

Introduction à l'assimilation de données non paramétrique

Projet Inter-Labex SEACS

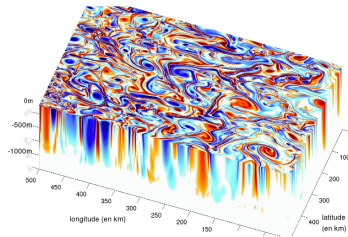
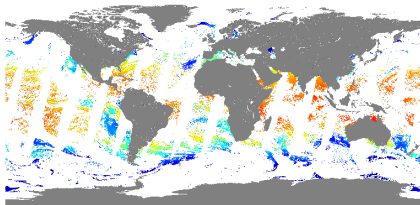
Pierre Tandeo

En collaboration avec :
P. Ailliot, A. Cuzol, V. Monbet, R. Fablet, R. Lguensat

Telecom Bretagne
Laboratoire Signal & Communications

27 mai 2015

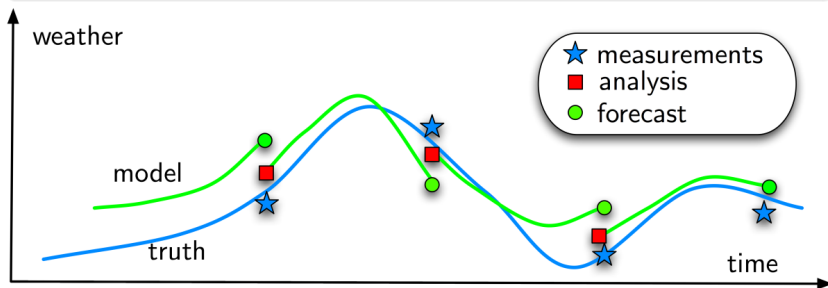
- Objectifs :
 - reconstruction de dynamique complexes
 - application à la surface des océans
- Moyens :
 - modèles dynamiques
 - observations satellite



Assimilation de données :

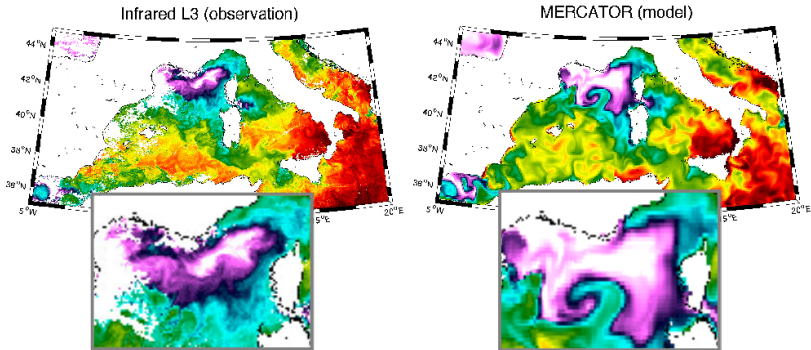
- estimer $p(\mathbf{x}|\mathbf{y})$, avec \mathbf{x} l'état caché, cf. [Evensen \(2009\)](#)
- mélange modèle \mathcal{M} et observations \mathbf{y} :

$$\begin{aligned}\frac{d\mathbf{x}(t)}{dt} &= \mathcal{M}(\mathbf{x}(t), \boldsymbol{\eta}(t)) \\ \mathbf{y}(t) &= \mathcal{H}(\mathbf{x}(t), \boldsymbol{\epsilon}(t))\end{aligned}$$



Problématiques de l'assimilation de données

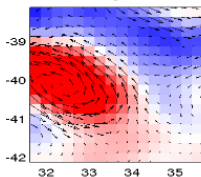
- 1 données VS modèle (haute résolution)
- 2 évaluation coûteuse du modèle (méthodes Monte Carlo)
- 3 paramétrisations incertaines



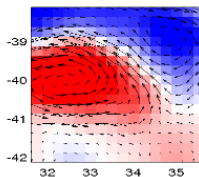
- Alternatives paramétriques
 - modèles réduits
 - modèles statistiques
- Alternative non paramétrique
 - répétabilité des évènements
 - données historiques

t

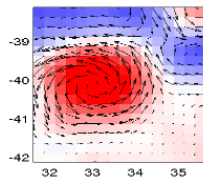
12-May-2014



24-Apr-2002

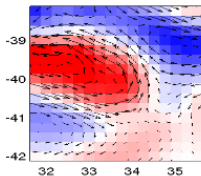


28-Jul-2011

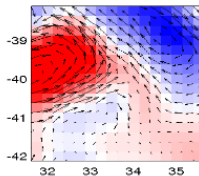


t+10

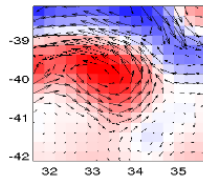
22-May-2014



04-May-2002



07-Aug-2011



Modèle espace d'état :

$$\begin{array}{lclclcl} \text{Etat caché :} & \mathbf{x}(t-1) & \xrightarrow{\mathcal{M}} & \mathbf{x}(t) & \longrightarrow & \dots \\ & & & \downarrow \mathcal{H} & & \\ \text{Observations :} & & & \mathbf{y}(t) & & \end{array}$$

Idée principale :

- émulation statistique de \mathcal{M} (machine learning)
- utilisation de données historiques (**catalogue**)

