Ensemble Variational Assimilation and Bayesian Estimation

Mohamed Jardak & Olivier Talagrand

DARC Meetings, 11^{th} Juin 2014, Reading, U.K.

Bayesianity

Under linearity and gaussianity, the following algorithm achieves Bayesian estimation

- Given the data

$$z = \overbrace{\Gamma}^{\text{linear operator}} \underbrace{x}_{\text{unknown}} + \zeta, \quad \zeta \in \mathcal{N}(\mu, \Sigma)$$

- The conditional posterior probability distribution is

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

with

$$x^a = \left(\Gamma^T \Sigma^{-1} \Gamma\right)^{-1} \Gamma^T \Sigma^{-1} (z - \mu) \text{ and } P^a = \left(\Gamma^T \Sigma^{-1} \Gamma\right)^{-1}$$

Ready recipe for producing sample of independent realizations of posterior probability distribution :

Perturb data vector additively according to error probability distribution $\mathcal{N}(0, \Sigma)$, and compute analysis x^a for each perturbed data vector.

Variational Assimilation : Linear and Gaussian case

The following algorithm produces a sample of independent realizations of the probability distribution of the state of the system, conditioned by the data x_0^b and y_k.

- Available data
  1. Background estimate at $t = 0,\ x_0^b = x_0 + \xi_0^b,\ \xi_0^b \in \mathcal{N}(0, P_0^b)$
  2. Observations at $t = t_k,\ y_k = \mathrm{H}_k x_k + \epsilon_k, \epsilon_k \in \mathcal{N}(0, \mathrm{R}_k)$
  3. Model (supposed to be exact) $x_{k+1} = \mathrm{M}_k x_k,$ and $k = 0, \cdots, K-1$
  4. Errors ξ_0^b and ϵ_k assumed to be unbiased and uncorrelated in time.
  5. H_k and M_k assumed linear
- The optimal state (mean of the Bayesian Gaussian pdf) at $t = 0$ minimizes the objective function

$$\left\{ \begin{array}{l} \mathcal{J}(\xi_0) = \dfrac{1}{2}(x_0^b - \xi_0)^T \left[P_0^b\right]^{-1} (x_0^b - \xi_0) + \dfrac{1}{2}\sum_k (y_k - \mathrm{H}_k\xi_k)^T \left[\mathrm{R}_k\right]^{-1} (y_k - \mathrm{H}_k\xi_k) \\ \xi_{k+1} = \mathrm{M}_k(\xi_k) = \mathrm{M}_k(\mathrm{M}_{k-1}(\xi_{k-1})) = \mathrm{M}_k(\mathrm{M}_{k-1}\cdots(\mathrm{M}_1(\mathrm{M}_0(\xi_0)))) \end{array} \right.$$

What happens under nonlinearity and non-Gaussianity ?

Objectives

- Objectively evaluate the EnsVAR as an ensemble estimator in the non-linear and non-Gaussian cases.
- Evaluate as far as possible the Bayesianity of the ensemble produced in the non-linear and non-Gaussian cases .
- Compare with other existent ensemble algorithm schemes (EnKF and PF) .

Objective evaluation of Bayesianity

How to objectively evaluate the Bayesian character of an ensemble estimation procedure ?

There is no general objective criterion for Bayesianity

weaker property of reliability

Bayesianity implies reliability therefore lack of reliability implies lack of Bayesianity.

reliability is the statistical consistency between the predicted probability of occurrence and the observed frequency of occurrence.

The Lorenz96 model

- Forward model

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

- Tangent linear model

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

- 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. $\frac{1}{\lambda_{max}} \simeq 2.5 days$, λ_{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 : 6000 to 9000 realizations.

Experimental procedure

- define a reference solution x_t^r by integrating the numerical nodel
- repeat the following steps over N_{real} successive assimilation windows
  1. produce observations at successive times $t_k, y_k = H_k x_k^r + \epsilon_k$
     ϵ_k is a random observation error not necessarily Gaussian, and H_k the observation operator not necessarily linear.
  2. for a given observation y_k, repeat N_{ens} times the following process
     1. perturb the observations $y_k, z_k = y_k + \delta_k$
        δ_k : is an independant realization of the probability distribution which has produced ϵ_k
     2. assimilate the perturbed observations z_k by variation assimilation.

This produces N_{ens} model solutions over the assimilation window, considered as a sample of the conditional probability distribution for the state of the observed system over the assimilation window.

- Validation & Verification of the set of N_{real} ensembe assimilations.

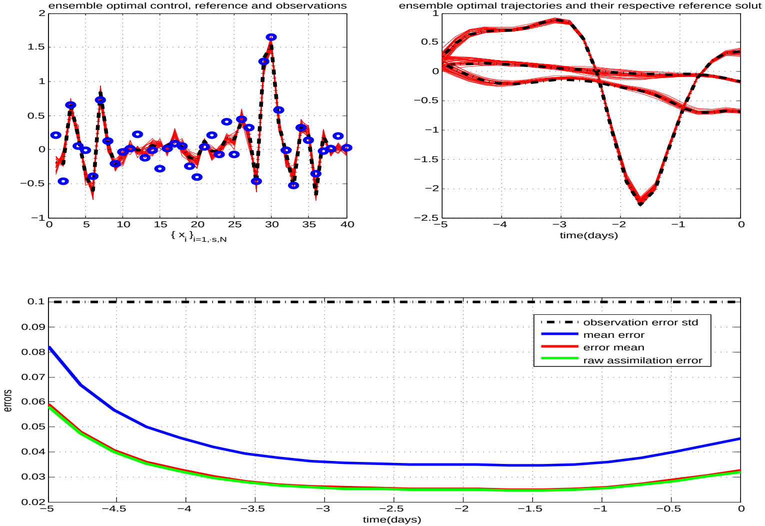
Validation &Verification diagnostic tools

- rank histogram.
- reliability diagram.
- Brier scores

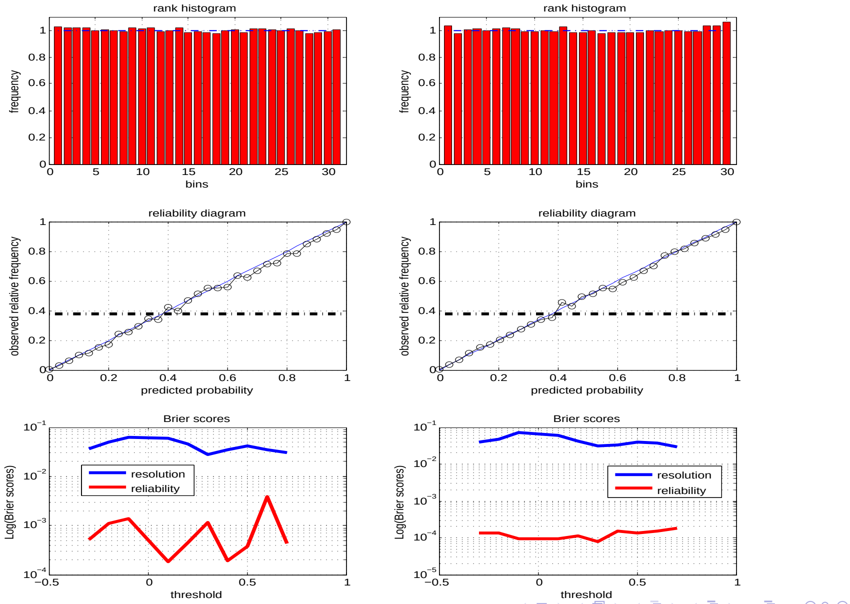
$$\mathbb{B} = \underbrace{\frac{1}{p_c(1-p_c)}}_{\text{uncertainty}} \left[\underbrace{\int_0^1 (p'-p)^2 g(p)dp}_{\text{reliability}} + \underbrace{\int_0^1 p'(1-p')g(p)dp}_{\text{resolution}} \right]$$

  - p predicted probability.
  - g the frequency with which p has been predicted.
  - $p'(p)$ observed frequency.
  - p_c the frequency of occurrence of the event $\mathscr{E}$ under observation.
- under linearity the expectation of the objective function at its minimum is half the number of observations p, $\mathbb{E}(\mathfrak{J}(x_{opt})) = \frac{p}{2}$.
- under Gaussianity we have : $\mathbb{V}ar(\mathfrak{J}(x_{opt})) = p$.

