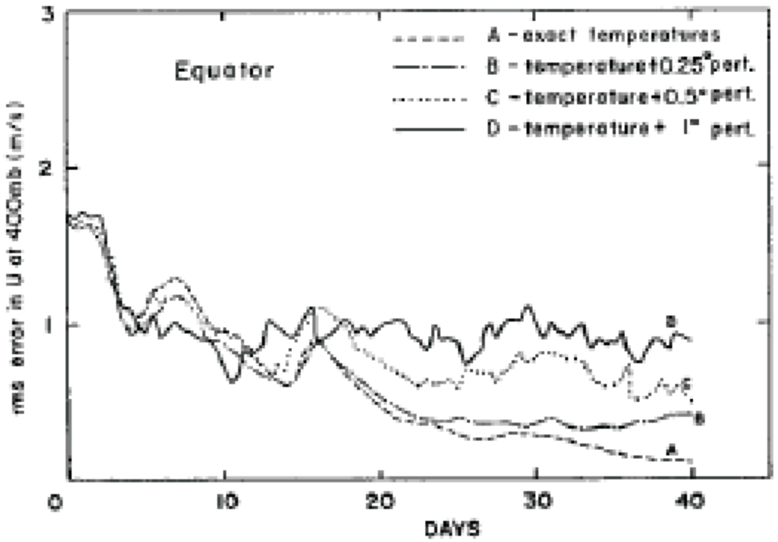


Fig. 2. Same as Fig. 1 for the equator.

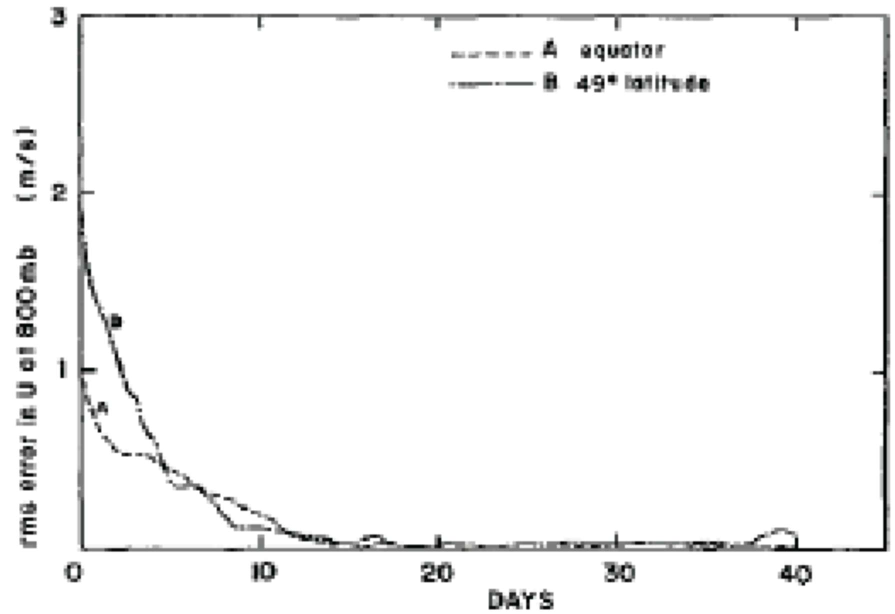


FIG. 4. The rms error in zonal wind (m sec^{-1}) at 800 mb for the equator and 49° latitude, for cases where the exact pressure at sea level and the temperature and wind at 400 mb are inserted every 12 hr at all grid points.

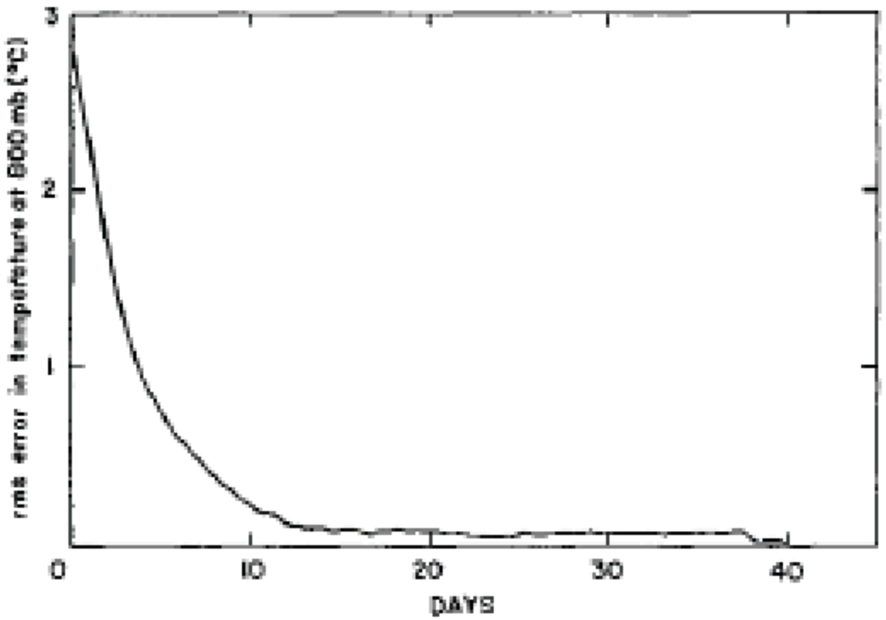


FIG. 5. Same as Fig. 4 except for temperature (°C) for the Northern Hemisphere.

Charney *et al*., *JAS*,1969

Question. Role of geostrophic adjustment ?

Evolution equation

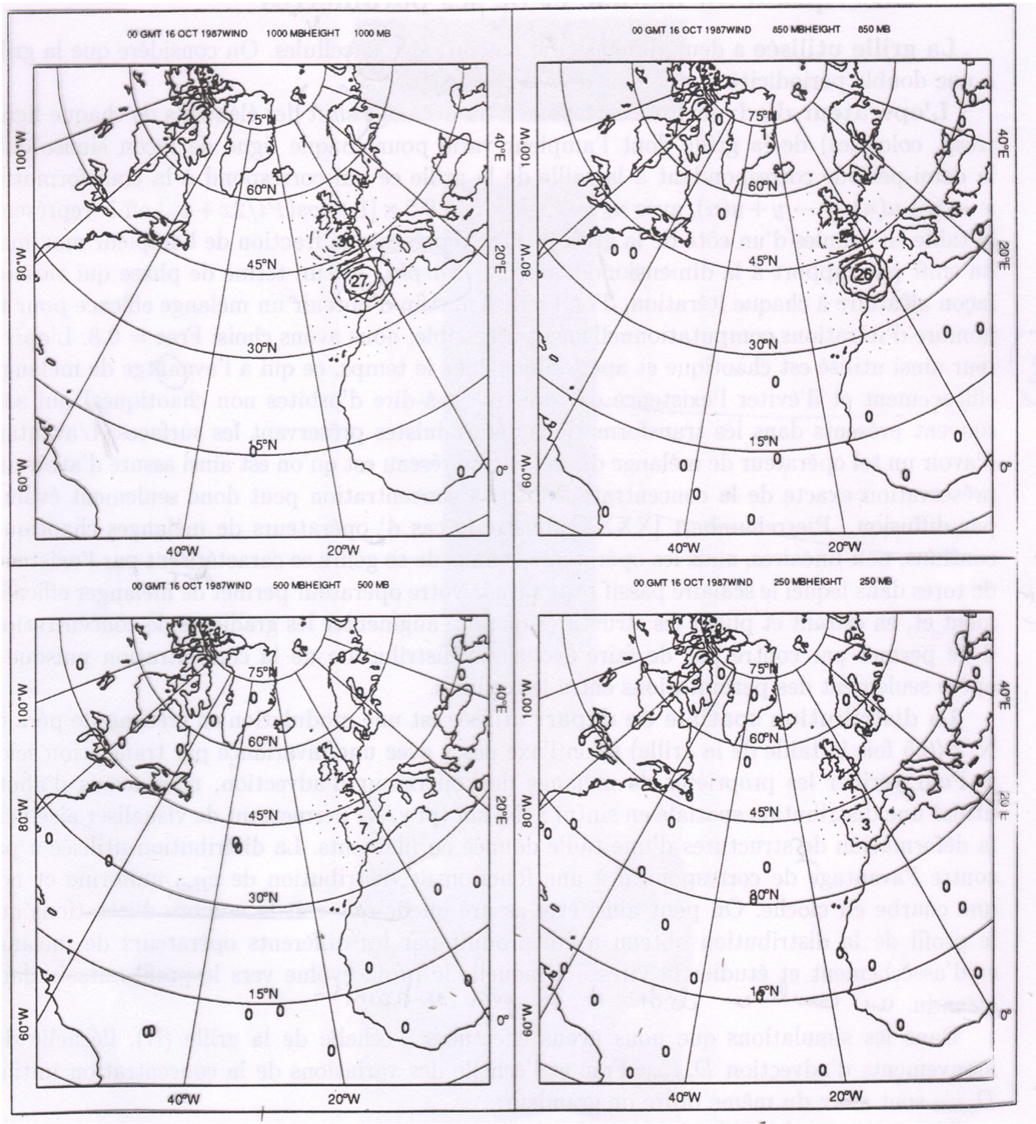
$$x_{k+1} = M_k(x_k) + \eta_k$$

M_k is (known) model, η_k is (unknown) model error

Observations at successive times

$$y_k = H_k(x_k) + \varepsilon_k \qquad\qquad k = 0, \ldots, K$$

For a period, assimilation was performed in operational applications through various forms of *Optimal Interpolation* (*OI*). At observation time, state predicted by assimilating model was updated with new observations, using *structure functions* for propagation of influence of the observations in space as well as to other physical fields.



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

Two classes of algorithms have come to dominate meteorological and oceanographical assimilation, at least in practical applications : *Ensemble Kalman Filter* and *Four-Dimensional Variational Assimilation*.

Kalman Filter (R. H. Jones, *JAS*, 1965, Petersen, *Tellus*, 1968, Ghil *et al.*) evolves in time an estimate of the state of the system and updates that estimate with new observations as they become available. Associated uncertainty is evolved in parallel, originally in the form of a covariance matrix

Analysis step

$$x^a_k = x^b_k + P^b_k H_k^{\mathrm{T}} [H_k P^b_k H_k^{\mathrm{T}} + R_k]^{-1} (y_k - H_k x^b_k)$$
$$P^a_k = P^b_k - P^b_k H_k^{\mathrm{T}} [H_k P^b_k H_k^{\mathrm{T}} + R_k]^{-1} H_k P^b_k$$

- Forecast step

$$x^b_{k+1} = M_k x^a_k$$
$$P^b_{k+1} = M_k P^a_k M_k^{\mathrm{T}} + Q_k$$

In geophysical applications, uncertainty is represented by a number (O(10-100)) of points in state space ⇒ *Ensemble Kalman Filter* (*EnKF*, Evensen *et al.*)

Variational Assimilation (Thompson, *Tellus*, 1961, Penenko and Obraztsov, *Soviet Meteorol. Hydrol*., 1976, Le Dimet and T., …) globally adjusts a model to observations distributed over a given time interval. Achieved through minimisation of a scalar objective function of the form

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

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

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

which measures the integrated misfit between a model solution and the observations. Model error can be taken into account through *Weak Constraint Variational Assimilation*.

