

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 8

18 Juin 2025

*Last course (June 11)*

- 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

## *This course*

- Particle filters (exactly bayesian)
- Ensemble Variational Assimilation
- Conclusions

## Exact bayesian estimation ?

### Particle filters

Predicted ensemble at time  $t$  :  $\{x^b_l, l = 1, \dots, L\}$ , each element with its own weight (probability)  $P(x^b_l)$

Observation vector at same time :  $y = H(x) + \varepsilon$

Bayes' formula

$$P(x^b_l|y) = P(y|x^b_l) P(x^b_l) / P(y)$$

Defines updating of weights

Bayes' formula

$$P(x^b_l|y) \sim P(y|x^b_l) P(x^b_l)$$

If error  $\varepsilon$  is independent of all previous data

$$P(y|x^b_l) = P[\varepsilon = y - H(x^b_l)]$$

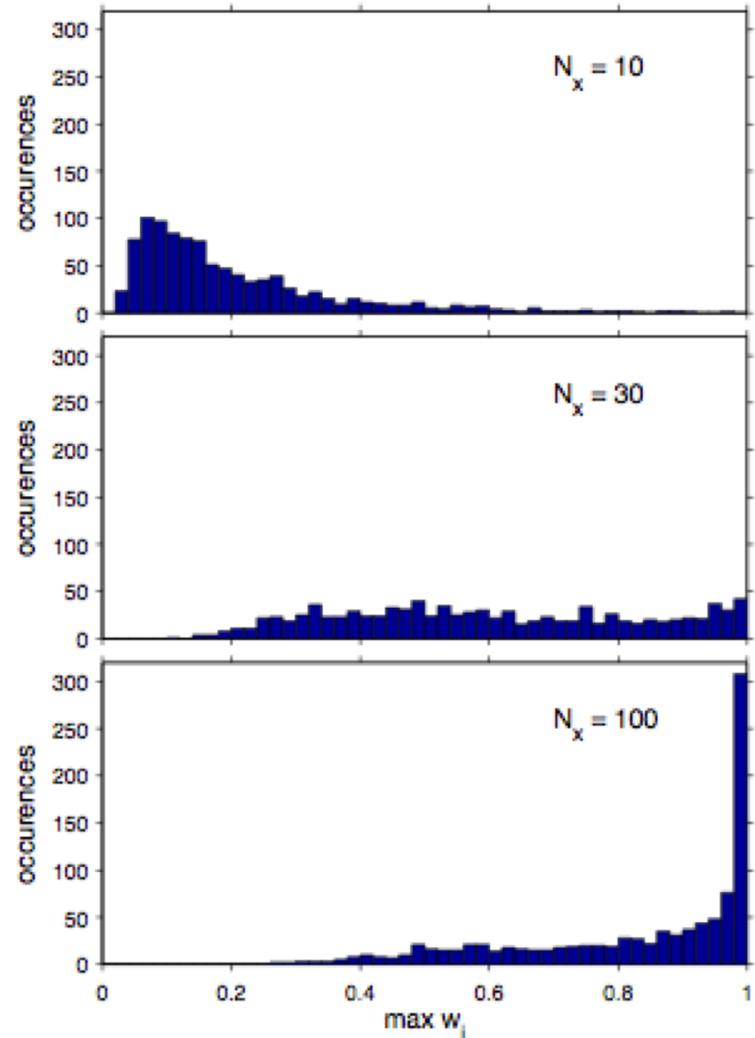
Defines updating of weights; particles are not modified. Asymptotically converges to bayesian pdf. Very easy to implement.

Observed fact. For large state dimension, ensemble tends to collapse.

# Behavior of $\max w^i$

---

▷  $N_e = 10^3$ ;  $N_x = 10, 30, 100$ ;  $10^3$  realizations



average squared error of posterior mean = 5.5

... = 25

... = 127

Problem originates in the ‘curse of dimensionality’. Large dimension pdf’s are very diffuse, so that very few particles (if any) are present in areas where conditional probability (‘*likelihood*’)  $P(y|x)$  is large.

## *Curse of dimensionality*

Standard one-dimensional gaussian random variable  $X$

$$P[|X| < \sigma] \approx 0.84$$

In dimension  $n = 100$ ,  $0.84^{100} = 3.10^{-8}$

$\chi^2$ -probability distribution of order  $p$

$$\chi^2(p) \sim \Sigma_p [\mathcal{N}(0, 1)]^2$$

Expectation  $m = p$ , variance  $\sigma^2 = 2p$

$$\sigma/m = \sqrt{(2/p)}$$

for large  $p$ , distribution is extremely peaked

Recall that, in gaussian variational assimilation,  $2E(\mathcal{J}_{min})$ , where  $\mathcal{J}_{min}$  is minimum of objective function, follows a  $\chi^2$ -probability distribution of order  $p$ , where  $p$  is the degree of overdetermination of the minimization

Bengtsson *et al.* (2008) and Snyder *et al.* (2008) evaluate that stability of filter requires the size of ensembles to increase exponentially with space dimension.

*Alternative possibilities* (review in van Leeuwen, 2017, *Annales de la faculté des sciences de Toulouse Mathématiques*, **26** (4), 1051-1085)

*Resampling*. Define new ensemble.

Simplest way. Draw new ensemble according to probability distribution defined by the updated weights. Give same weight to all particles. Particles are not modified, but particles with low weights are likely to be eliminated, while particles with large weights are likely to be drawn repeatedly. For multiple particles, add noise, either from the start, or in the form of ‘model noise’ in ensuing temporal integration.

Random character of the sampling introduces noise. Alternatives exist, such as *residual sampling* (Lui and Chen, 1998, van Leeuwen, 2003). Updated weights  $w_l$  are multiplied by ensemble dimension  $L$ . Then  $p$  copies of each particle  $l$  are taken, where  $p$  is the integer part of  $Lw_l$ . Remaining particles, if needed, are taken randomly from the resulting distribution.

However, resampling is not sufficient to avoid degeneracy of filters.

## *Markov chain Monte Carlo (MCMC) Methods*

*Sequence of random vectors*  $\{x^n, n = 0, \dots\}$

Assume  $P(x^n | x^{n-1}, \dots, x^0) = P(x^n | x^{n-1})$

**Markovianity.** Verified in particular if  $x^n = F(x^{n-1}, \eta)$ , where  $F$  is deterministic, and  $\eta$  is random with *a priori* known probability distribution.

*Sequence of observations*  $\{y^n, n = 0, \dots\}$

Assume  $P(y^n | x^n, x^{n-1}, \dots, x^0) = P(y^n | x^n)$

Verified in particular if  $y^n = G(x^n, \varepsilon)$ , where  $G$  is deterministic, and  $\varepsilon$  is random with *a priori* known probability distribution.

We want to estimate  $P(x^n | y^n, \dots, y^0) \equiv P(x^n | y^{0:n})$

$$\begin{aligned} P(x^n | y^{0:n}) &= P(x^n | y^n, y^{0:n-1}) = P(y^n | x^n, y^{0:n-1}) P(x^n | y^{0:n-1}) / P(y^n | y^{0:n-1}) \\ &= P(y^n | x^n) P(x^n | y^{0:n-1}) / P(y^n | y^{0:n-1}) \end{aligned}$$

$$P(x^n | y^{0:n-1}) = \int P(x^n | x^{n-1}) P(x^{n-1} | y^{0:n-1}) dx^{n-1}$$

### ***Chapman-Kolmogorov equation***

This defines a recursive equation for  $P(x^n | y^{0:n})$

Particular case

$$\begin{aligned} x^n &= M_n x^{n-1} + \eta_n & M_n \text{ linear, } \eta_n \text{ Gaussian with } a \text{ priori known pdf} \\ y^n &= H_n x^n + \varepsilon_n & H_n \text{ linear, } \varepsilon_n \text{ Gaussian with } a \text{ priori known pdf} \end{aligned}$$

$\Rightarrow$  Kalman filter

Idea :

Use a *proposal density* that is closer to the new observations than the density defined by the predicted particles (for instance the density defined by EnKF, after the latter has used the new observations).

We are now to discuss a very interesting property of particle filters that has received little attention in the geophysical community. We start from Bayes:

$$p(x^{0:n}|y^{0:n}) = \frac{p(y^n|x^n)p(x^n|x^{n-1})}{p(y^n)}p(x^{0:n-1}|y^{0:n-1}). \quad (5.1)$$

To simplify the analysis, and since we concentrate on a filter here, let us first integrate out the past, to get:

$$p(x^n|y^{0:n}) = \frac{p(y^n|x^n)}{p(y^n)} \int p(x^n|x^{n-1})p(x^{n-1}|y^{0:n-1}) dx^{n-1}. \quad (5.2)$$

This expression does not change when we multiply and divide by a so-called proposal transition density  $q(x^n|x^{n-1}, y^n)$ , so:

$$\begin{aligned} p(x^n|y^{0:n}) \\ &= \frac{p(y^n|x^n)}{p(y^n)} \int \frac{p(x^n|x^{n-1})}{q(x^n|x^{n-1}, y^n)} q(x^n|x^{n-1}, y^n) p(x^{n-1}|y^{0:n-1}) dx^{n-1}. \end{aligned} \quad (5.3)$$

As long as the support of  $q(x^n|x^{n-1}, y^n)$  is equal to or larger than that of  $p(x^n|x^{n-1})$  we can always do this. This last condition makes sure we don't divide by zero. Let us now assume that we have an equal-weight ensemble of particles from the previous analysis at time  $n - 1$ , so

$$p(x^{n-1}|y^{0:n-1}) = \sum_{i=1}^N \frac{1}{N} \delta_{x_i^{n-1}}. \quad (5.4)$$

Using this in the equation above gives:

$$p(x^n|y^{0:n}) = \sum_{i=1}^N \frac{1}{N} \frac{p(y^n|x^n)}{p(y^n)} \frac{p(x^n|x_i^{n-1})}{q(x^n|x_i^{n-1}, y^n)} q(x^n|x_i^{n-1}, y^n). \quad (5.5)$$

As a last step, we run the particles from time  $n - 1$  to  $n$ , i.e. we sample from the transition density. However, instead of drawing from  $p(x^n|x_i^{n-1})$ , so running the original model, we sample from  $q(x^n|x_i^{n-1}, y^n)$ , so from a modified model. Let us write this modified model as

$$x^n = g(x^{n-1}, y^n) + \hat{\beta}^n \quad (5.6)$$

so that we can write for the transition density, assuming  $\hat{\beta}^n$  is Gaussian distributed with covariance  $\hat{Q}$ :

$$q(x^n|x^{n-1}, y^n) = N(g(x^{n-1}, y^n), \hat{Q}). \quad (5.7)$$

Drawing from this density leads to:

$$p(x^n | y^{0:n}) = \sum_{i=1}^N \frac{1}{N} \frac{p(y^n | x_i^n)}{p(y^n)} \frac{p(x_i^n | x_i^{n-1})}{q(x_i^n | x_i^{n-1}, y^n)} \delta(x^n - x_i^n) \quad (5.8)$$

so the posterior pdf at time  $n$  can be written as:

$$p(x^n | y^{0:n}) = \sum_{i=1}^N w_i \delta_{x_i^n} \quad (5.9)$$

with weights  $w_i$  given by:

$$w_i = \frac{1}{N} \frac{p(y^n | x_i^n)}{p(y^n)} \frac{p(x_i^n | x_i^{n-1})}{q(x_i^n | x_i^{n-1}, y^n)}. \quad (5.10)$$

We recognise the first factor in this expression as the likelihood, and the second as a factor related to using the proposal transition density instead of the original transition density to propagate from time  $n - 1$  to  $n$ , so it is related to the use of the proposed model instead of the original model. Note that because the factor  $1/N$  and  $p(y^n)$  are the same for each particle and we are only interested in relative weights, we will drop them from now on, so

$$w_i = p(y^n | x_i^n) \frac{p(x_i^n | x_i^{n-1})}{q(x_i^n | x_i^{n-1}, y^n)}. \quad (5.11)$$

Several variants of proposal densities have been defined and studied : perform an EnKF up to observation time, and then use the obtained ensemble as proposal density, *nudge* the model integration between times  $n-1$  and  $n$  towards the observations at time  $n$ , perform a 4D-Var on each particle, *optimal proposal density*, ...

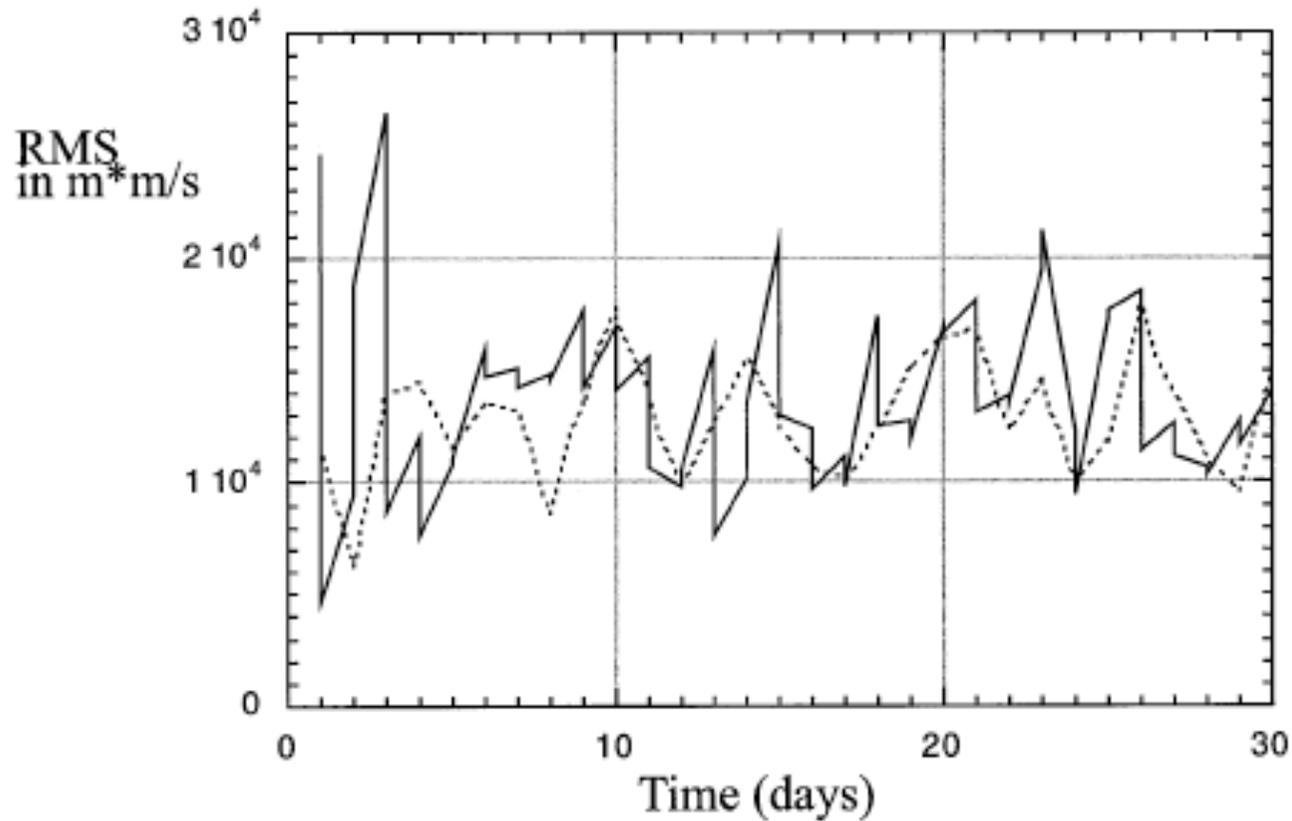


FIG. 12. Comparison of rms error ( $\text{m}^2 \text{s}^{-1}$ ) between ensemble mean and independent observations (dotted line) and the std dev in the ensemble (solid line). The excellent agreement shows that the SIRF is working correctly.