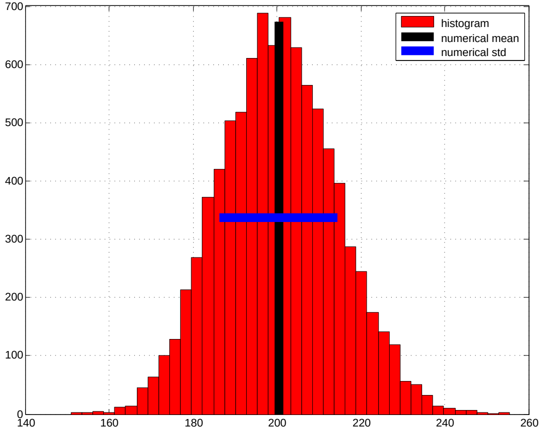
EnsVAR : the Lorenz96 model linear case (5 days)



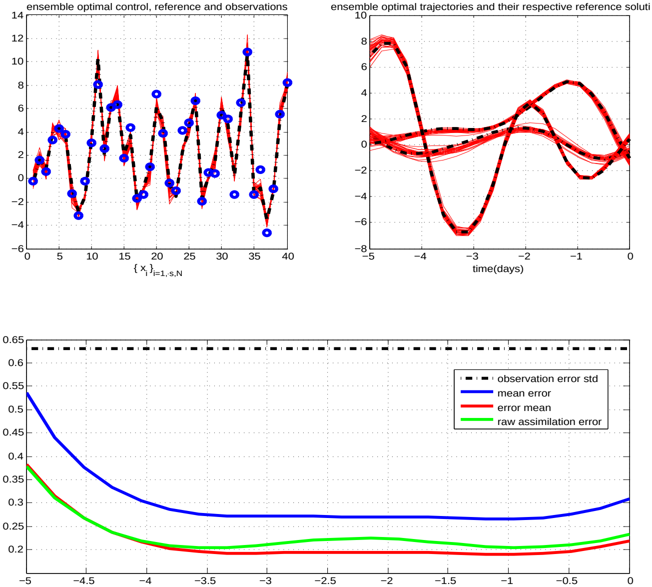
EnsVAR : the Lorenz96 model linear case (5 days)



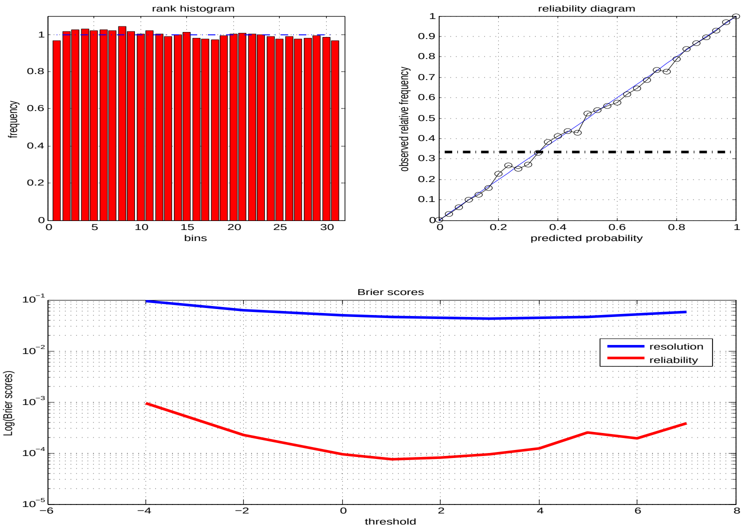
EnsVAR : the Lorenz96 model linear case (5 days)



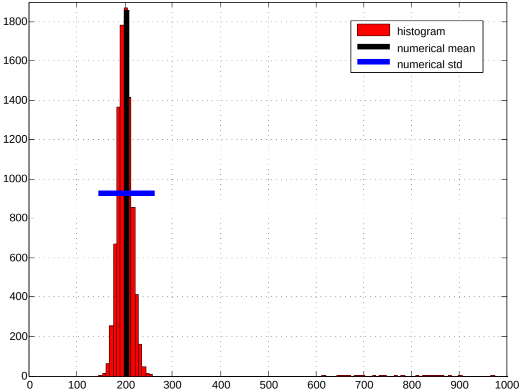
EnsVAR : the Lorenz96 model nonlinear case (5 days)



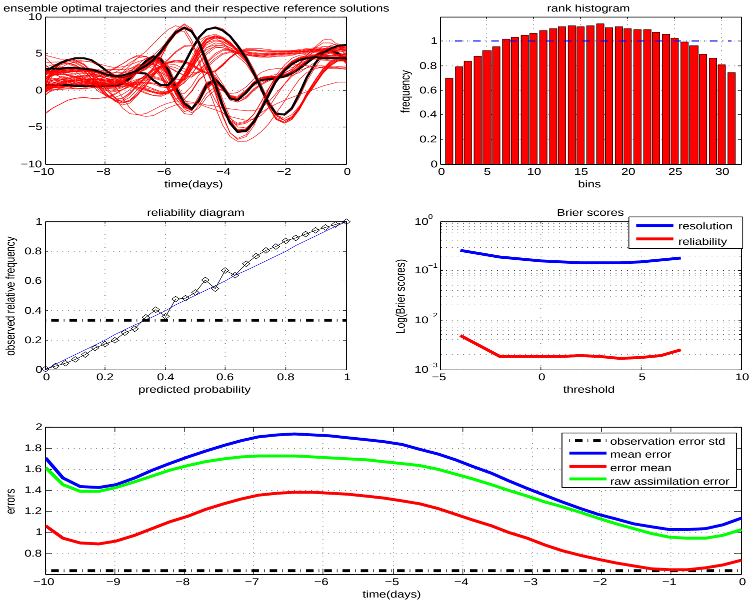
EnsVAR : the Lorenz96 model nonlinear case (5 days)



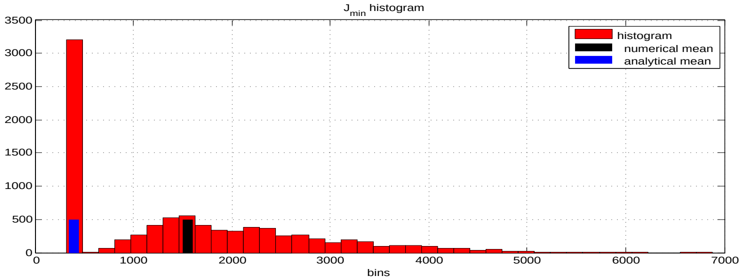
EnsVAR : the Lorenz96 model nonlinear case (5 days)

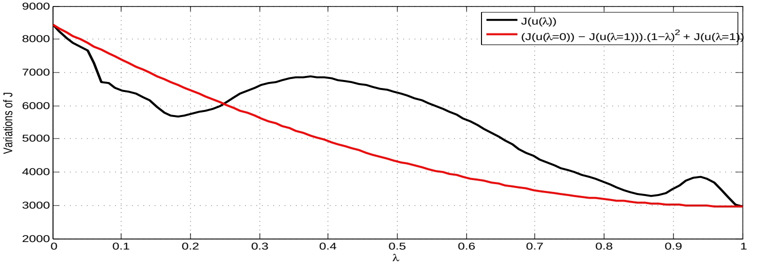


EnsVAR : the Lorenz96 model nonlinear case (10 days)