Minimization made possible through the *adjoint* of dynamical model $x_{k+1} = M_k(x_k)$ (as well as of observation operator $y_k = H_k(x_k)$)

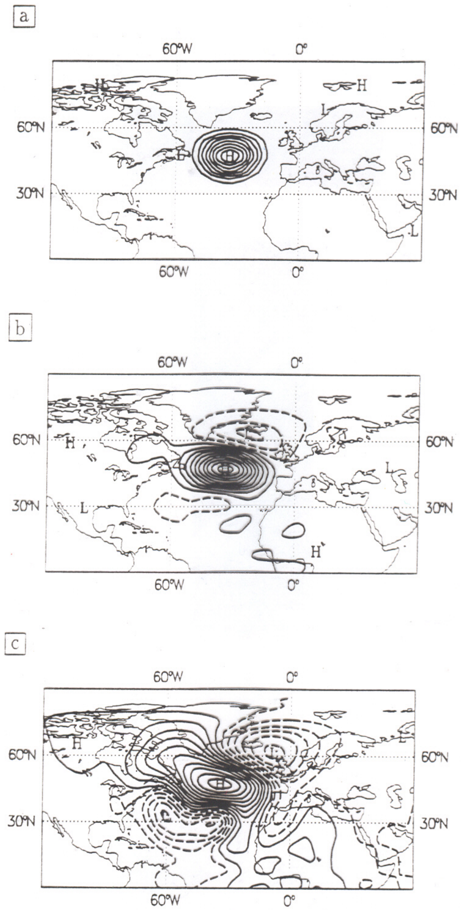
'4D-Var'

Both algorithms are empirical extensions to weakly nonlinear and non-Gaussian situations of algorithms which produce the same results in linear situations. And achieve bayesian estimation in linear and additive gaussian situations.

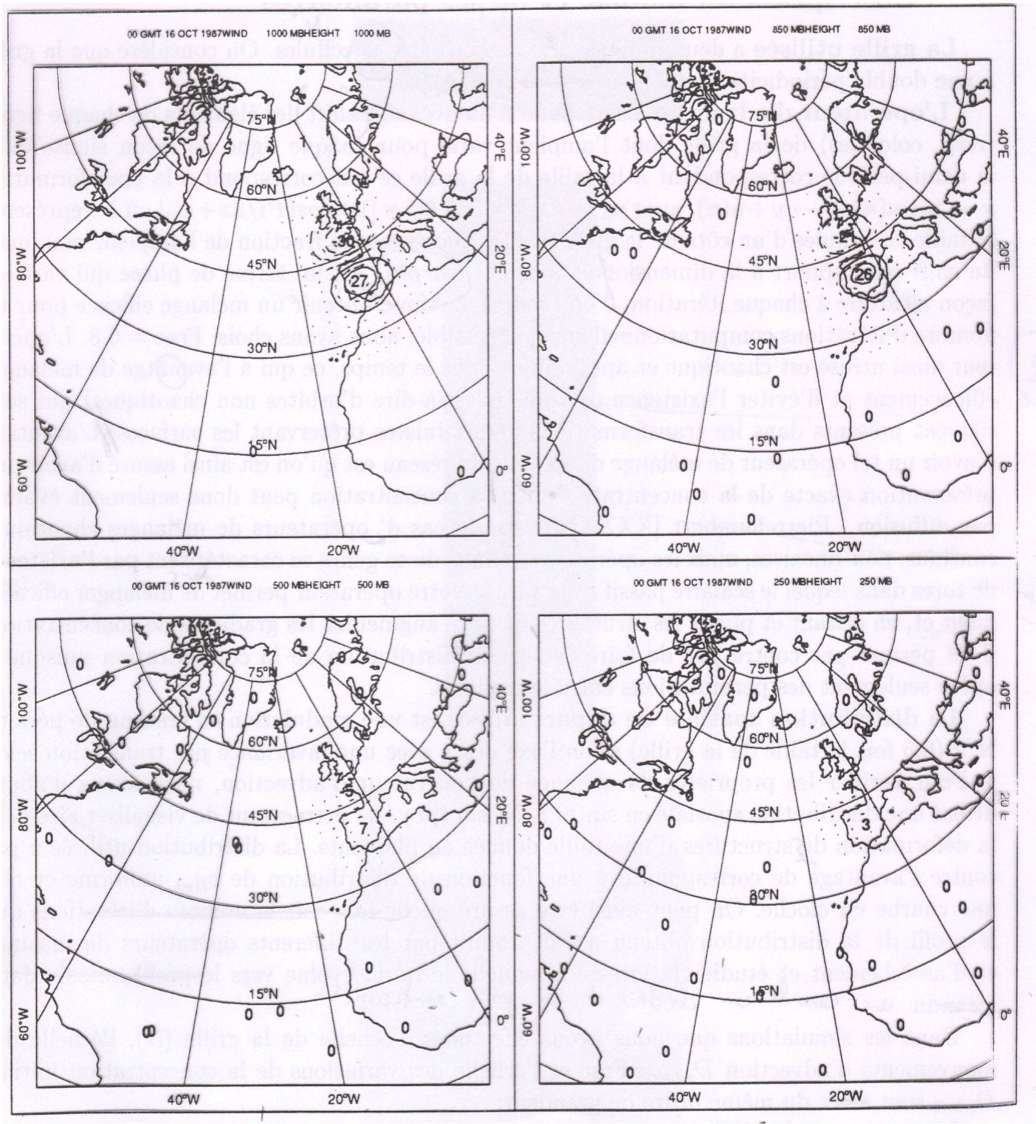
Their success in meteorological and oceanographical situation lies in the fact that they are able to take into account the temporal evolution of the uncertainty on the state of the flow, and in particular of the growth in that uncertainty caused by the 'instabilities of the day'.

Kalman Filter does explicitly compute the evolution of the uncertainty, either through the evolution of the covariance matrix of the estimation error (standard linear filter) or through the evolution of the ensemble elements (EnKF). Variational Assimilation 'knows' of the instabilities of the day through the dynamical model. If an adjoint code is used, it is through that adjoint that the information on the instabilities is carried backwards to the control variable at initial time.

Particle Filters (P. J. van Leeuwen and colleagues) are independent of any gaussian or linear hypothesis.

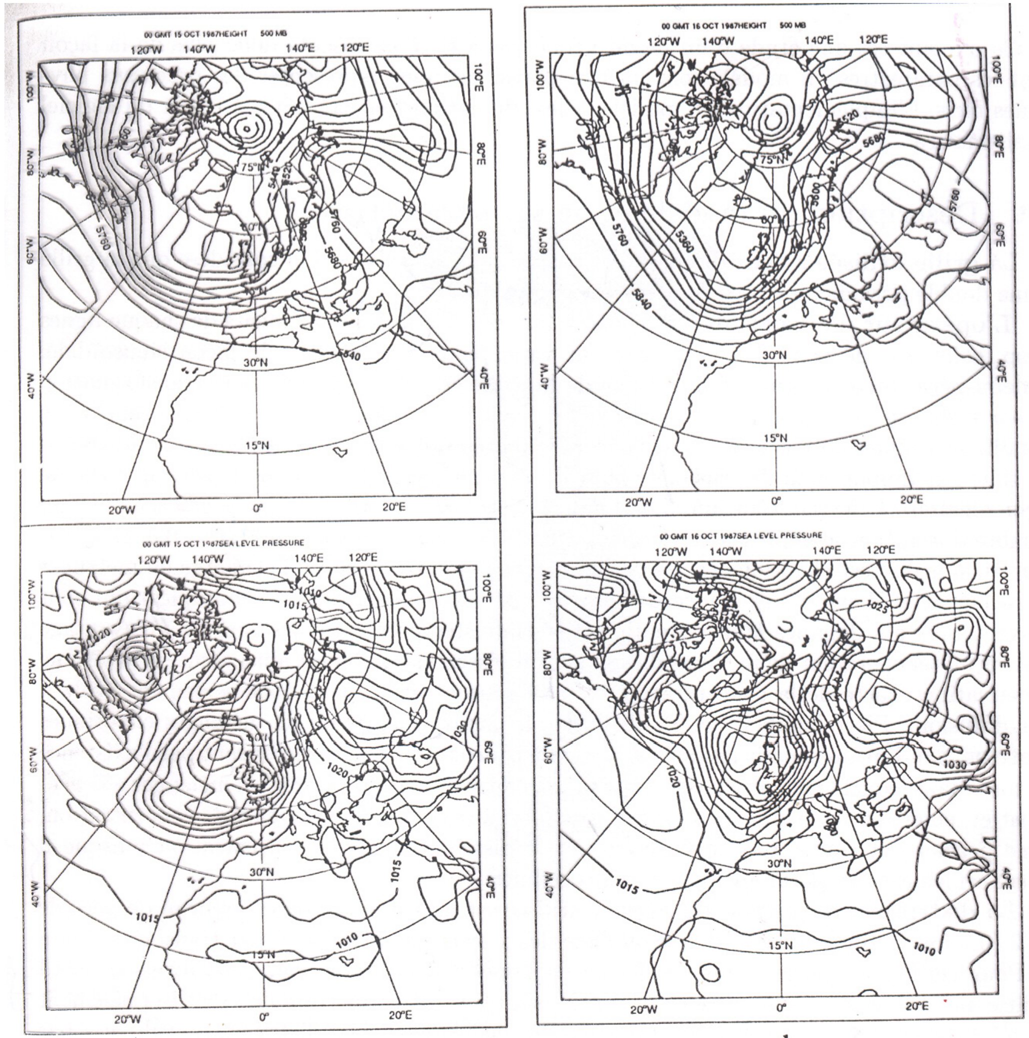


Temporal evolution of the 500-hPa geopotential autocorrelation with respect to point located at 45N, 35W. From top to bottom: initial time, 6- and 24-hour range. Contour interval 0.1. After F. Boutier.



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)

