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

Année Universitaire 2019-2020

# Modélisation Numérique de l'Écoulement Atmosphérique et Assimilation de Données

Olivier Talagrand

Cours 7

12 Mai 2020

- Assimilation in the Unstable Subspace (AUS)

- Ensemble variational assimilation (EnsVAR)

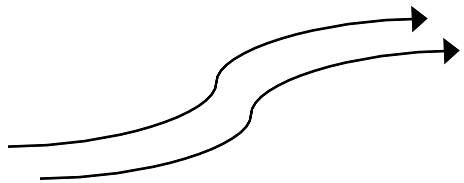
- Particle Filters *or* History of Numerical Weather Prediction

## Return on a few basics

- Basic (nonlinear) model

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

- Perturbation $\delta x_0$ at time $0$. Resulting perturbation $\delta x_k$ evolves in time according to



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

$$= M_k'(x_k)\ \delta x_k + o(\delta x_0)$$

where $M_k'(x_k)$ is jacobian of $M_k$ at point $x_k$.

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

is *tangent linear model* along solution $x_k$.

## **Return on a few basics** (continuation)

*Tangent linear model*

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

*Adjoint model*

$$\lambda_k = [M_k'(x_k)]^{\mathrm{T}}\,\lambda_{k+1}$$

Describes evolution with respect to $k$ of gradient of a scalar function $\mathcal{J}$ with respect to $x_k$.

## And a little more …

*Tangent linear model* $\delta\xi_{k+1} = M_k'(x_k)\, \delta\xi_k$

How does $\delta\xi_k$ vary as $k$ tends to infinity ? If system is unstable (globally stable), one can expect $\delta\xi_k$ to increase to infinity (to decrease to 0).

If $|\delta\xi_k| \sim \exp(\lambda k)$ then $\lim (1/k) \ln |\delta\xi_k| = \lambda$

Theorem (Oseledec)

For a large class of systems which possess an asymptotic attractor, the quantity $(1/k) \ln |\delta\xi_k|$ tends to one of a finite number of values

$$\lambda_n \leq \ \ldots \leq \lambda_2 \leq \lambda_1$$

(where $n$ is the dimension of the system under consideration)

The $\lambda_i$'s are the *Lyapunov exponents* of the system

## And still a little more …

$$\lambda_n \leq \ \ldots \leq \lambda_2 \leq \lambda_1$$

The Lyapunov exponents characterize the behaviour of infinitesimally small perturbations imposed on the state of the system.

There exists a sequence of linear subsets of the whole state space $S$

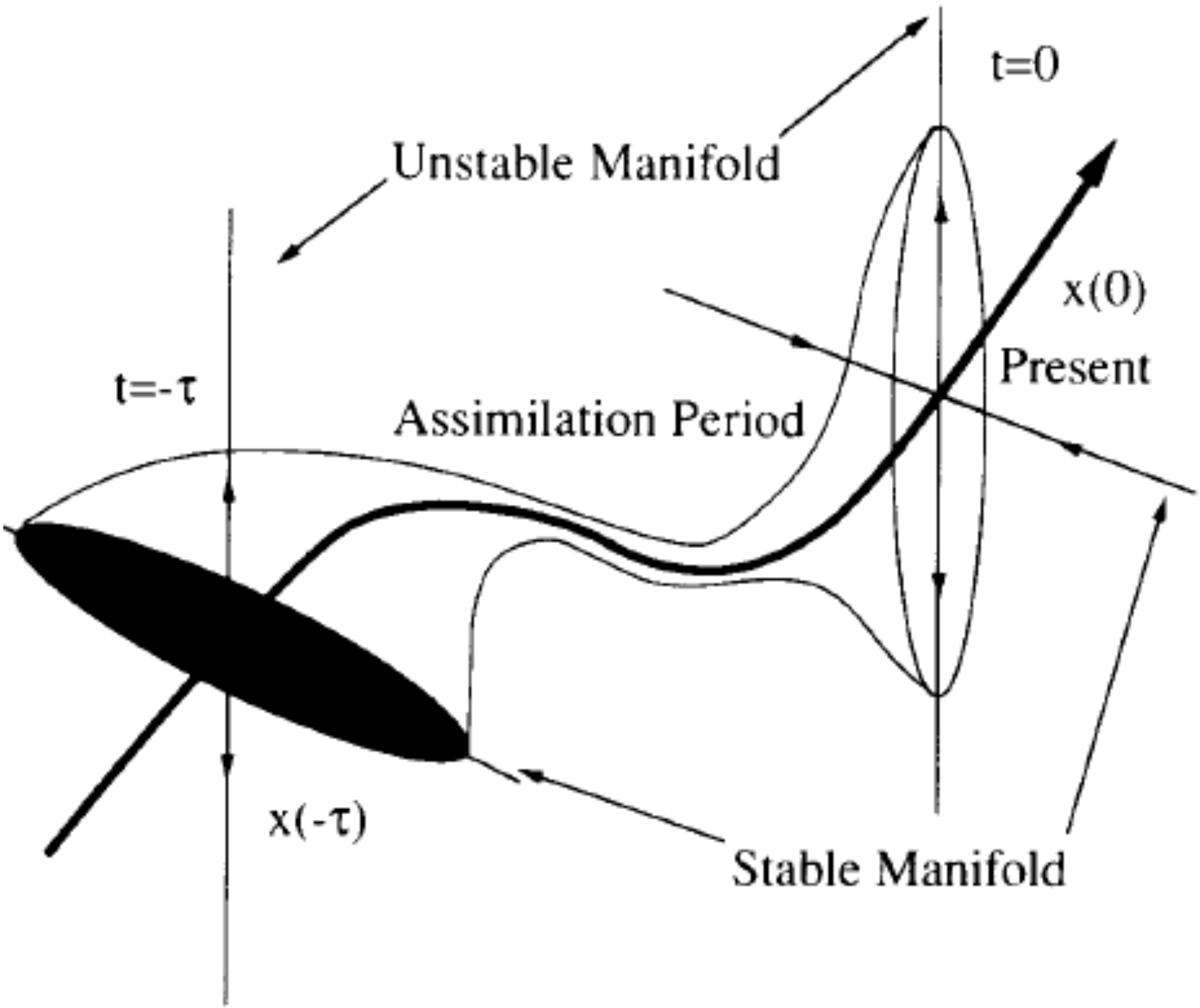
$$\emptyset = \mathcal{E}_{n+1} \subset \mathcal{E}_n \subset \ldots \subset \mathcal{E}_2 \subset \mathcal{E}_1 = S$$

such that the quantity $(1/k)\ \ln\left|\,\delta\xi_k\right|$ tends to $\lambda_i$ when $\delta\xi_0$ belongs to $\mathcal{E}_i - \mathcal{E}_{i+1}$.

The presence of (at least) one positive Lyapunov exponent is a sign of sensitivity to initial conditions, *i. e. chaos*.

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.

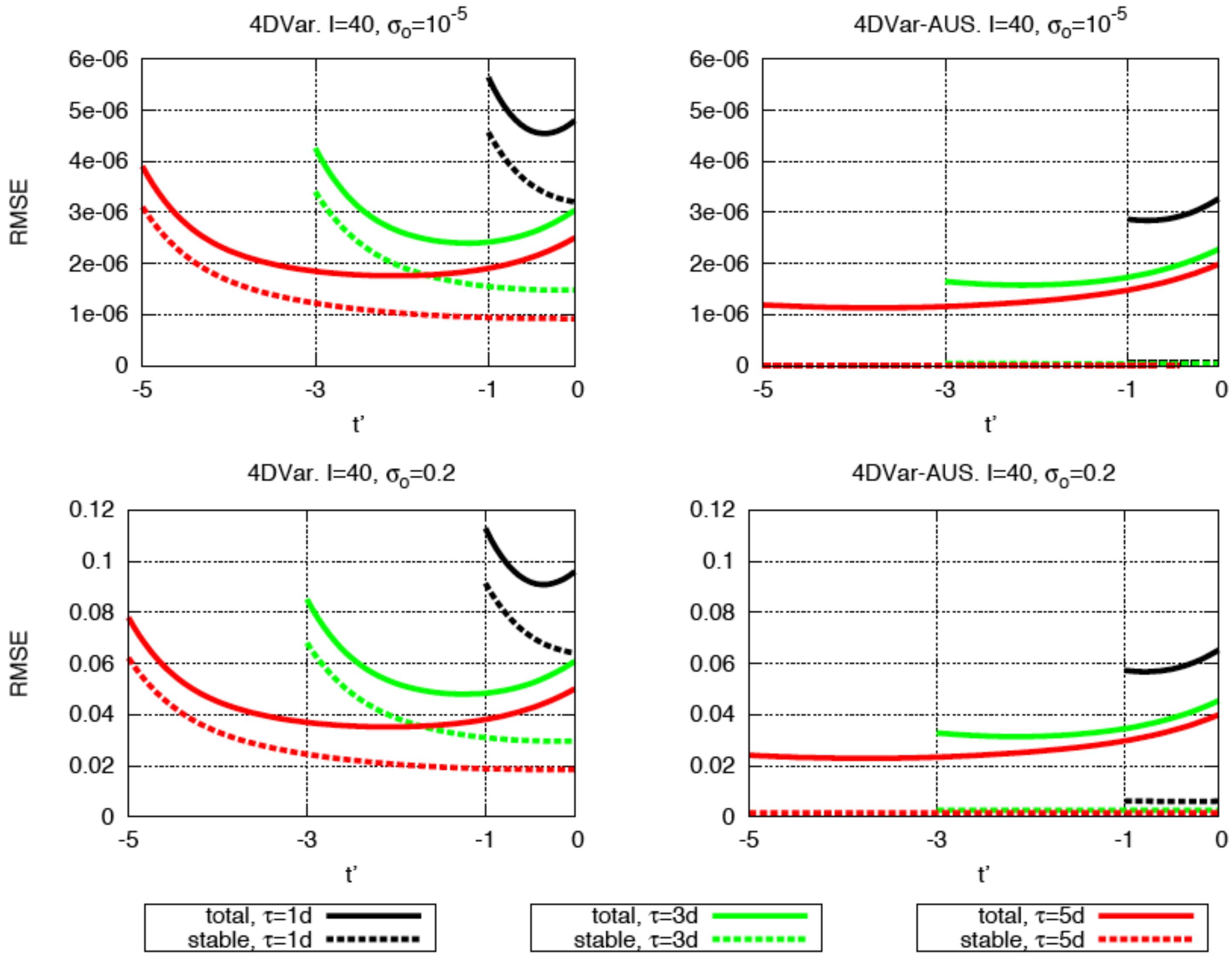



Figure 3. Time average RMS error within 1, 3, 5 days assimilation windows as a function of $t' = t - \tau$, with $\sigma_o = .2, 10^{-5}$ for the model configuration $I = 40$. Left panel: 4DVar. Right panel: 4DVar-AUS with $N = 15$. Solid lines refer to total assimilation error, dashed lines refer to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

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

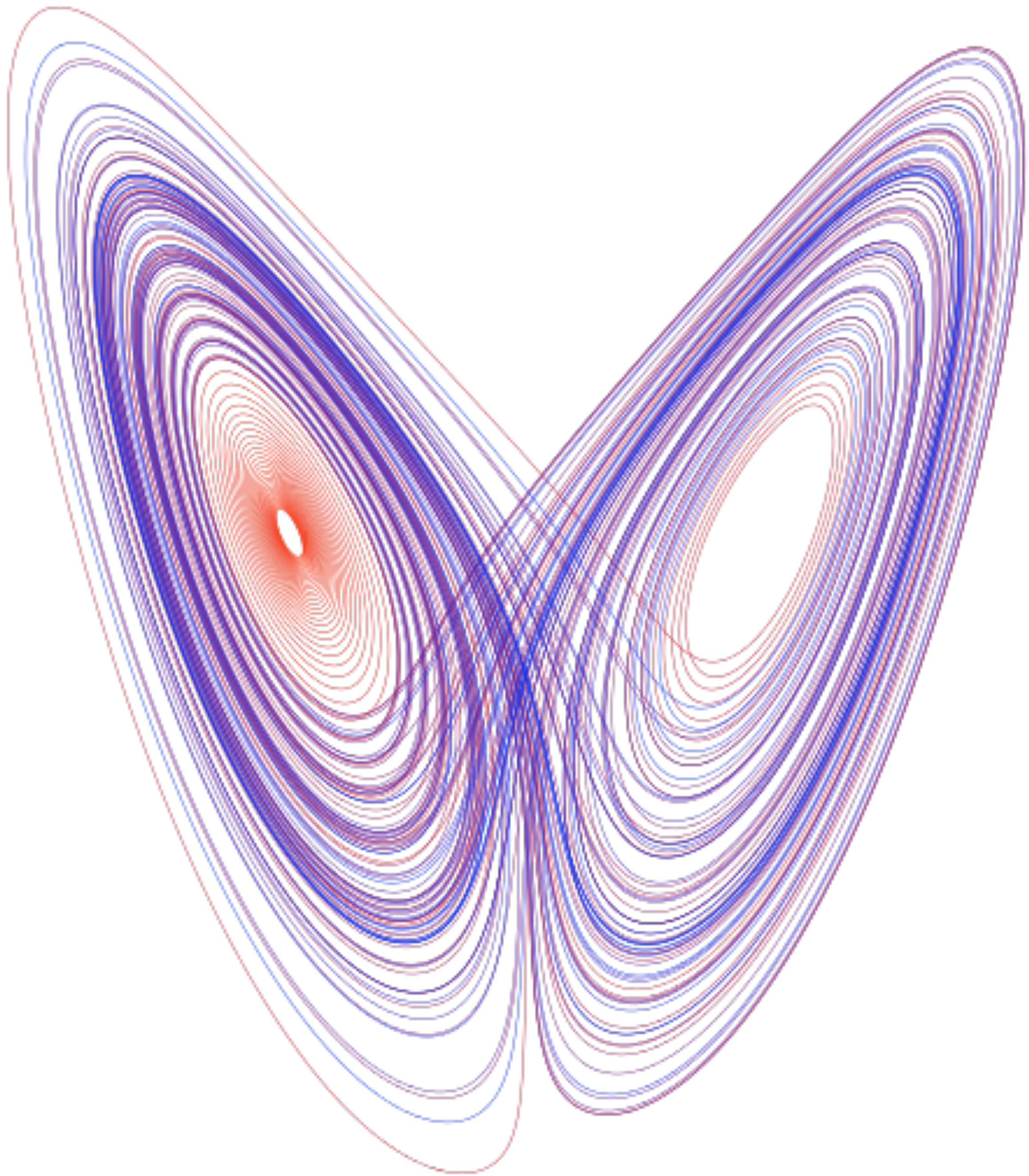
Lorenz (1963)

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

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

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

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



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

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

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

*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*.)