*The Equivalent-Weights Particle Filter (Ades and van Leeuwen, QJRMS, 2013).*

Make the proposal density depend on the whole ensemble at time  $n-1$ , and not only on  $x_l^{n-1}$ , use density of the form  $q(x^n | x^{n-1}_{1,L}, y^n)$ , where  $1,L$  denotes all ensemble indices, rather than of the more restrictive form  $q(x^n | x_l^{n-1}, y^n)$ . This gives many degrees of freedom which can be exploited for obtaining at time  $n$  an ensemble with almost equal weights.

*Example* Vorticity equation model with random error  $\eta$ .

$$\frac{D(\zeta+f)}{Dt} = \eta$$

State-vector dimension  $\approx 65,000$

Decorrelation time: 25 timesteps

One complete noisy model field observed every 50 timesteps

24 particles

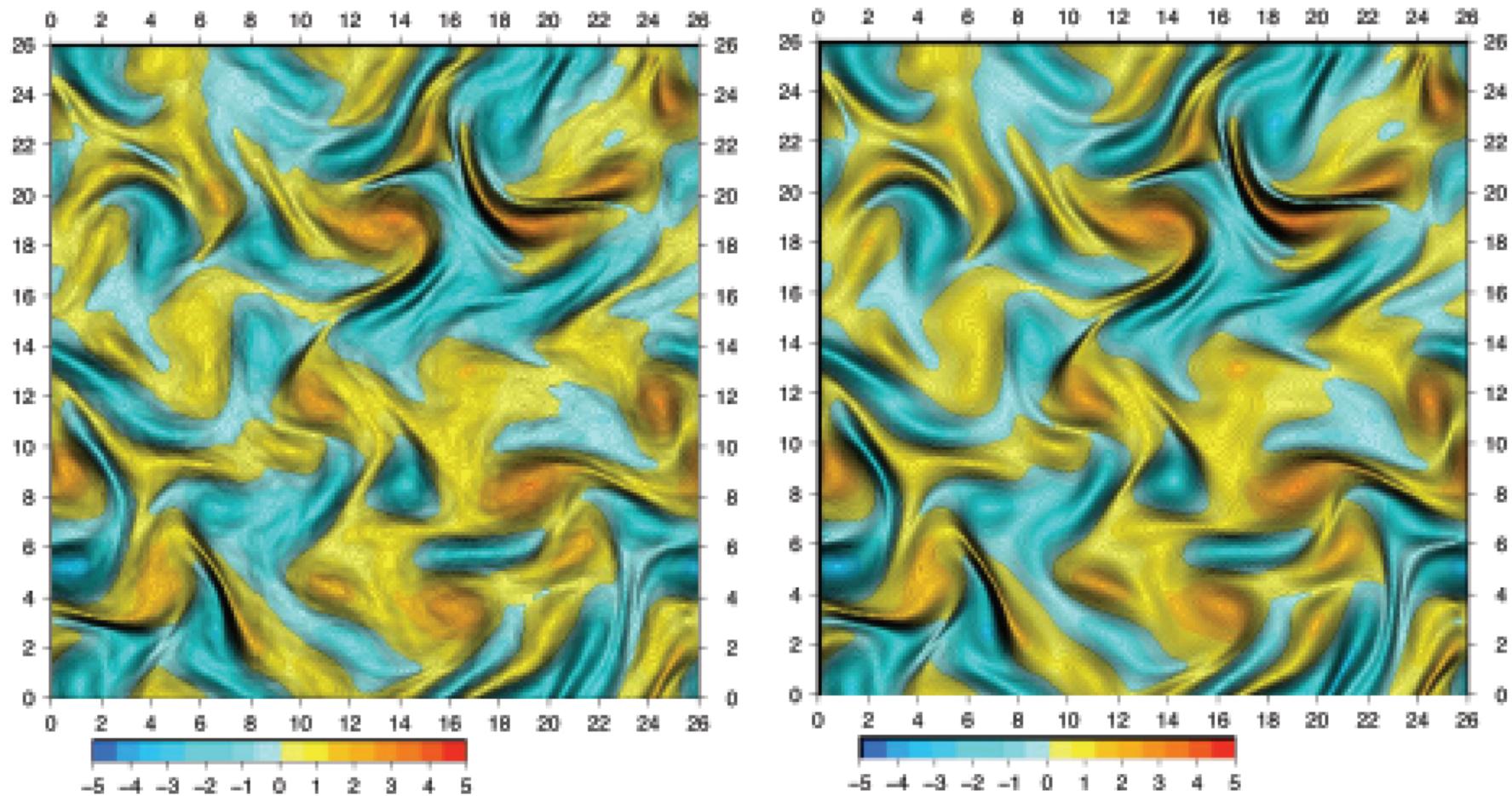


Figure 5.3. Snap shot of the vorticity field of the truth (right) and the particle filter mean (left) at time 25. Note the highly chaotic state of the fields, and the close to perfect tracking. (12 observations)

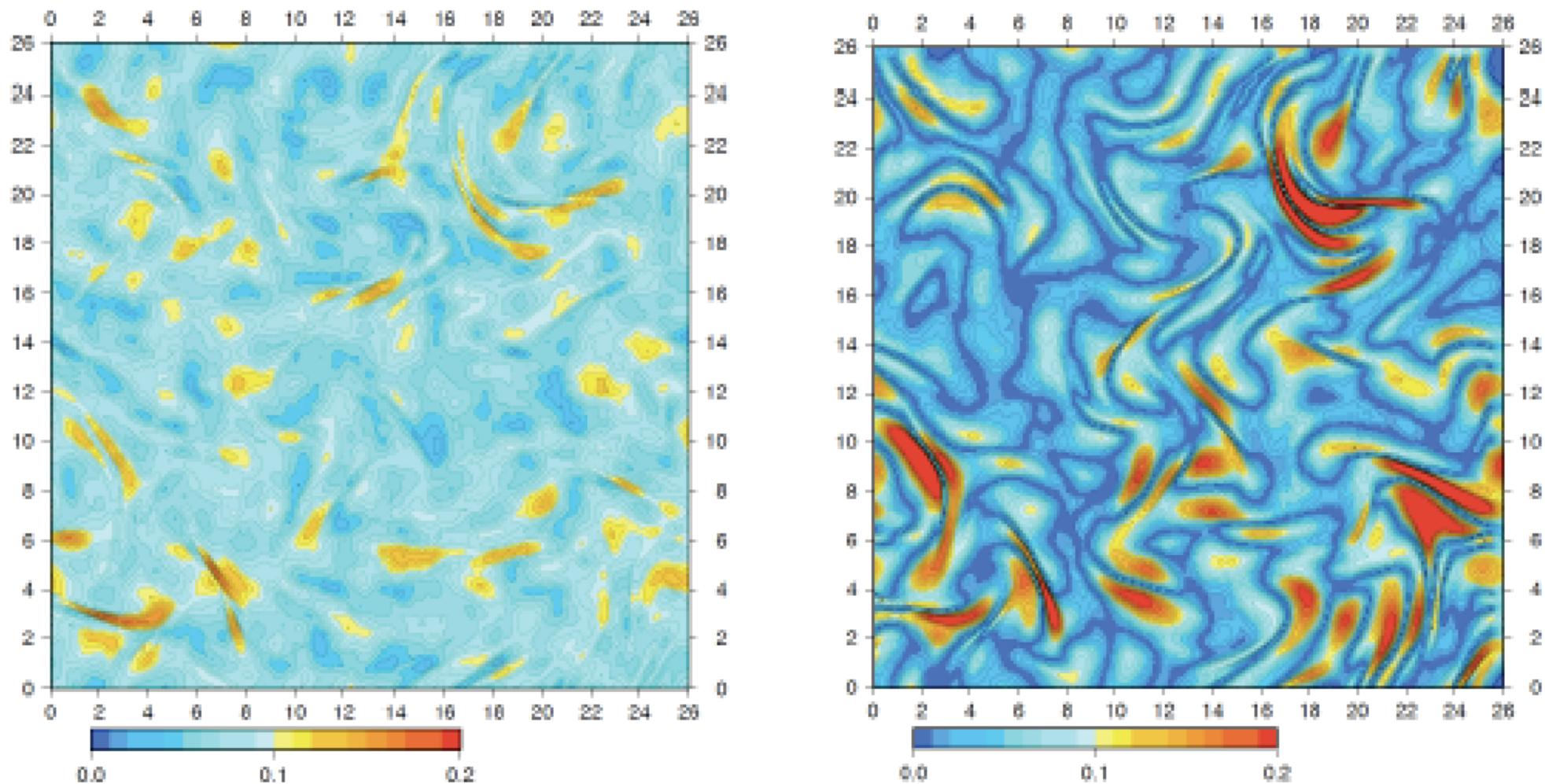


Figure 5.4. Snap shot of the absolute value of the mean-truth misfit and the standard deviation in the ensemble. The ensemble underestimates the spread at several locations, but averaged over the field it is slightly higher, 0.074 versus 0.056.

Bayesianity : experts say all these filters are bayesian  
(in the limit of infinite ensemble size)

Possible difficulties : numerical implementation,  
numerical cost

Particle filters are actively studied (van Leeuwen,  
Morzfeld, ...)

- Ensemble Variational Assimilation (*EnsVAR*).

(work with M. Jardak, 2018)

## *Ensemble Variational Assimilation*

**Data of the form**

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

Conditional probability distribution is

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

with

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

## Variational form

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

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

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

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

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

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

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

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.

Used at ECMWF and Météo-France (under the name *Ensemble of Data Assimilations, EDA*) for defining initial conditions of ensemble prediction, and also for defining background error covariance matrix in 4D-Var, but not for assimilation *per se*.

## 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

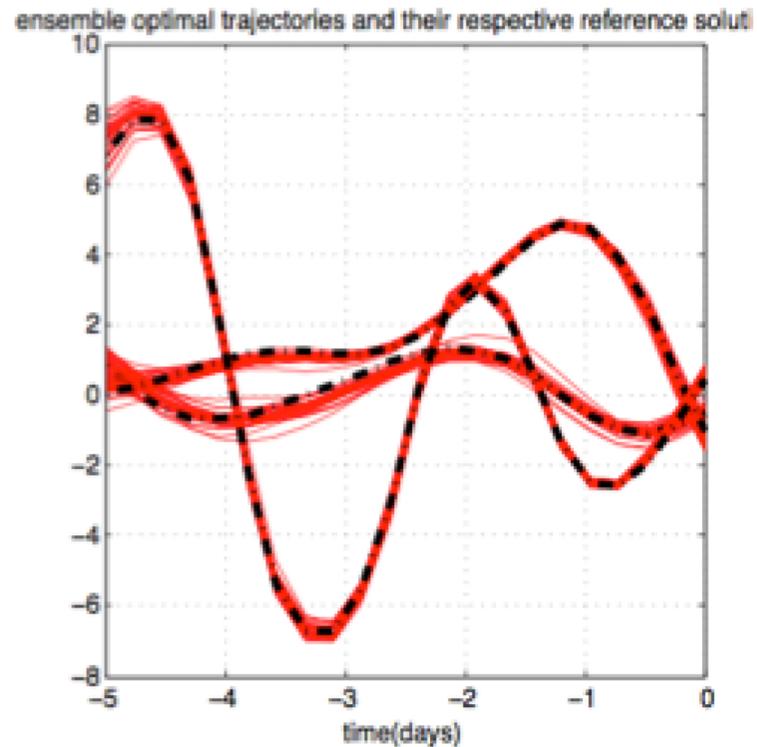
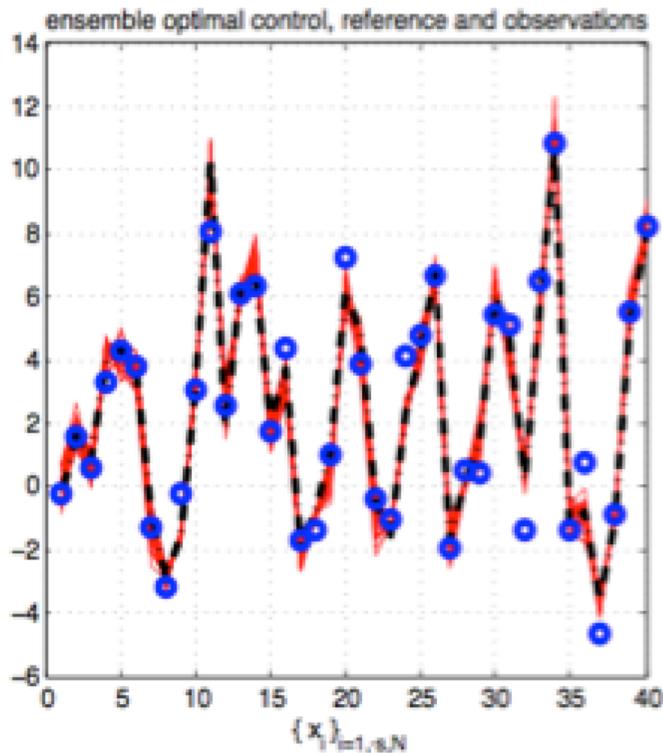
$$\frac{dx_k}{dt} = (x_{k+1} - x_{k-2})x_{k-1} - x_k + F \quad \text{for } k = 1, \dots, N$$

- 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  $N_{ens} = 30$ , number of ensemble members.
- ⑥  $\frac{1}{\lambda_{max}} \simeq 2.5days$ ,  $\lambda_{max}$  the largest Lyapunov exponent.
- ⑦  $\Delta t = 0.05 = 6hours$ , 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  $t_k$  (every 12 hours) 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

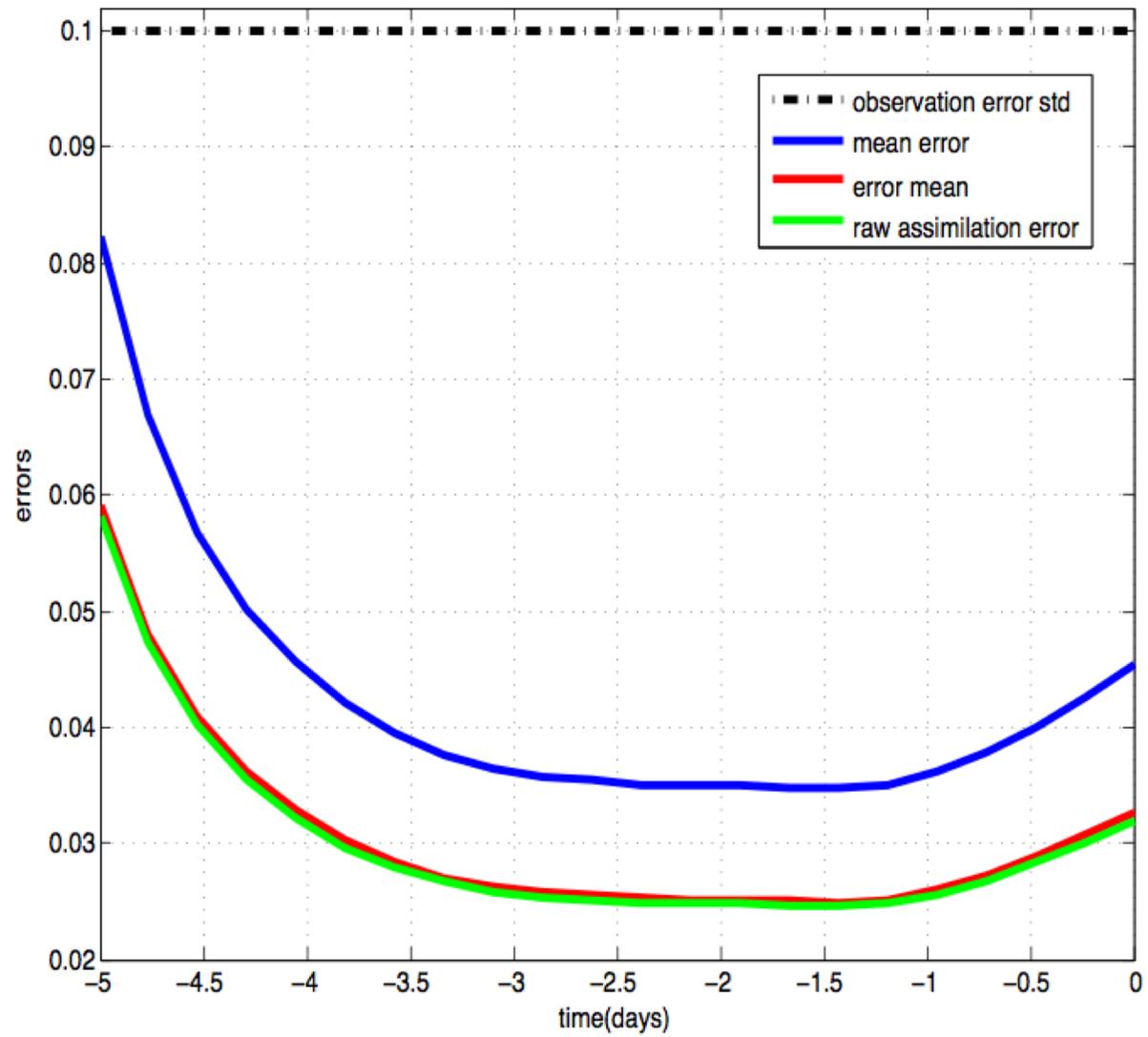
$$y_k \rightarrow z_k = y_k + \delta_k$$