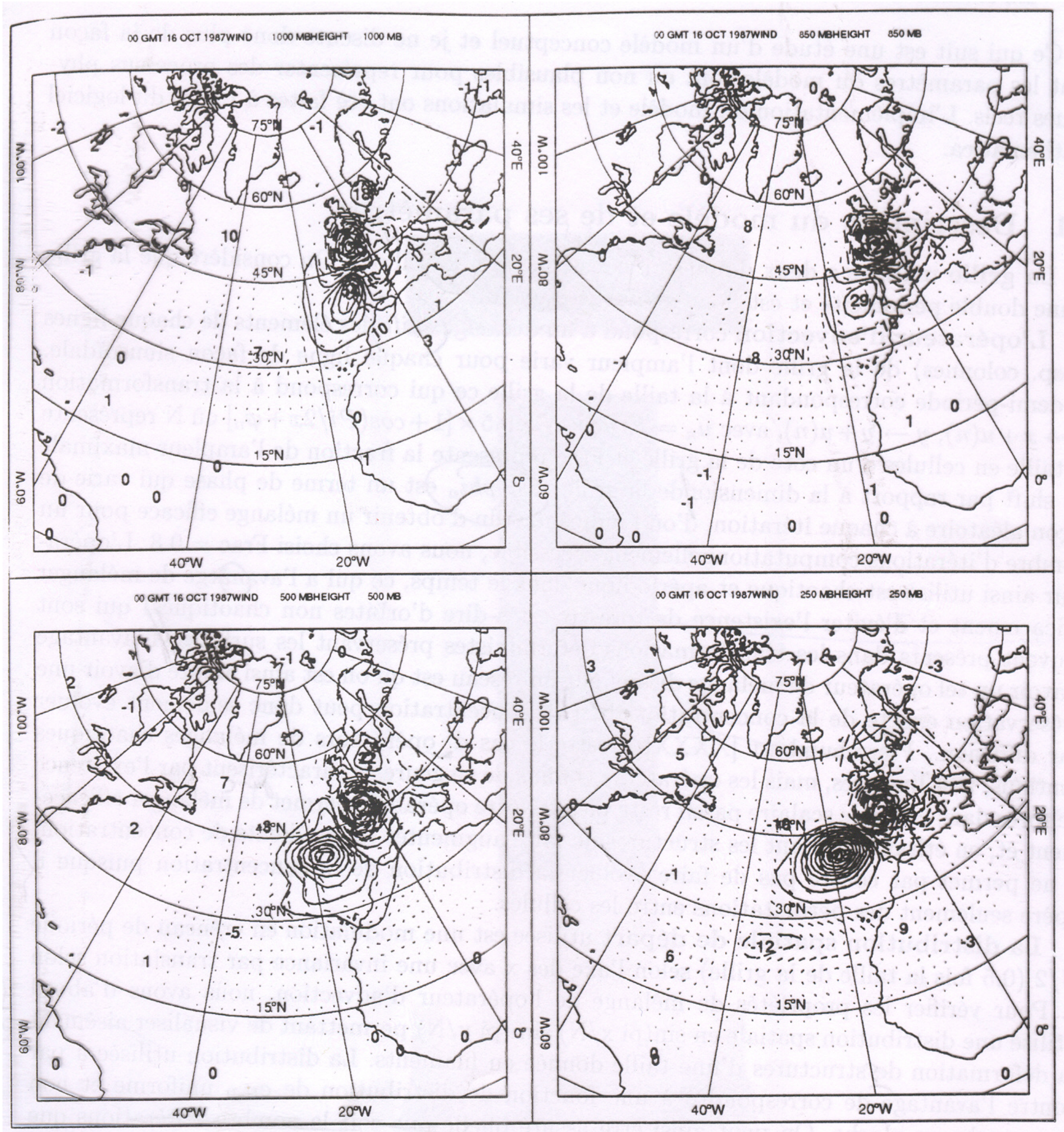
Thépaut *et al.*, 1993, *Mon. Wea. Rev.*, **121**, 3393-3414



b d

FIG. 1. Background fields for 0000 UTC 15 October–0000 UTC 16 October 1987. Shown here are the Northern Hemisphere (a) 500-hPa geopotential height and (b) mean sea level pressure for 15 October and the (c) 500-hPa geopotential height and (d) mean sea level pressure for 16 October. The fields for 15 October are from the initial estimate of the initial conditions for the 4DVAR minimization. The fields for 16 October are from the 24-h T63 adiabatic model forecast from the initial conditions. Contour intervals are 80 m and 5 hPa.

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

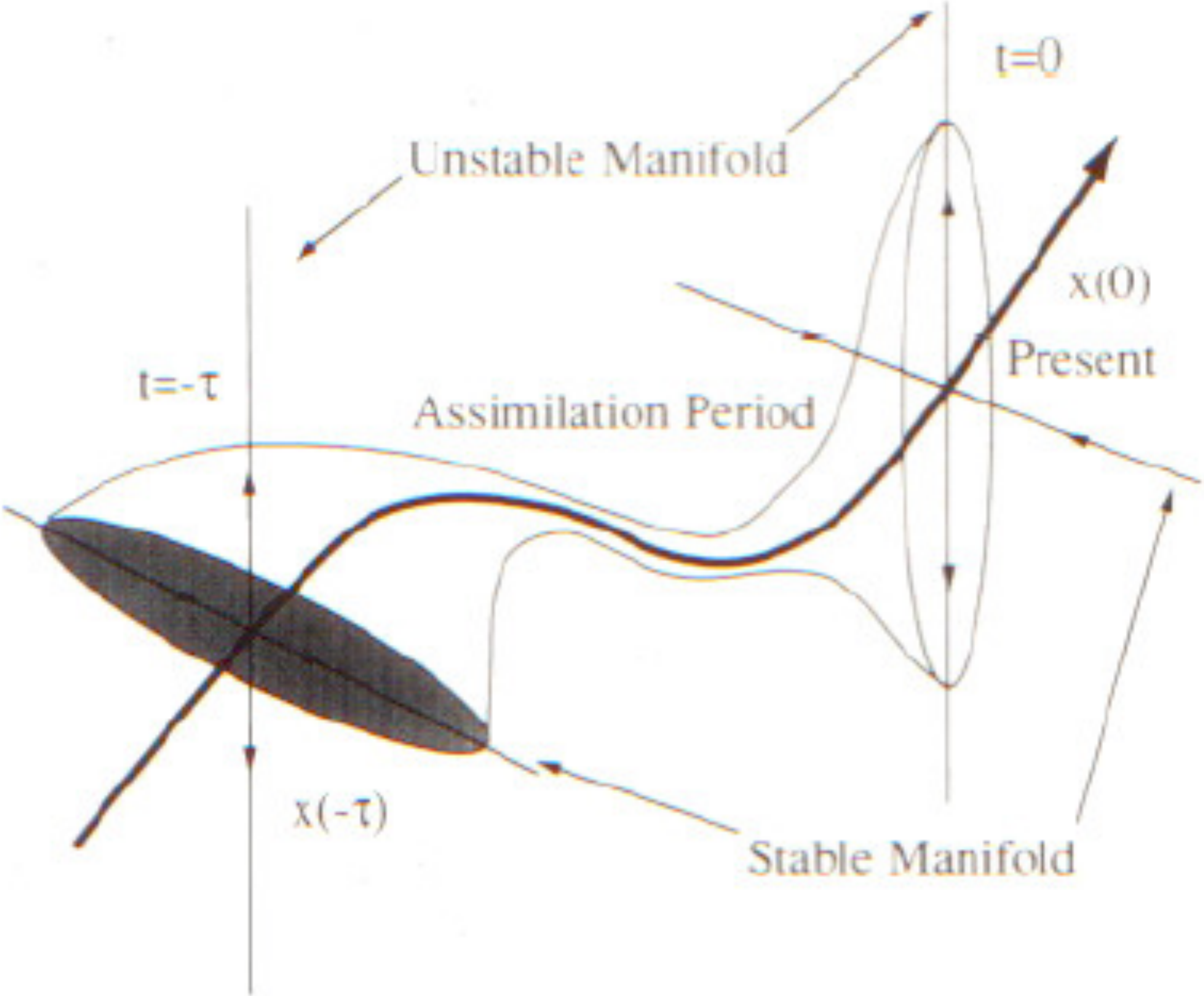


Same as before, but at the end of a 24-hr 4D-Var

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

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 (Pires, Vautard and T.,, *Tellus*, 1996).

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



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

Consequence : 4D-Var assimilation, 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.

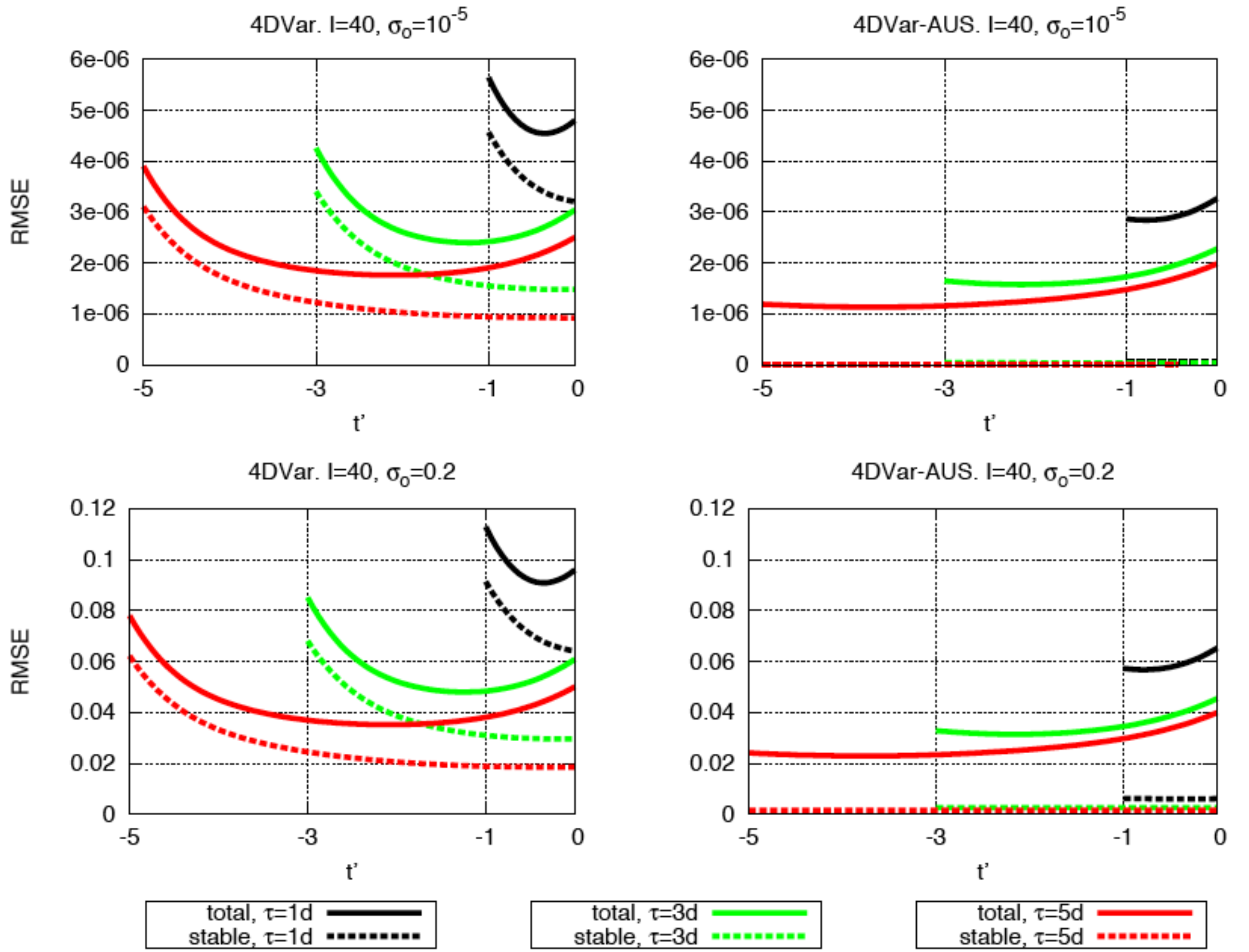
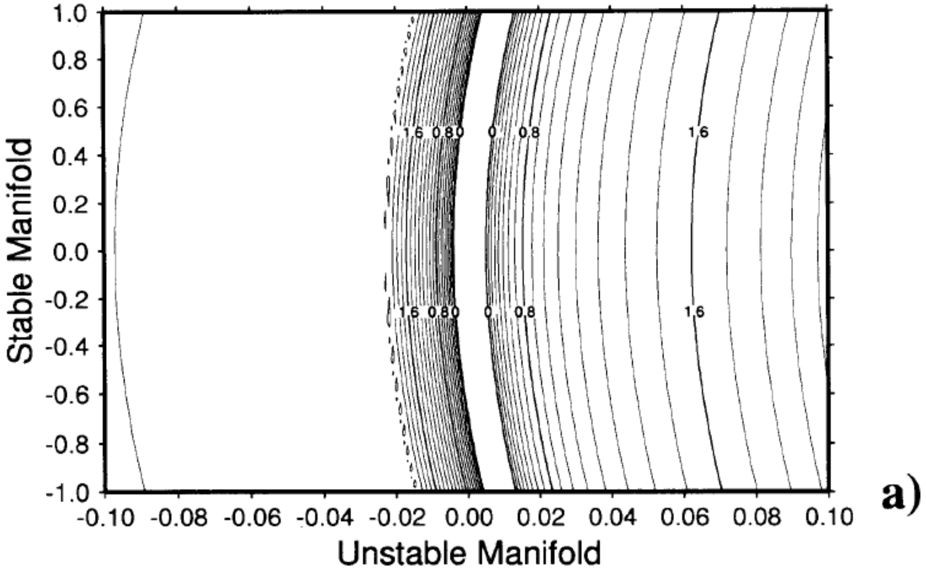