.

| $\mu(C(\tau, x))$ | 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 |
| $\tau = 2$ | $5.9 \times 10^{-2}$ | 5.74 | $4.5 \times 10^{-2}$ | 0.401 |
| $\tau = 3$ | $3.3 \times 10^{-2}$ | 29.4 | $2.9 \times 10^{-2}$ | 0.397 |
| $\tau = 8$ | $1.4 \times 10^{-2}$ | 59.9 | * | 0.396 |

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

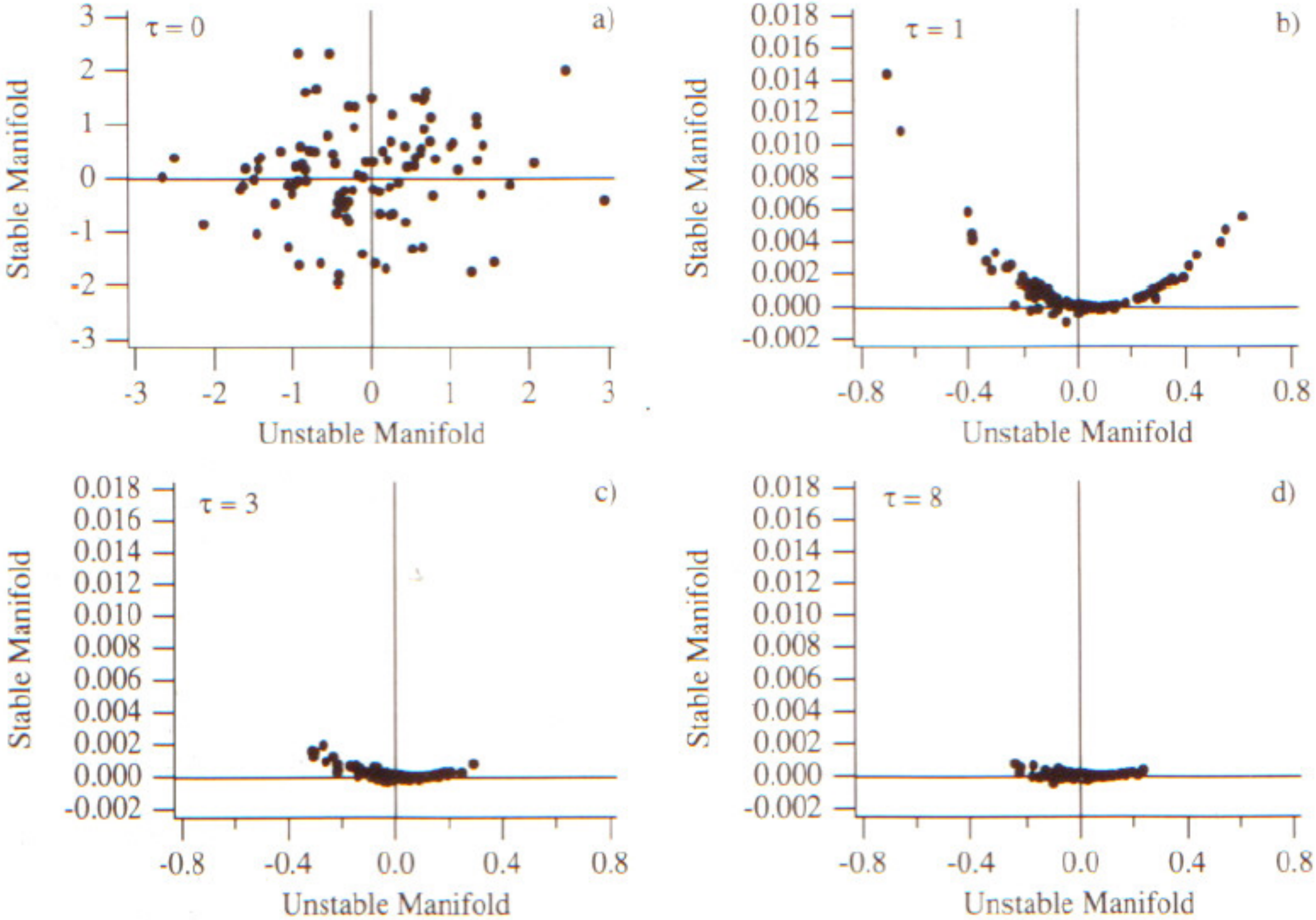



*Fig. 7.* Projection of the 100 minimizing solutions, at the end of the assimilation period, onto the plane spanned by the stable and unstable directions, defined as in Fig. 3. Values of $\tau$ are indicated on the panels. The projection is not an orthogonal projection, but a projection parallel to the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

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

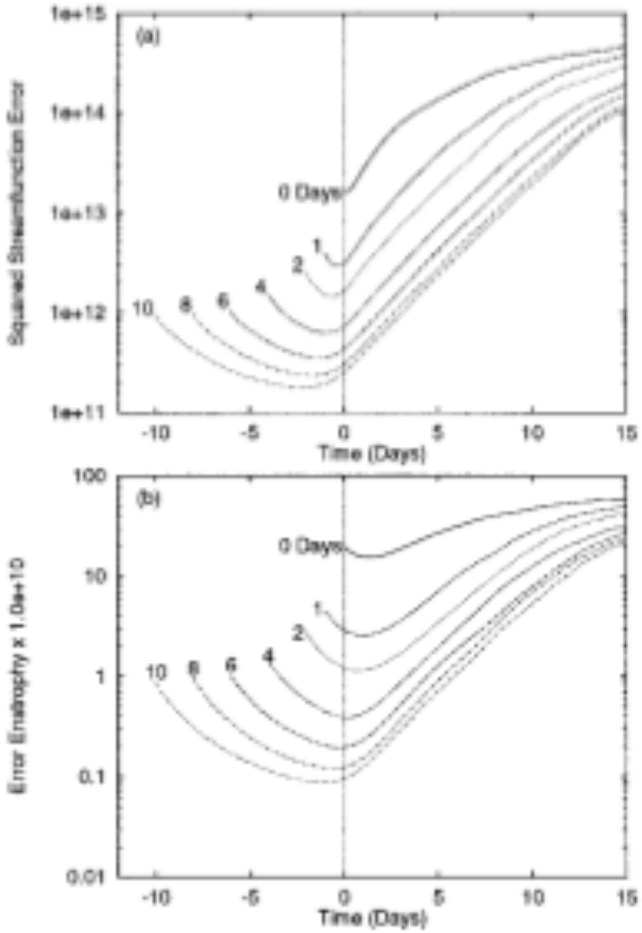



*Fig. 5.* Median values of the (a) streamfunction squared error, and (b) enstrophy error for the 200 forecast set as a function of forecast time and of the assimilation time $T_a$.

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

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.

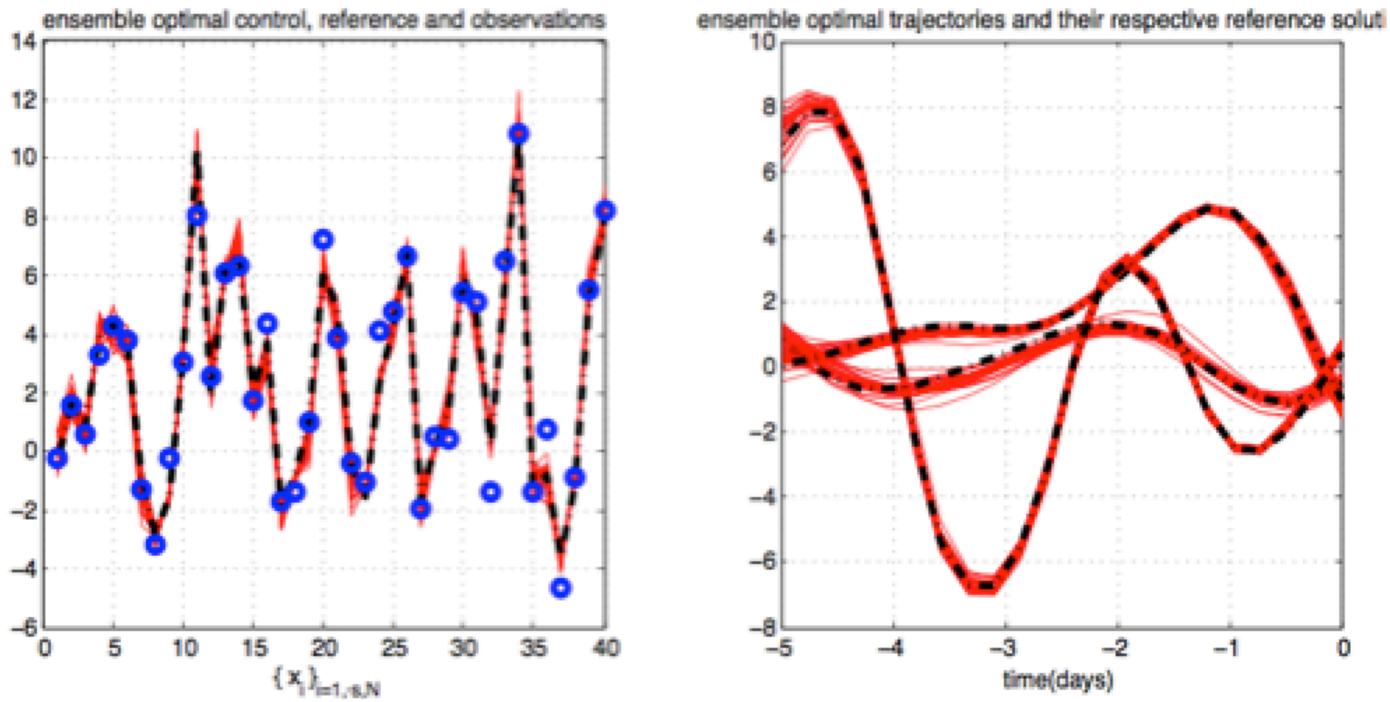
Three values of $I$ have been used, namely $I = 40, 60, 80$, which correspond to respectively $N^+ = 13$, 19 and 26 positive Lyapunov exponents.

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.

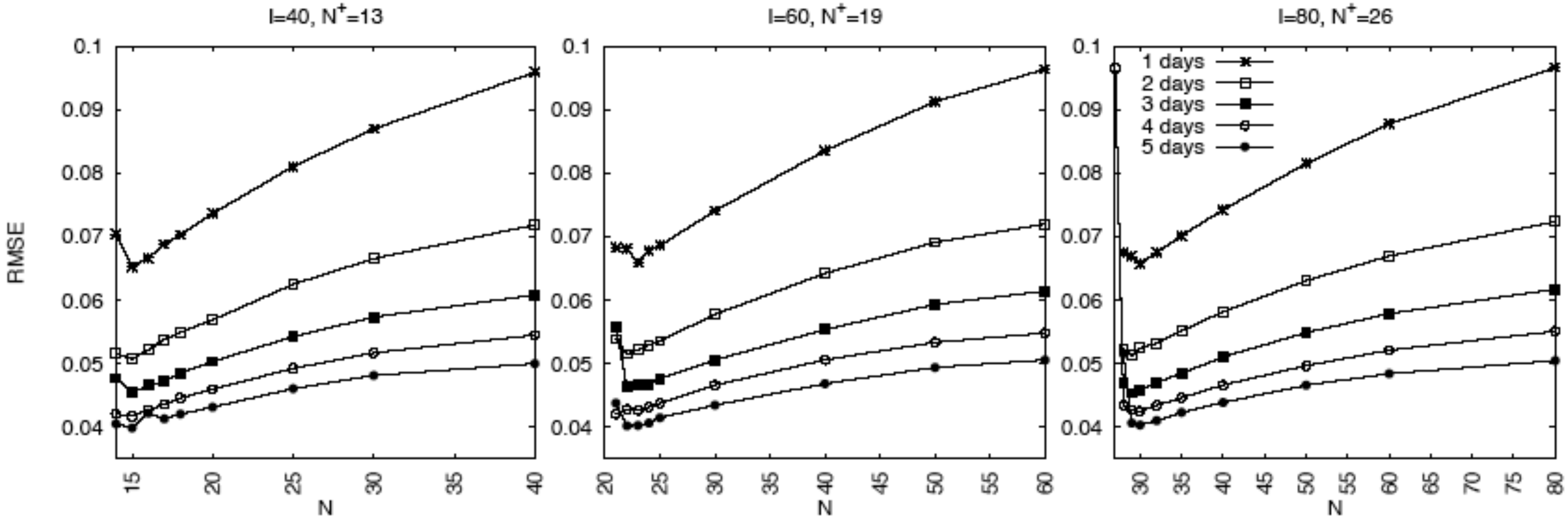



Figure 1. Time average RMS analysis error at $t = \tau$ as a function of the subspace dimension $N$ for three model configurations: $I$=40, 60, 80. Different curves in the same panel refer to different assimilation windows from 1 to 5 days. The observation error standard deviation is $\sigma_o = 0.2$.

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.

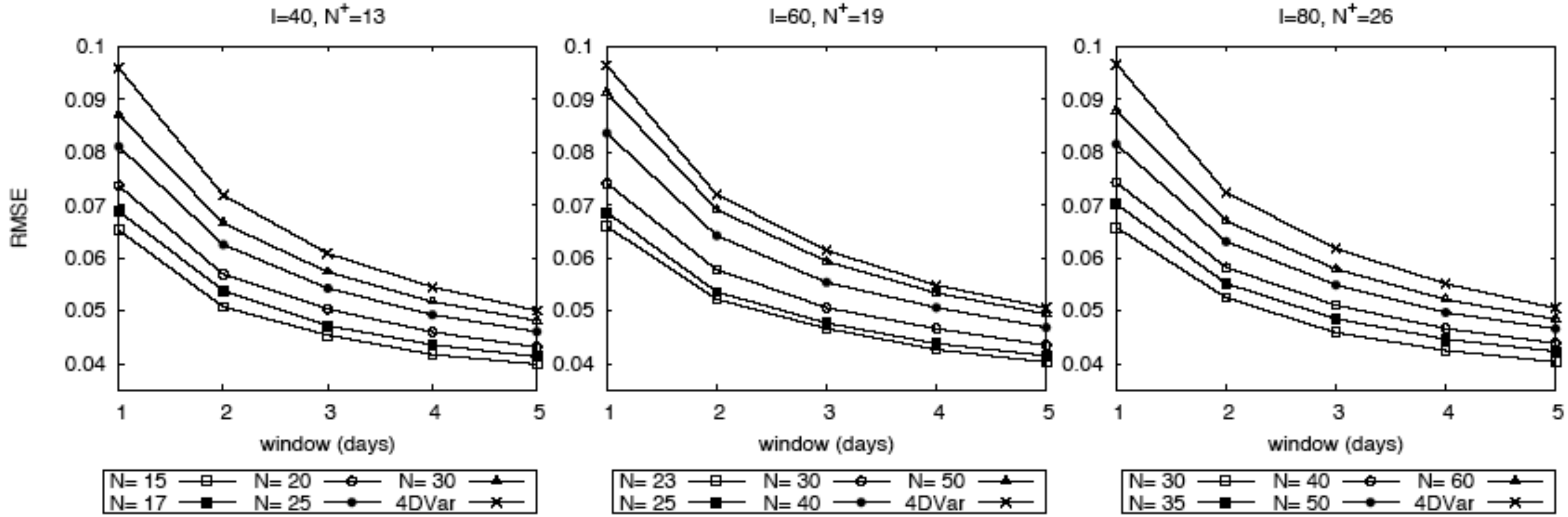



Figure 2. Time average RMS analysis error at $t = \tau$ as a function of the length of the assimilation window for three model configurations: $I$=40, 60, 80. Different curves in the same panel refer to a different subspace dimension $N$ of 4DVar-AUS and to standard 4DVar. $\sigma_o = 0.2$.

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 $\tau$.