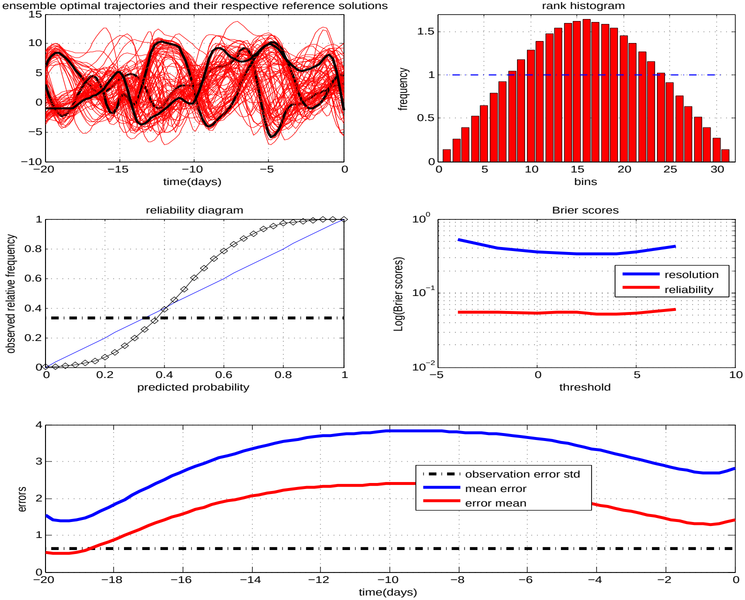


EnsVAR : the Lorenz96 model nonlinear case (10 days)

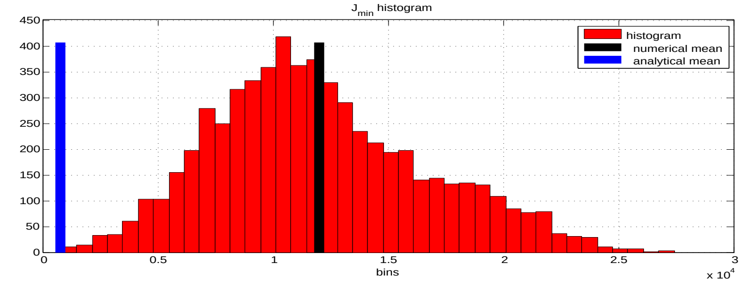


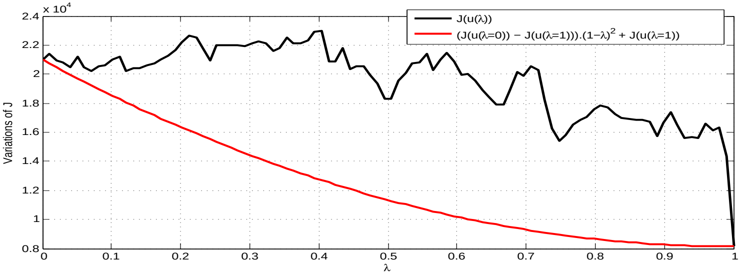


EnsVAR : the Lorenz96 model nonlinear case (20 days)

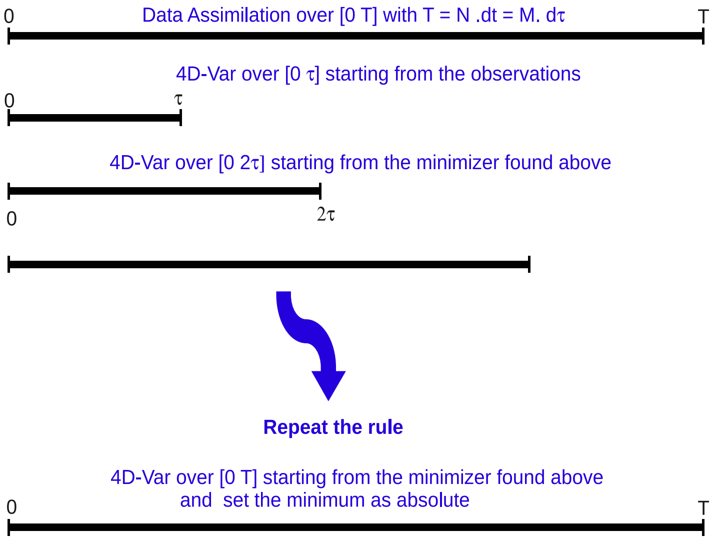


EnsVAR : the Lorenz96 model nonlinear case (20 days)

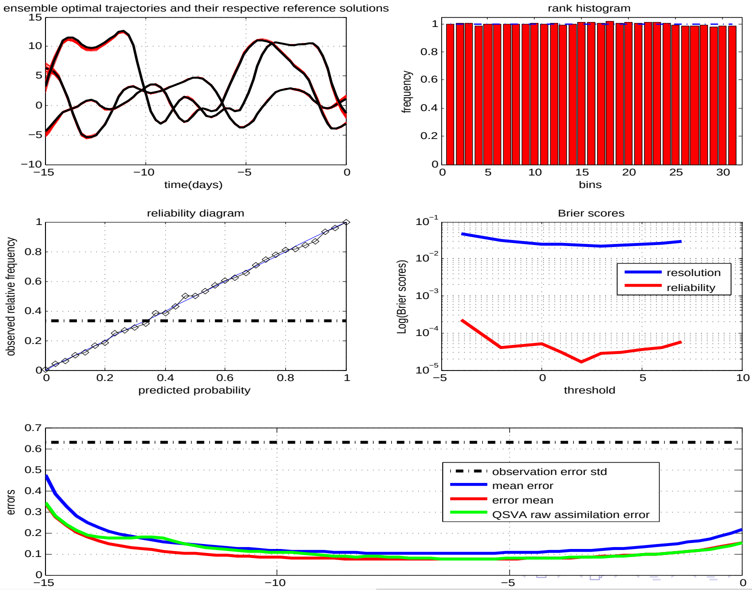




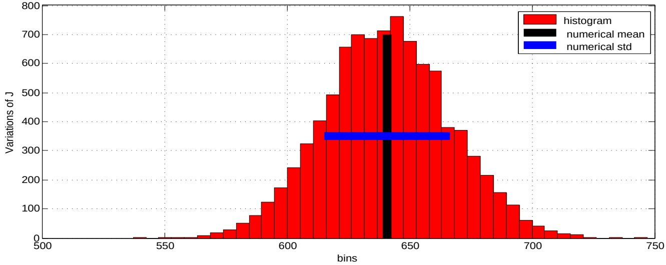
Quasi-Static Variational Assimilation (QSVA)

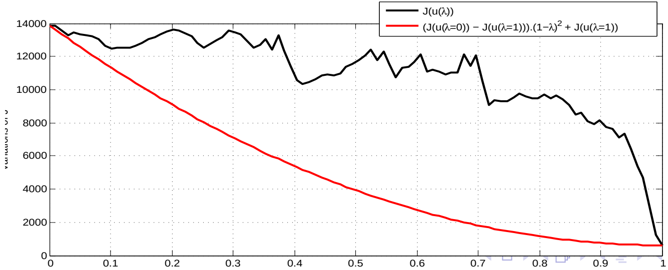


EnsVAR : the Lorenz96 model nonlinear case (15 days) with QSVA

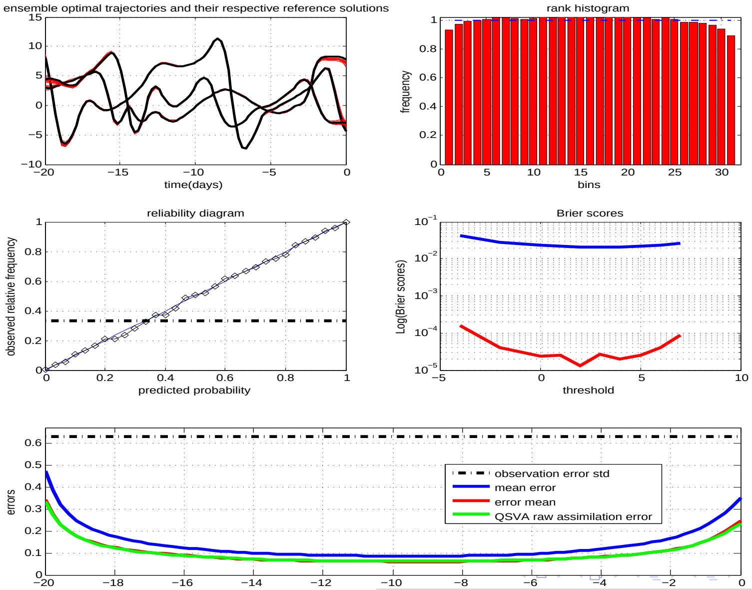


EnsVAR : the Lorenz96 model nonlinear case (15 days with QSVA)