Figure 3. Time average RMS error within 1, 3, 5 days assimilation windows as a function of $t' = t - \tau$, with $\sigma_o = .2, 10^{-5}$ for the model configuration $I = 40$. Left panel: 4DVar. Right panel: 4DVar-AUS with $N = 15$. Solid lines refer to total assimilation error, dashed lines refer to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



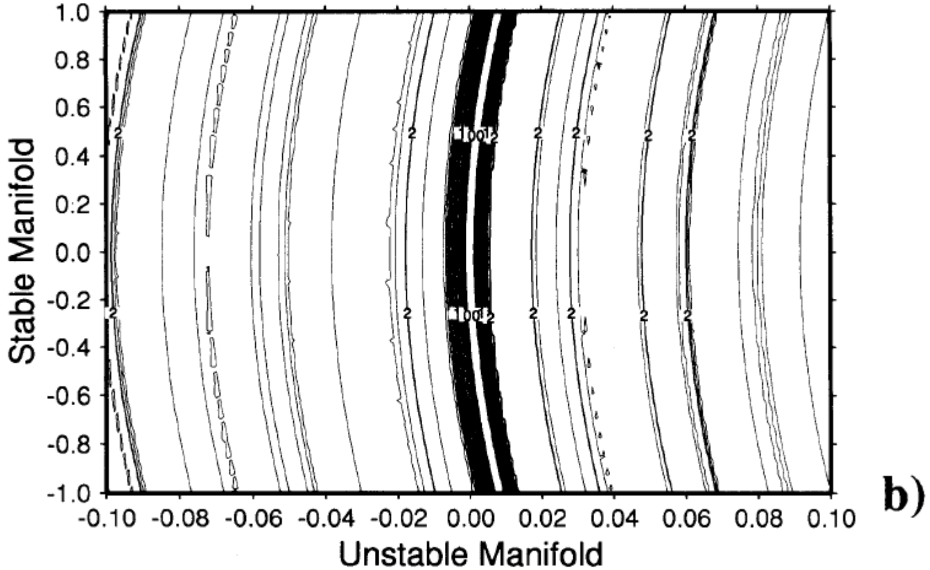


Fig. 3. Variations of the error-free forward cost-function $J'_e(\tau, \hat{x}, x)$ (Lorenz system) in the plane spanned by the stable and unstable directions, as determined from the tangent linear system (see text), and for $\tau = 6$ (panel (a)) and $\tau = 8$ (panel (b)) respectively. The metric has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

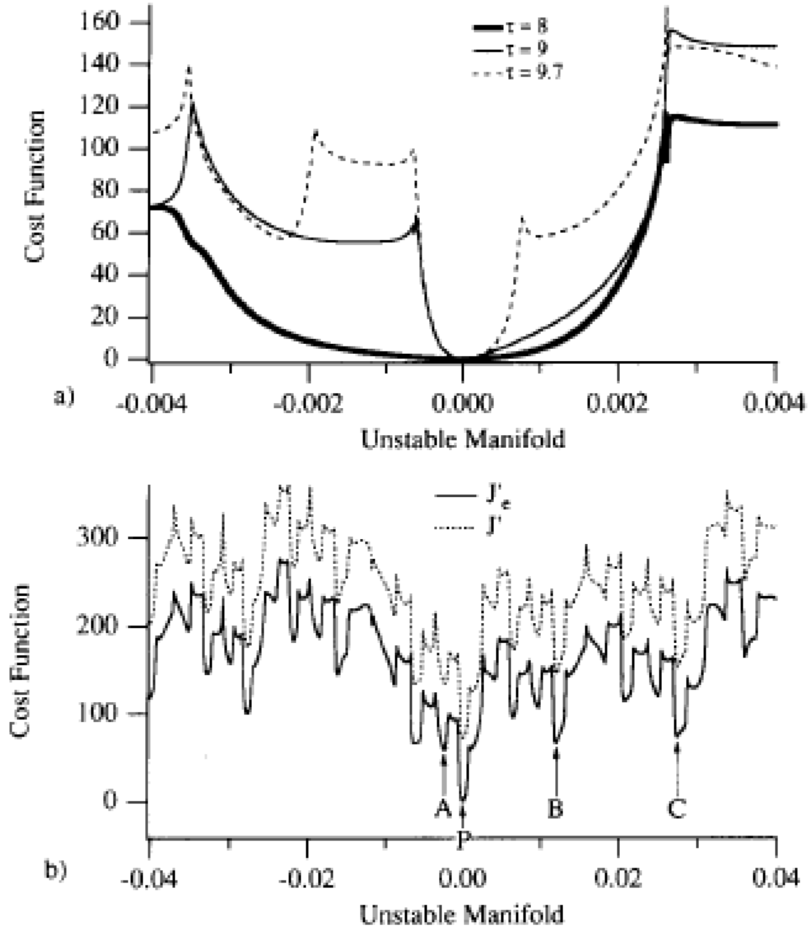
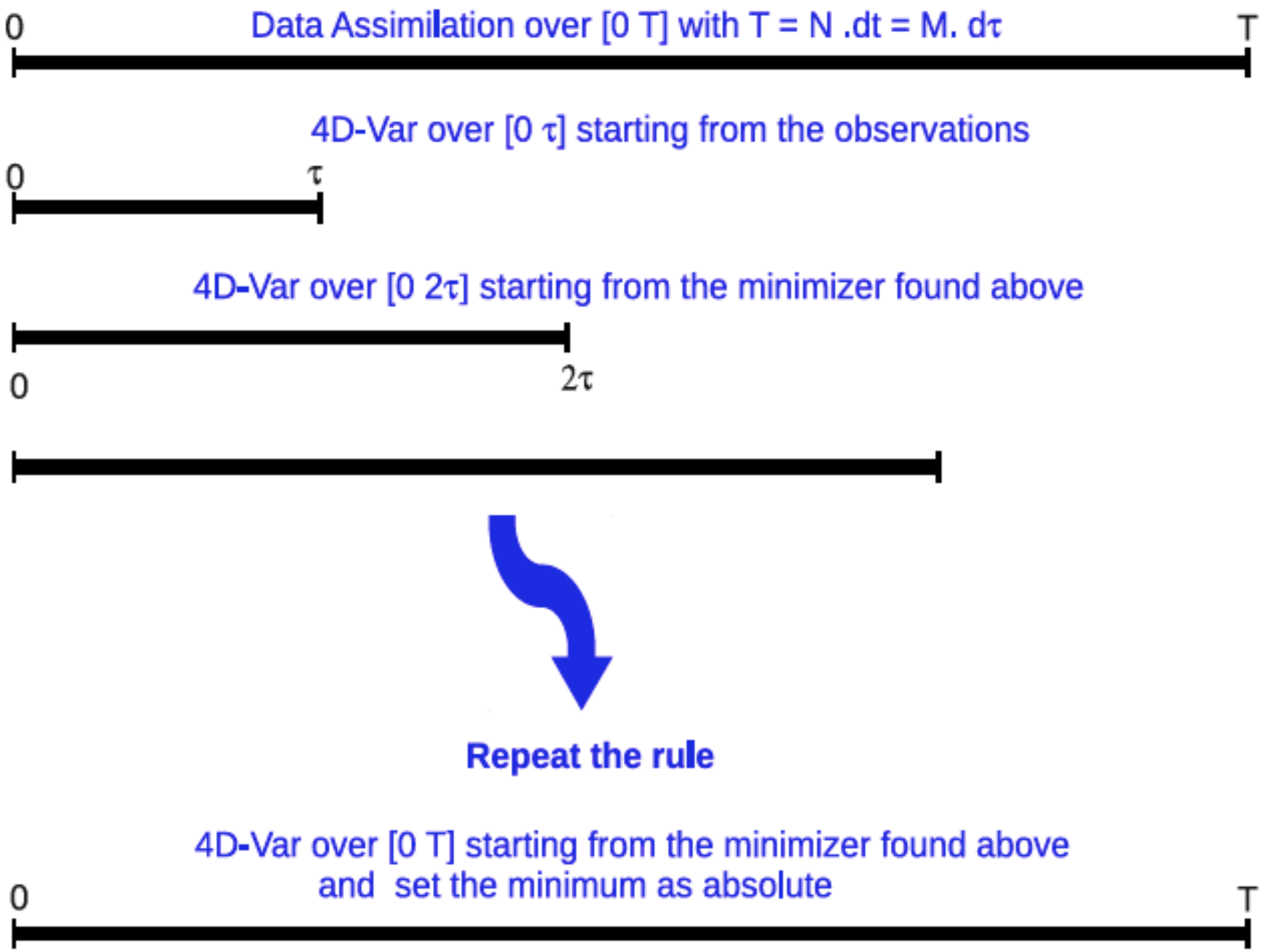


Fig. 4. Panel (a): Cross-section of the error-free forward cost-function $J'_e(\tau, \hat{x}, x)$ along the unstable manifold, for various values of τ. Panel (b). As in panel (a), for $\tau = 9.7$, and with a display interval ten times as large, respectively for the error-free forward cost-function $J'_e(\tau,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

Quasi-Static Variational Assimilation (QSVA)