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 circumstances 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

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

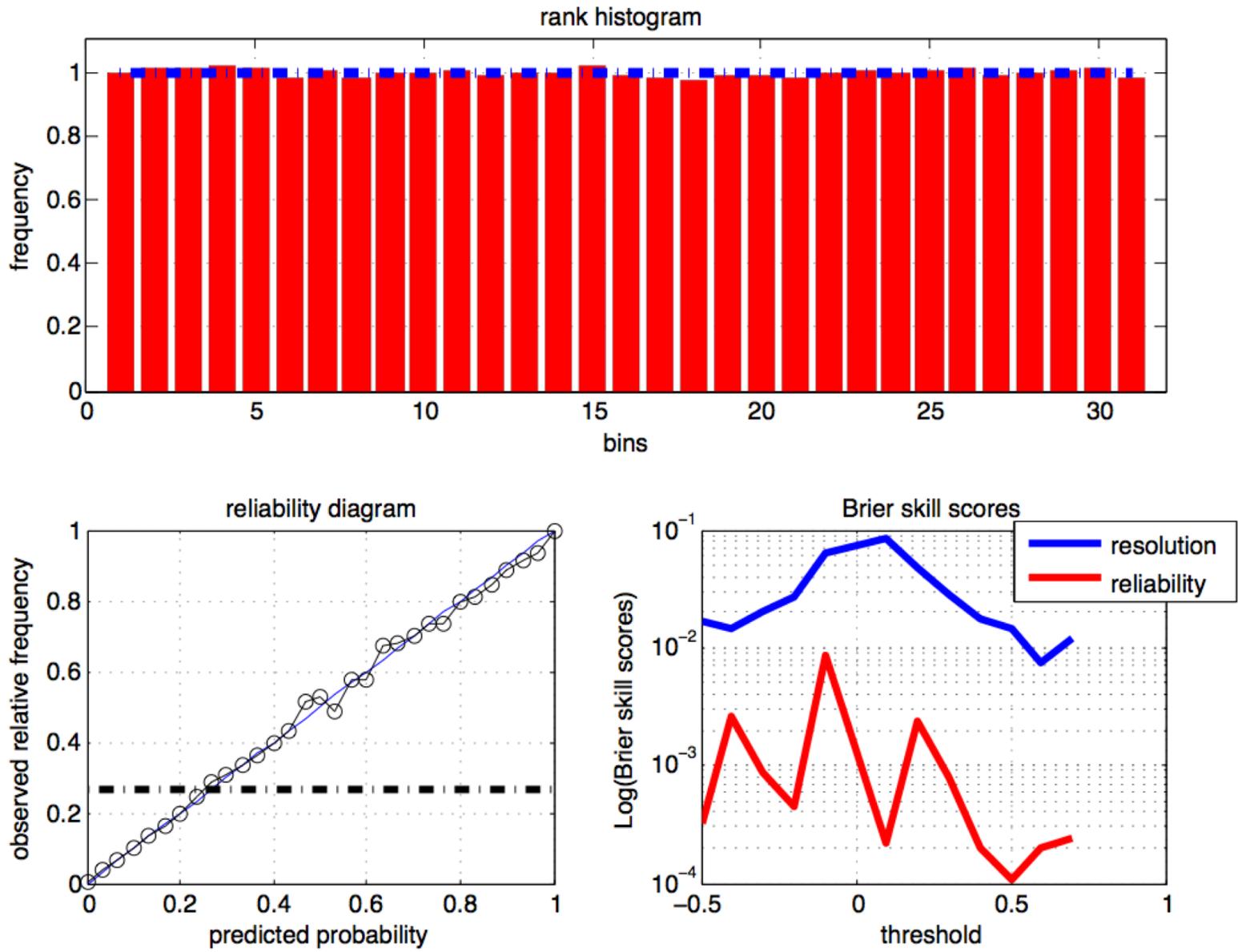
Reliability can be objectively validated, provided a large enough sample of realizations of the estimation system is available.

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 the capability of the estimation system to *a priori* distinguish between different situations. It 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.

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

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



Linearized Lorenz'96. 5 days

## Objective function

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

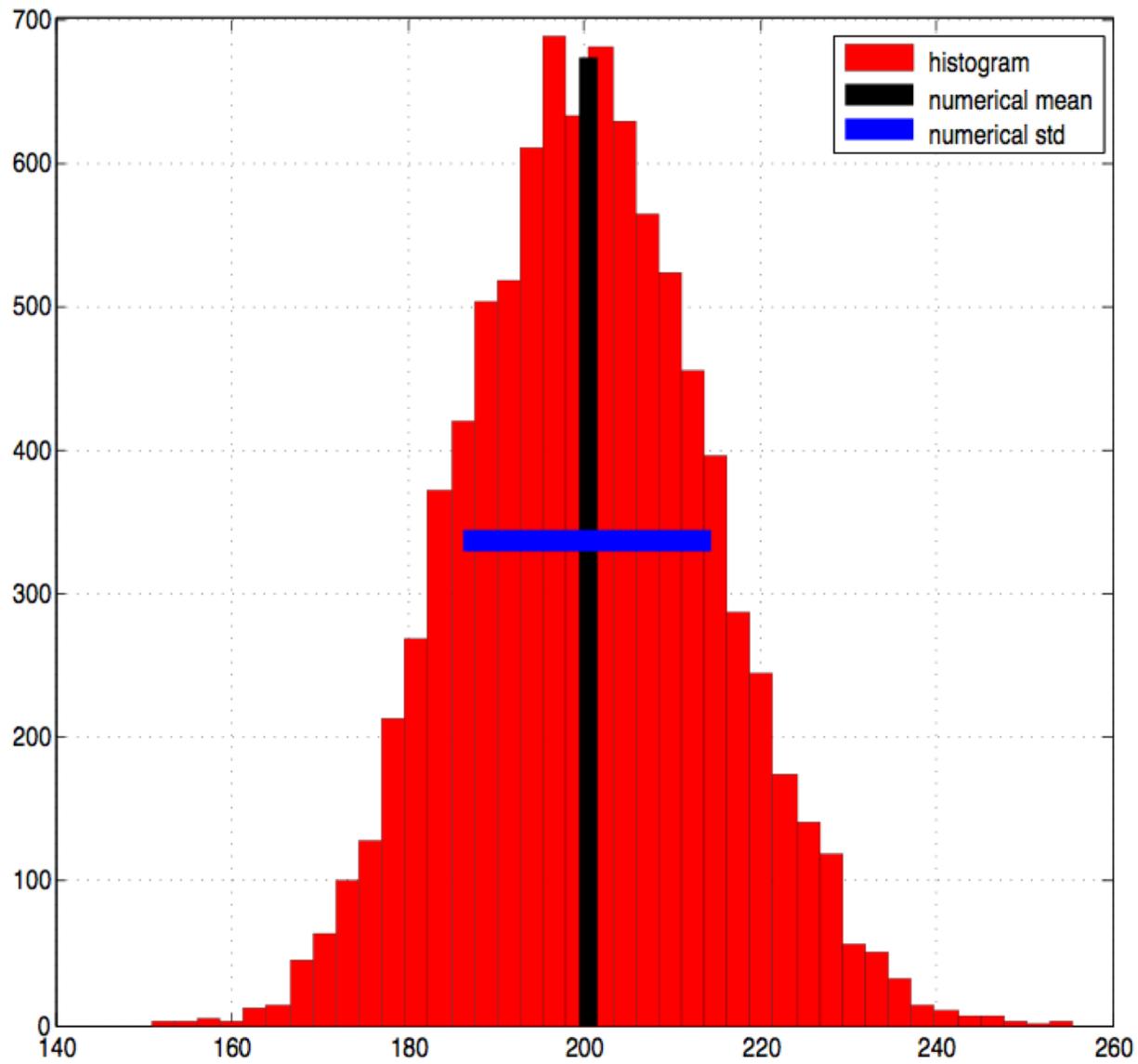
$$\begin{aligned} \mathcal{J}_{min} \equiv \mathcal{J}(\mathbf{x}^a) &= (1/2) [\Gamma\mathbf{x}^a - \mathbf{z}]^T S^{-1} [\Gamma\mathbf{x}^a - \mathbf{z}] \\ &= (1/2) \mathbf{d}^T [E(\mathbf{d}\mathbf{d}^T)]^{-1} \mathbf{d} \end{aligned}$$

where  $\mathbf{d}$  is innovation

$$\Rightarrow E(\mathcal{J}_{min}) = p/2 \quad (p = \dim \mathbf{y} = \dim \mathbf{d})$$

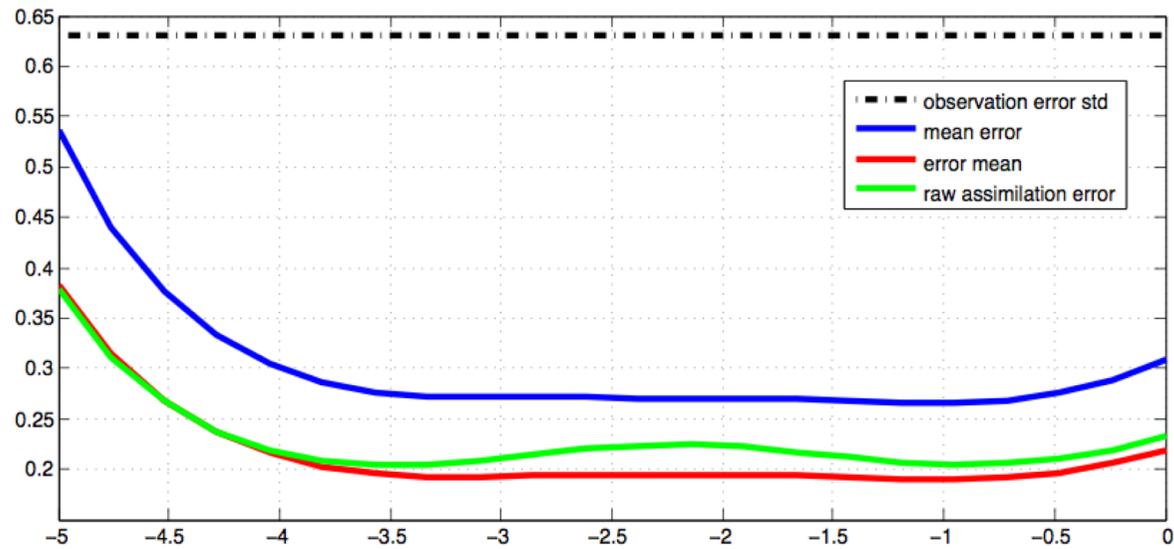
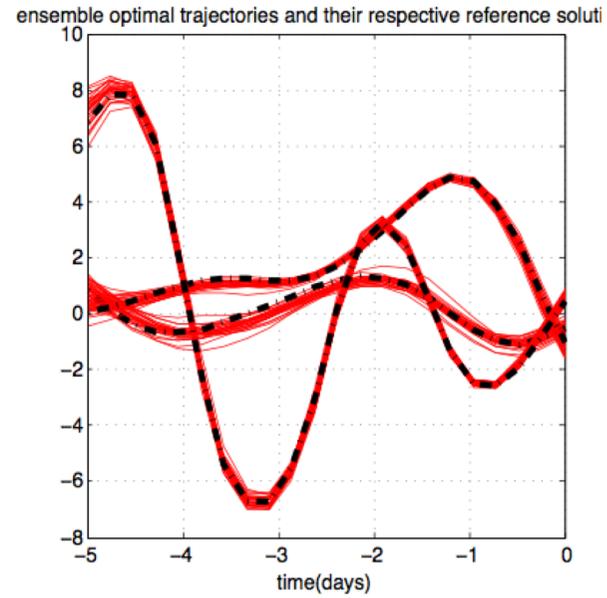
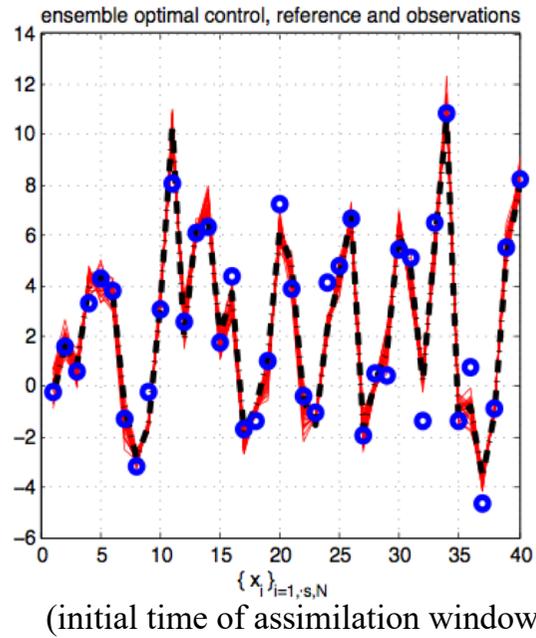
If  $p$  is large, a few realizations are sufficient for determining  $E(\mathcal{J}_{min})$

*Remark.* If in addition errors are gaussian, the quantity  $2E(\mathcal{J}_{min})$  follows a  $\chi^2$ -probability distribution of order  $p$ . For that reason the criterion  $E(\mathcal{J}_{min}) = p/2$  is often called the  $\chi^2$  criterion. Also  $Var(\mathcal{J}_{min}) = p/2$  in the gaussian case.