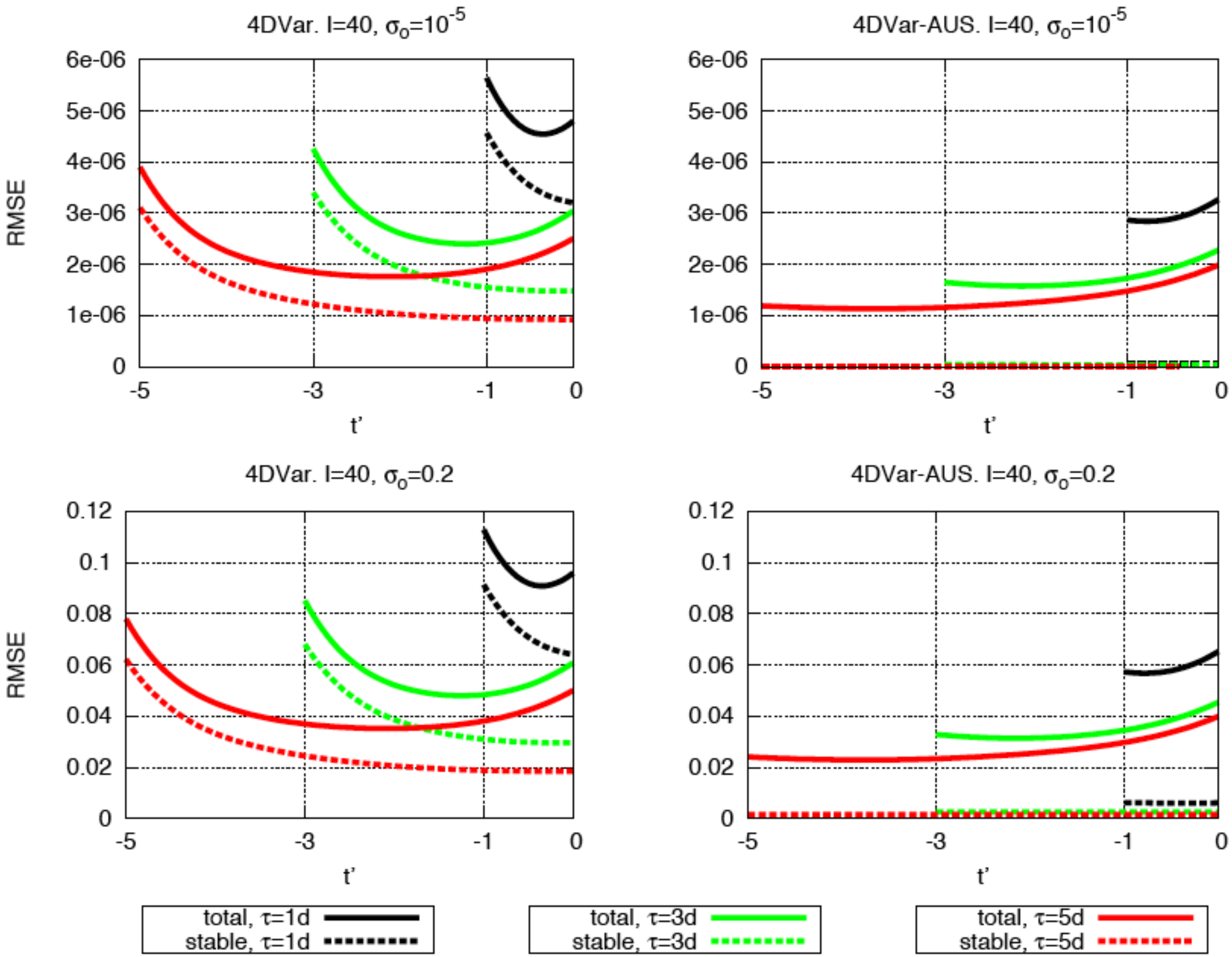



Figure 3. Time average RMS error within $1, 3, 5$ days assimilation windows as a function of $t' = t - \tau$, with $\sigma_o = .2, 10^{-5}$ for the model configuration $I = 40$. Left panel: 4DVar. Right panel: 4DVar-AUS with $N = 15$. Solid lines refer to total assimilation error, dashed lines refer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

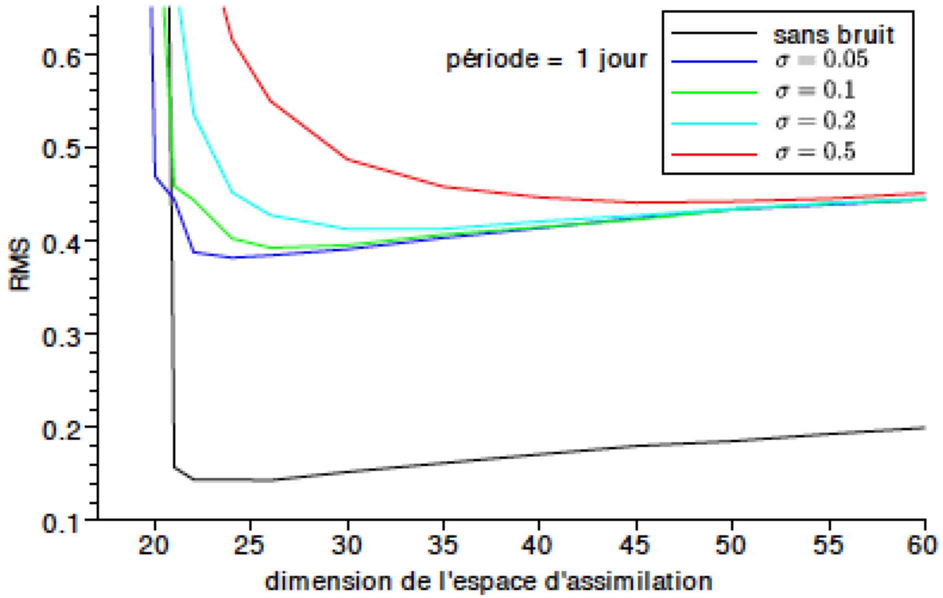
These results apply because no explicit background is available at the initial time of the assimilation window (only the unstable subspace is known). A proper background (obtained for instance from a properly implemented Kalman Filter, or from an Ensemble Variational Assimilation) would not only say 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.

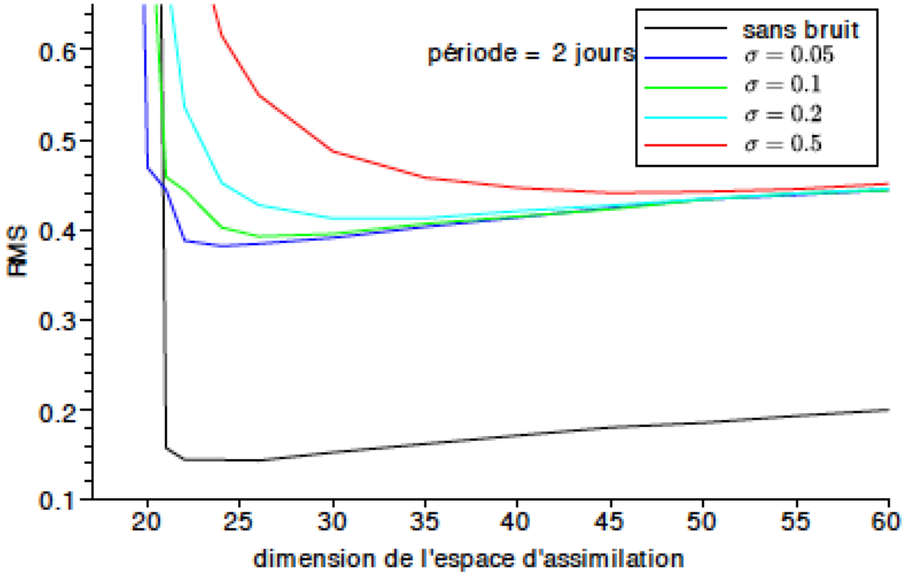
Questions.

- 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 $\tau$) 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).

- Ensemble Variational Assimilation (*EnsVAR*).

  (work with M. Jardak, 2018)

## *Ensemble Variational Assimilation*

**Data of the form**

$$z = \Gamma x + \zeta, \qquad \zeta \sim \mathcal{N}[0, S]$$

Conditional probability distribution is

$$P(x \mid z) = \mathcal{N}[x^a, P^a]$$

with

$$x^a = (\Gamma^{\mathrm{T}} S^{-1} \Gamma)^{-1} \Gamma^{\mathrm{T}} S^{-1} z$$
$$P^a = (\Gamma^{\mathrm{T}} S^{-1} \Gamma)^{-1}$$

**Variational form**

$$P(x \mid z) \propto \exp[\, -(z - \Gamma\xi)^{\mathrm{T}} S^{-1} (z - \Gamma\xi)/2 \,] \propto \exp[\, -(\xi - x^a)^{\mathrm{T}} (P^a)^{-1} (\xi - x^a)/2 \,]$$

Conditional expectation $x^a$ minimizes following scalar *objective function*, defined on state space $\mathcal{X}$

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

$$P^a = [\partial^2 \mathcal{J} / \partial\xi^2]^{-1}$$

Ready recipe for determining Monte-Carlo sample of conditional pdf $P(x \mid z)$ :

- Perturb data vector $z$ according to its own error probability distribution

$$z \rightarrow z' = z + \delta, \qquad \delta \sim \mathcal{N}[0, S]$$

and compute

$$x'^a = (\Gamma^T S^{-1} \Gamma)^{-1} \Gamma^T S^{-1} z'$$

$x'^a$ is distributed according to $\mathcal{N}[x^a, P^a]$

*Ensemble Variational Assimilation* (*EnsVar*) implements that algorithm, the expectations $x'^a$ being computed by standard variational assimilation.

**Present purpose**

Evaluate EnsVar as a probabilistic estimator when implemented in nonlinear and/or non-Gaussian cases, *i. e*., through minimization of

$$\xi \in \mathcal{X} \rightarrow \mathcal{J}(\xi) \equiv (1/2)\,[\Gamma(\xi) - z']^{T}\, S^{-1}\, [\Gamma(\xi) - z']$$

where $\Gamma$ may be nonlinear, and errors affecting data $z$ may be non-Gaussian.

- Objectively compare with other existing ensemble assimilation algorithms : *Ensemble Kalman Filter* (*EnKF*), *Particle Filters* (*PF*)

- Simulations performed on two small-dimensional chaotic systems, the Lorenz'96 model and the Kuramoto-Sivashinsky equation

# The Lorenz96 model

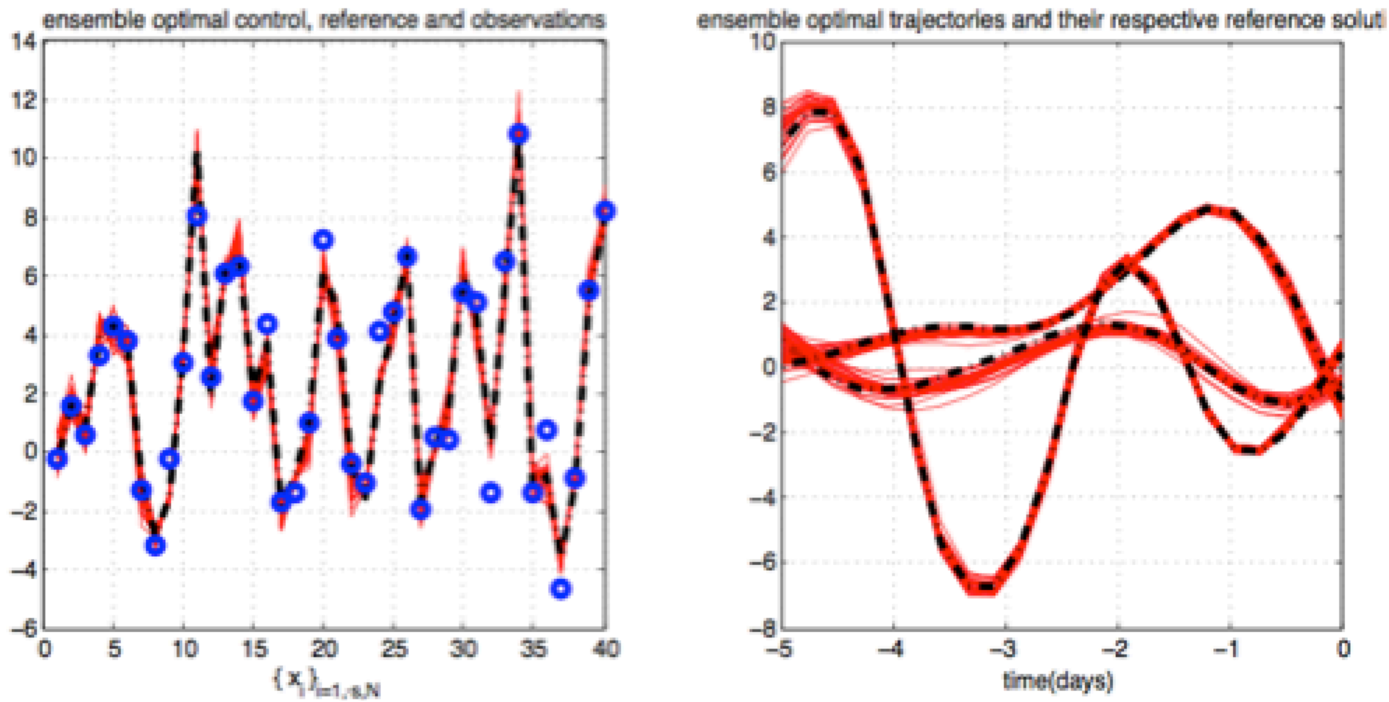
- Forward model

$$\frac{dx_k}{dt} = (x_{k+1} - x_{k-2})x_{k-1} - x_k + F \quad \text{for} \quad k = 1, \cdots, N$$

- Set-up parameters :
  1. the index $k$ is cyclic so that $x_{k-N} = x_{k+N} = x_k$.
  2. $F = 8$, external driving force.
  3. $-x_k$, a damping term.
  4. $N = 40$, the system size.
  5. $Nens = 30$, number of ensemble members.
  6. $\dfrac{1}{\lambda_{max}} \simeq 2.5 days, \lambda_{max}$ the largest Lyapunov exponent.
  7. $\Delta t = 0.05 = 6 hours$, the time step.
  8. frequency of observations : every 12 hours.
  9. number of realizations : 9000 realizations.

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’



**Experimental procedure (1)**

0. Define a *reference solution* $x_t^r$ by integration of the numerical model

1. Produce 'observations' at successive times $t_k$ of the form

$$y_k = H_k x_k^r + \varepsilon_k$$

where $H_k$ is (usually, but not necessarily) the unit operator, and $\varepsilon_k$ is a random (usually, but not necessarily, Gaussian) 'observation error'.

**Experimental procedure (2)**

2. For given observations $y_k$, repeat $N_{ens}$ times the following process

- 'Perturb' the observations $y_k$ as follows

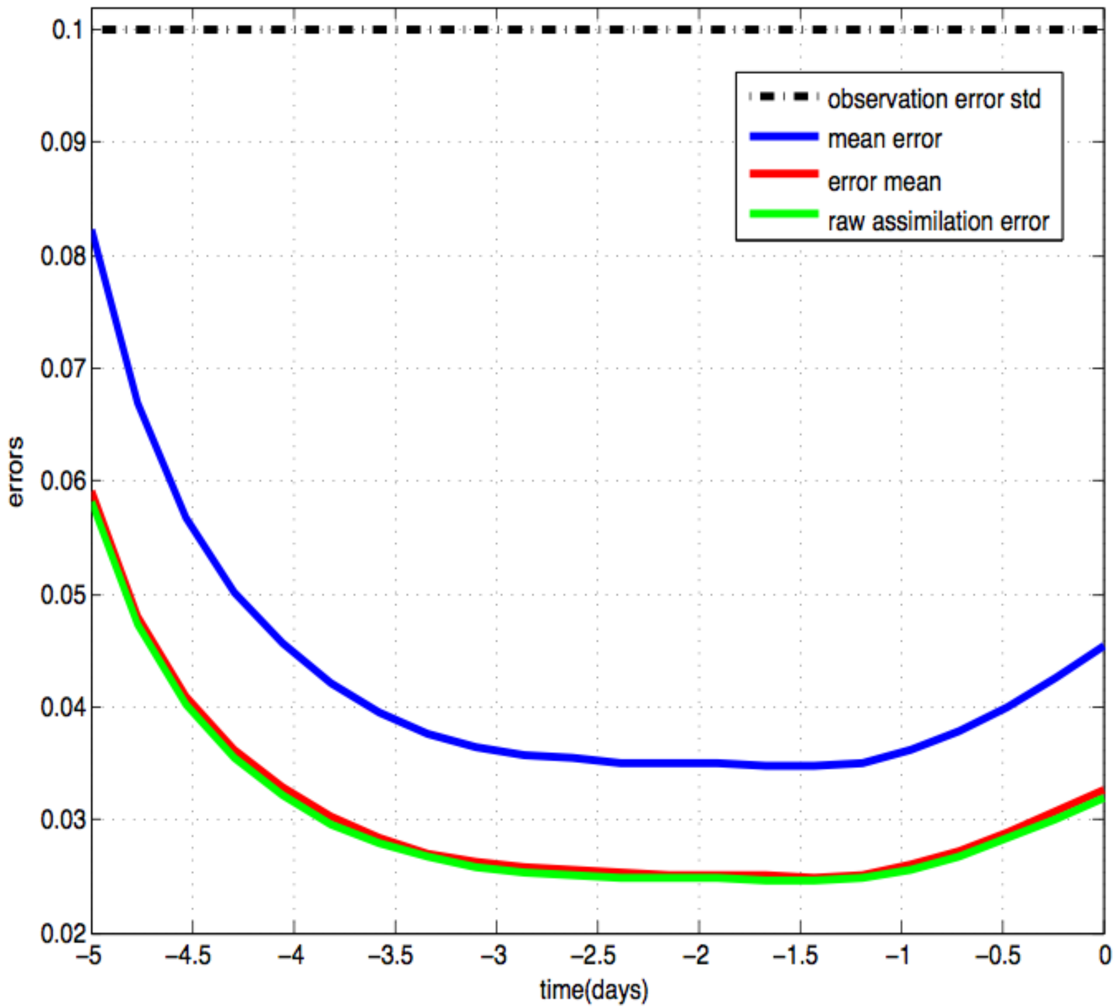
$$y_k \rightarrow z_k = y_k + \delta_k$$

where $\delta_k$ is an independent realization of the probability distribution which has produced $\varepsilon_k$.