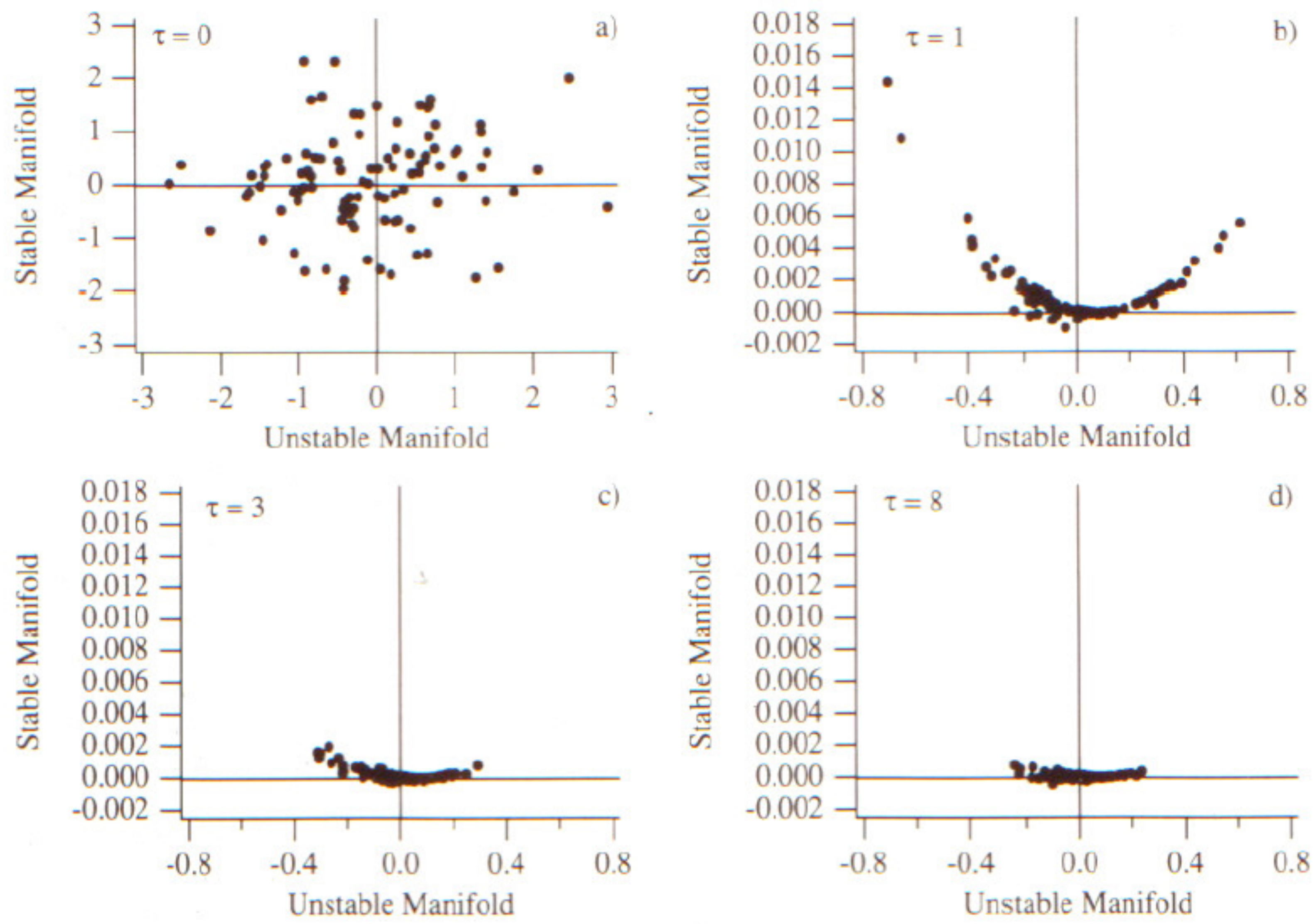


Fig. 7. Projection of the 100 minimizing solutions, at the end of the assimilation period, onto the plane spanned by the stable and unstable directions, defined as in Fig. 3. Values of τ are indicated on the panels. The projection is not an orthogonal projection, but a projection parallel to the local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

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

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

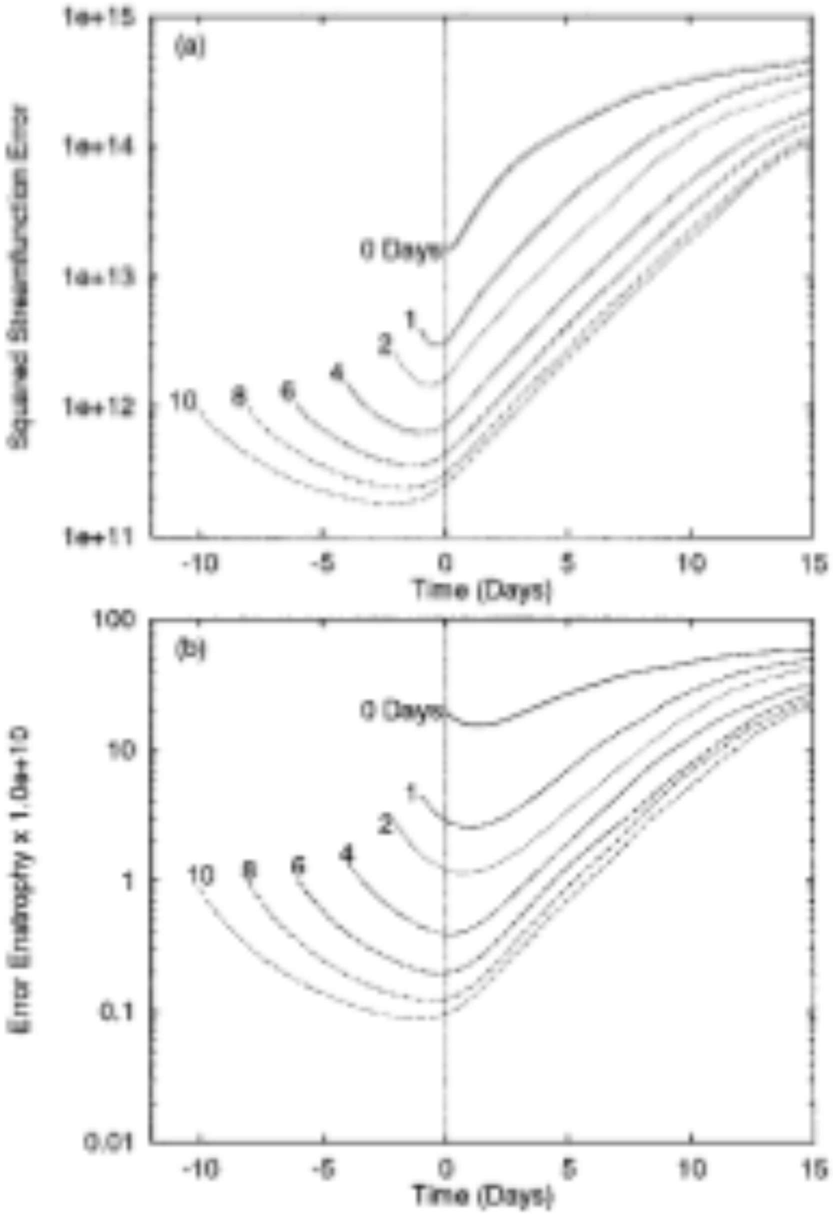


Fig. 5. Median values of the (a) streamfunction squared error, and (b) enstrophy error for the 200 forecast set as a function of forecast time and of the assimilation time T_a.

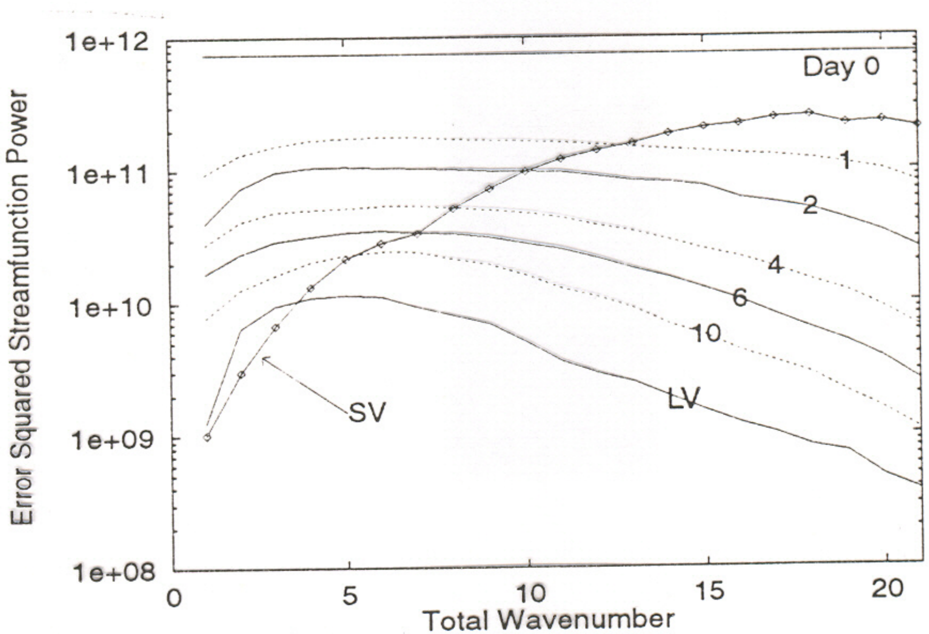


Fig. 5 Time averaged power spectra of the day 0 error in the squared streamfunction norm for the 200 forecast experiments. Individual curves are labeled by their respective values of T_a. Also included are the spectral structure of the leading Lyapunov vector (LV) and the spectral structure of the leading 2-day singular vector (SV) averaged over the 200 forecast experiments. The amplitude for these vectors is arbitrary.

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

(Trevisan and Palatella, 2011, *NPG)*

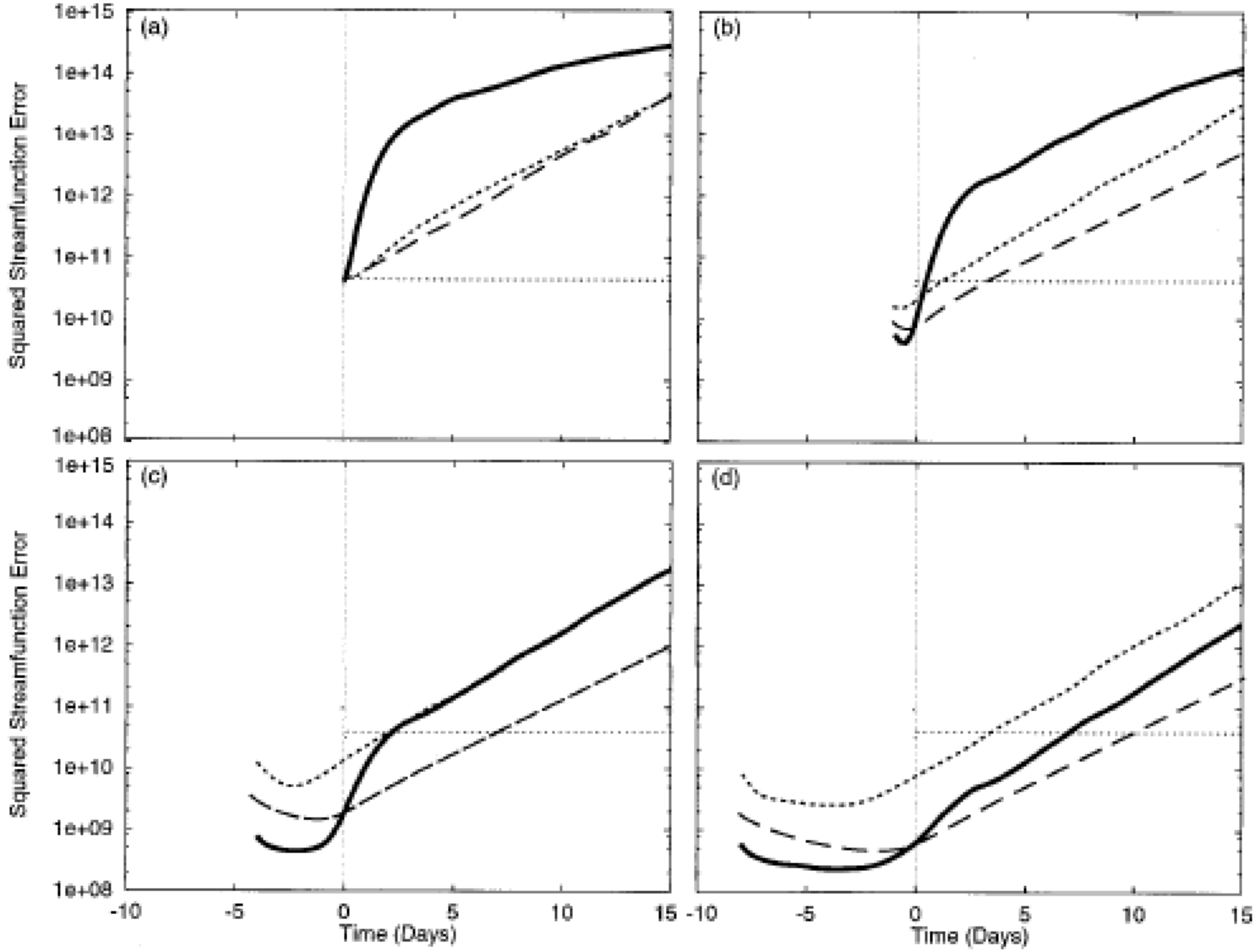


FIG. 1. Assimilation and forecast error in the squared streamfunction norm (units in $m^4 s^{-2}$) for the distributed error experiments described in the text. Assimilation periods are (a) no assimilation, (b) 1 day, (c) 4 days, and (d) 8 days. The solid, dashed, and heavy dotted curves denote SV, WN, and LV error structures, respectively, while the level of observational error is indicated by the light dotted curve.