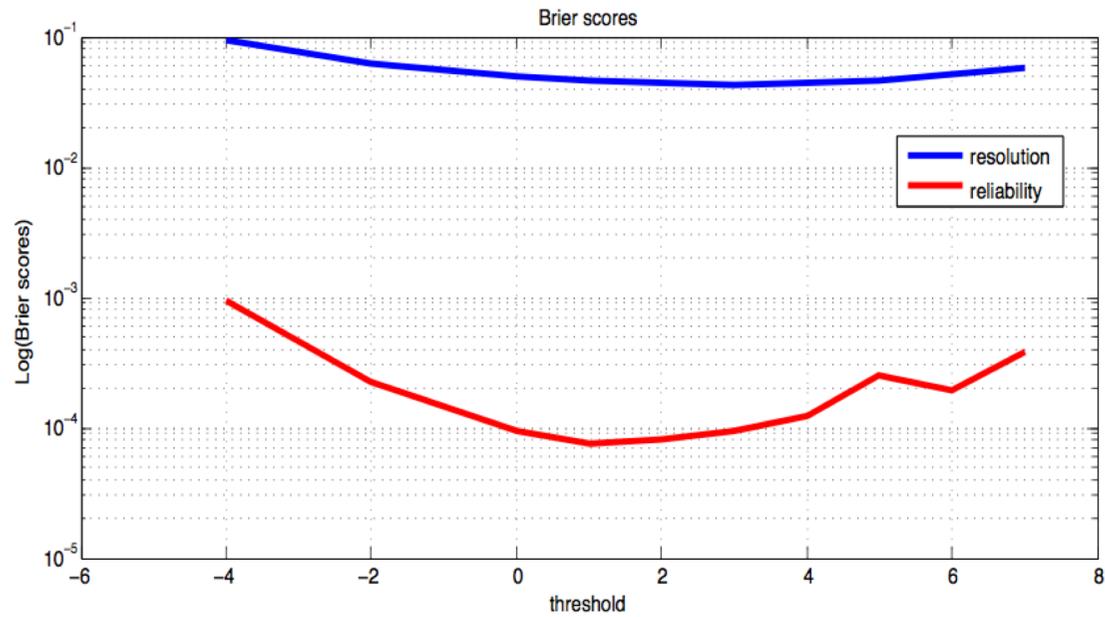
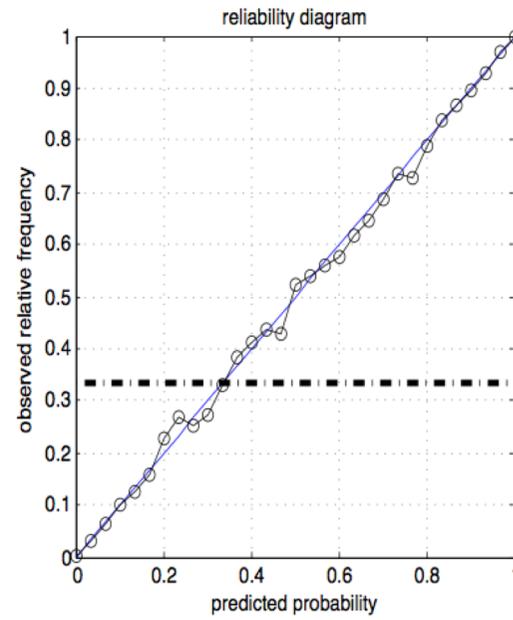
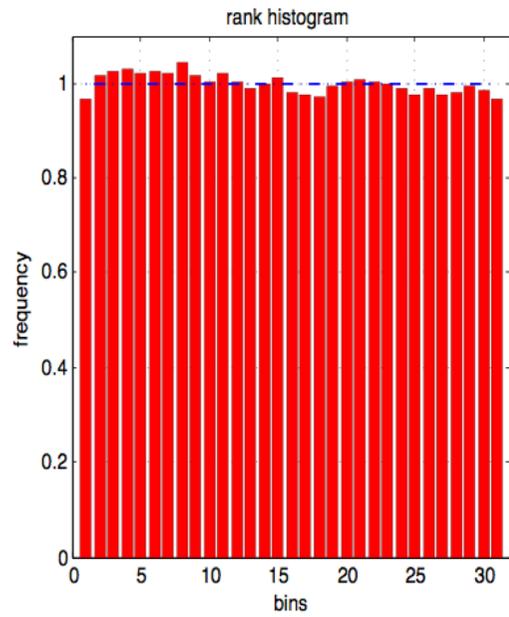


Linearized Lorenz'96. 5 days. Histogram of  $\mathcal{J}_{min}$   
 $E(\mathcal{J}_{min}) = p/2 (=200)$  ;  $\sigma(\mathcal{J}_{min}) = \sqrt{(p/2)} (\approx 14.14)$

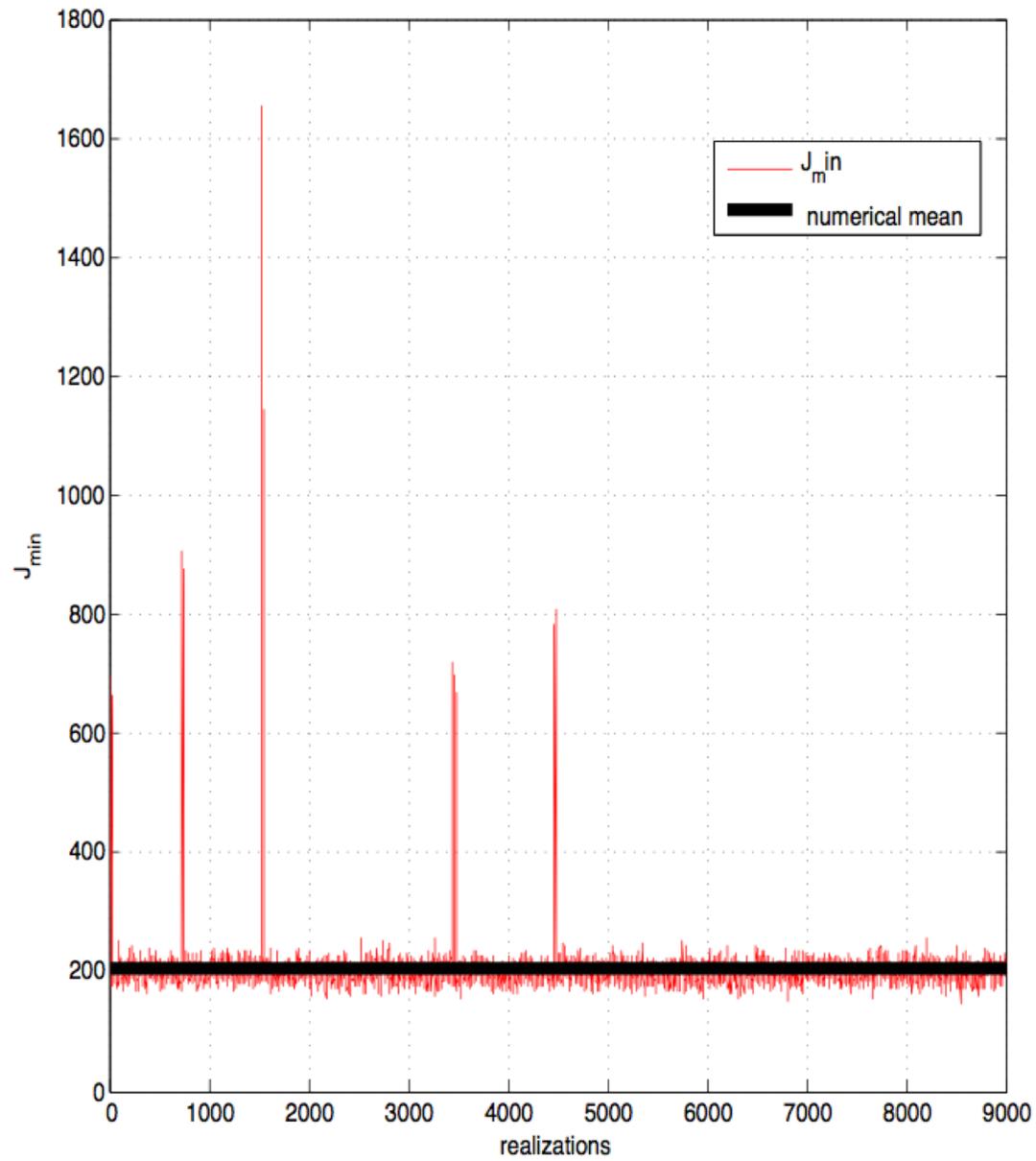
Observed values 199.39 and 14.27



Nonlinear Lorenz'96. 5 days

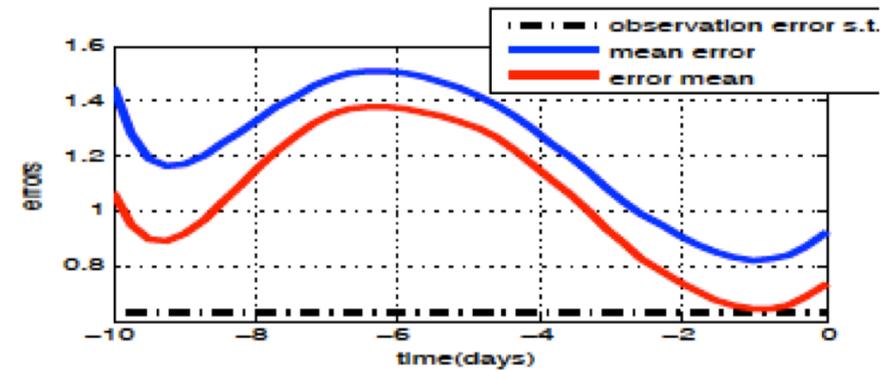
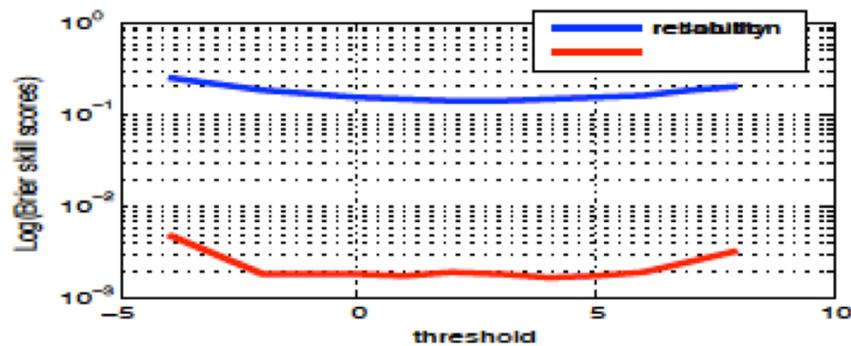
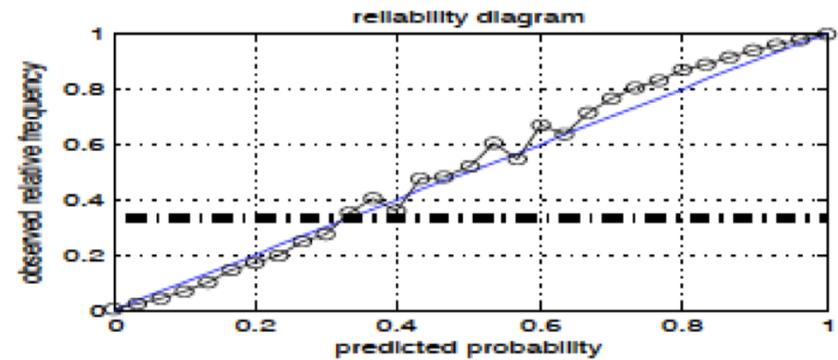
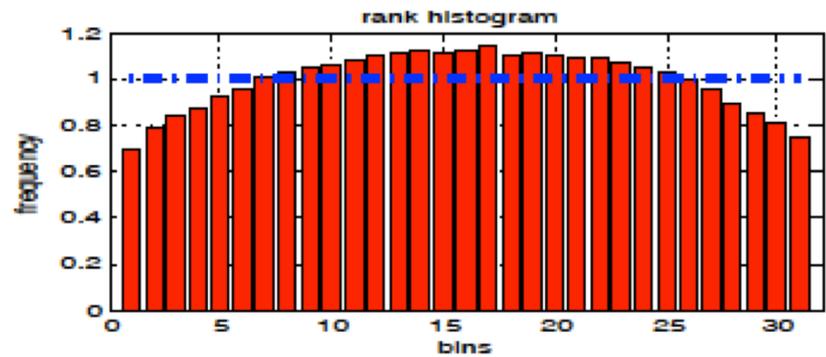
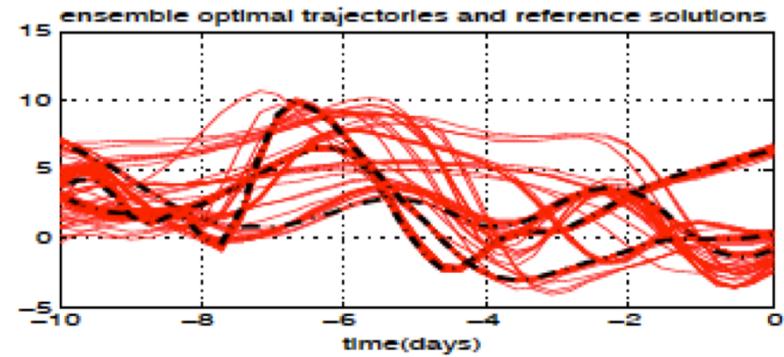
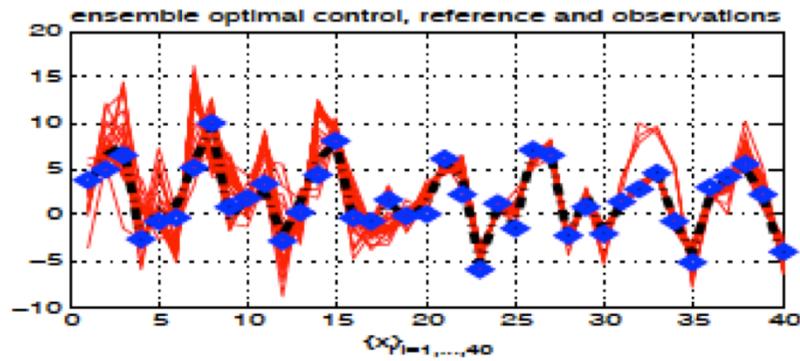