- Assimilate the 'perturbed' observations $z_k$ by variational assimilation

This produces $N_{ens}$ (=30) model solutions over the assimilation window, considered as making up a tentative sample of the conditional probability distribution for the state of the observed system over the assimilation window.

The process 1-2 is then repeated over $N_{real}$ successive assimilation windows. Validation is performed on the set of $N_{real}$ (=9000) ensemble assimilations thus obtained.




Linearized Lorenz’96. 5 days

**How to objectively evaluate the performance of an ensemble (or more generally probabilistic) estimation system ?**

- There is no general objective criterion for Bayesianity

- We use instead the weaker property of *reliability*, *i. e.* statistical consistency between predicted probabilities and observed frequencies of occurrence (it rains with frequency 40% in the circonstances where I have predicted 40% probability for rain).

Denote $Y$ the predicted probability distribution, and $X$ the verifying reality. Consider the probability distribution for the couples $(X, Y)$ (that probability distribution can be obtained empirically). Reliability is the property that

$$P(X \mid Y) = Y \text{ for any } Y$$

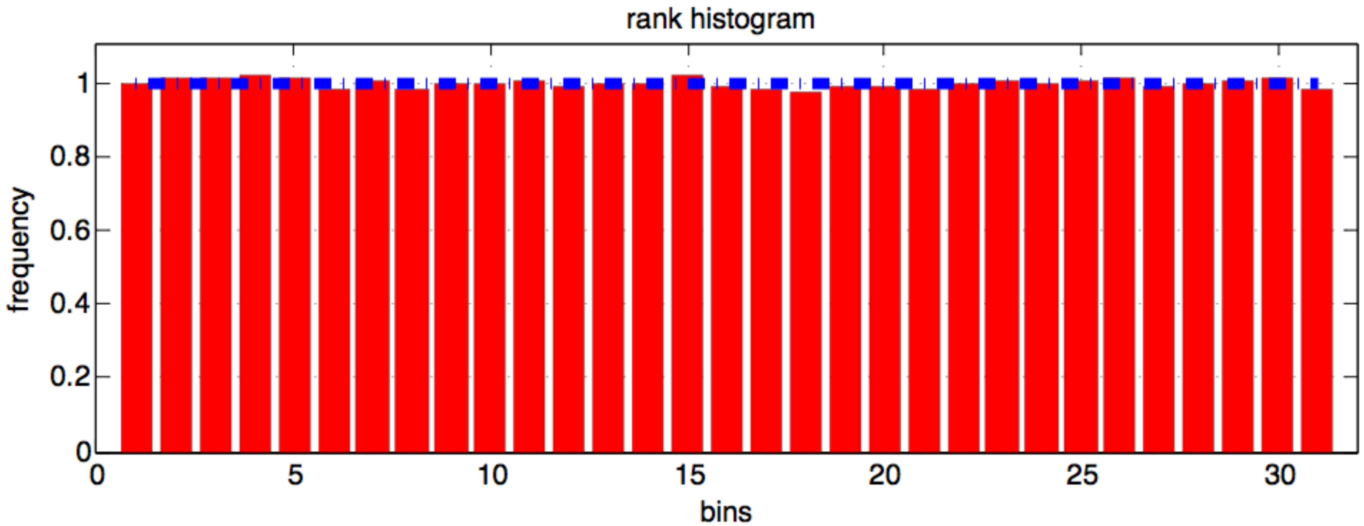
Reliability can be objectively validated, provided a large enough sample of realizations of the estimation system is available.

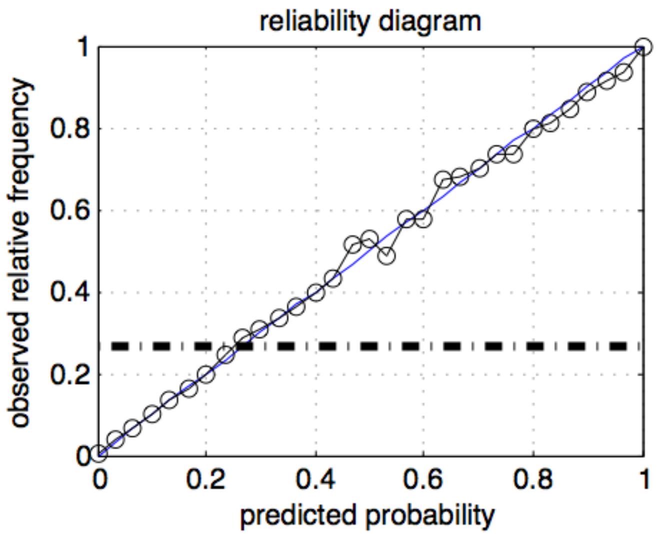
Bayesianity implies reliability, the converse not being true.

In addition, we evaluate *resolution* (also called *sharpness*), which bears no direct relation to bayesianity, and is best defined as the degree of statistical dependence between $X$ and $Y$ (J. Bröcker). Total absence of resolution is independence between $X$ and $Y$, *viz.*

$$P(X \mid Y) = P(X) \text{ for any } Y$$

Resolution, beyond reliability, measures the degree of usefulness of the ensembles.





Linearized Lorenz'96. 5 days

**Objective function**

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

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

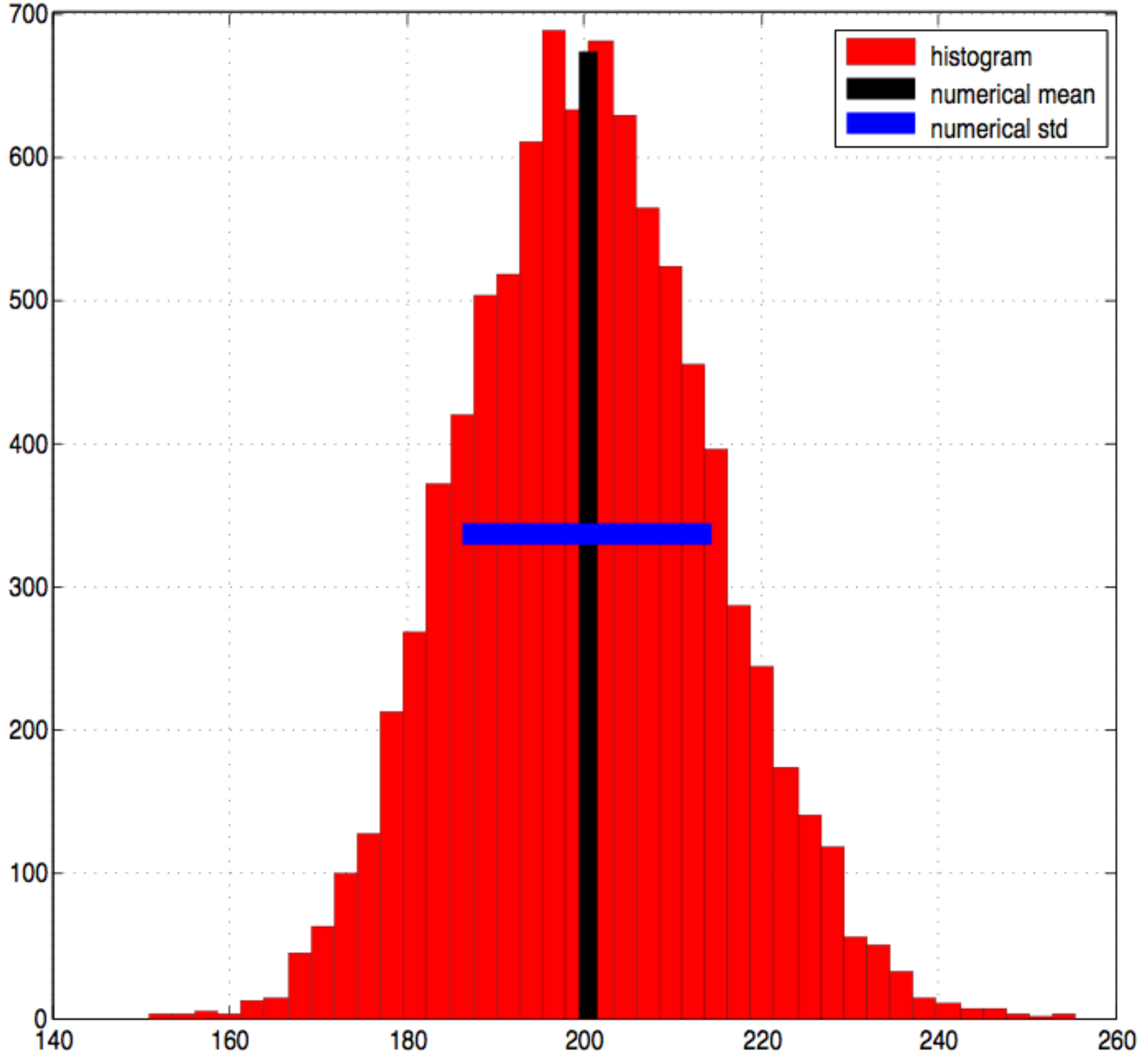
$$=\; (1/2)\ \boldsymbol{d}^{\mathrm{T}}\ [E(\boldsymbol{d}\boldsymbol{d}^{\mathrm{T}})]^{-1}\ \boldsymbol{d}$$

where $\boldsymbol{d}$ is innovation

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

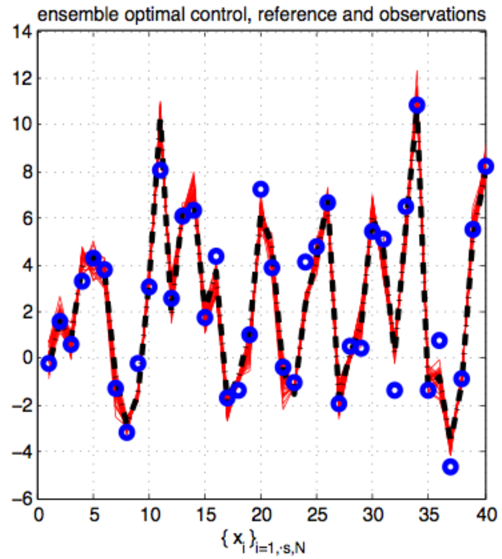
If $p$ is large, a few realizations are sufficient for determining $E(\mathcal{J}_{min})$
Often called $\chi^2$ criterion.

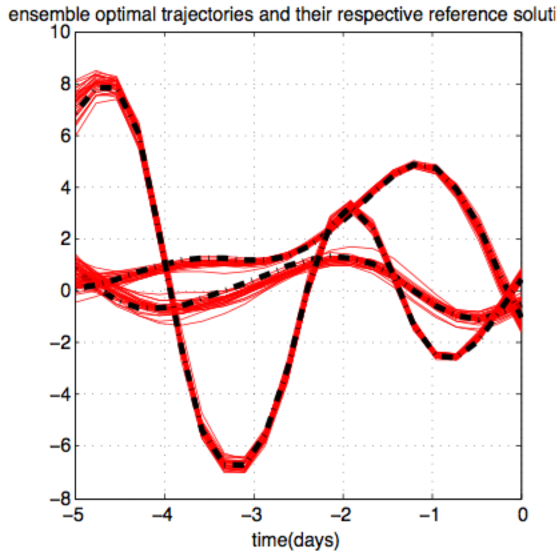
*Remark*. If in addition errors are gaussian $Var(\mathcal{J}_{min}) \;=\; p/2$

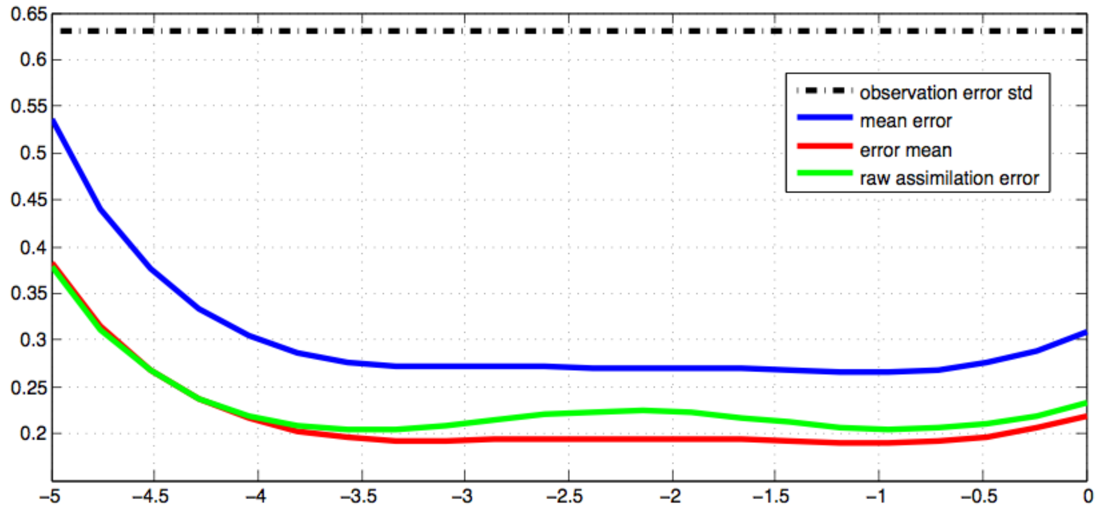



Linearized Lorenz'96. 5 days. Histogram of $\mathcal{J}_{min}$

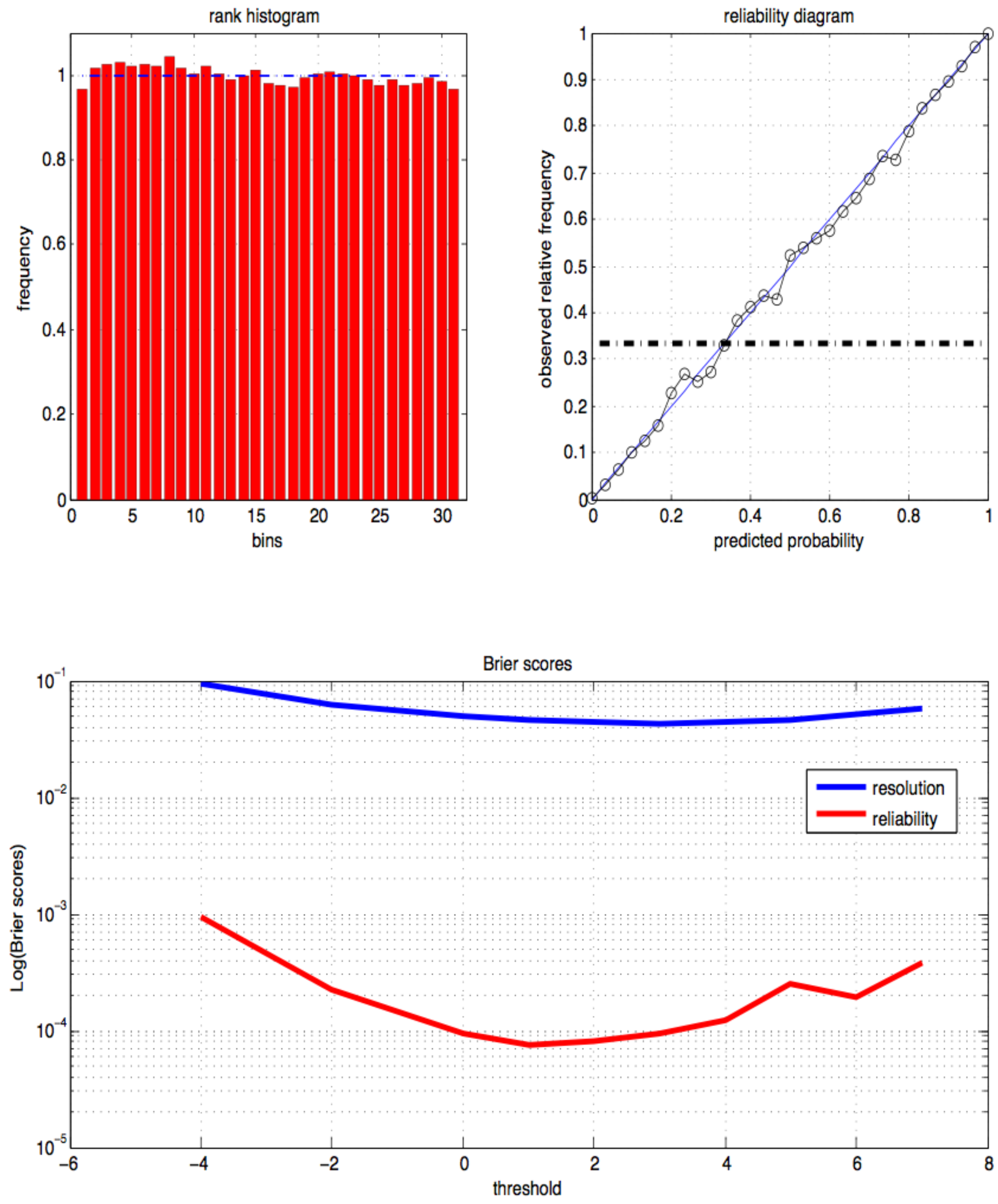
$E(\mathcal{J}_{min}) = p/2$ (=200) ; $\sigma(\mathcal{J}_{min}) = \sqrt{(p/2)}$ (≈14.14)