Swanson, Palmer and Vautard, 2000, *JAS*

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, at least in terms of cost efficiency, to concentrate assimilation in modes that have been unstable in the recent past, where uncertainty is likely to be largest.

Also, 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, D'Isidoro and T., 2010, Four-dimensional variational assimilation in the unstable subspace and the optimal subspace dimension, *Q. J. R. Meteorol. Soc.*.

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).

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 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)

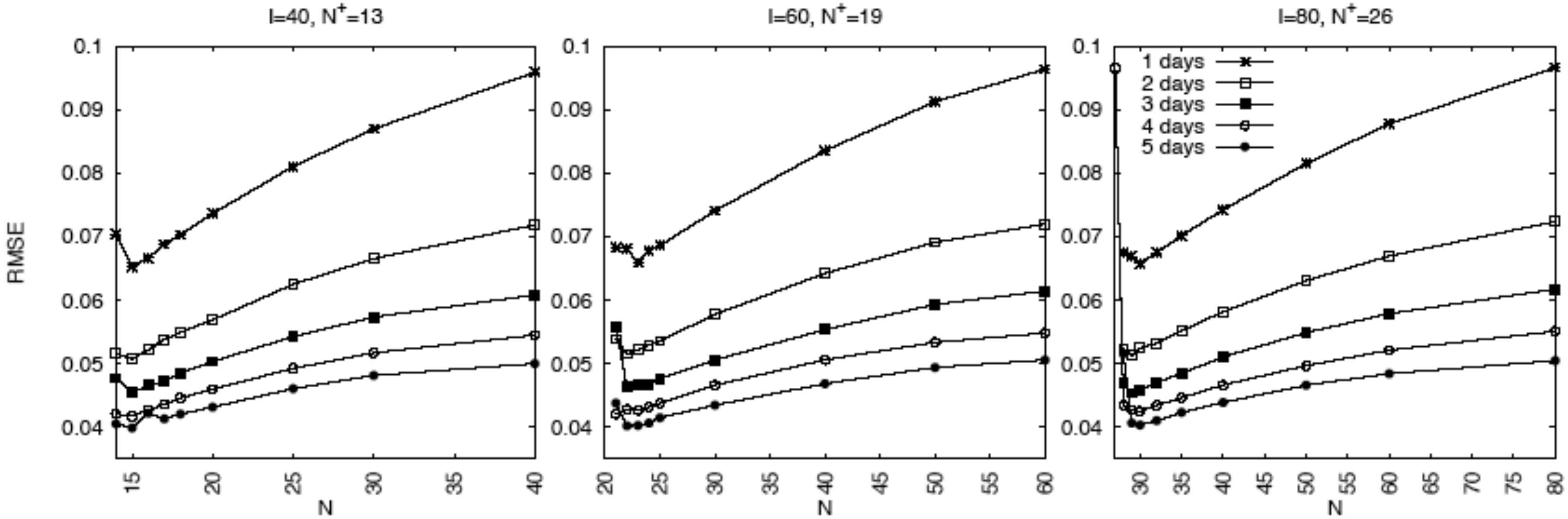


Figure 1. Time average RMS analysis error at $t = \tau$ as a function of the subspace dimension N for three model configurations: I=40, 60, 80. Different curves in the same panel refer to different assimilation windows from 1 to 5 days. The observation error standard deviation is $\sigma_o = 0.2$.

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.

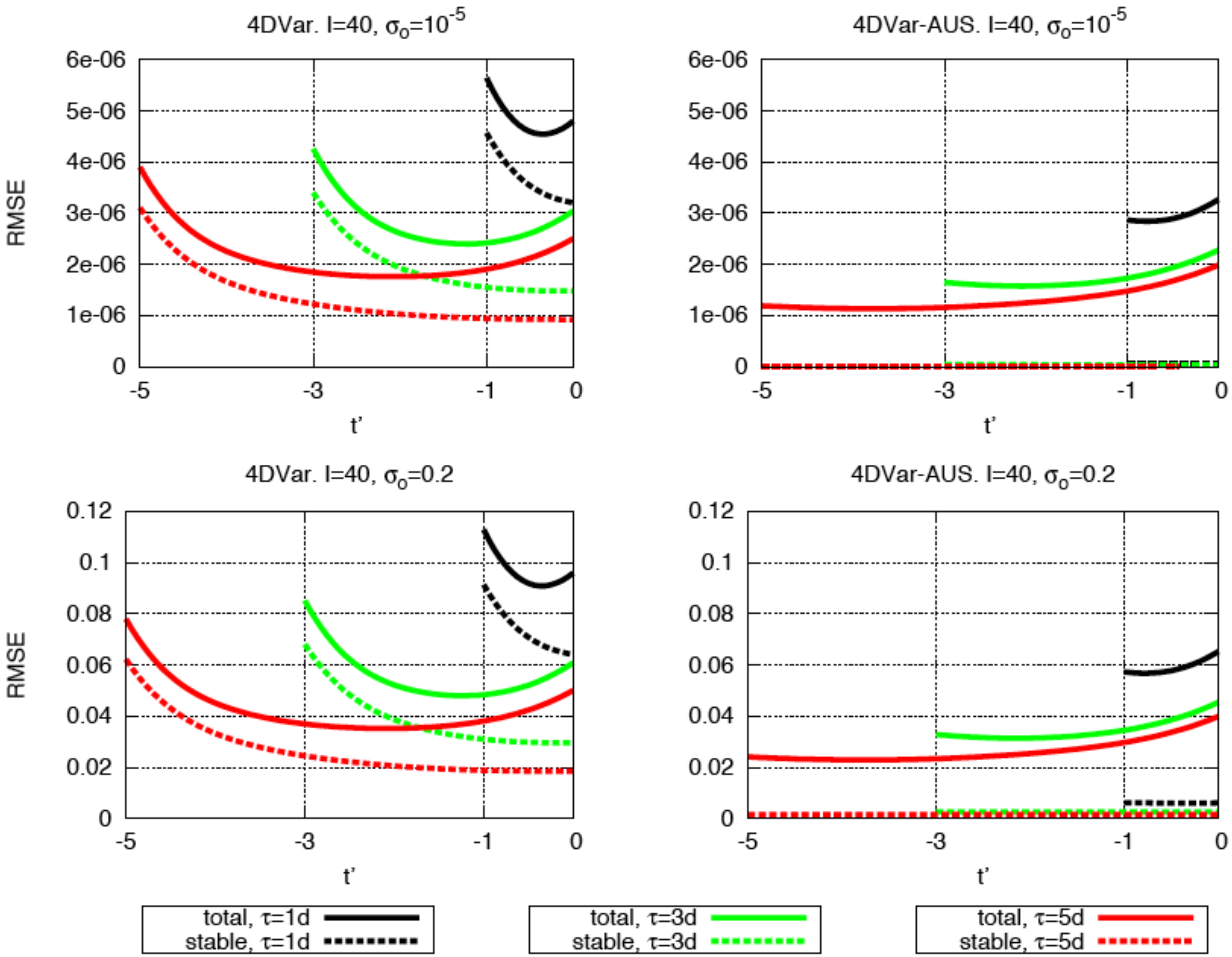


Figure 3. Time average RMS error within 1, 3, 5 days assimilation windows as a function of $t' = t - \tau$, with $\sigma_o = .2, 10^{-5}$ for the model configuration $I = 40$. Left panel: 4DVar. Right panel: 4DVar-AUS with $N = 15$. Solid lines refer to total assimilation error, dashed lines refer to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

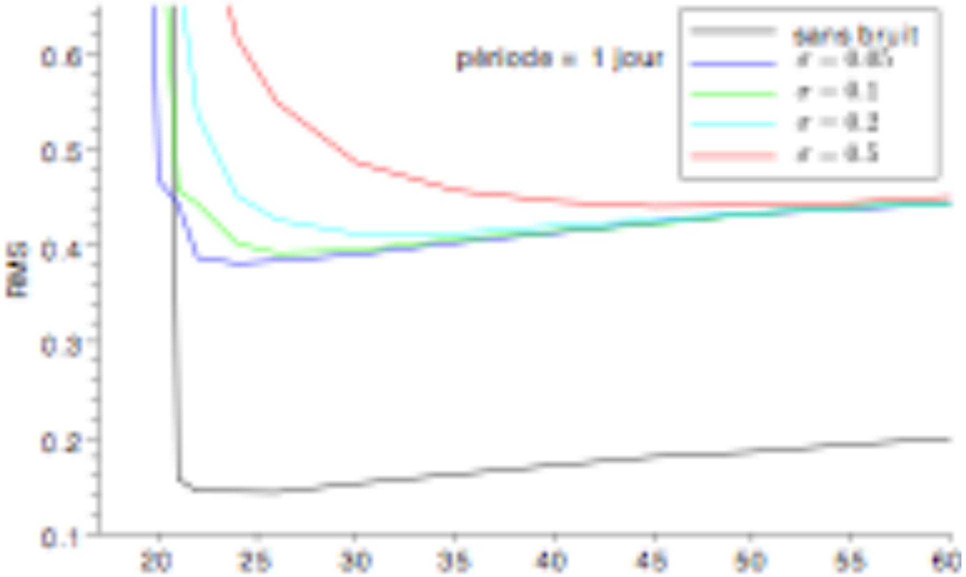
Results show that, if all degrees of freedom that are available to the model are used, the minimization process introduces components along the stable modes of the system, in which no error is present, in order to ensure a closer fit to the observations. This degrades the closeness of the fit to reality.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

Now, consider for instance a Kalman Filter, which carries in time an explicit estimate of the uncertainty on the current state of the system. If it is properly implemented (*i.e.*, if it produces a reliable estimate of the uncertainty), it will know that the uncertainty is concentrated in the unstable subspace, and will not need (except maybe for economy of computation) to explicitly restrict the analysis to that subspace.