Nonlinear Lorenz'96. 5 days

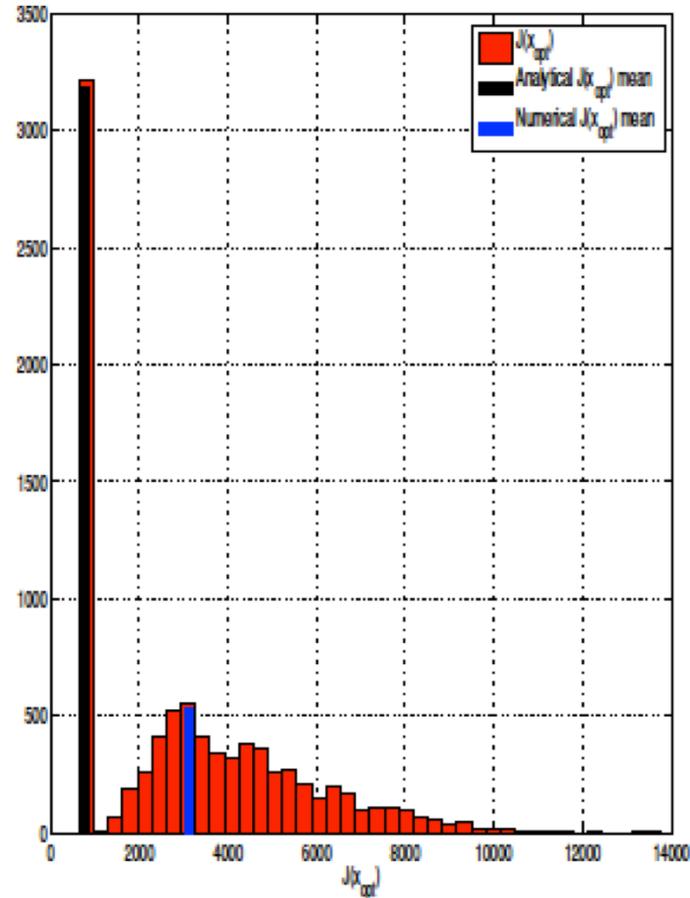


Nonlinear Lorenz'96. 5 days. Histogram of  $J_{\min}$

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

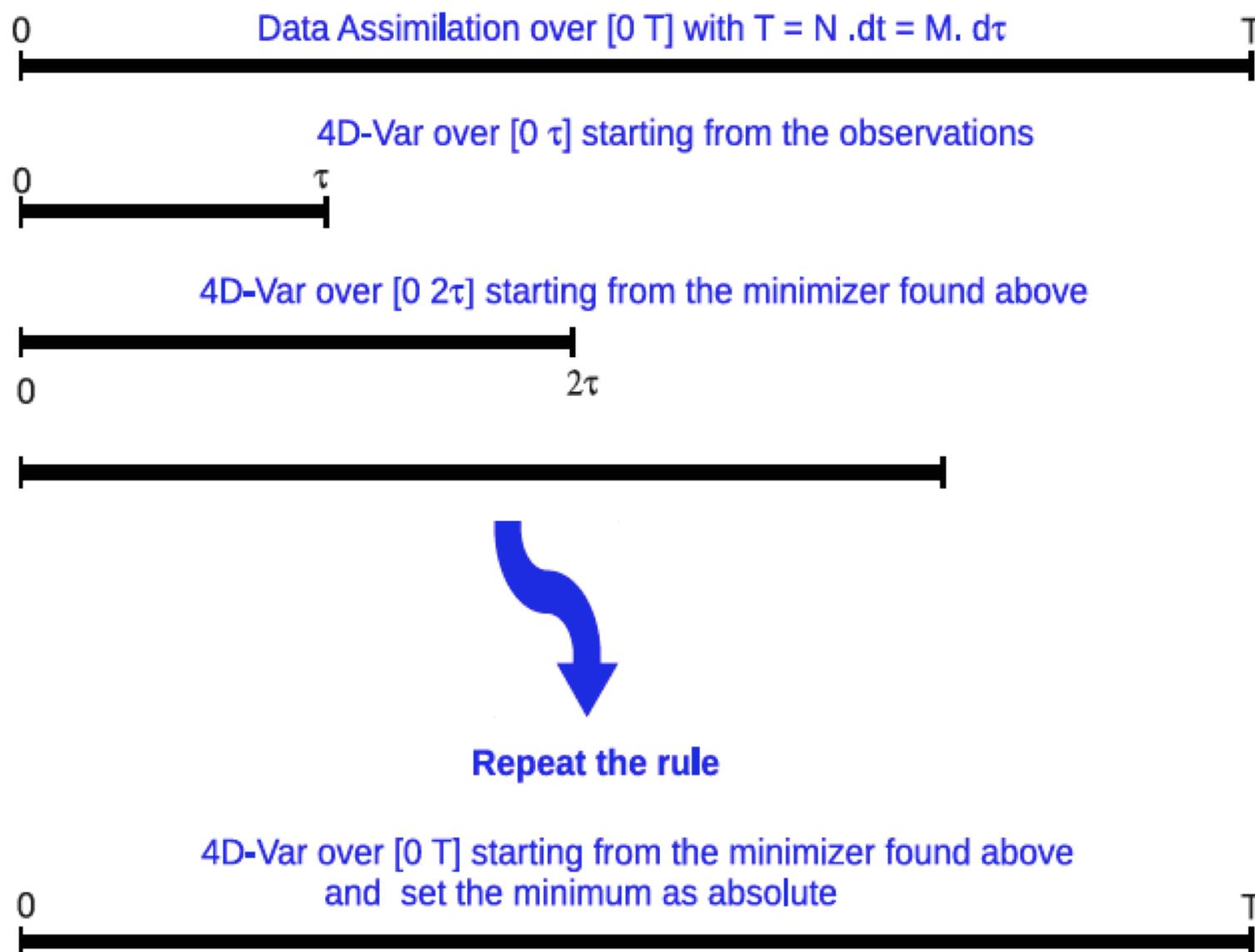


# EnsVar : consistency

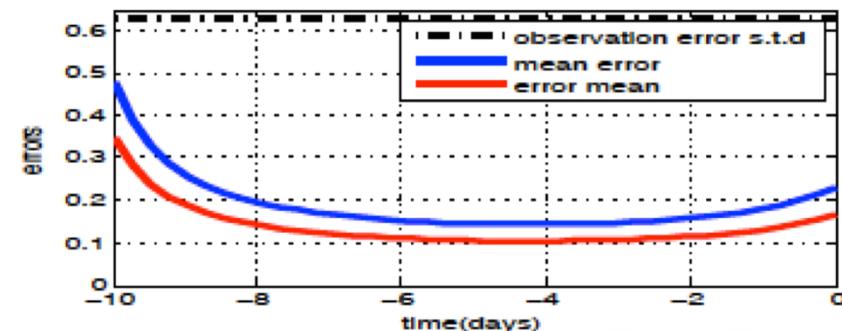
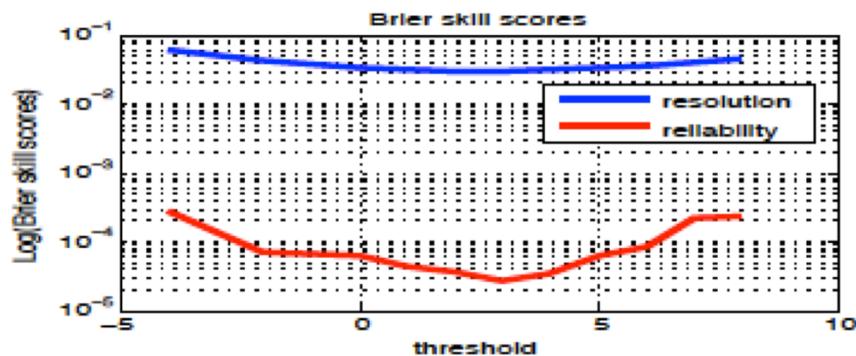
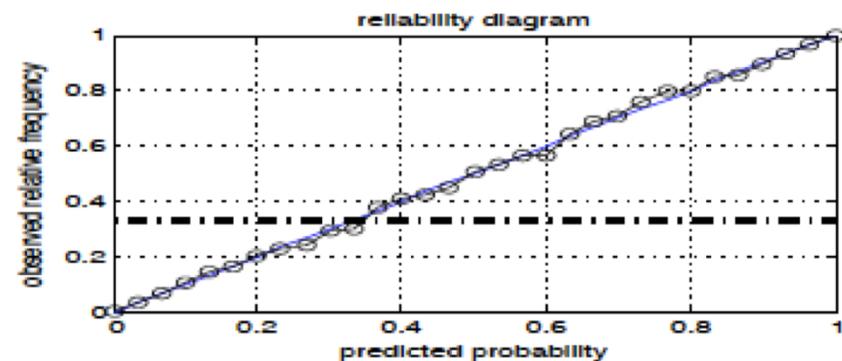
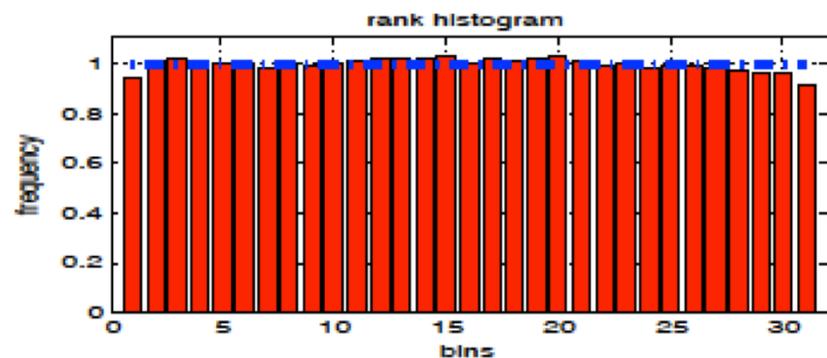
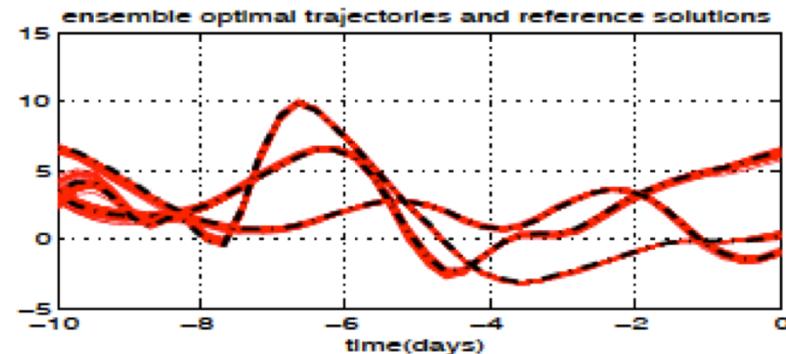
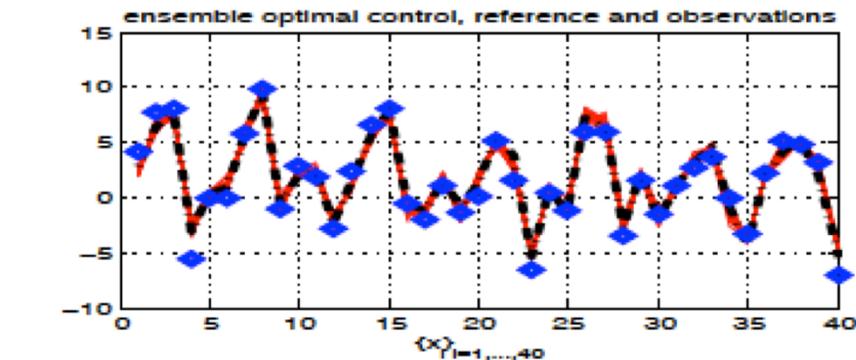


Nonlinear Lorenz'96. 10 days. Histogram of  $J_{min}$

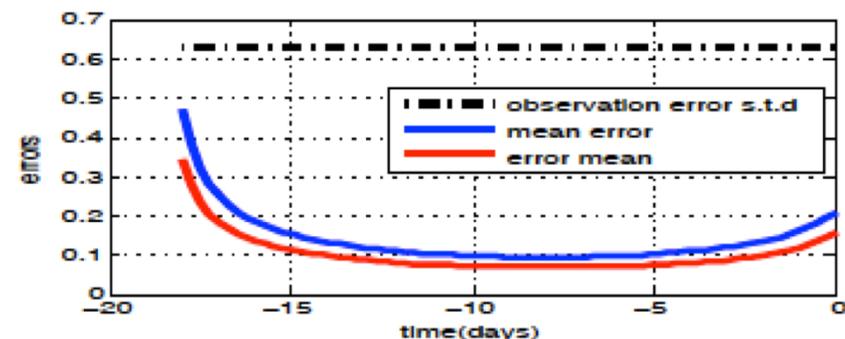
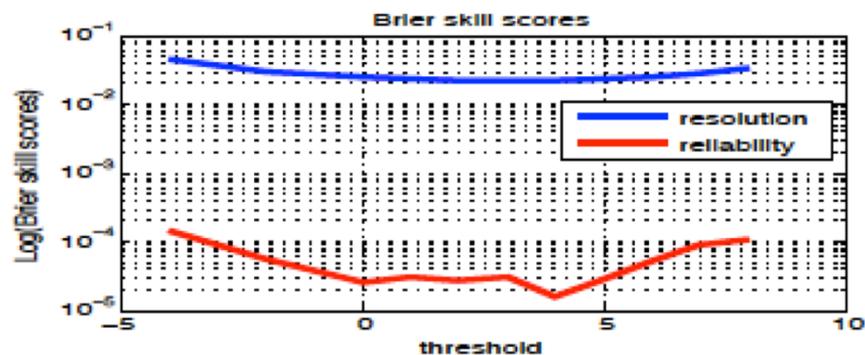
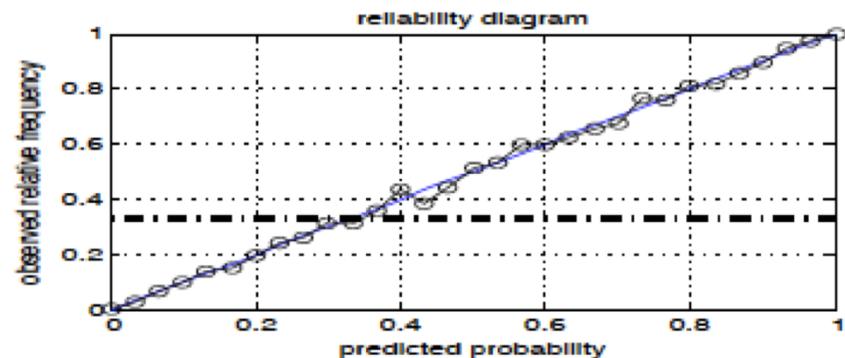
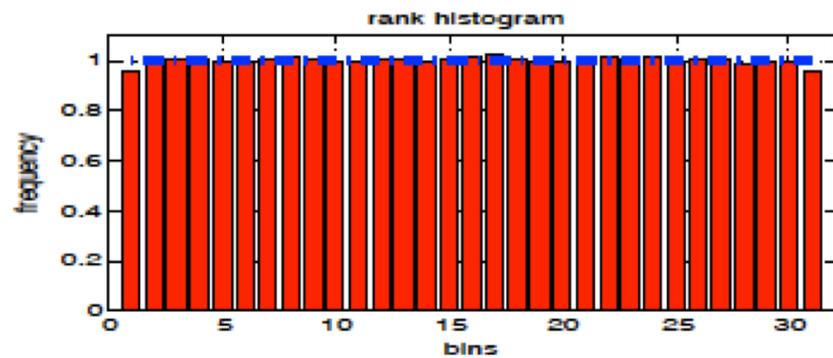
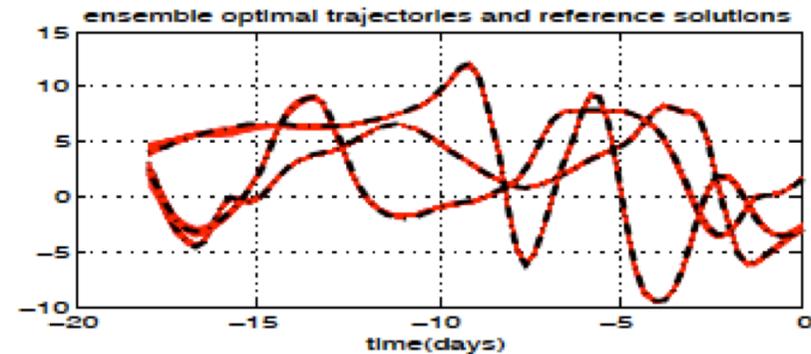
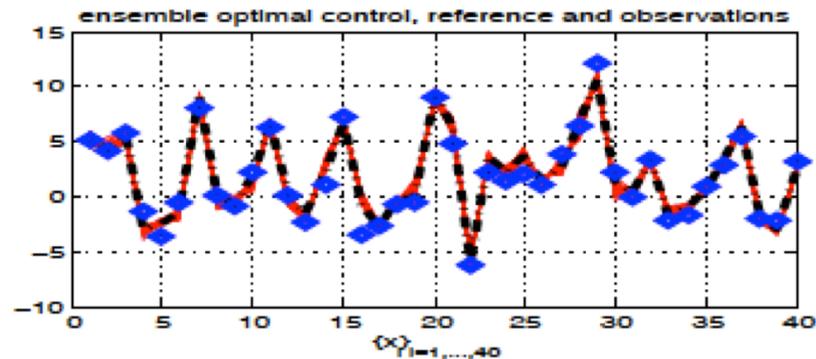
# Quasi-Static Variational Assimilation (QSVA)