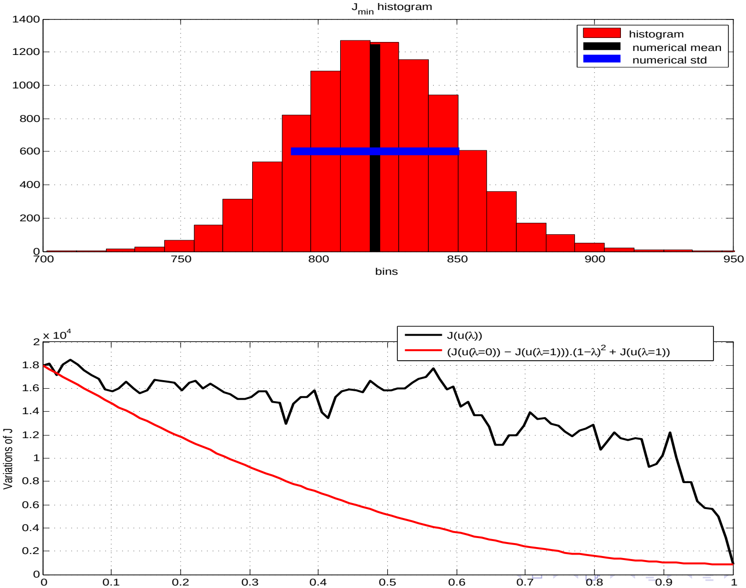




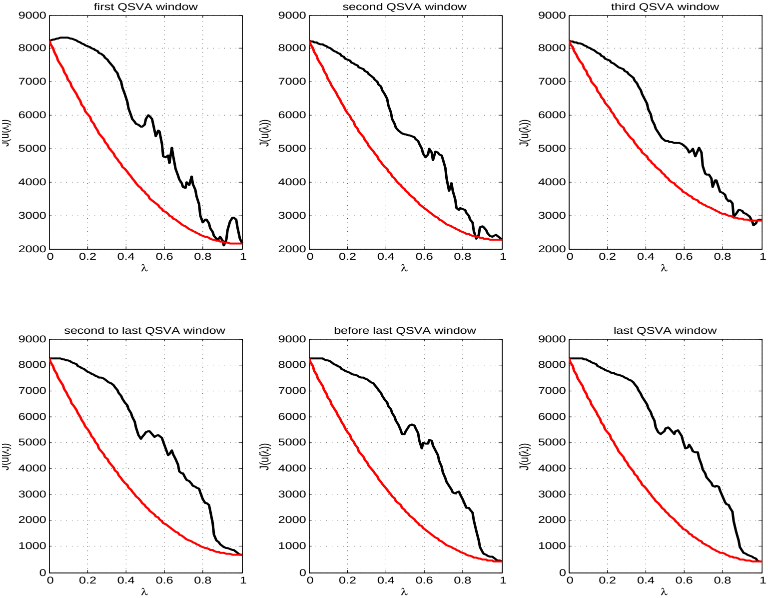
EnsVAR : the Lorenz96 model nonlinear case (20 days with QSVA)



EnsVAR : the Lorenz96 model nonlinear case (20 days with QSVA)



EnsVAR : variations of $\mathcal{J}$ during a QSVA procedure



EnsVAR : measuring the Gaussianity

For a random variable y

- **Skewness and Kurtosis :**

$$\mathbb{S}kew(y) = \mathbb{E}\left[\left(\frac{y-\mu}{\sigma}\right)^3\right], \ \mathbb{K}urt(y) = \mathbb{E}(y^4) - 3\mathbb{E}(y^2)$$

- **Negentropy :**
  - Entropy : the degree of information that the observation of the variable gives.

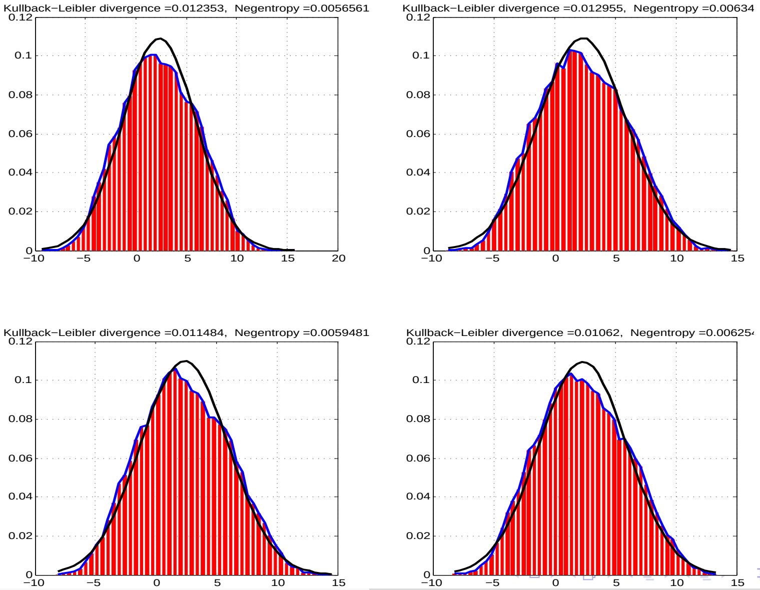
$$H(y) = -\int \underbrace{f(y)}_{\text{density function of y}} \log f(y) dy$$

  - Negentropy

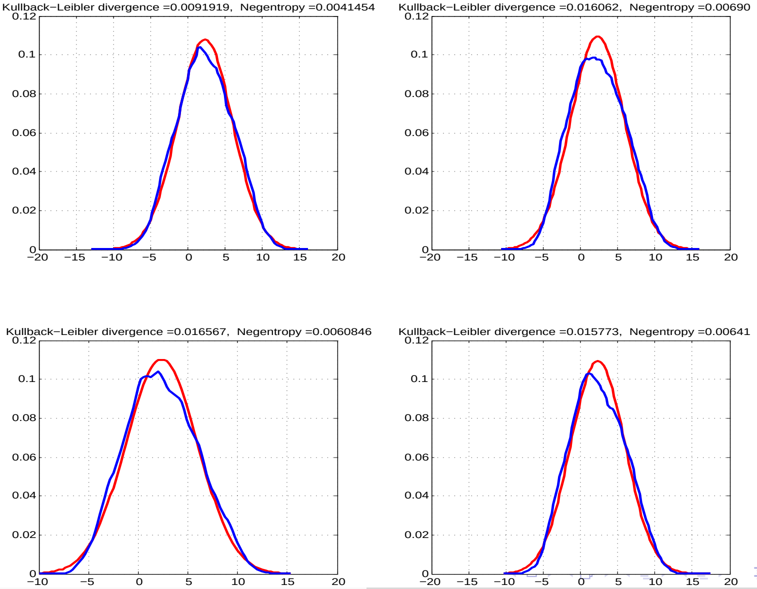
$$J(y) = H(y_{\text{gaussian}}) - H(y) \approx \frac{1}{12}\left[\mathbb{E}(y^3)\right]^2 + \frac{1}{48}\left[\mathbb{K}urt(y)\right],$$

y_{gaussian} is a Gaussian variable of the same covaraince matrix as y.