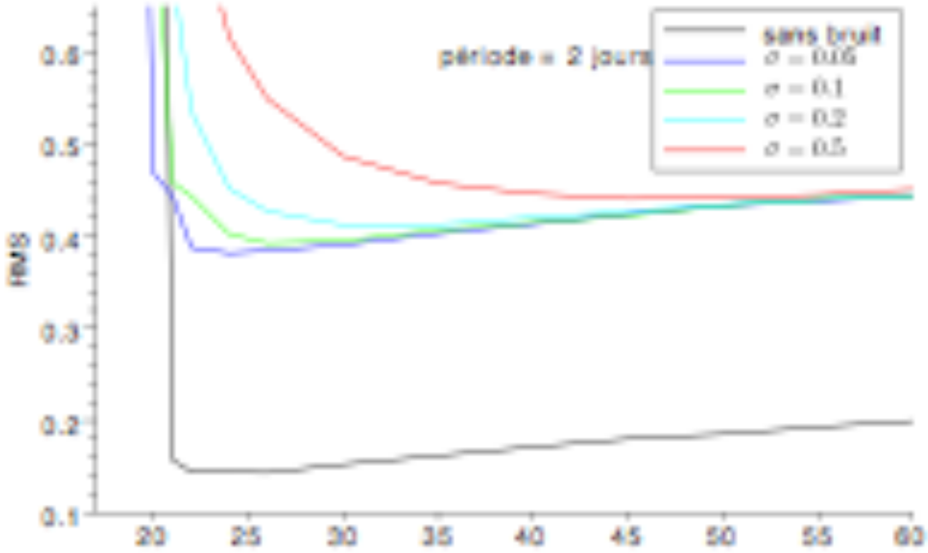
Trevisan and Palatella (*NPG*, 2011). In the Extended Kalman Filter, and in the case of a perfect model, uncertainty on the state of the system concentrates on the subspace spanned by the backwards Lyapunov vectors associated with non-negative exponents. Restricting the assimilation to that subspace is asymptotically equivalent to the full Filter.

Bocquet and Carrassi (*Tellus*, 2017)

Presentation by J. M. López yesterday, and other presentations in this session (Uboldi, Palatella, Bocquet *et al.*, Grudzien, …)



τ = 1 day



τ = 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).

Impact of model errors

- Impact of model errors (other talks) ?

- All the above has been done in the context of the tangent linear approximation, and of the associated machinery of Lyapunov exponents and vectors. That is sucessful because those exponents and vectors tend rapidly to their asymptotic values (*i.e.,* the systems under consideration have ‘good’ ergodicity).

- What happens when the tangent linear approximation is not valid (thresholds, bounded variables) ? Gaussian anamorphosis ?

Swanson, Palmer and Vautard, 2000, *JAS*

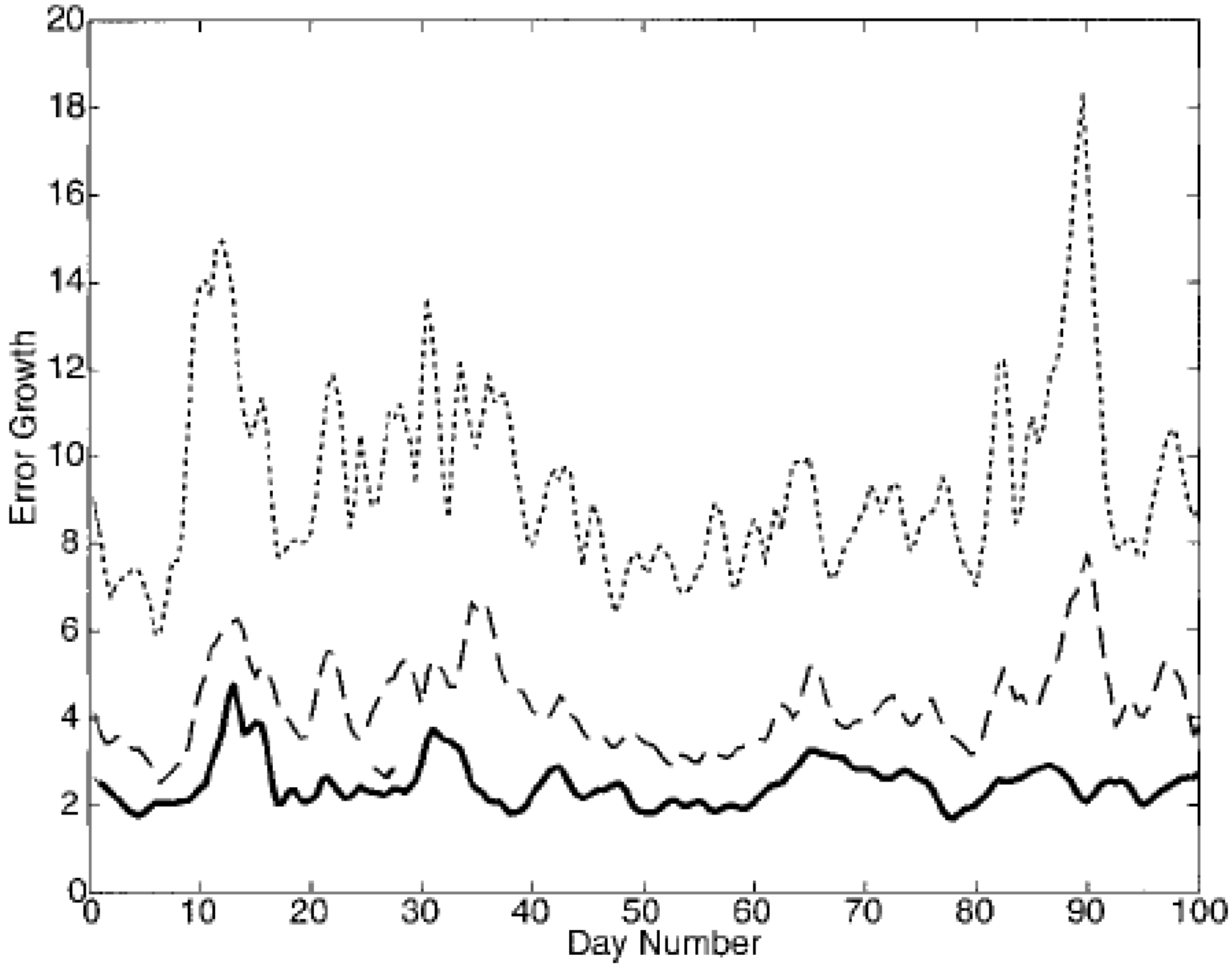


FIG. 3. Error growth over a 4-day period as a function of time in the MM93 model for the leading LV (solid), and the most rapidly growing perturbations in the subspace of 10 LVs (dashed) and 100 LVs (dotted).

A few of the (many) remaining problems

- *Observability*

What can we know from which observations ?

Which space-time distributions of observations completely define the state of the system ?

Does the knowledge of the history of the mass field defines the velocity field (or vice-versa) ? Geostrophic adjustment.

But : system has (infinitely) many scales, is chaotic, and observations are noisy !

Linear case

Data available in the form

$$z = \Gamma x + \zeta$$

where x is unknown, Γ is a known matrix (and ζ is 'error')

Observability $\Leftrightarrow \Gamma$ is one-to-one $\Leftrightarrow$ rank$\Gamma =$ dim$x$ $\Leftrightarrow$ data vector can be decomposed into 'background' plus additional 'observations'

Fukumori *et al*. (1993): oceanic circulation is observable from surface height.

Infinite dimension : Linear first-order EDPs : characteristics

In general, solver may be unbounded (diffusion)

But, in all cases, observability depends only on space-time distribution of observations.

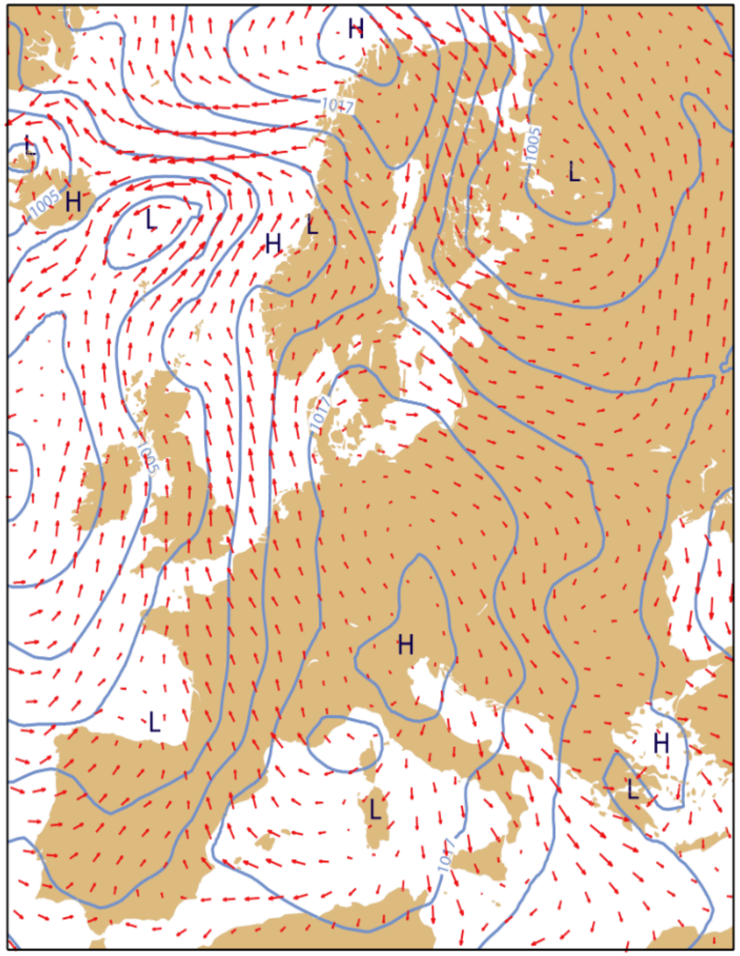
No more true in nonlinear case : observability depends on observed values (advection).

Observability can be studied numerically.

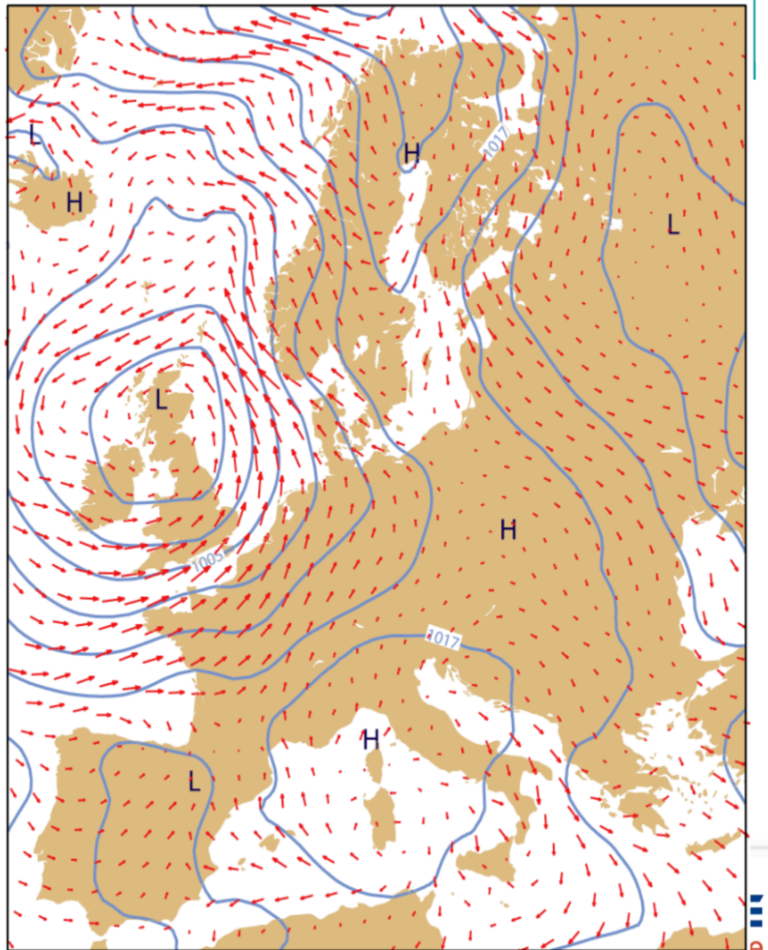
ERA-20C (ECMWF): Ensemble variational reassimilation of surface observations (surface pressure and 10m-wind) over the period 1900-2010