# EnsVar : the non-linear Lorenz96 model 10 days with QSVA ( $\tau = 1$ day)



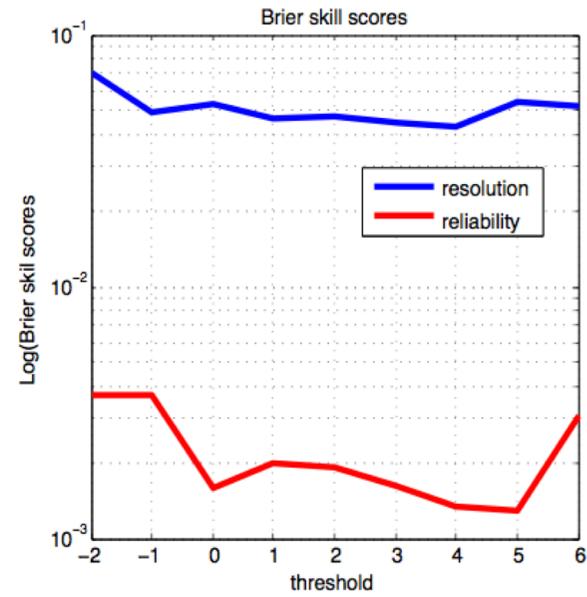
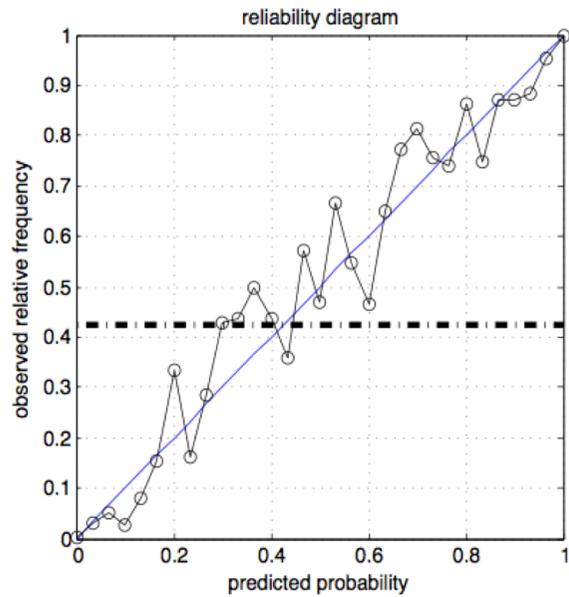
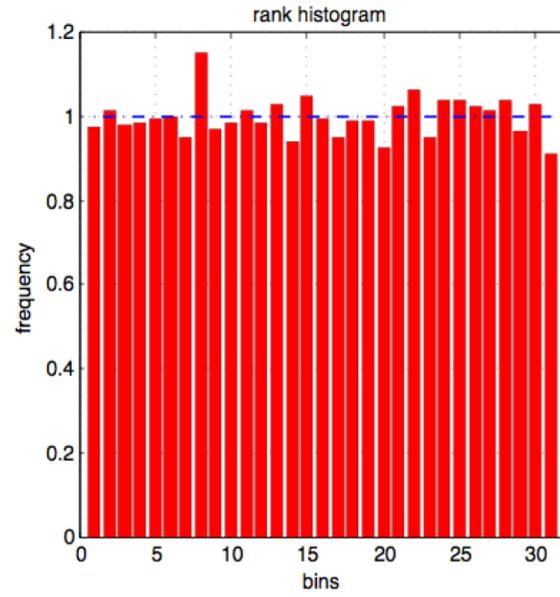
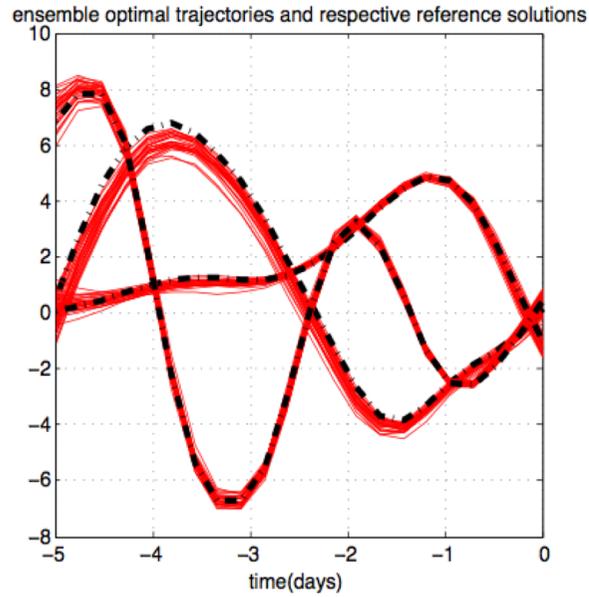
# EnsVar : the non-linear Lorenz96 model 18 days with QSVA ( $\tau = 1$ day)

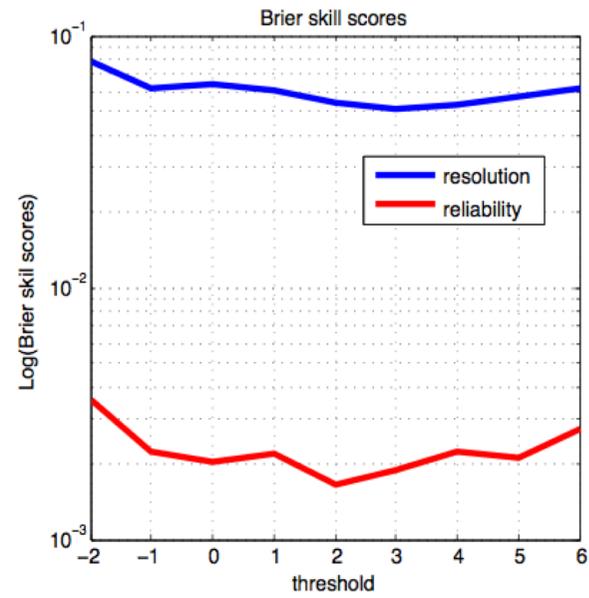
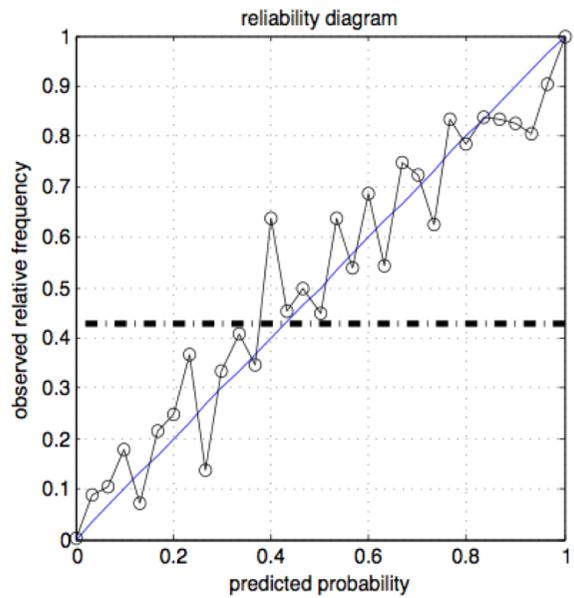
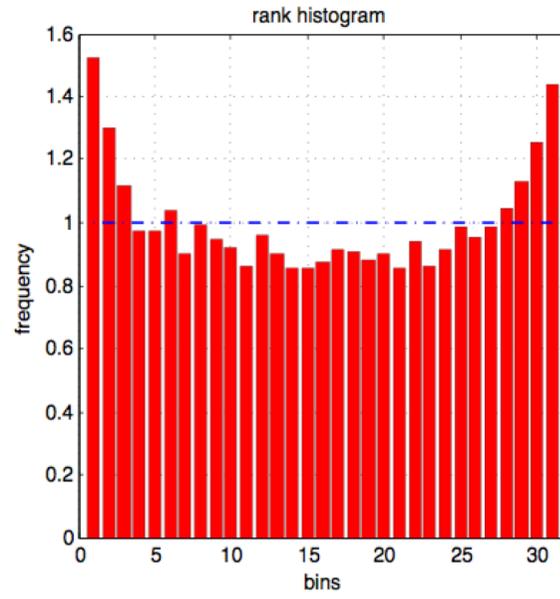
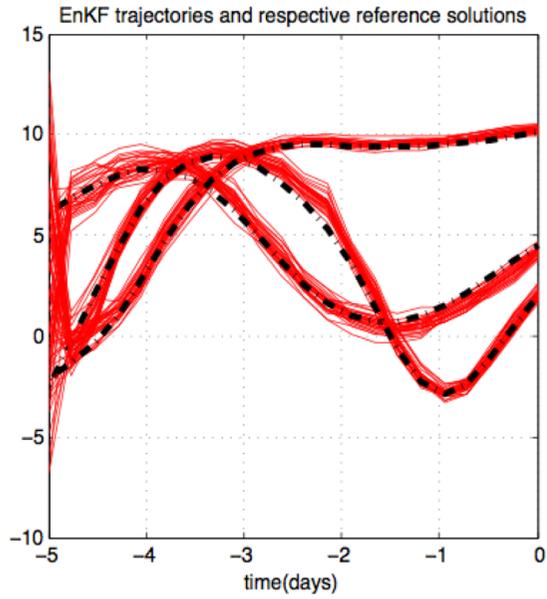


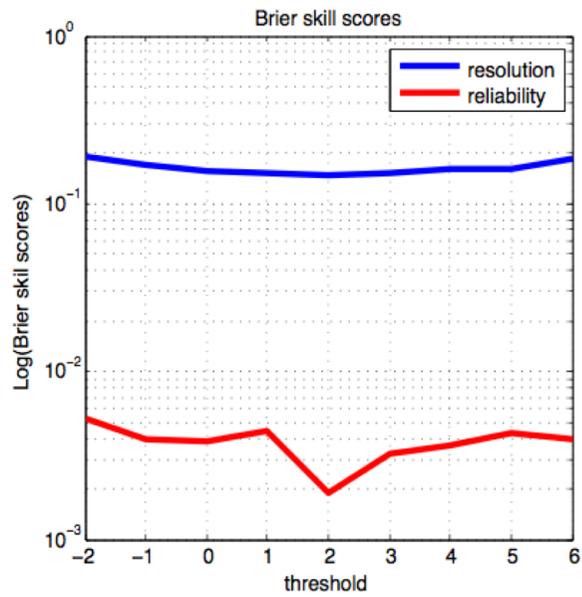
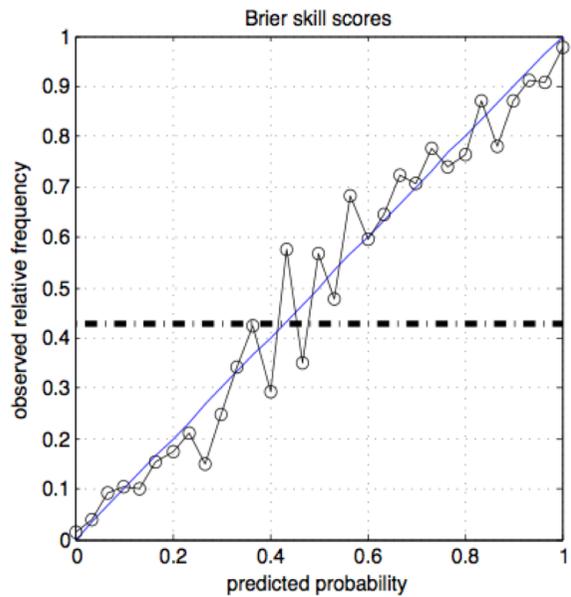
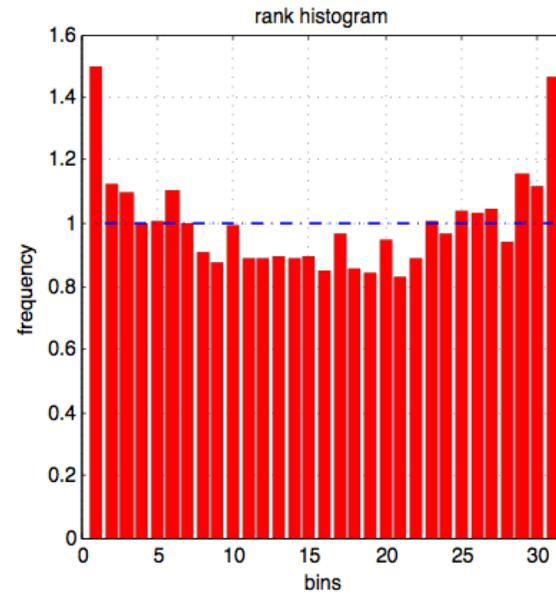
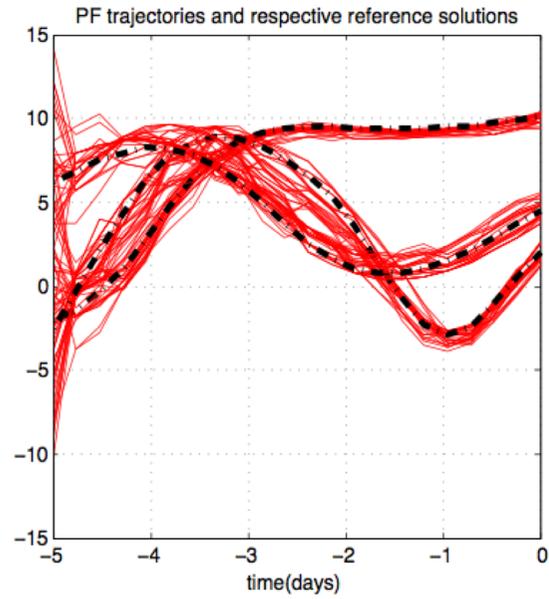
- Results are independent of the Gaussian character of the observation errors (trials have been made with various probability distributions)
- Ensembles produced by EnsVar are very close to Gaussian, even in strongly nonlinear cases.

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

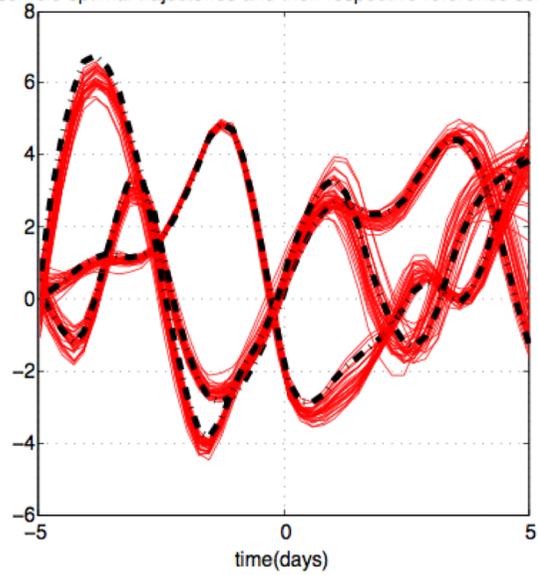
Both of these algorithms being sequential, comparison is fair only at end of assimilation window



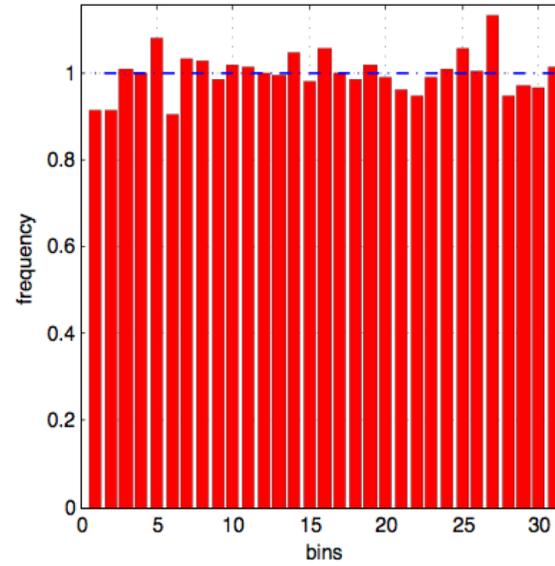




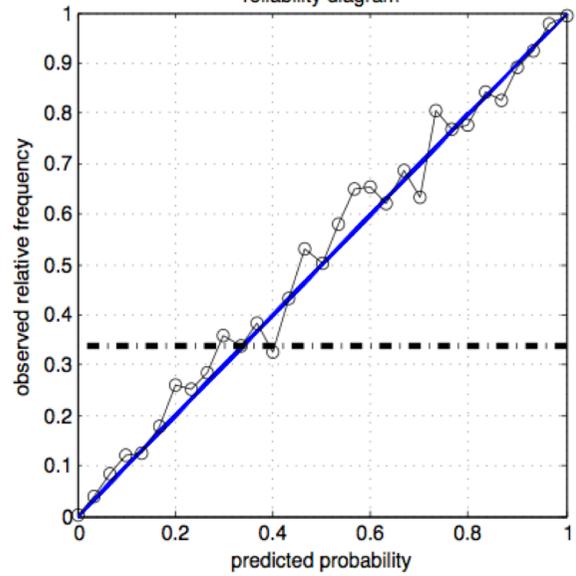
ensemble optimal trajectories and their respective reference solutions