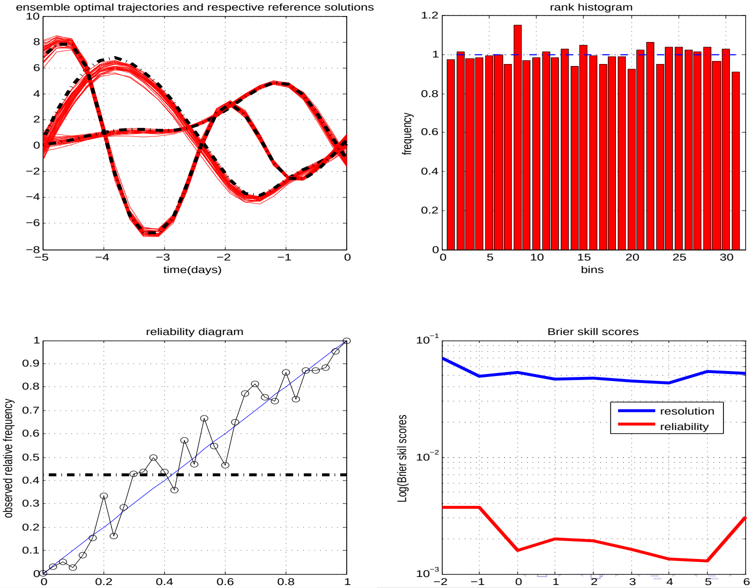
EnsVAR : the Lorenz96 model nonlinear case 5 days : Gaussianity



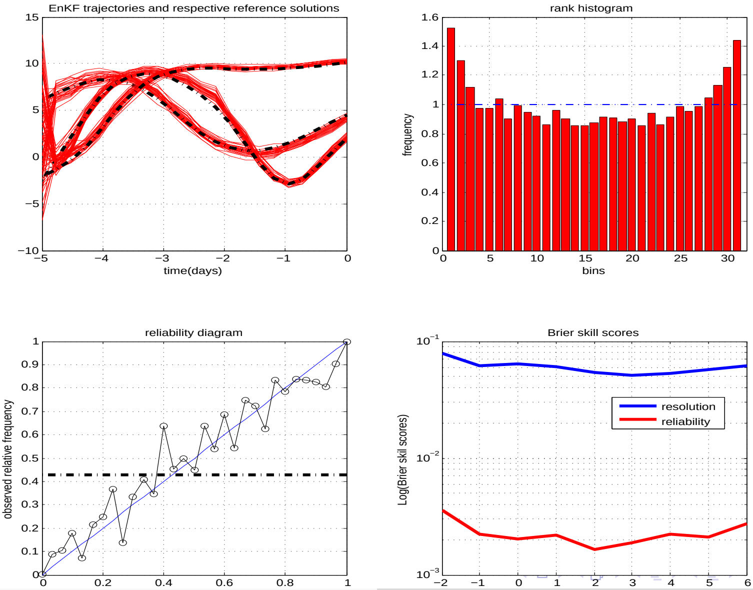
EnsVAR : the Lorenz96 model nonlinear case 10 days : Gaussianity



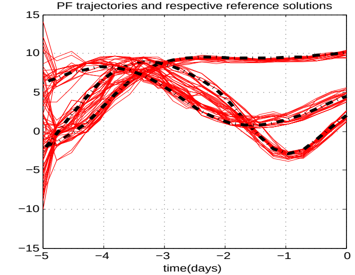
EnsVAR : the Lorenz96 model nonlinear case 5days, end of the assimiltaion window

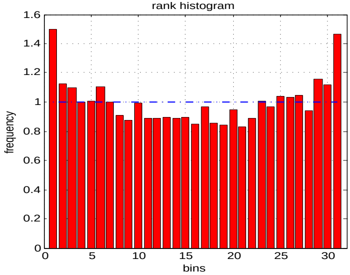


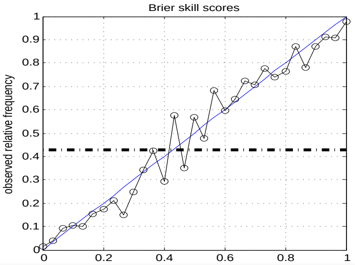
EnKF : the Lorenz96 model nonlinear case 5days, end of the assimiltaion window



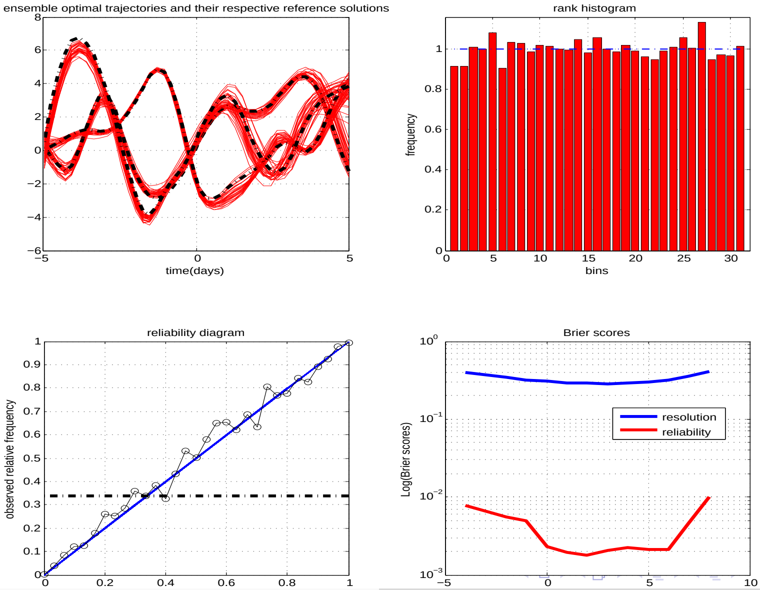
PF : the Lorenz96 model nonlinear case 5days, end of the assimiltaion window



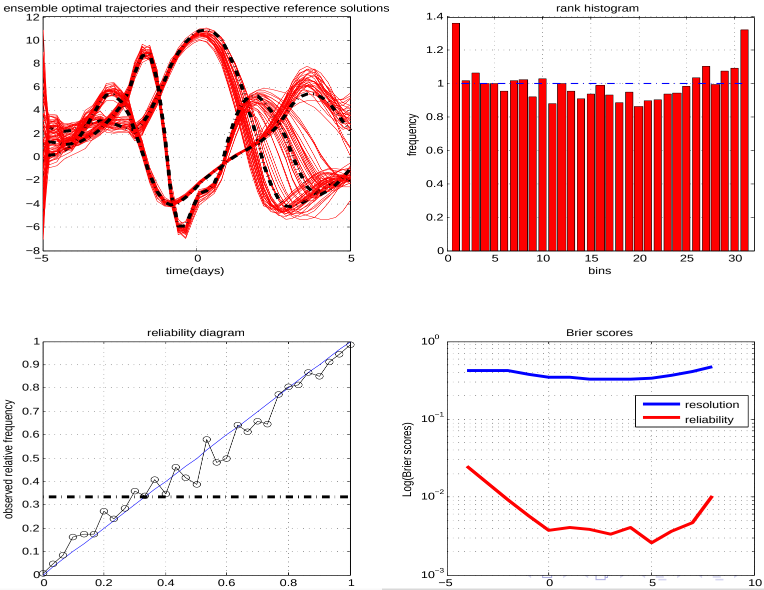




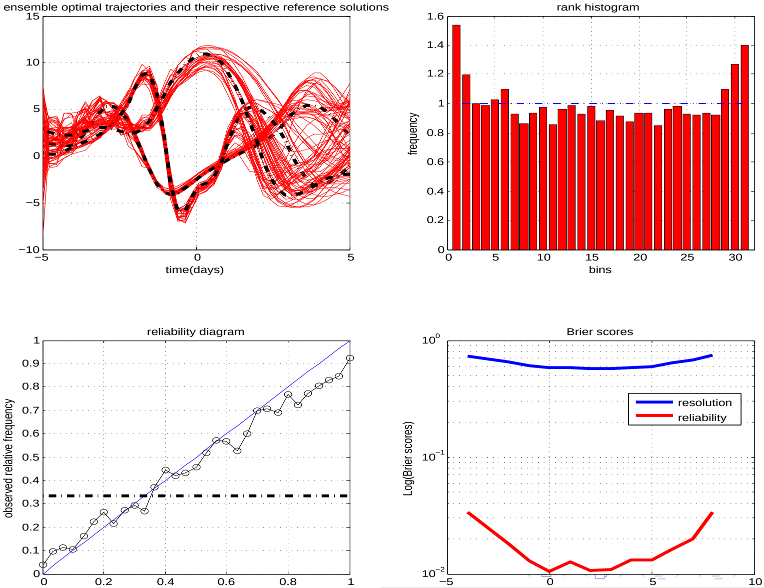
EnsVAR : the Lorenz96 model nonlinear case 5days, end of the forecast