Réanalyse 1er-2 août 1914

PMSL and 10m wind
12UTC 1 August 1914

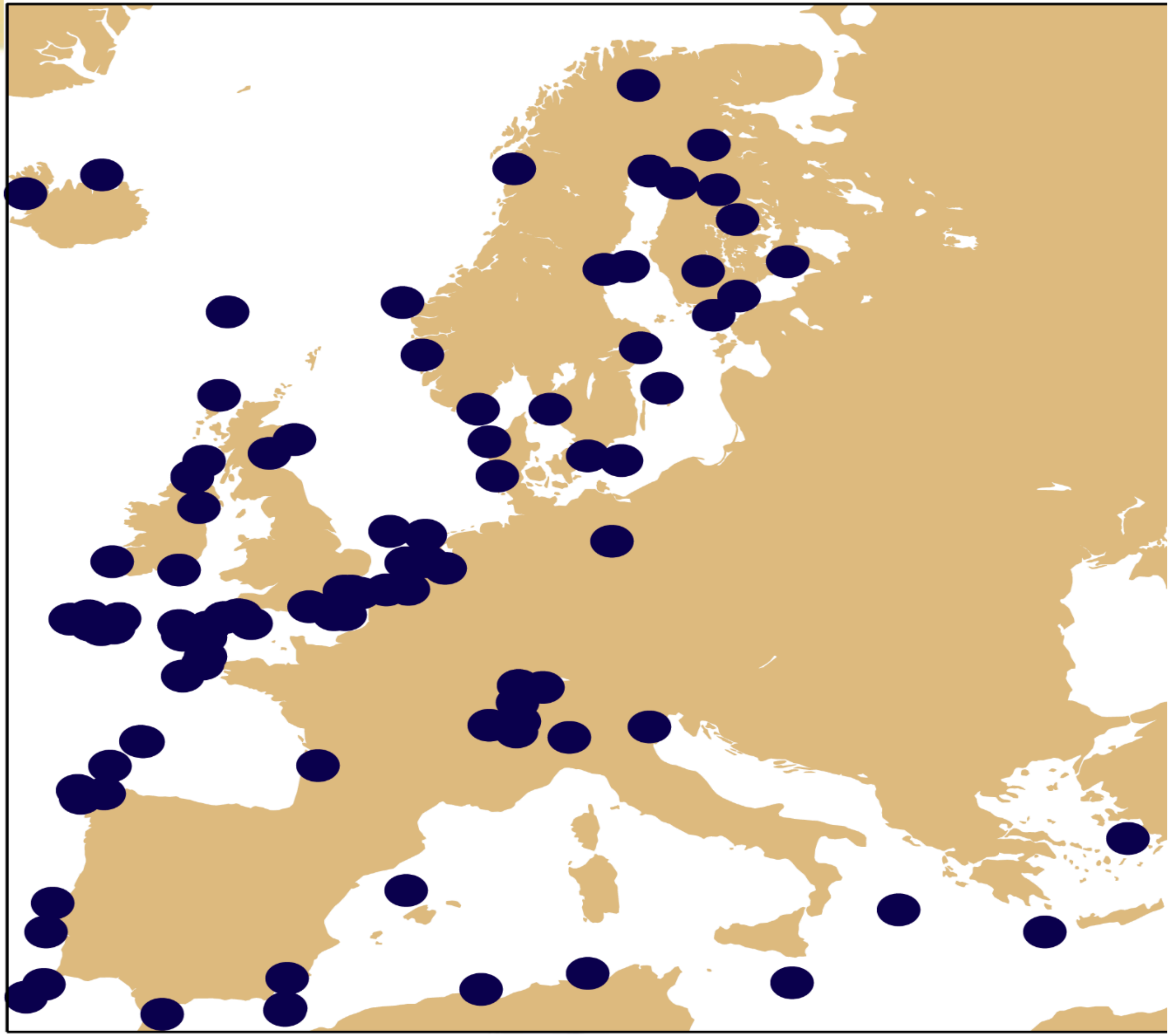


PMSL and 10m wind
12UTC 2 August 1914



Credit Poli and Simmons (ECMWF)

Observation locations
09-15UTC 1 August 1914



Credit
Poli and
Simmons
(ECMWF)

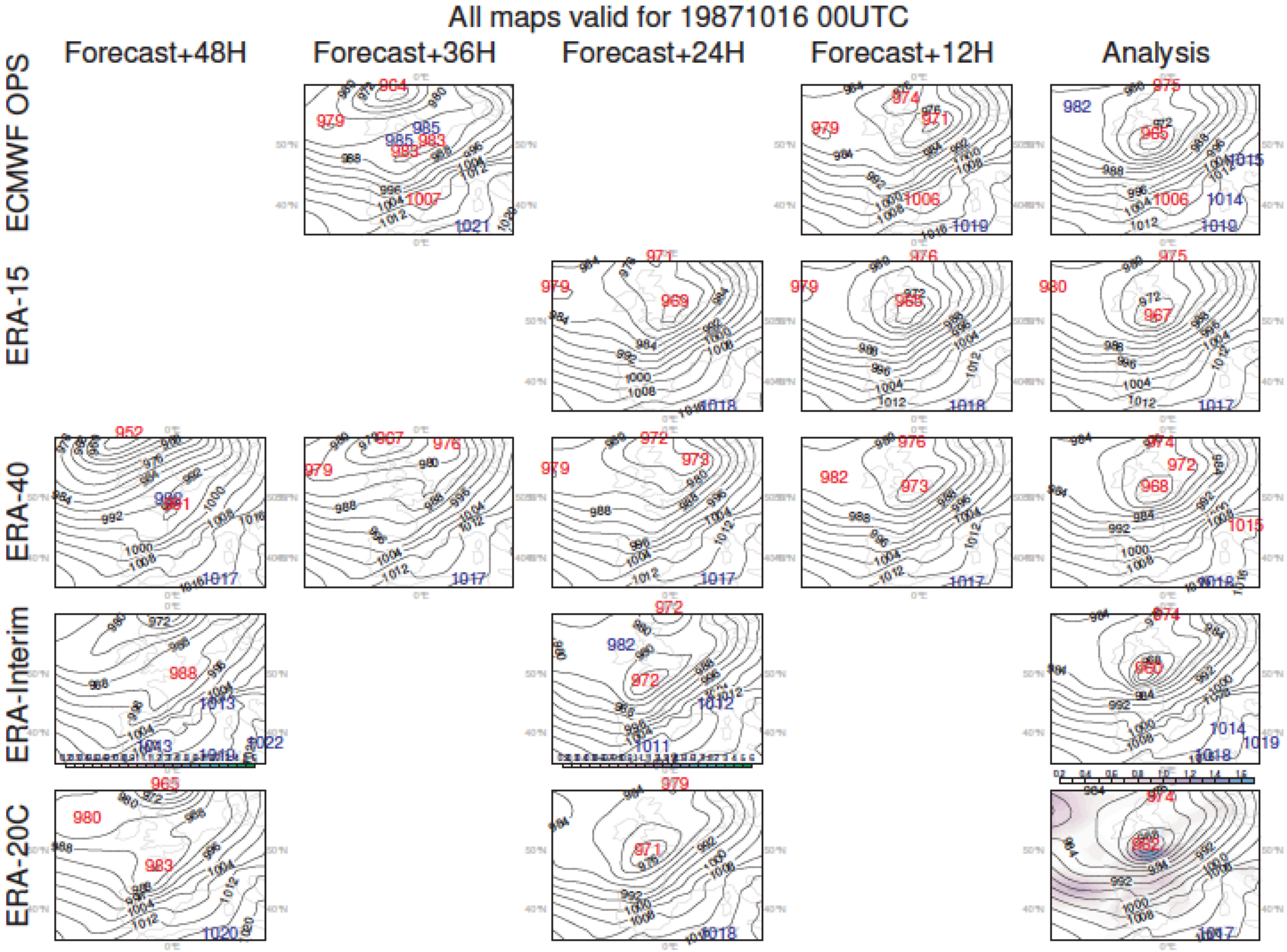


Figure 42: Maps of mean sea level pressure for 16 October 1987, 00 UTC. For ERA-20C analysis, the colour shading shows the ensemble spread.

Poli *et al.*, 2013, ECMWF, ERA Report No 14

Observability. Works by E. Titi (Texas A&M University) and colleagues

Farhat, Lunasin and Titi, 2016, ***CONTINUOUS DATA ASSIMILATION FOR A 2D BÉNARD CONVECTION SYSTEM THROUGH HORIZONTAL VELOCITY MEASUREMENTS ALONE,*** *Journal of Nonlinear Science*, **27**(2017)

2-D incompressible fluid between two horizontal walls maintained at fixed different temperatures $T_{bottom} > T_{top} \Rightarrow$ convective motions transporting heat from bottom to top.

Boussinesq equations for horizontal vector velocity $\boldsymbol{u}$ and normalized temperature θ

$$\boldsymbol{u}_t - \nu\, \Delta \boldsymbol{u} + (\boldsymbol{u}.\nabla)\boldsymbol{u} + \nabla p' - \theta\, \mathbf{e}_2 = 0 \qquad (1a)$$

$$\theta_t - \kappa\, \Delta\theta + (\boldsymbol{u}.\nabla)\theta - \boldsymbol{u}.\mathbf{e}_2 = 0 \qquad (1b)$$

$$\nabla.\boldsymbol{u} = 0 \qquad (1c)$$

+ BC and periodicity in horizontal direction

Assume horizontal component v_1 of velocity of some solution $(\boldsymbol{v}(t),\ \eta(t))$ is observed exactly and corresponding observations are introduced into a numerical integration of eq. (1) modified through following scheme (‘*nudging*’)

$$\boldsymbol{u}_t - \nu\,\Delta\boldsymbol{u} + (\boldsymbol{u}.\nabla)\boldsymbol{u} + \nabla p' - \theta\,\mathbf{e}_2 = -\mu\,(u_1 - v_1)\,\mathbf{e}_1 \qquad \text{with } \mu > 0 \qquad (1'\text{a})$$

Theorem. If the spatial density of the observations is high enough, and nudging coefficient μ is large enough, the ‘nudged’ solution $(\boldsymbol{u}(t),\ \theta(t))$ converges exponentially to observed solution $(\boldsymbol{v}(t),\ \eta(t))$ when t tends to infinity.

Observation of horizontal component of velocity only ensures complete asymptotic observability, and algorithm is available for asymptotic reconstruction of observed solution (but infinite temporal density of observations seems to be necessary for the trheorem to hold).

If there are errors in the observations, error on the nudging solution can be estimated in terms of the errors on the observations.

Assimilation, which originated from the need of defining initial conditions for numerical weather forecasts, has gradually extended to many diverse applications

- Oceanography
- Atmospheric chemistry (both troposphere and stratosphere)
- Oceanic biogeochemistry
- Ground hydrology
- Terrestrial biosphere and vegetation cover
- Glaciology
- Magnetism (both planetary and stellar)
- Plate tectonics
- Planetary atmospheres (Mars, …)
- Reassimilation of past observations (mostly for climatological purposes, ECMWF, NCEP/NCAR)
- Identification of source of tracers
- Parameter identification
- *A priori* evaluation of anticipated new instruments
- Definition of observing systems *(Observing Systems Simulation Experiments)*
- Validation of models
- Sensitivity studies (adjoints)
- …

It has now become a major tool of numerical environmental science