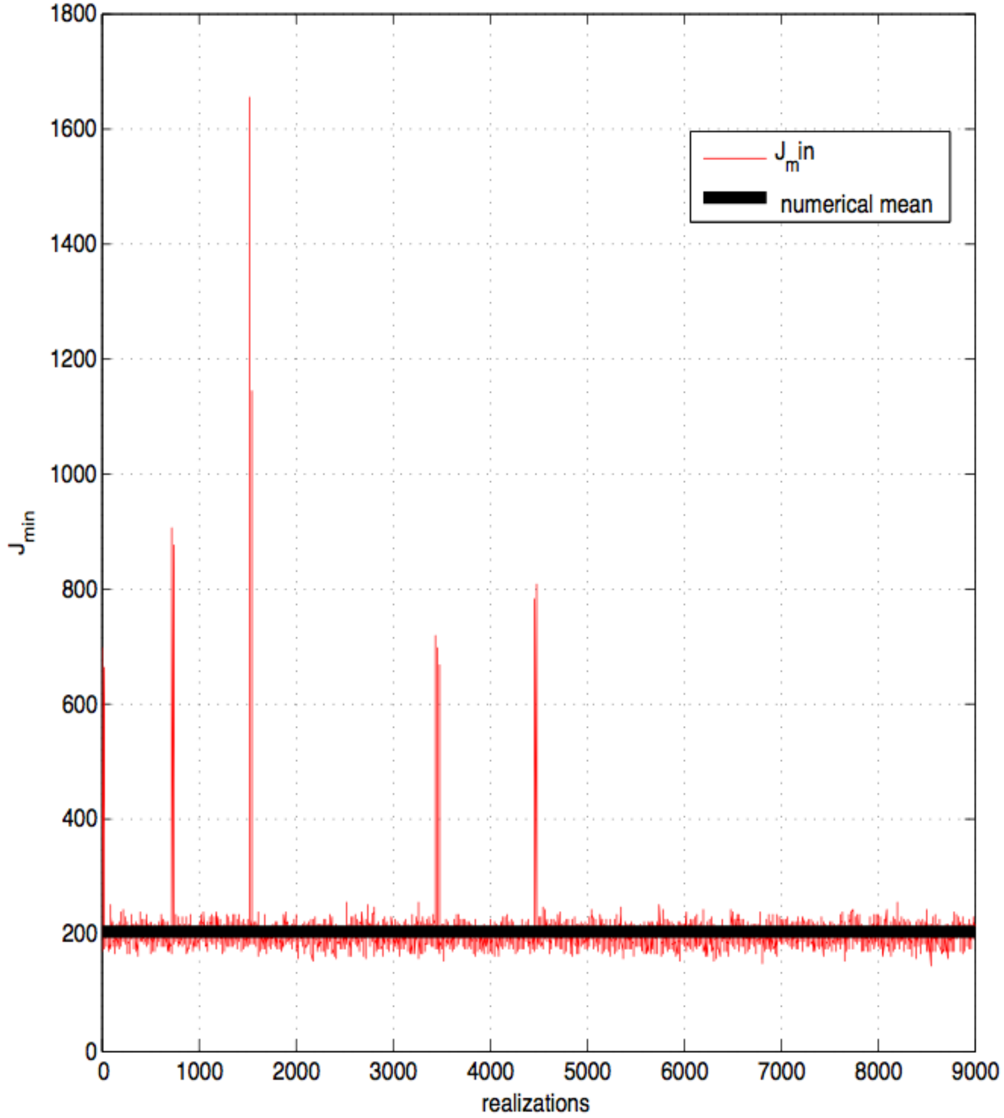




Nonlinear Lorenz’96. 5 days

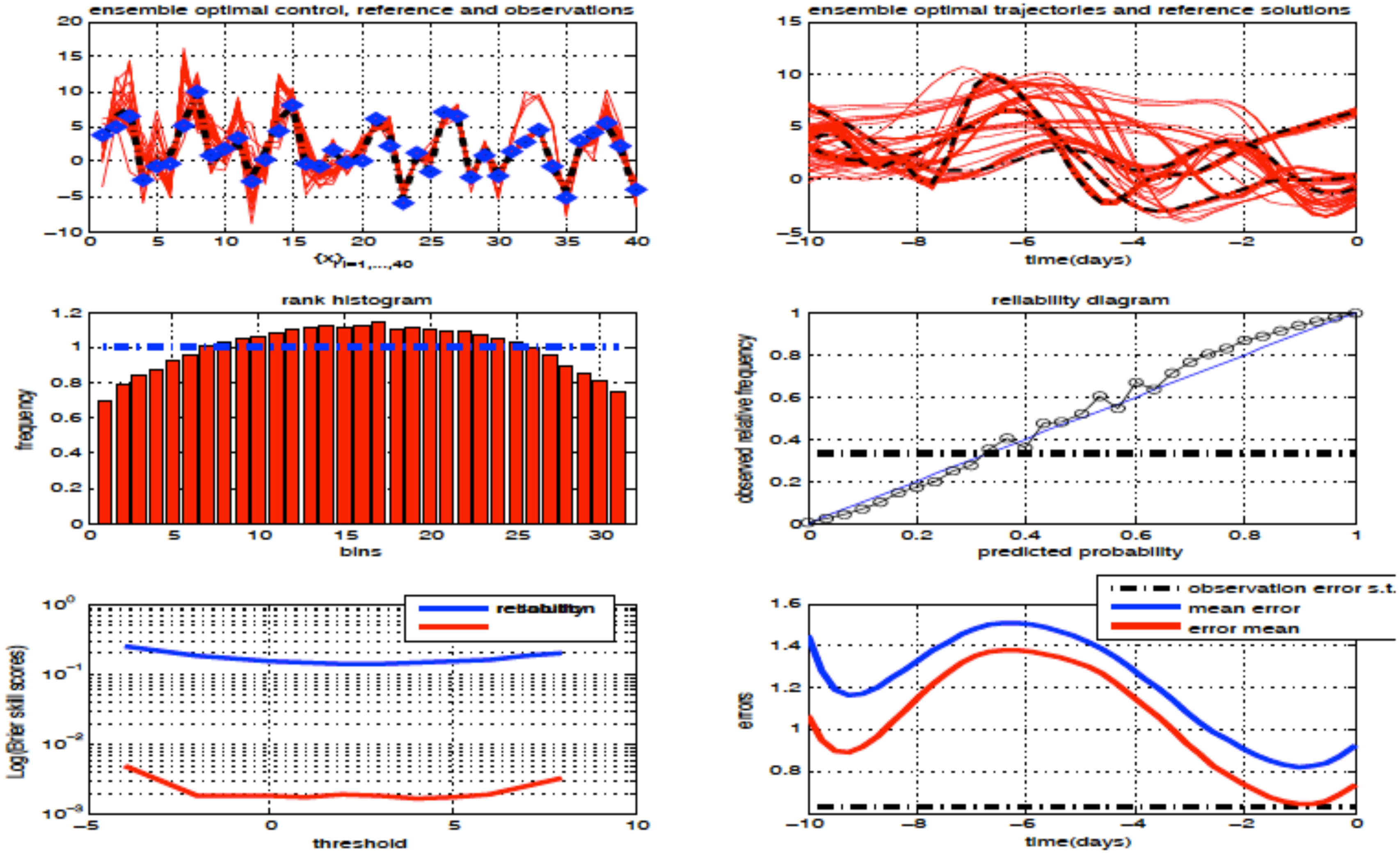



Nonlinear Lorenz'96. 5 days

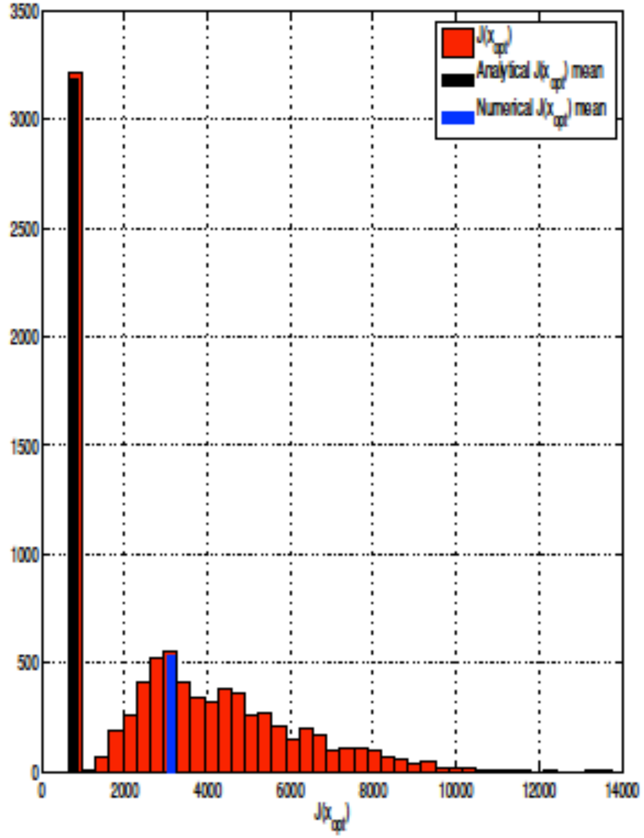



Nonlinear Lorenz'96. 5 days. Histogram of $\mathcal{J}_{min}$

# EnsVar : the non-linear Lorenz96 model (10 days $\simeq$ 2 TU)

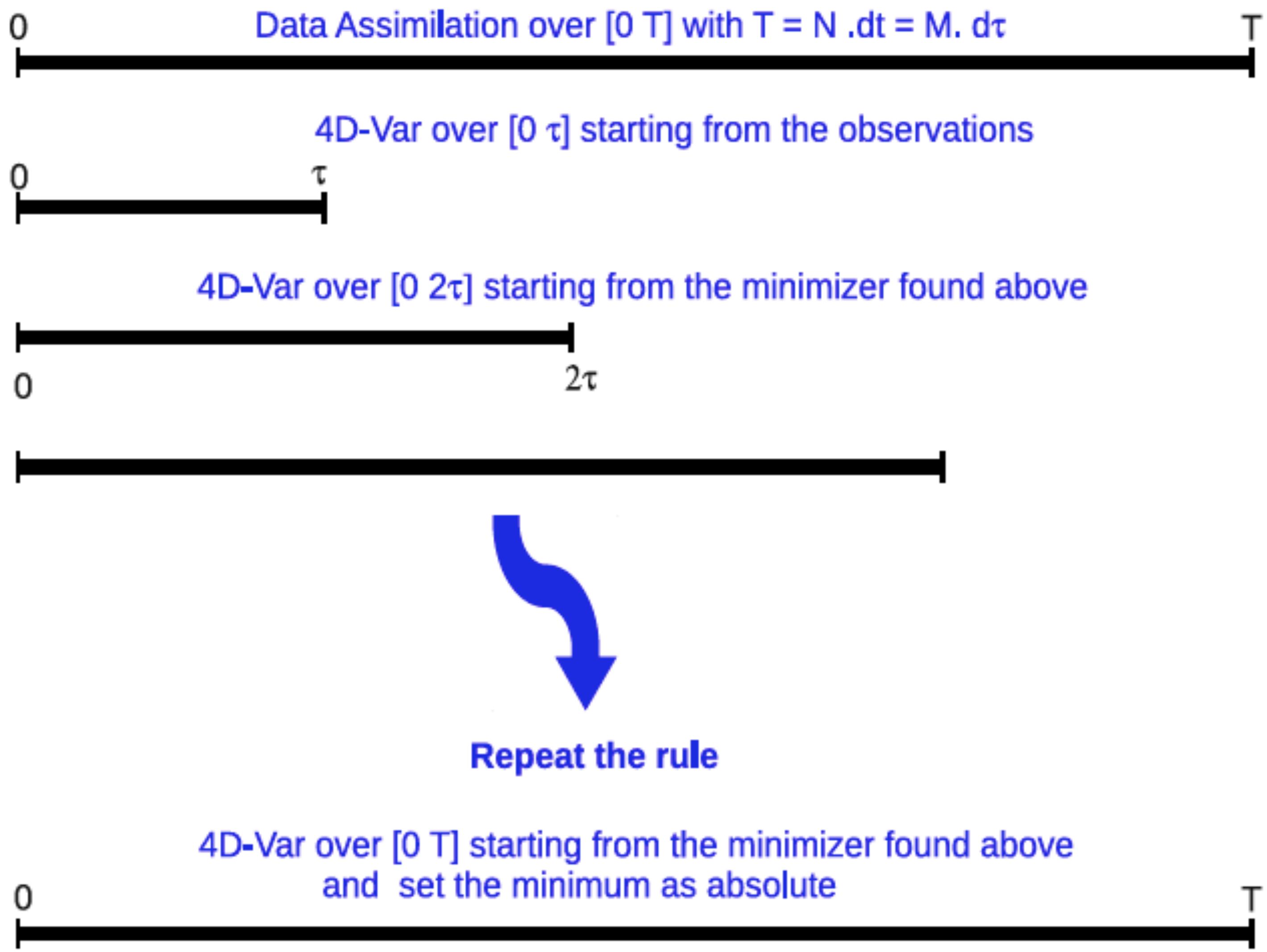


# EnsVar : consistency

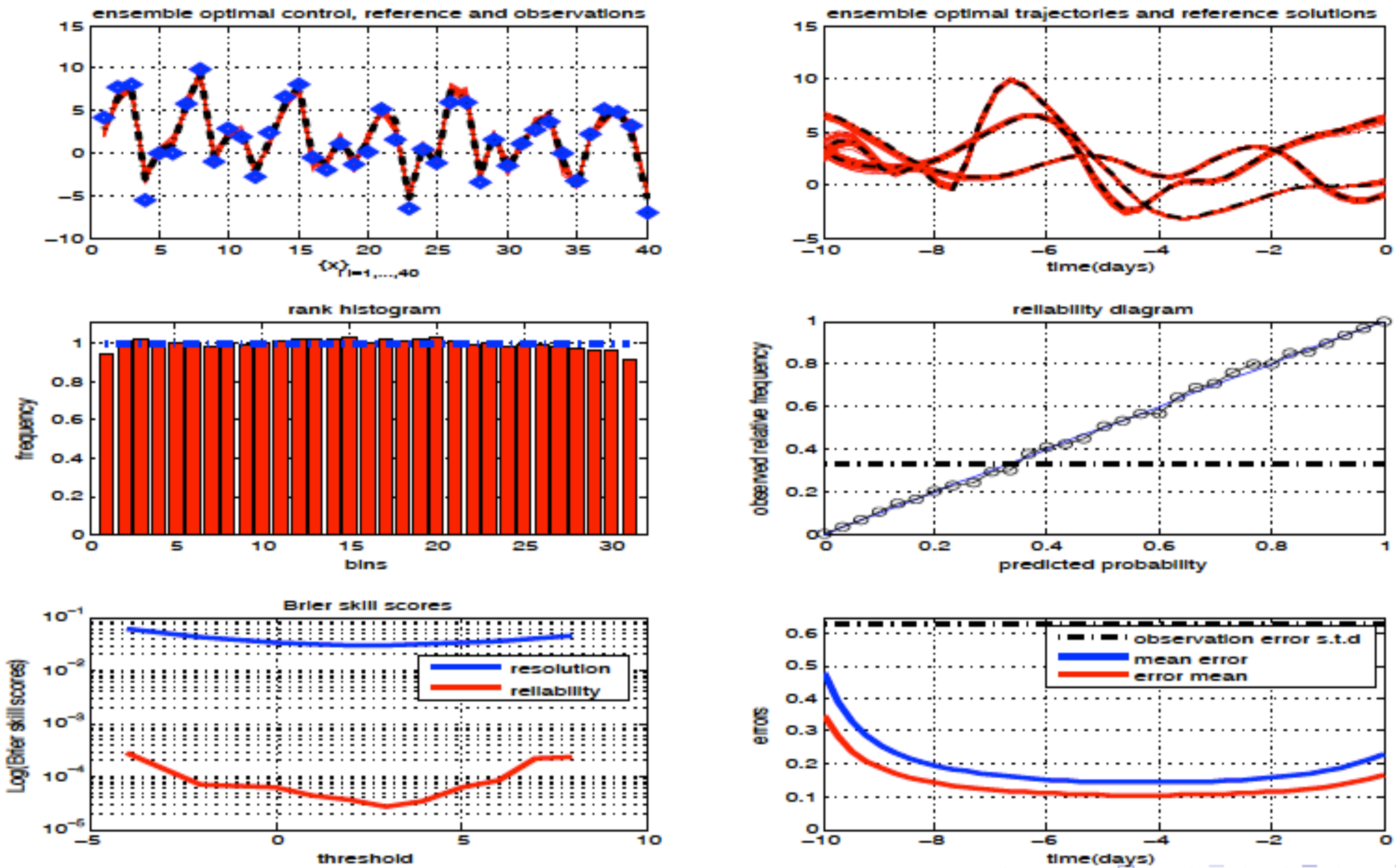



Nonlinear Lorenz'96. 10 days. Histogram of $\mathcal{J}_{min}$

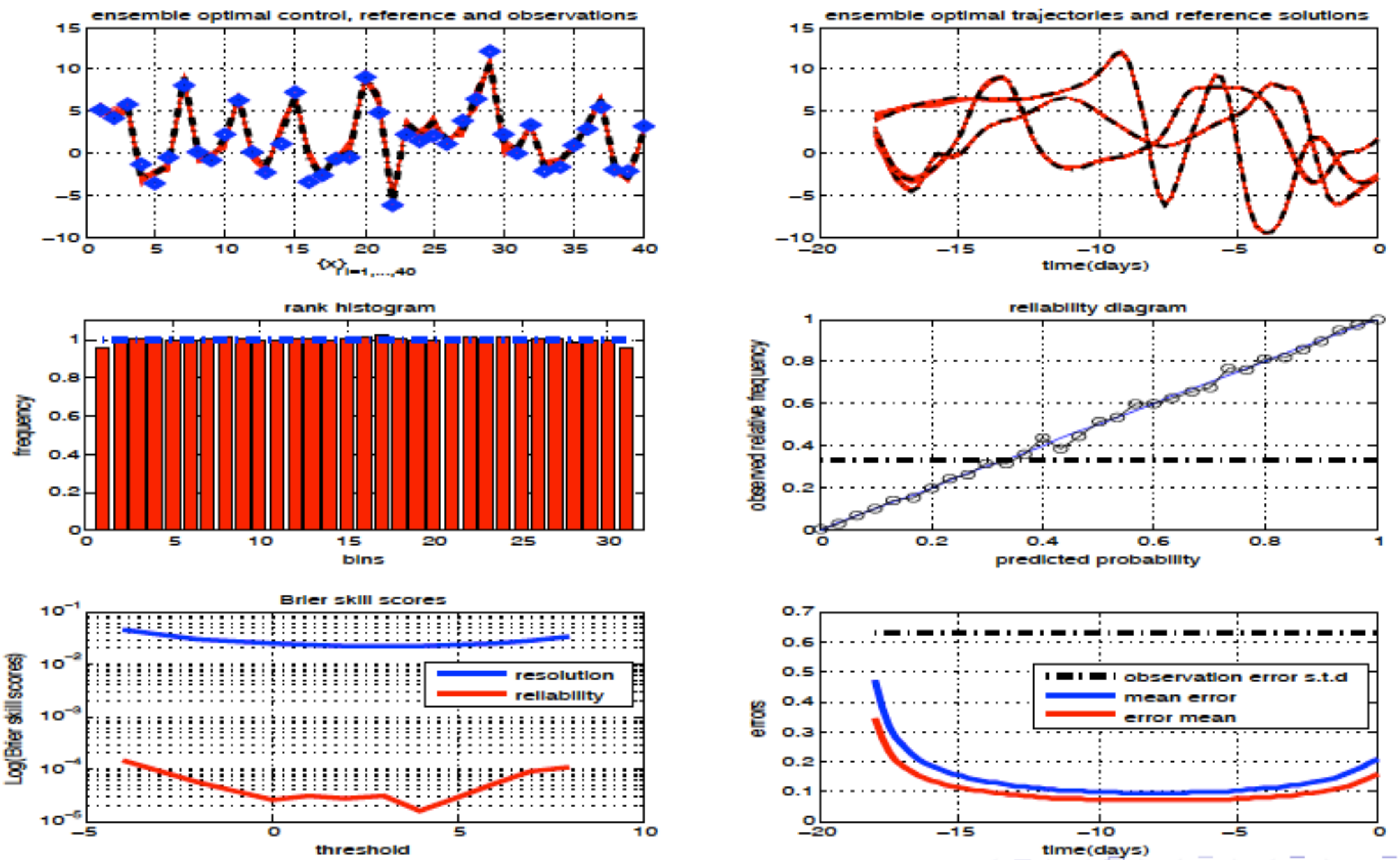