EnKF : the Lorenz96 model nonlinear case 5days, end of the forecast



PF : the Lorenz96 model nonlinear case 5days, end of the forecast



Assimilation and Forecasting RMSE at the end of 5 days

DA procedure / *method*	*Assimilation*	Forecasting
EnsVAR	0.2193510	1.49403506
EnKF	0.2449690	1.67176110
PF	0.7579790	2.62461295

Weak constraint EnsVAR

- define the objective function.

$$\mathfrak{J}(x,\eta_1,\eta_2,\cdots.\eta_{N-1},\eta_N) = \frac{1}{2}\left\{(x-x_b)^T\mathrm{B}^{-1}(x-x_b)\right\} \quad +$$

$$\frac{1}{2}\sum_{i=0}^{N}\left\{(y_i-H_i(x_i))^T\mathrm{R}_{\mathbf{i}}^{-\mathbf{1}}(y_i-H_i(x_i))\right\} + \frac{1}{2}\sum_{i=1}^{N}\eta_i^T\mathrm{Q}_i^{-1}\eta_i$$

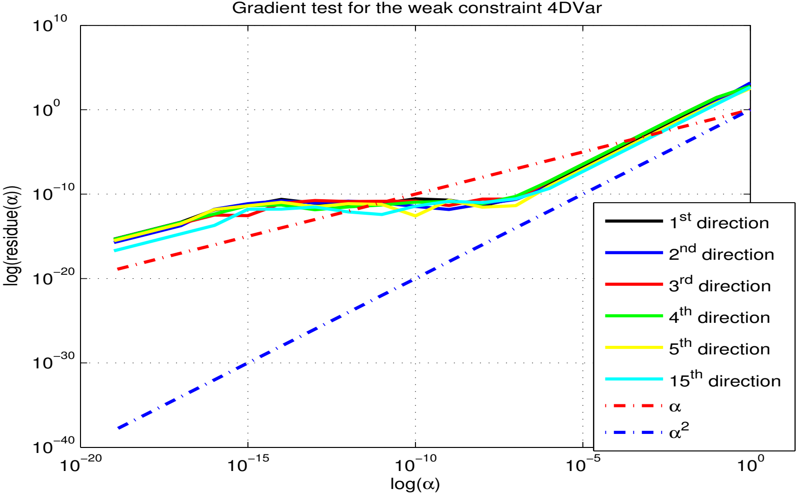
1. B background error covariance matrix and R observation error covariance matrix.
2. Q model error covariance matrix.
3. $H:\mathbb{R}^{state}\longrightarrow\mathbb{R}^{obs}$ observation operator.
4. x_b background state vector and y_i observation vector at time $t=t_i$.
5. η_i model error vector at $t=t_i$ with $x(t_i)=\mathfrak{M}_{t_i\leftarrow t_{i-1}}(x(t_{i-1}))+\eta_i$

- find the optimal control variable $(x_0^{opt},\eta_1^{opt},\eta_2^{opt},\cdots.\eta_N^{opt})$ and the optimal trajectory x^{opt}.

$$(x_0^{opt},\eta_1^{opt},\eta_2^{opt},\cdots.\eta_N^{opt}) = \min_{x,\eta_1,\eta_2,\cdots.\eta_N\in\mathfrak{A}}\mathfrak{J}(x,\eta_1,\eta_2,\cdots.\eta_N)$$

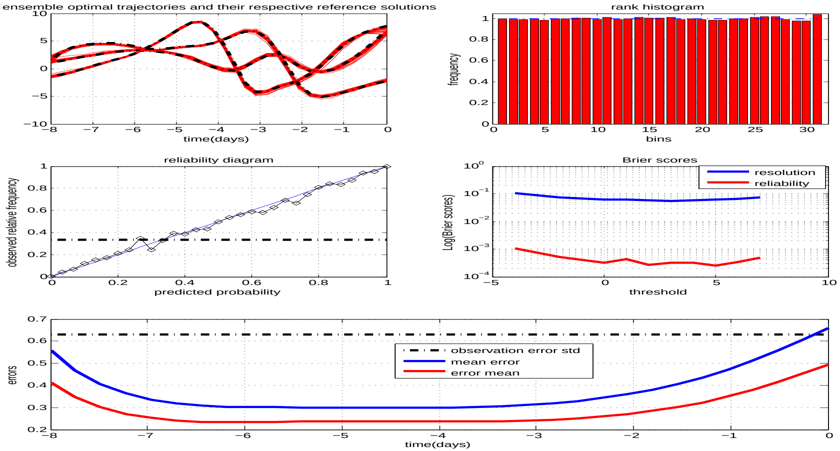
$$x_i^{opt} = \mathfrak{M}_{t_i\leftarrow t_{i-1}}(\mathfrak{M}_{t_{i-1}\leftarrow t_{i-2}}\cdot\cdot(\mathfrak{M}_{t_2\leftarrow t_1}(\mathfrak{M}_{t_1\leftarrow t_0}(x_0^{opt})+\eta_1^{opt})+\eta_2^{opt})\cdot\cdot+\eta_{i-1}^{opt})+\eta_i^{opt}$$

Weak EnsVAR : testing the gradient

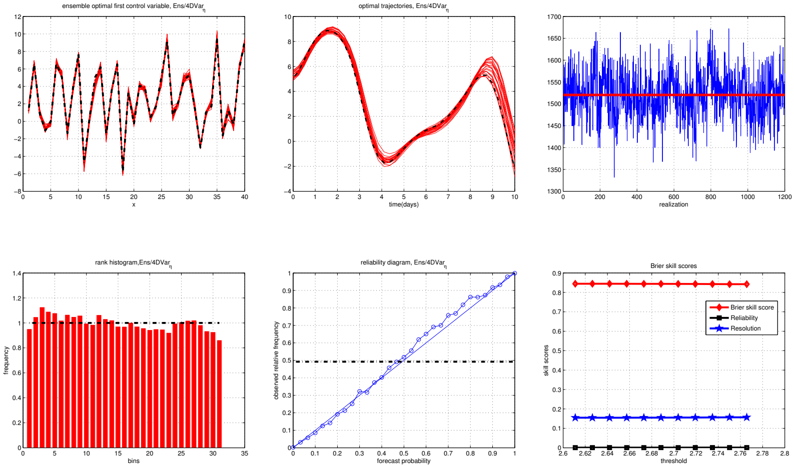


$$residue(\alpha) = \mathcal{J}(x + \alpha\delta x) - \mathcal{J}(x) - \nabla\mathcal{J}(x) \cdot \delta x$$