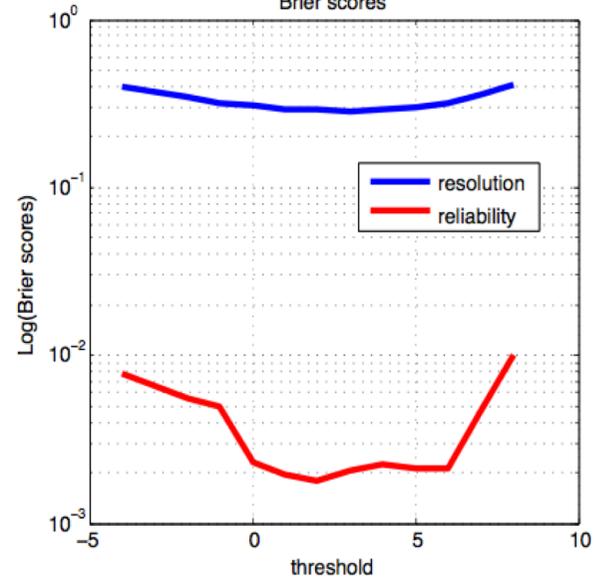
rank histogram



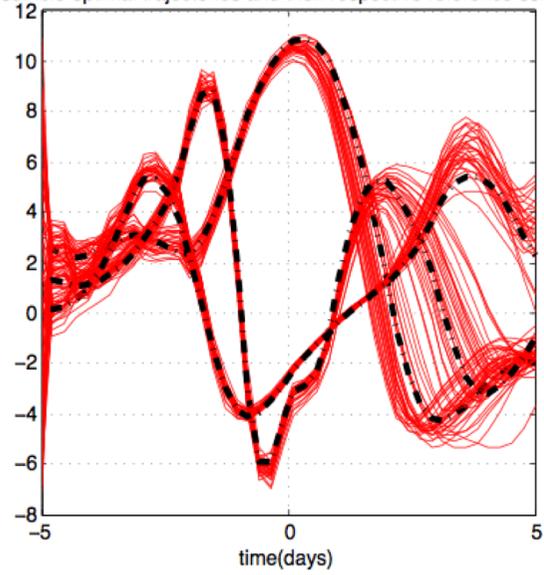
reliability diagram



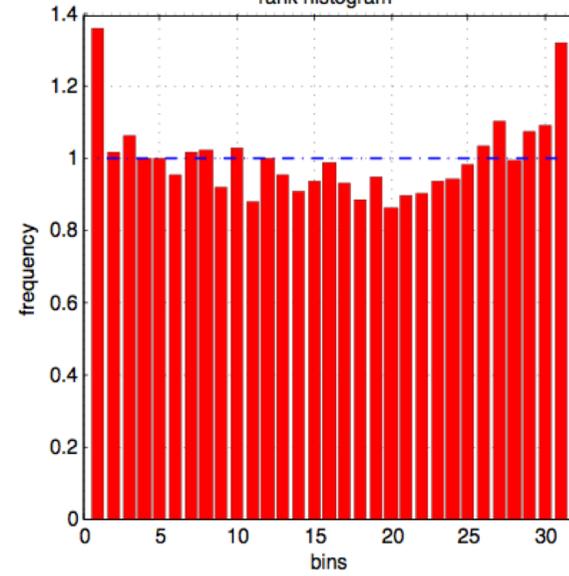
Brier scores



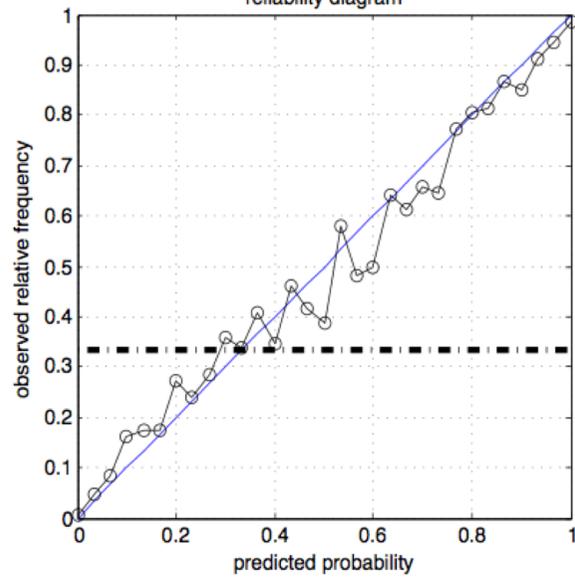
ensemble optimal trajectories and their respective reference solutions



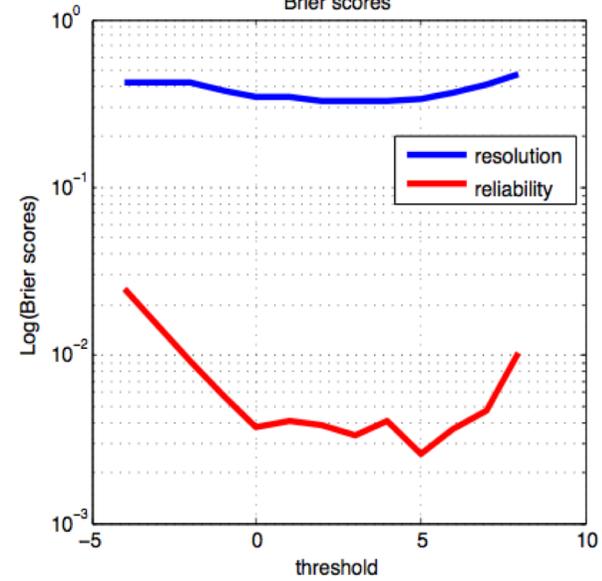
rank histogram



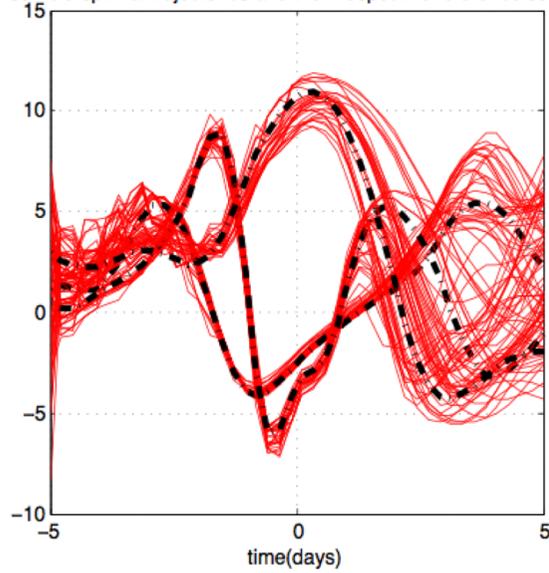
reliability diagram



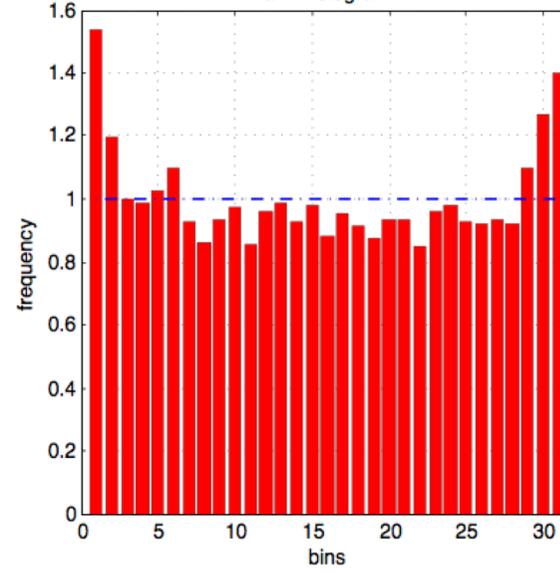
Brier scores



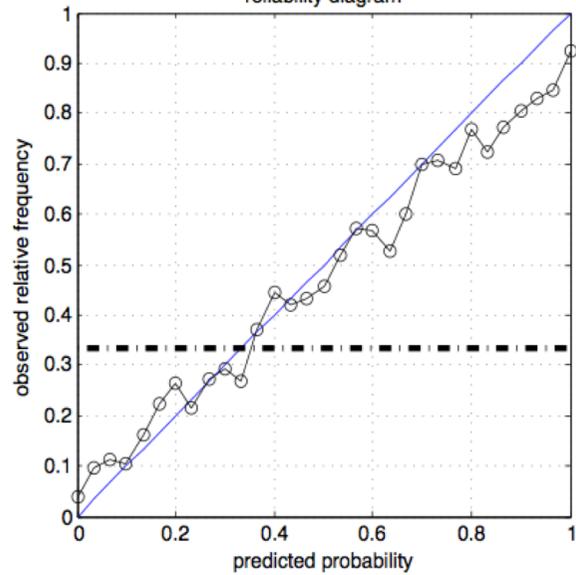
ensemble optimal trajectories and their respective reference solutions



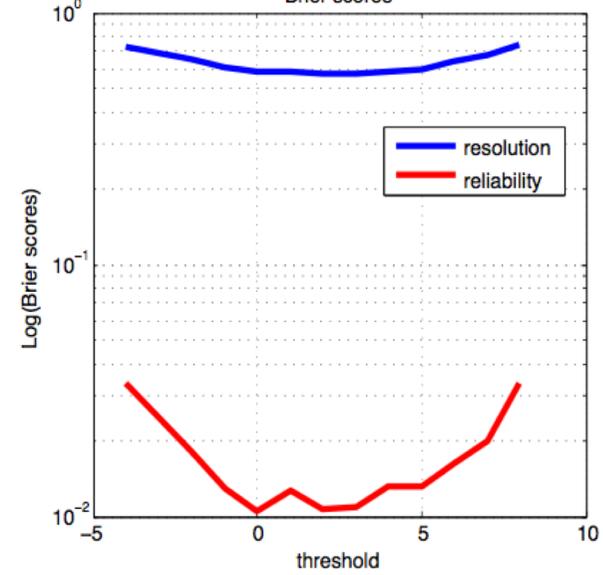
rank histogram



reliability diagram



Brier scores



<i>method</i> \ <i>DA procedure</i>	<i>Assimilation</i>	<i>Forecasting</i>
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

## *Weak constraint* variational assimilation

Allows for errors in the assimilating model

In the present case, data

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

$$y_k = H_k x_k + \varepsilon_k \quad E(\varepsilon_k \varepsilon_k^T) = R_k \delta_{kk}$$

(no background, but observation, at time 0)

- Nonlinear model (Lorenz' 96)

$$x_{k+1} = M_k(x_k) + \eta_k \quad E(\eta_k \eta_k^T) = Q_k \delta_{kk}, \quad k = 0, \dots, K-1$$

Errors assumed to be unbiased and uncorrelated in time

Control variables : initial state  $\xi_0$ , and  $K$  model errors  $\{\eta_k, k = 0, \dots, K-1\}$

Objective function of the form

$$(\xi_0, \eta_0, \dots, \eta_{K-1}) \rightarrow$$

$$\mathcal{J}(\xi_0, \eta_0, \dots, \eta_{K-1})$$

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

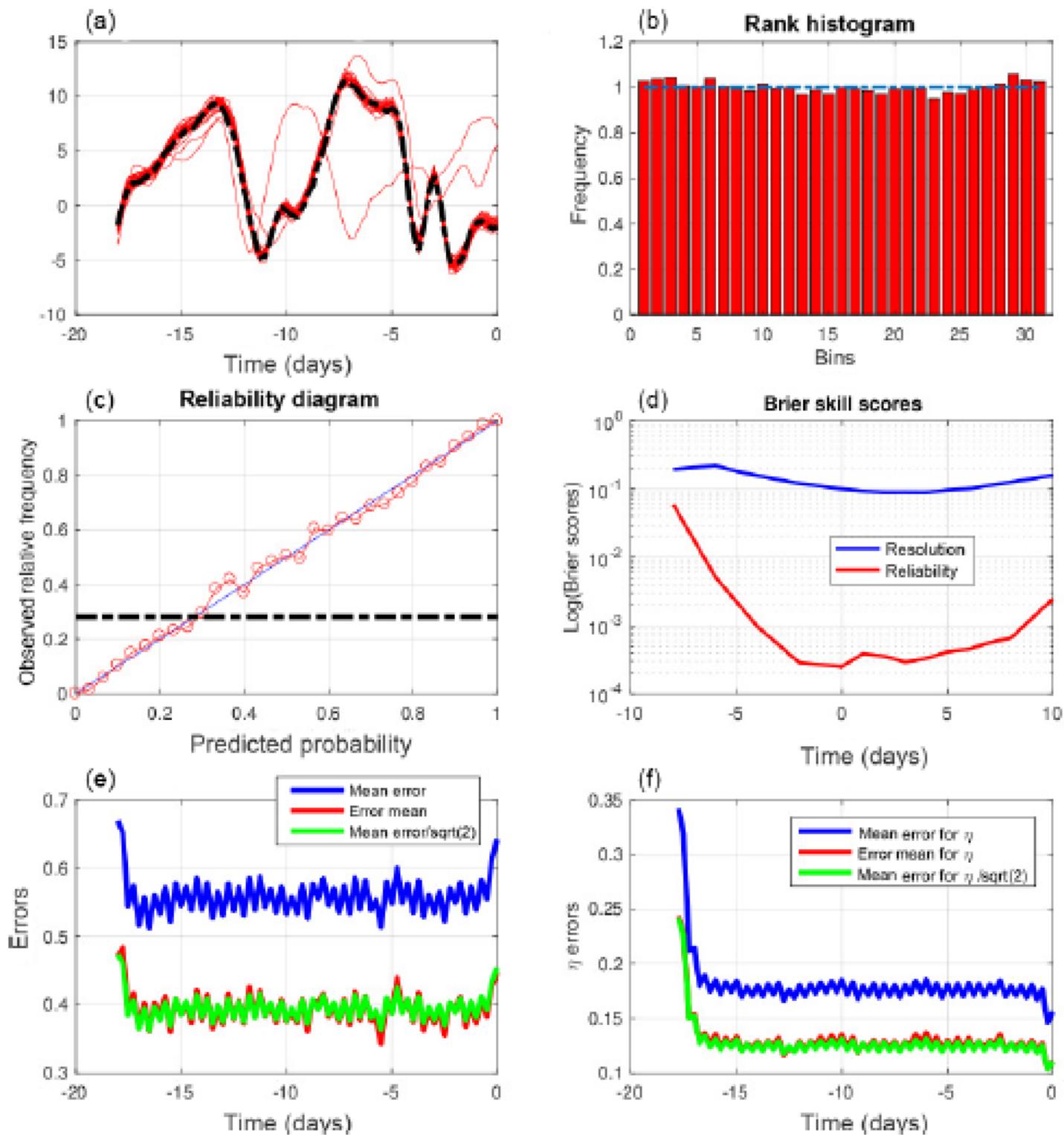
$$+ (1/2) \sum_{k=0, \dots, K-1} \eta_k^T Q_k^{-1} \eta_k$$