# Quasi-Static Variational Assimilation (QSVA)



# EnsVar : the non-linear Lorenz96 model 10 days with QSVA



# EnsVar : the non-linear Lorenz96 model 18 days with QSVA

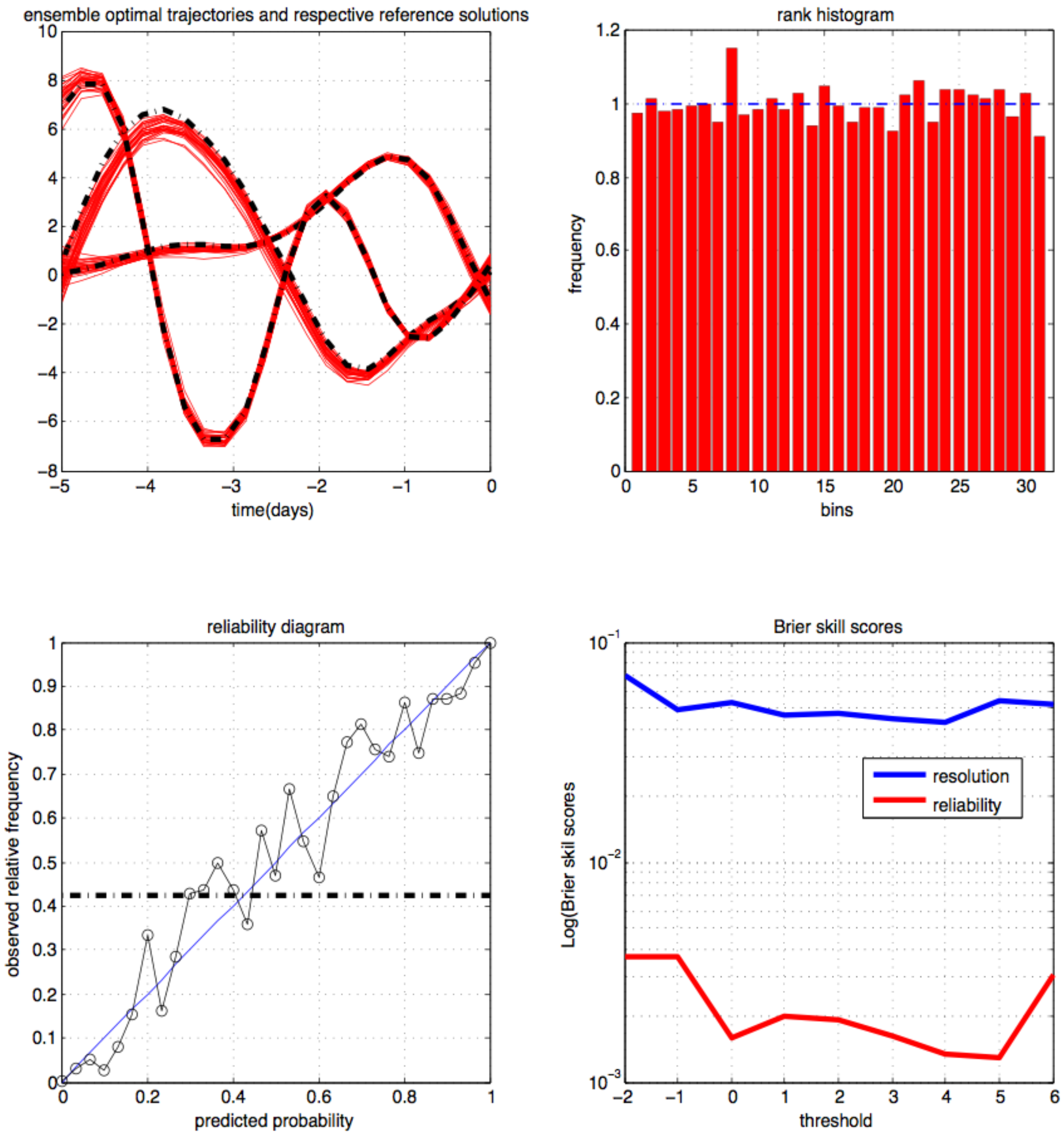


- Results are independent of the Gaussian character of the observation errors (trials have been made with various probability distributions)

- Ensembles produced by EnsVar are very close to Gaussian, even in strongly nonlinear cases.

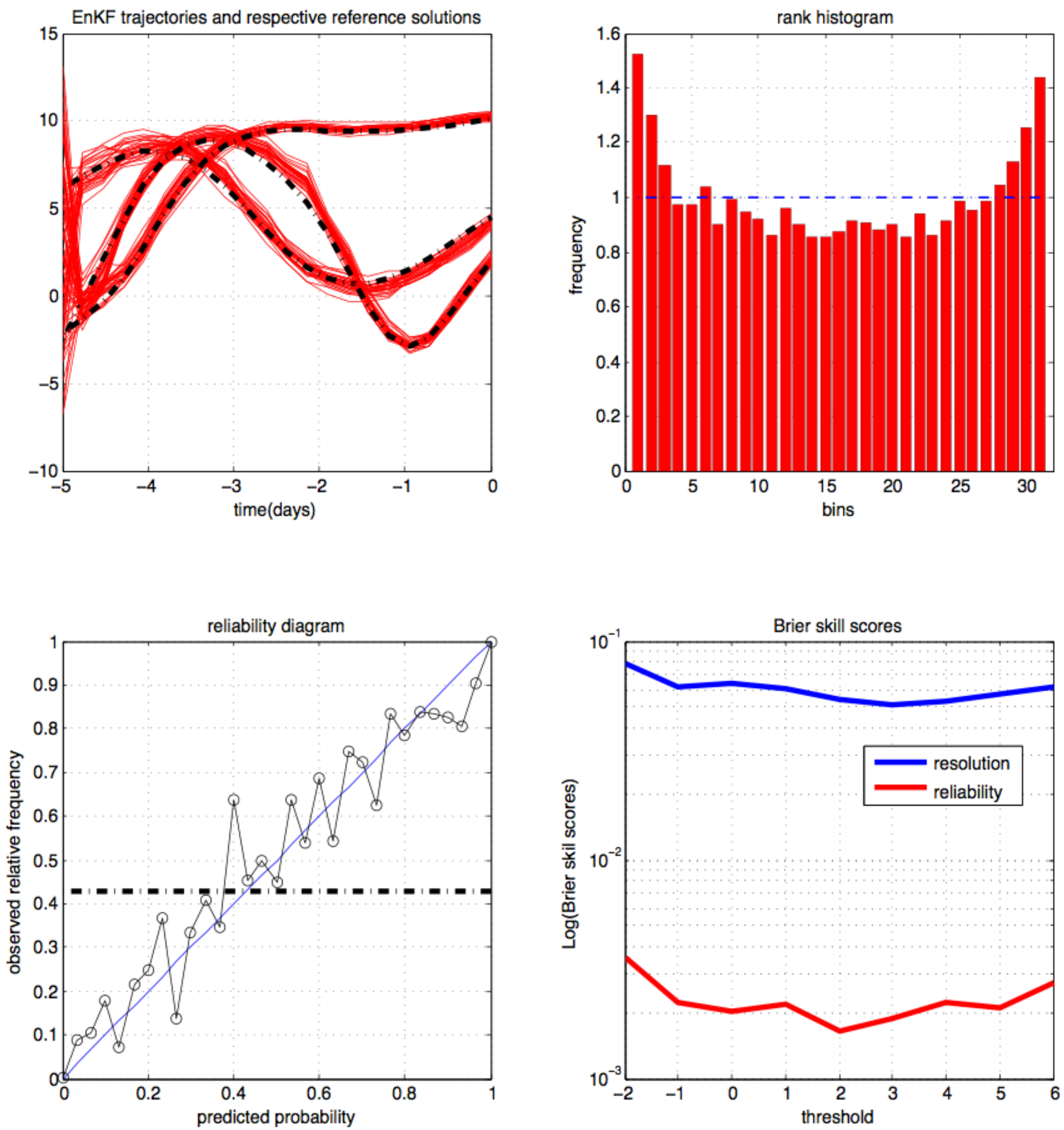
- Comparison *Ensemble Kalman Filter* (*EnKF*) and *Particle Filters* (*PF*)

  Both of these algorithms being sequential, comparison is fair only at end of assimilation window