Weak EnsVAR : the Lorenz96 model 8 days



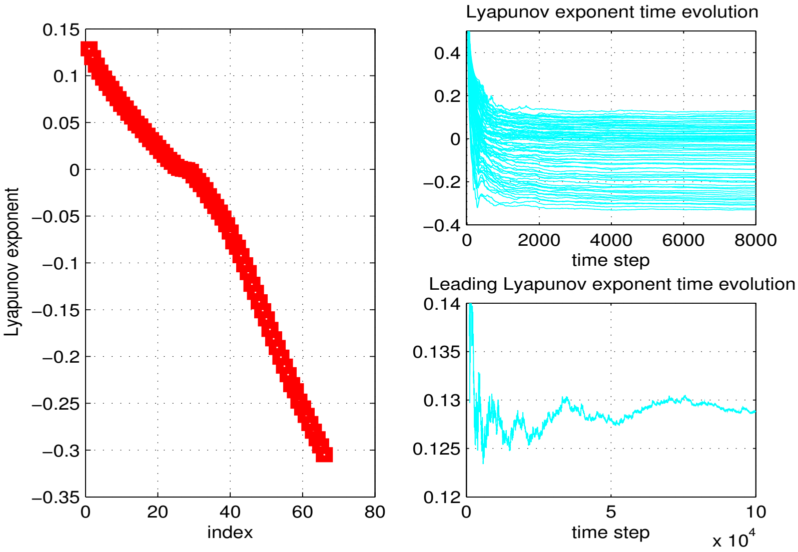
EnsVAR$_\eta$: 10 days time length



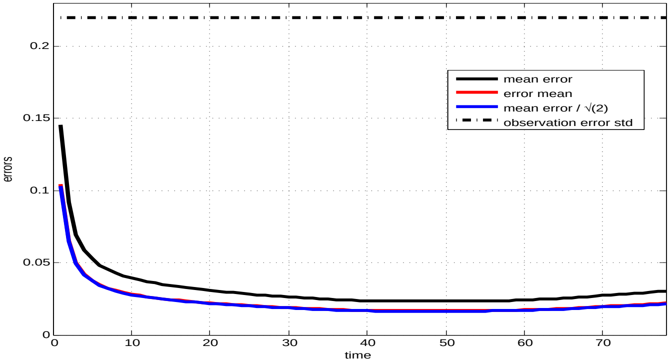
EnsVAR : the Kuramoto-Sivashinsky model

$$
\begin{cases}
\dfrac{\partial u}{\partial t} + \dfrac{\partial^4 u}{\partial x^4} + \dfrac{\partial^2 u}{\partial x^2} + u\dfrac{\partial u}{\partial x} = 0, \quad x \in [0, L] \\[2ex]
\dfrac{\partial^i u}{\partial x^i}(x+L, t) = \dfrac{\partial^i u}{\partial x^i}(x, t) \text{ for } i = 0, 1, \cdots, 4, \ \forall t > 0 \\[2ex]
u(x, 0) = u_0(x)
\end{cases}
$$

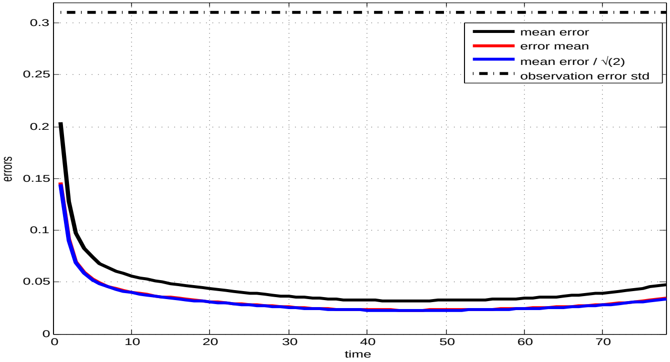
EnsVAR : the Kuramuto-Sivashinsky model



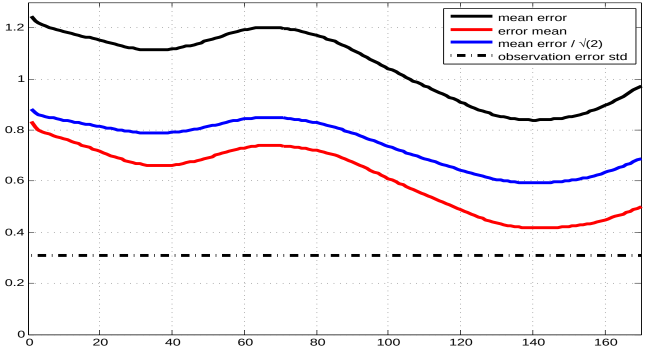
EnsVAR : the Kuramoto-Sivashinsky model, linear case



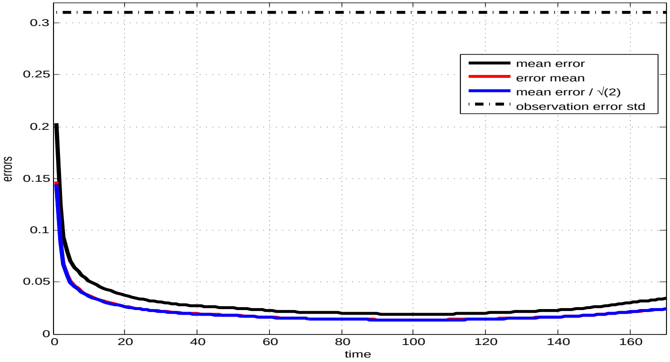
EnsVAR : the Kuramuto-Sivashinsky model, nonlinear case 1TU



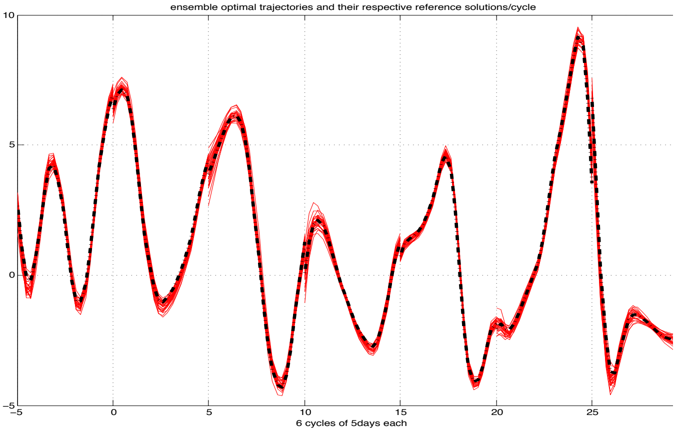
EnsVAR : the Kuramuto-Sivashinsky model, nonlinear case 2TU



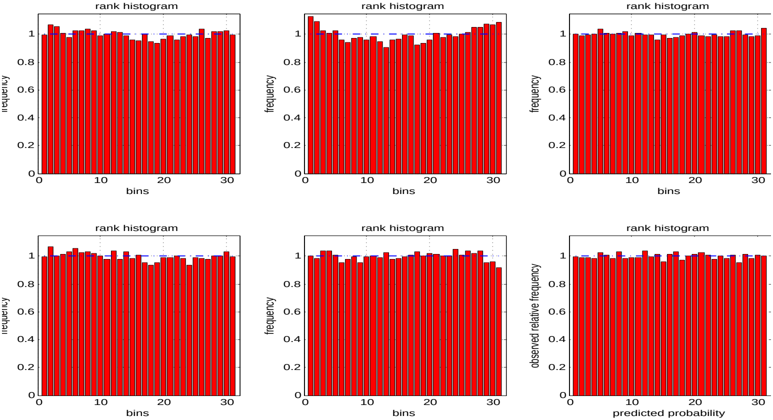
EnsVAR : the Kuramuto-Sivashinsky model, nonlinear case 2TU with QSVA



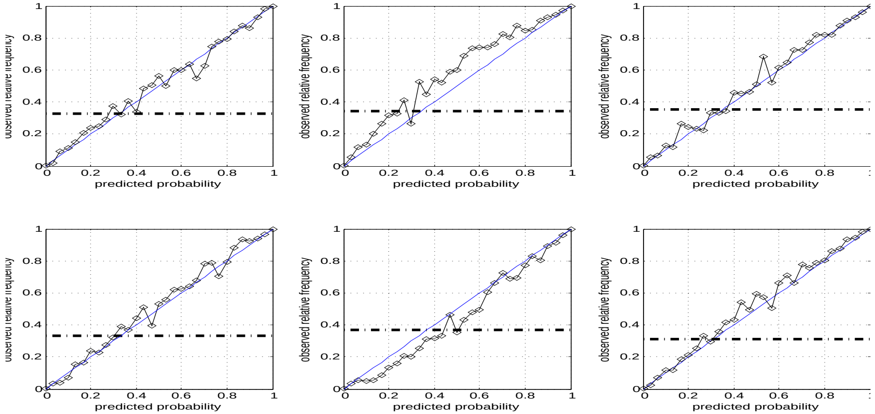
EnsVAR Cycling : use of overlapping DA windows



EnsVAR Cycling : use of overlapping DA windows



EnsVAR Cycling : use of overlapping DA windows



EnsVAR Cycling : follow the EnsVAR-AUS

- $\mathbb{E}_0$ the matrix whose columns are the N most unstable orthonormal tangent vectors.
- The evolution of $\mathbb{E}_0$over$[0,\tau]$: $\mathbb{M}_{t_i \leftarrow 0}\mathbb{E}_0 = \mathbb{E}_i \mathbf{\Lambda}_i, \quad i = 0, \cdots, n$ with