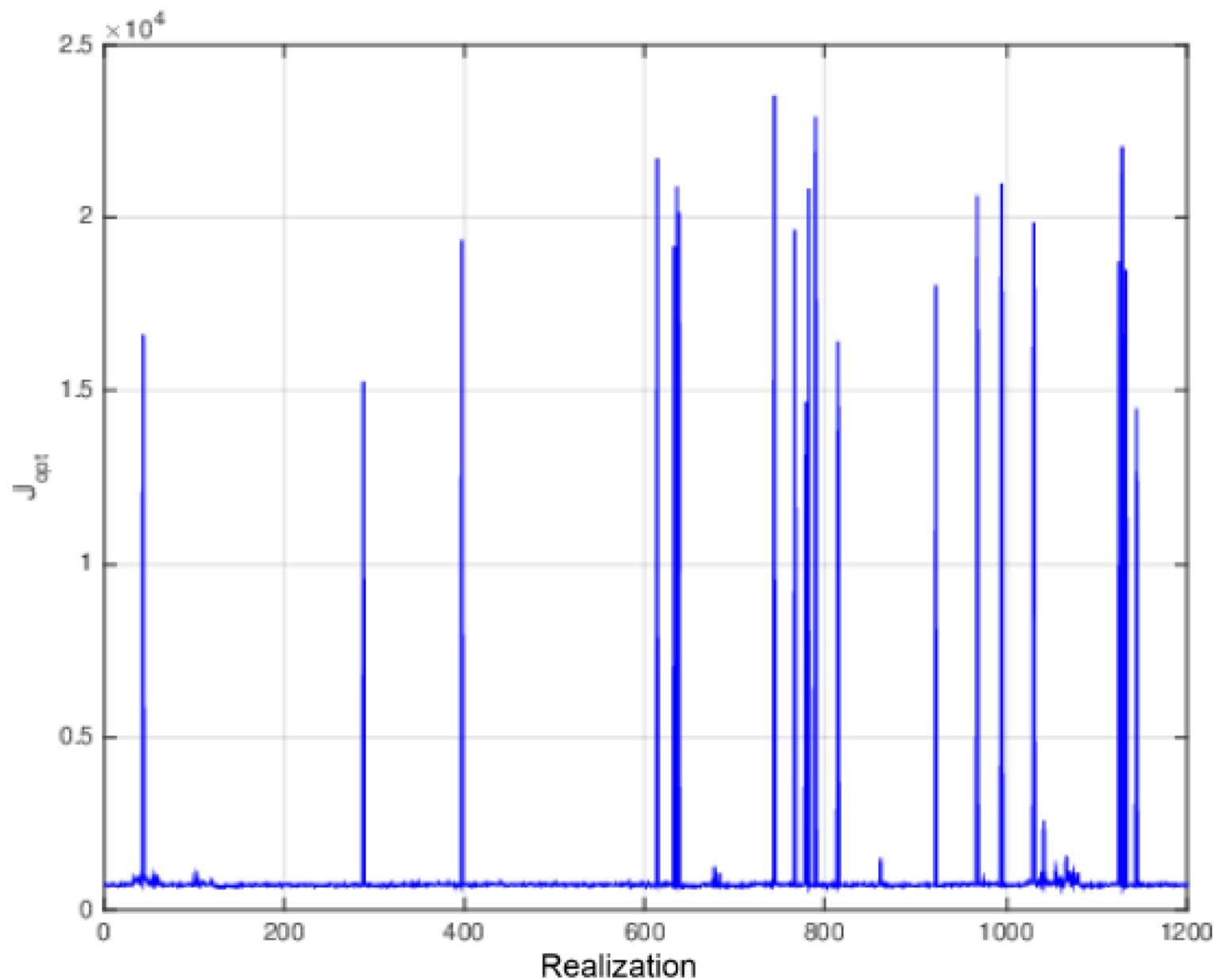
subject to

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

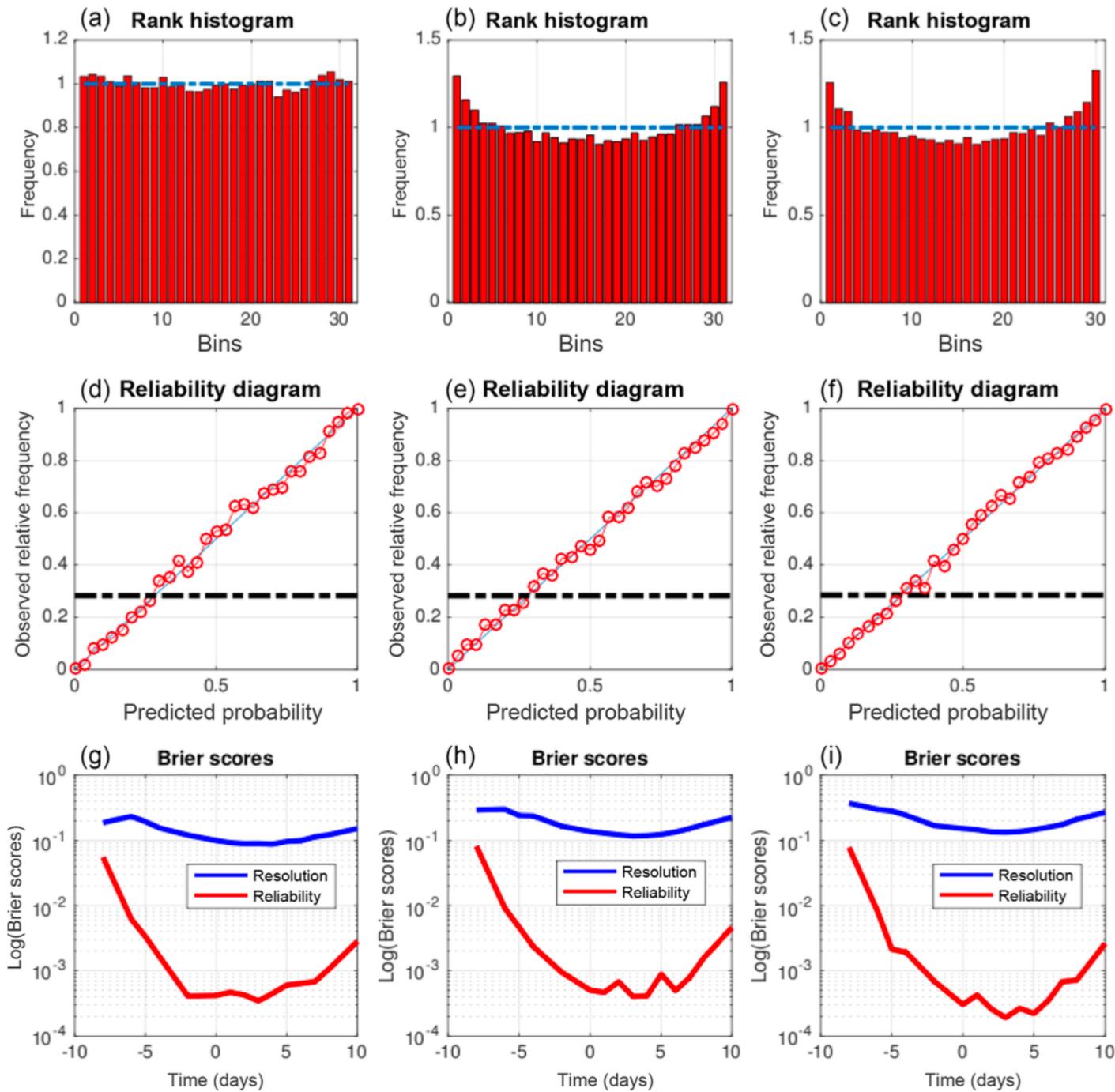
It turns out that QSVA is no more necessary. The model error term in the objective function has a regularizing effect which makes the function much smoother.

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





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

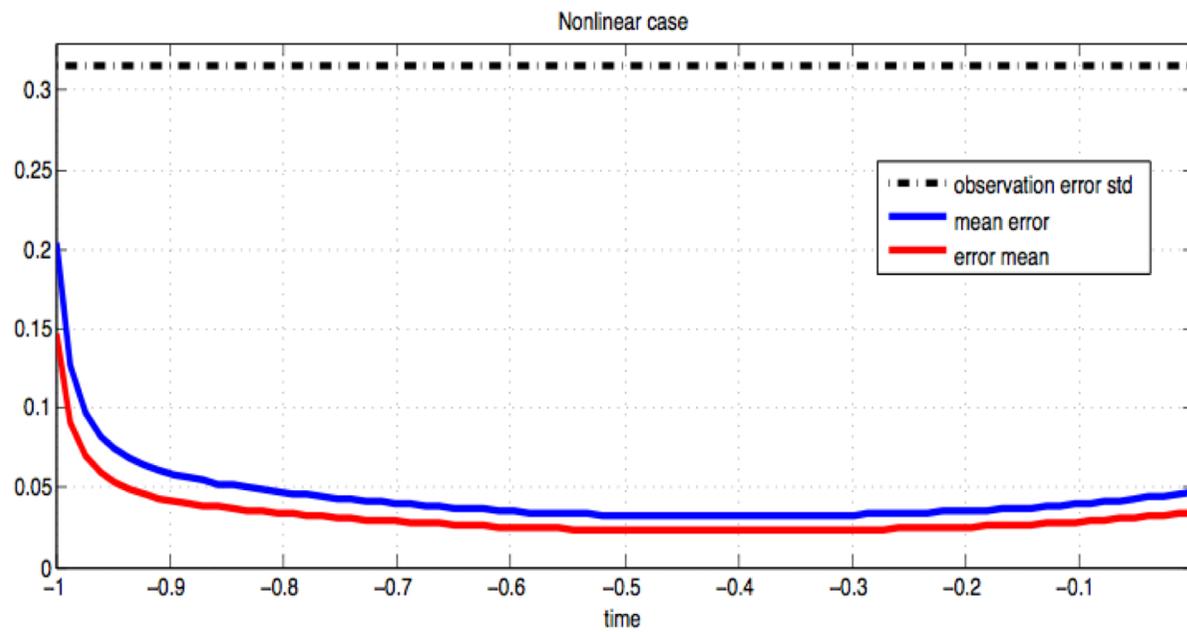
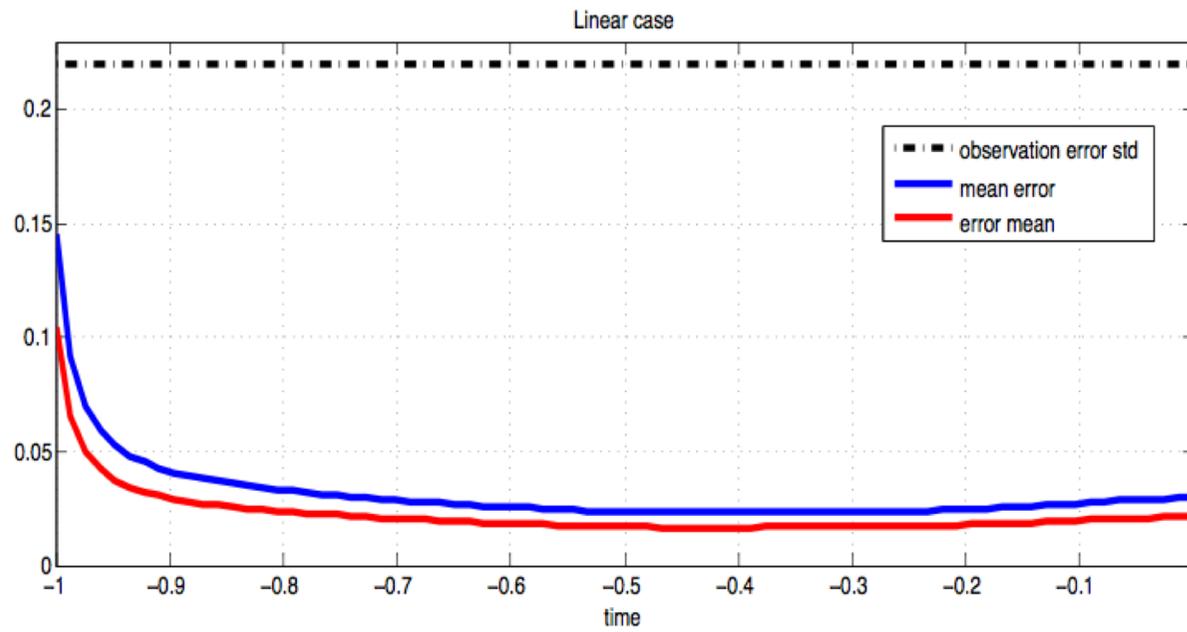


**Figure 12.** Compared performance of EnVAR, EnKF, and PF (columns from left to right respectively) over the last 13 days of 18-day assimilation windows. **(a, b, c)** Rank histogram. **(d, e, f)** Reliability diagram for the event  $\mathcal{E} = \{x > 1.02\}$ , which occurs with frequency 0.27. **(g, h, i)** Reliability and resolution components of the Brier score for events  $x < \tau$  as functions of the threshold  $\tau$ .

## 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  $x$ ,  $L = 32\pi$



## 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 Kuramoto-Sivashinsky model.

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

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

## 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**).

## *Summary and wrapping up*

- *Assimilation originated from the need of providing initial conditions for Numerical Weather Forecasts.*
- *Its purpose can be defined as estimation of the state of the atmospheric flow from all the available relevant information (i.e., observations and numerical model, or even differential equations, describing the dynamics of the flow).*
- *General bayesian approach. Determine the probability distribution for the state of the system, conditioned to everything we know (impossible to achieve in practice, but nevertheless very useful as providing consistent conceptual principle and guidelines as to which directions to take).*

## *Best Linear Unbiased Estimator (BLUE).*

*Data of the form*

$$\begin{array}{ll} \text{Background} & \mathbf{x}^b = \mathbf{x} + \boldsymbol{\zeta}^b \\ \text{Observations} & \mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\varepsilon} \end{array}$$

*Analysis*

$$\underbrace{\mathbf{x}^a}_{\text{analysis}} = \underbrace{\mathbf{x}^b}_{\text{background}} + \underbrace{\mathbf{P}^b \mathbf{H}^T [\mathbf{H}\mathbf{P}^b \mathbf{H}^T + \mathbf{R}]^{-1}}_{\text{gain matrix}} \underbrace{(\mathbf{y} - \mathbf{H}\mathbf{x}^b)}_{\text{innovation}}$$

$$\underbrace{\mathbf{P}^a}_{\text{error covariance matrices}} = \underbrace{\mathbf{P}^b} - \mathbf{P}^b \mathbf{H}^T [\mathbf{H}\mathbf{P}^b \mathbf{H}^T + \mathbf{R}]^{-1} \mathbf{H}\mathbf{P}^b$$

*BLUE achieves bayesian estimation in linear cases with (globally) gaussian additive errors)*

*Implemented in practice, more or less heuristically and empirically, in non-gaussian and (not too strongly) nonlinear cases in the form of*

*- Ensemble Kalman Filter (EnKF)*

*- Variational Assimilation (VA)*

*for both of which there exist innumerable variants*

*Particle Filters independent of either linearity or gaussianity hypotheses*

*In assimilation of meteorological and oceanographic data, powerful methods require taking into account the temporal evolution of the state of the flow over the assimilation window. Achieved by Kalman Filter and Variational Assimilation, as well as by Particle Filters.*

*In the course of assimilation, uncertainty on the state of the flow is concentrated in the modes that have been recently unstable (because of the dynamics or of the model error)*

*⇒ Assimilation in the Unstable Subspace (AUS), which is the subject of active research*

*The expected development of machine learning methods will certainly have a strong impact on assimilation, but which one ?*

Assimilation is related to many different theoretical and practical aspects (is multidisciplinary !)

- *Estimation theory*
- *Probability theory*
- *Theory of dynamical systems*
- *Atmospheric and oceanic dynamics*
- *Atmospheric and oceanic predictability*
- *Instrumental physics*
- *Optimisation theory*
- *Control theory*
- *Algorithmics and computer science*
- ...

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

- Oceanography
- Palaeoclimatology
- 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)
- Mathematical studies, independently of direct real life applications
- ...

It has now become a major tool of numerical environmental science, and a subject of mathematical study in its own right.

A few of the (many) remaining problems :

- Reduce amount of computations by restricting assimilation to appropriate subspace
- Observability (what to observe in order to know what we want to know ? Data are noisy, system is chaotic !)
- More accurate identification and quantification of errors affecting data particularly the assimilating model (will always require independent hypotheses)
- Assimilation of images
- ...

. HDFLook project (LOA-USTL) (MODIS October 2 2002 [18h10] ((Hurricane Hernan (Baja Cali



*La Fin du Cours ...*