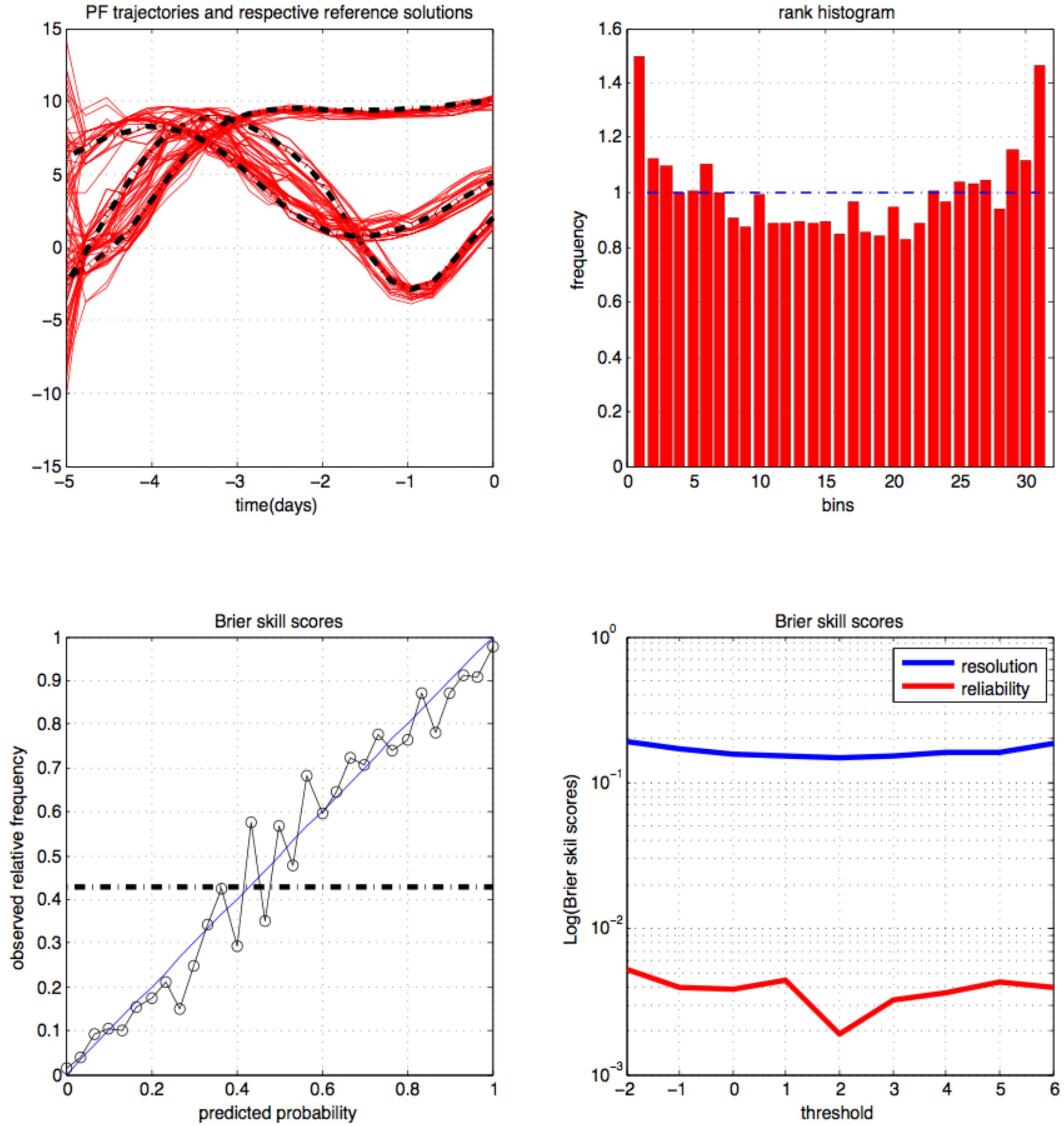





Nonlinear Lorenz’96. 5 days. Diagnostics at end of assimilation window

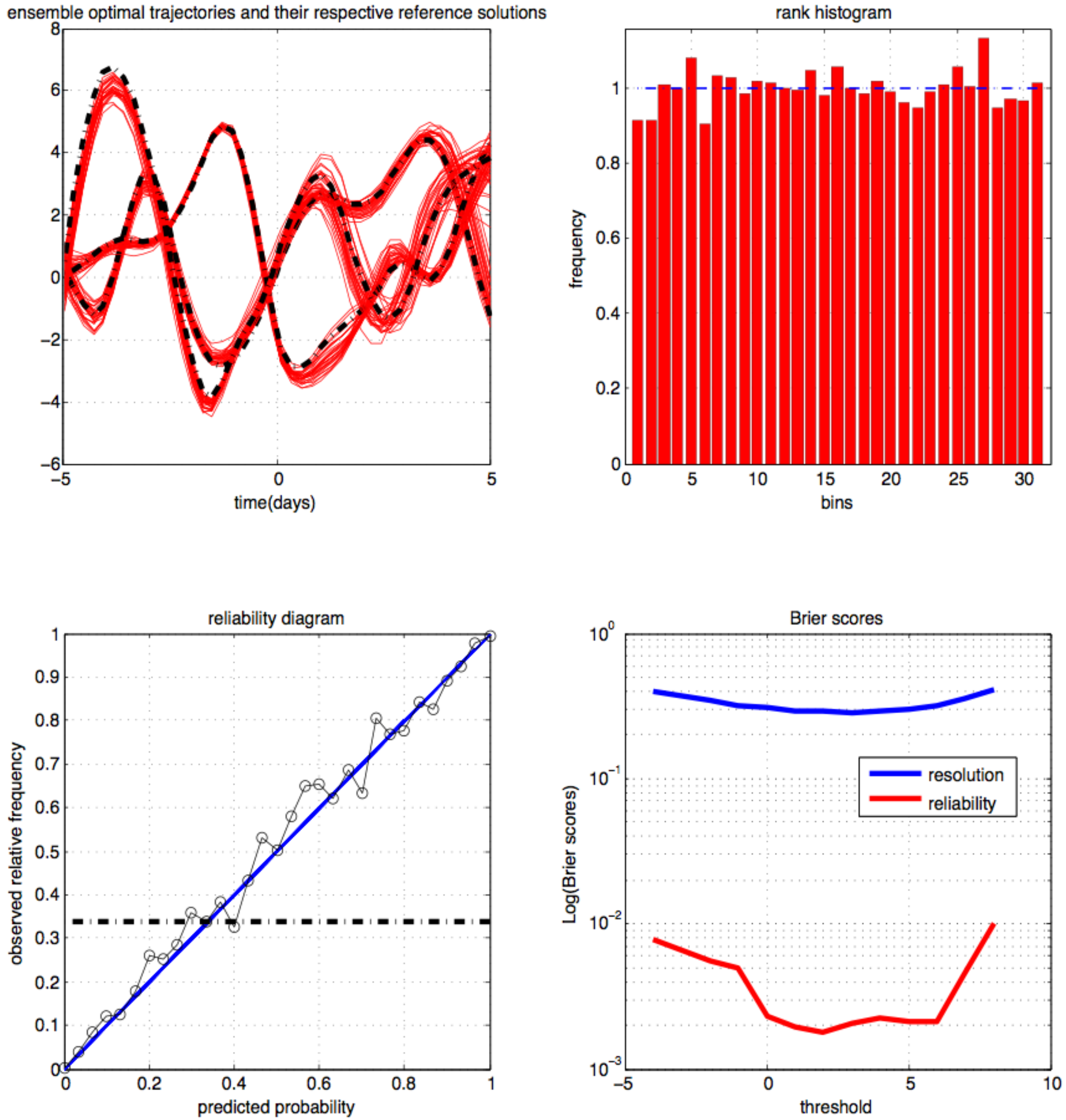



Nonlinear Lorenz'96. EnKF. Diagnostics after 5 days of assimilation



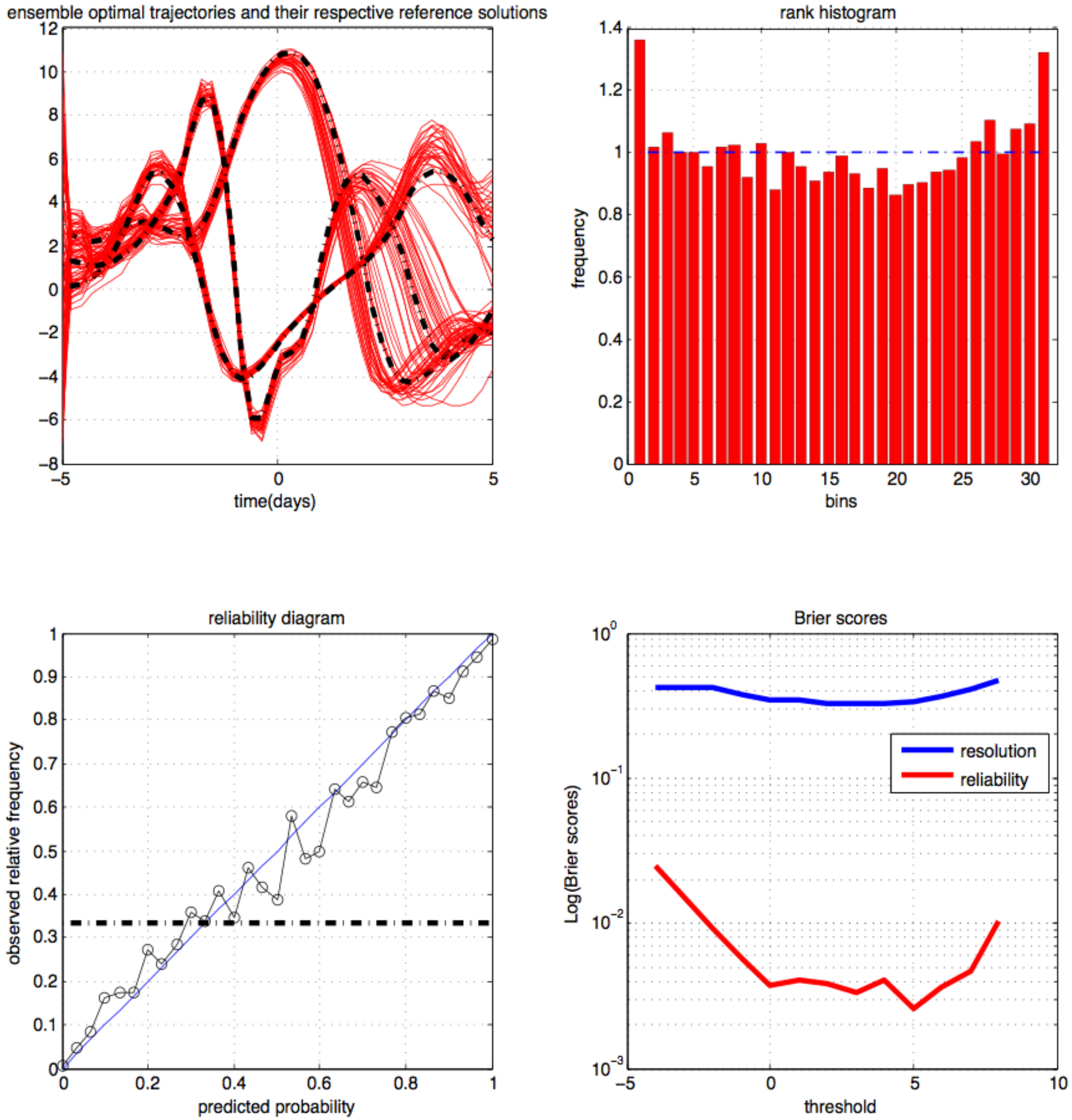



Nonlinear Lorenz'96. PF. Diagnostics after 5 days of assimilation

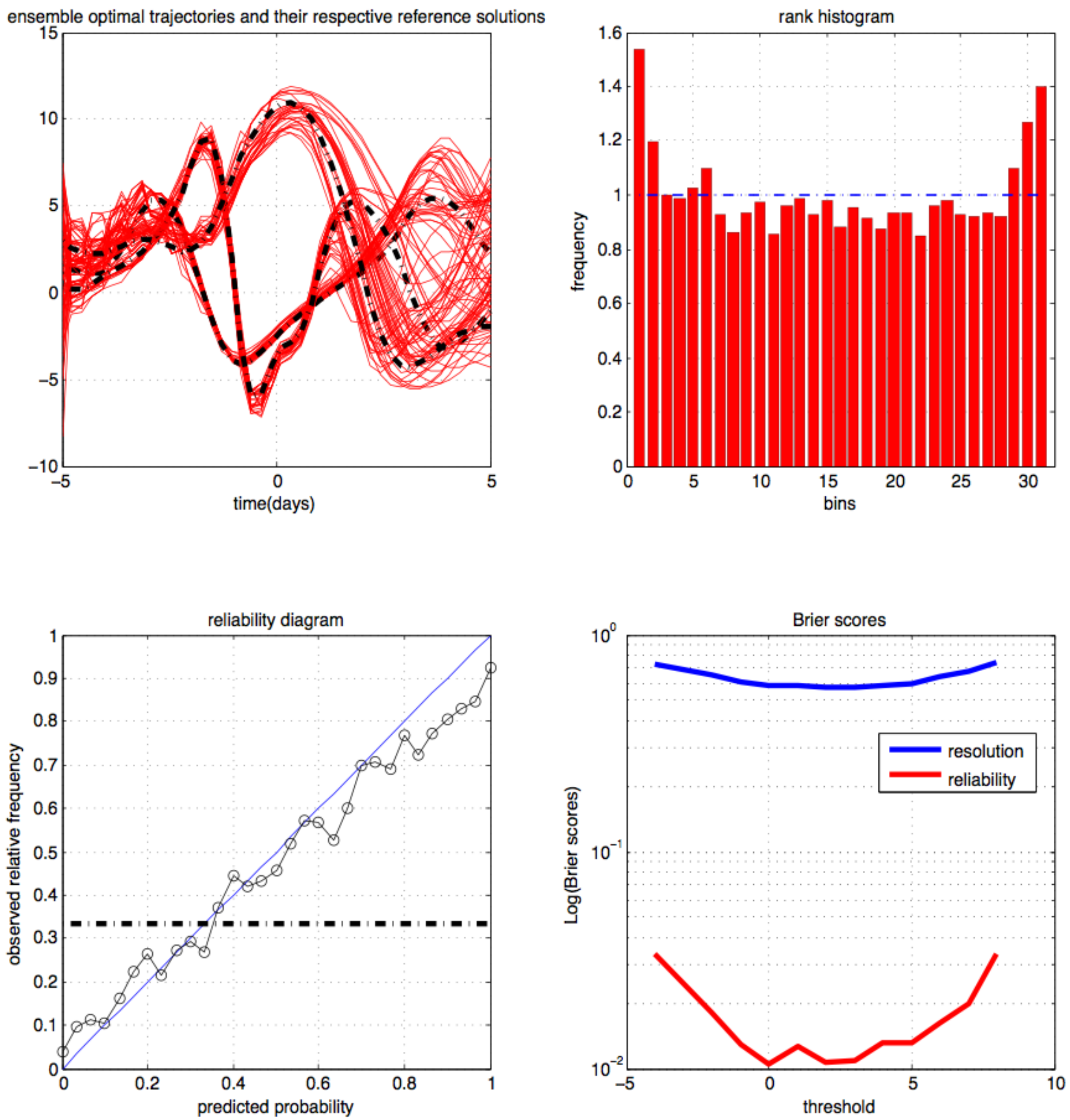





EnsVAR. Diagnostics for 5-day forecasts




EnKF. Diagnostics for 5-day forecasts




PF. Diagnostics for 5-day forecasts

| method \ DA procedure | *Assimilation* | Forecasting |
|---|---|---|
| EnsVAR | 0.2193510 | 1.49403506 |
| EnKF | 0.2449690 | 1.67176110 |
| PF | 0.7579790 | 2.62461295 |

RMS errors at the end of 5-day assimilations and 5-day forecasts

From course 6

# *Weak constraint* variational assimilation

Allows for errors in the assimilating model

- Data

  - Background estimate at time $0$

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

  - 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

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

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

In the present case, objective function of the form

$$(\xi_0, \eta_1, \ldots, \eta_{K\text{-}1}) \rightarrow$$

$$\mathcal{J}(\xi_0, \eta_1, \ldots, \eta_{K\text{-}1})$$

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

$$+ (1/2)\ \Sigma_{k=0,\ldots,K\text{-}1}\, \eta_k^{\mathrm{T}} Q_k^{-1} \eta_k$$