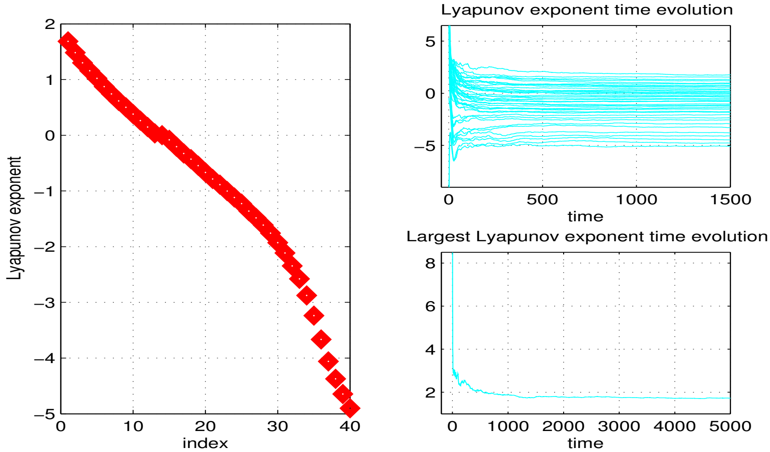
$$\mathbf{\Lambda}_i = diag \left[exp \int_0^{t_i} \lambda^{(1)}(t)dt, exp \int_0^{t_i} \lambda^{(2)}(t)dt, \cdots, exp \int_0^{t_i} \lambda^{(N)}(t)dt \right]$$

- For an increment $\delta x_0 \in \mathbb{E}_0$, its projection $\widetilde{\delta x_0} = \mathbb{E}_0 \mathbb{E}_0^T \delta x_0$
- The time evolution $\widetilde{\delta x_i} = \mathbb{M}_{t_i \leftarrow 0}\mathbb{E}_0 \mathbb{E}_0^T \delta x_0 = \mathbb{E}_i \mathbf{\Lambda}_i \mathbb{E}_0^T \delta x_0$
- The cost function gradient in the reduced space is $\widetilde{\nabla \mathfrak{J}}$

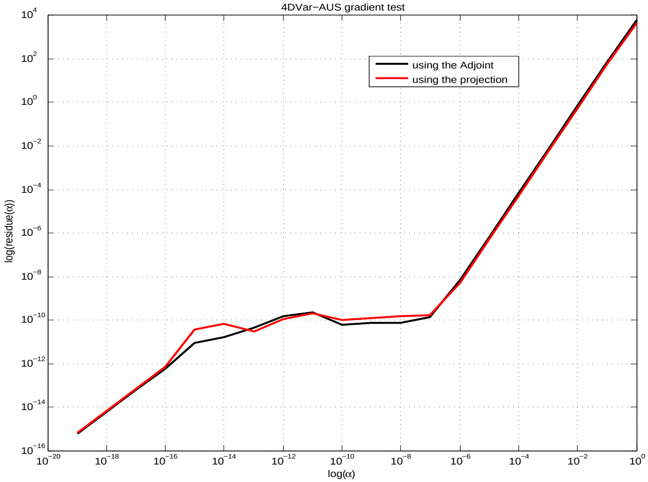
$$\widetilde{\nabla \mathfrak{J}}(X) = \left[\sum_{i=0}^{N} \mathbf{\Lambda}_i \mathbb{E}_i^T \mathbf{H}_i^T \mathbf{R}_i^{-1} (Y_i - \mathcal{H}_i(X_i)) \right]$$

- Update from cycle (k) to cycle $(k+1)$
  - $\mathbb{M}_{\tau \leftarrow 0}\mathbb{E}_0^{(k)} = \mathbb{E}_\tau^{(k)} \mathbf{\Lambda}_\tau = \mathbb{E}_0^{(k+1)} \mathbf{T}$ Gram-Schmidt orthogonalization
  - update $x_0^{(k+1)} = \mathfrak{M}_{\tau \leftarrow 0} x_0^{(k)}$

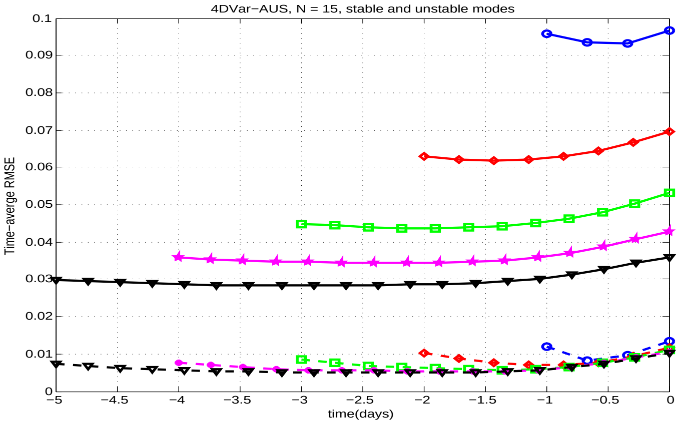
EnsVAR-AUS : Lyapunov exponents and directions



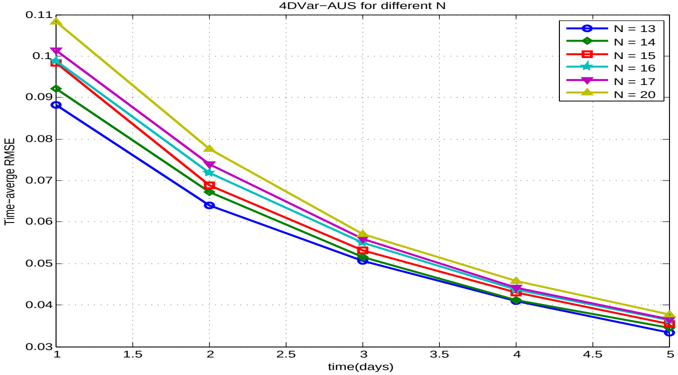
EnsVAR-AUS : gradient test



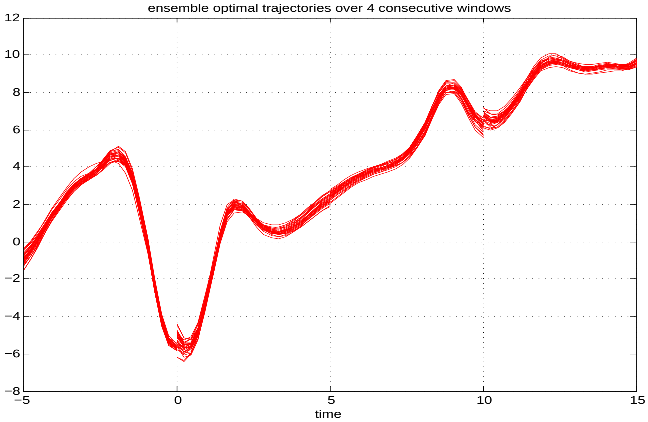
EnsVAR-AUS : Stable and unstable RMSEs



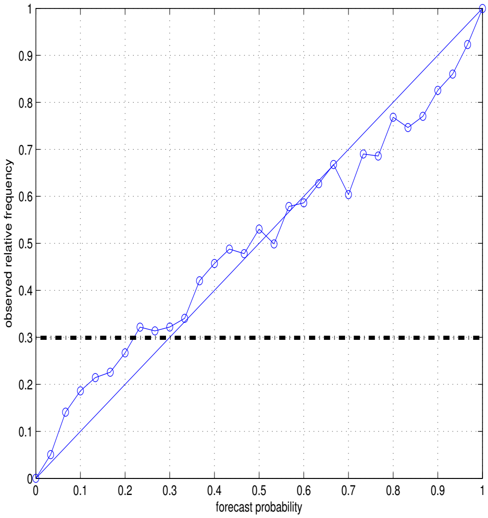
EnsVAR-AUS : Stable and unstable RMSEs



EnsVAR-AUS : Optimal trajectories



EnsVAR-AUS : Reliability



Summary, pros and cons

- 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.

- 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

- Costly (Nens 4D-Var assimilations).
- Empirical.

Thank You