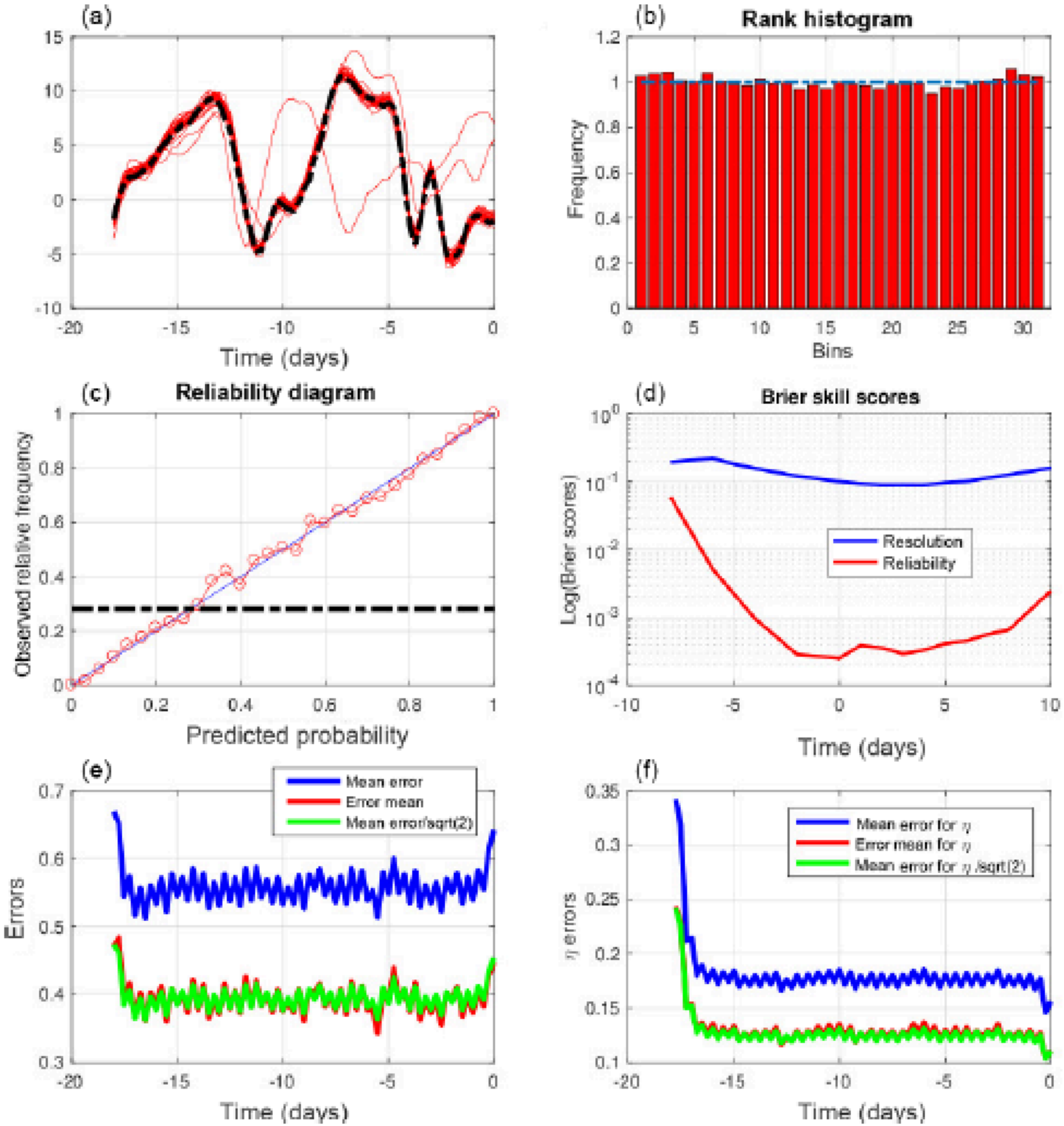
subject to

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

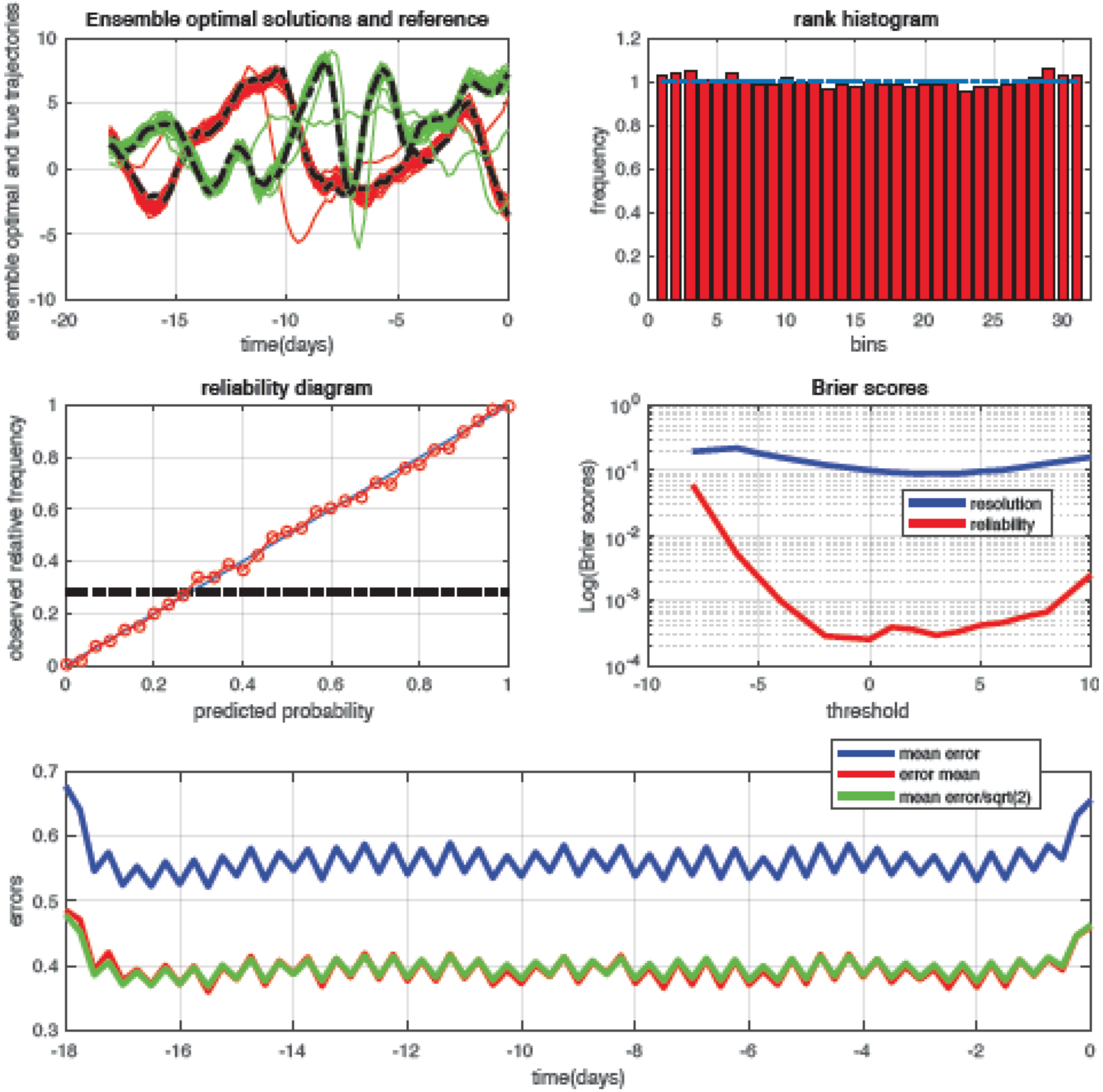
'Observations' consist of

- sequence $\{y_k\}$ , $\quad k = 0, \ldots, K$ (with unit observation operator $H_k$)
- observations $0$ for $\eta_k$ , $\quad k = 0, \ldots, K\text{-}1$

Weak-constraint ensemble variational assimilation
18 days, Q = 0.1
1200 realizations



Weak constraint EnsVar 18 days assimilation, Q=0.1 and 1200 realisations



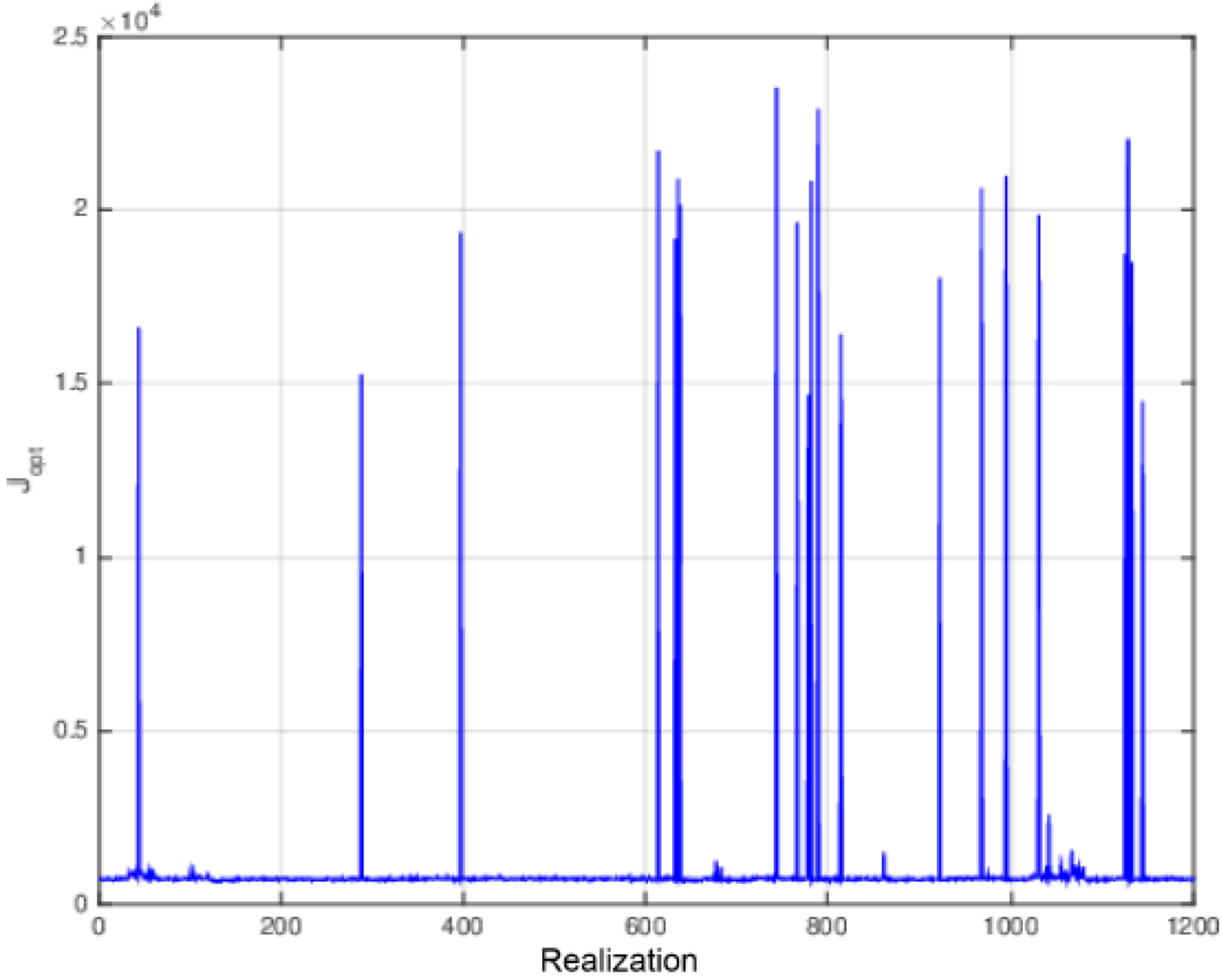
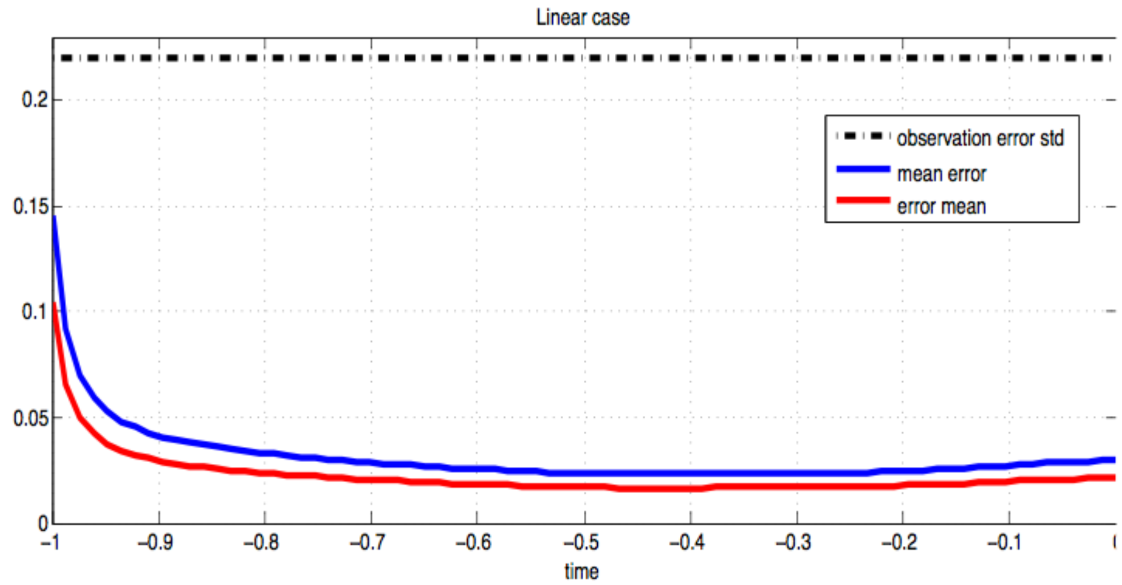



**Figure 11.** Values of (half) the minima of the objective function for all realizations of the weak-constraint assimilations over 18-day windows.

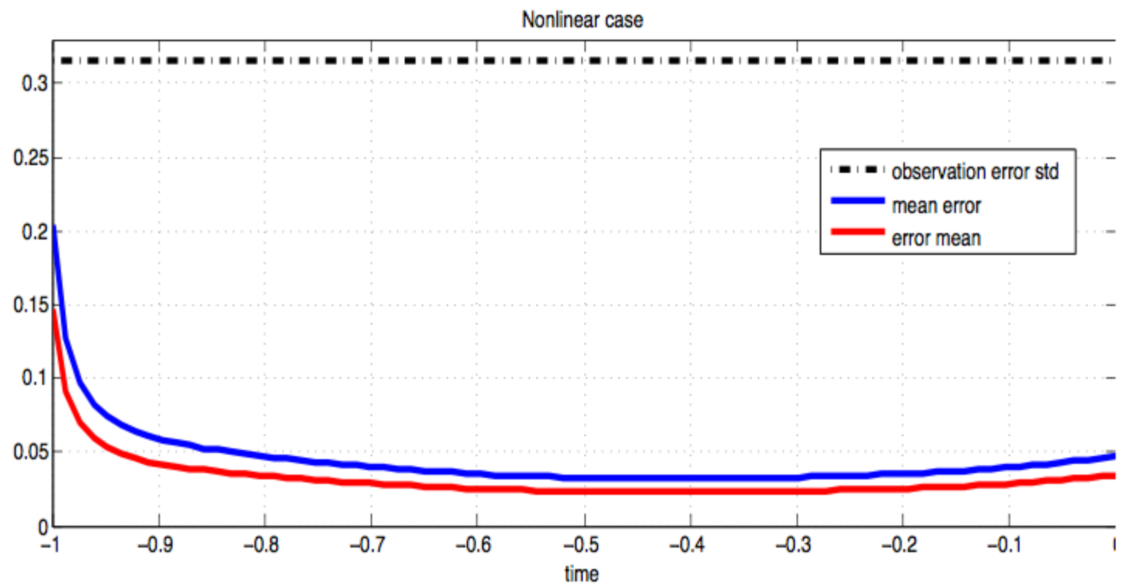
Kuramoto-Sivashinsky equation

$$\frac{\partial u}{\partial t} + \frac{\partial^4 u}{\partial x^4} + \frac{\partial^2 u}{\partial x^2} + u\frac{\partial u}{\partial x} = 0, \quad x \in [0, L]$$

with periodicity in $x$, $L = 32\pi$



Linear case
observation error std
mean error
error mean
time

Nonlinear case
observation error std
mean error
error mean
time

**Summary**

- Under non-linearity and non-Gaussianity the EnsVar is a reliable and consistent ensemble estimator (provided the QSVA is used for long DA windows) .
- EnsVar is at least as good an estimator as EnKF and PF.
- Similar results have been obtained for the Kuramuto-Sivashinsky model.

Ensembles obtained are Gaussian, even if errors in data are not

Produces Monte-Carlo sample of (probably not) bayesian pdf

# EnsVar : Pros and cons

## Pros

- Easy to implement when having a 4D-Var code
- Highly parallelizable
- No problems with algorithm stability (i.e. no ensemble collapse, no need for localization and inflation, no need for weight resampling)
- Propagates information in both ways and takes into account temporally correlated errors

## Cons

- Costly (Nens 4D-Var assimilations).
- Empirical.
- Cycling of the process (work in progress).

# *History of Numerical Weather Prediction*

**Wilhelm Bjerknes**

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

École de Météorologie de Bergen

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

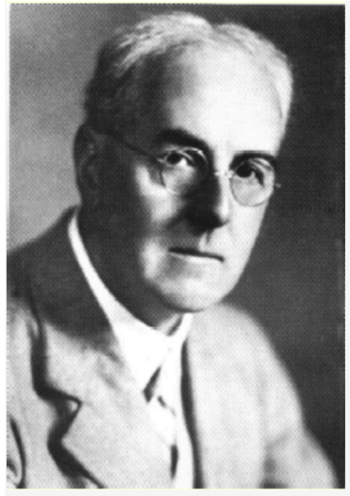
  $f(p, \rho, e) = 0$ $(p/\rho = rT, e = C_v T)$

- 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

## *History of Numerical Weather Prediction* (continuation 2)

**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, …)

Charney developed barotropic model

Jule Gregory Charney en 1978.

First operational numerical forecast around 1955 in Sweden

(C. G. Rossby)

## *History of Numerical Weather Prediction* (continuation 3)

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

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

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

Ensemble prediction

## *History of Numerical Weather Prediction* (continuation 4)

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 to be published in 2021-2022

# Cours à venir

~~Jeudi 19 Mars~~

~~Jeudi 26 mars~~

~~Jeudi 02 avril~~

~~Jeudi 09 avril~~

~~Mardi 21 avril, 14h00~~

~~Mardi 28 avril, 14h00~~

~~Mardi 12 mai, 14h00~~

Mardi 26 mai, 14h00