Analogues (t)	Successeurs (t+dt)
(-0.3268, +3.2644, +25.5134)	(+0.0131, +3.2278, +24.8371)
(+0.0131, +3.2278, +24.8371)	(+0.3177, +3.2017, +24.1889)
⋮	⋮
(-2.7587, -4.5007, +19.1790)	(-2.9344, -4.7112, +18.8037)
(-2.9344, -4.7112, +18.8037)	(-3.1147, -4.9464, +18.4530)

Modèle espace d'état :

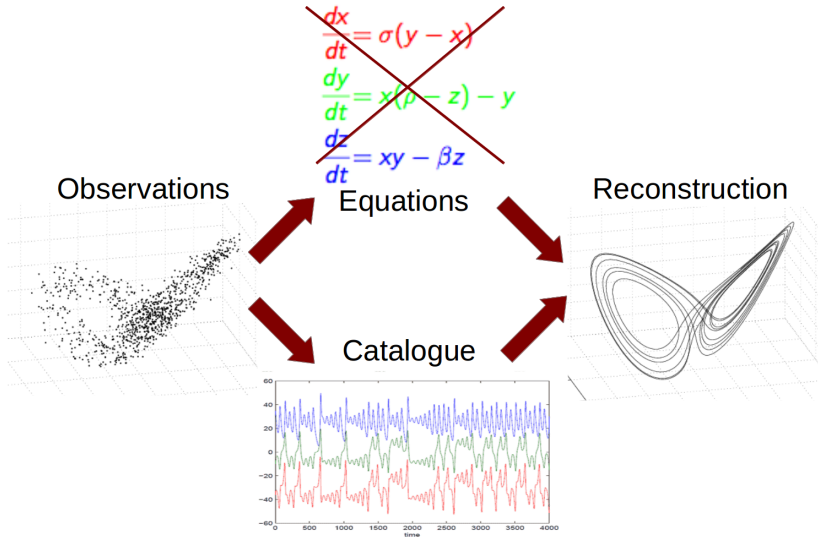
$$\begin{array}{lclclcl} \text{Etat caché :} & \mathbf{x}(t-1) & \xrightarrow{\mathcal{M}} & \mathbf{x}(t) & \longrightarrow & \dots \\ & & & \downarrow \mathcal{H} & & \\ \text{Observations :} & & & \mathbf{y}(t) & & \end{array}$$

Plusieurs implémentations :

- état continu Gaussien (EnKF) \rightarrow Pierre Tandeo
- état continu non paramétrique (PF) \rightarrow Anne Cuzol
- état discret (HMM) \rightarrow Redouane Lguensat

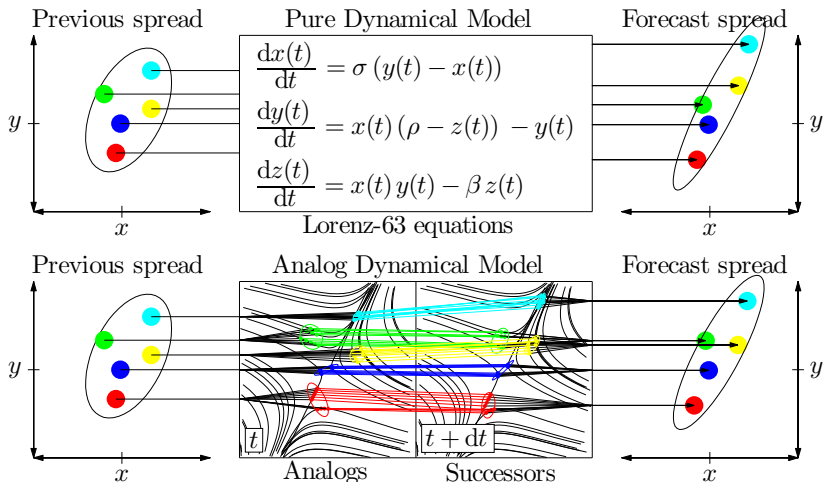
Application sur Lorenz-63

Objectif



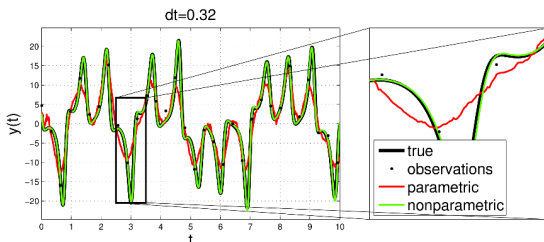
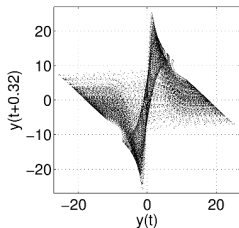
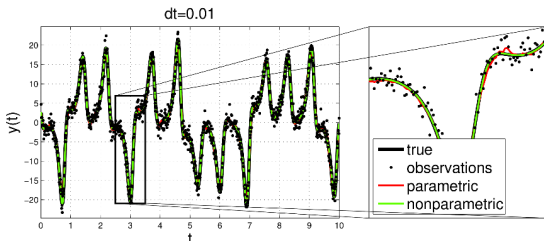
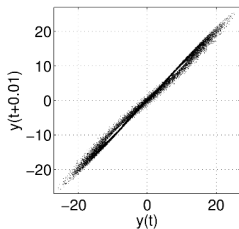
Application sur Lorenz-63

Mise en oeuvre (état Gaussien)



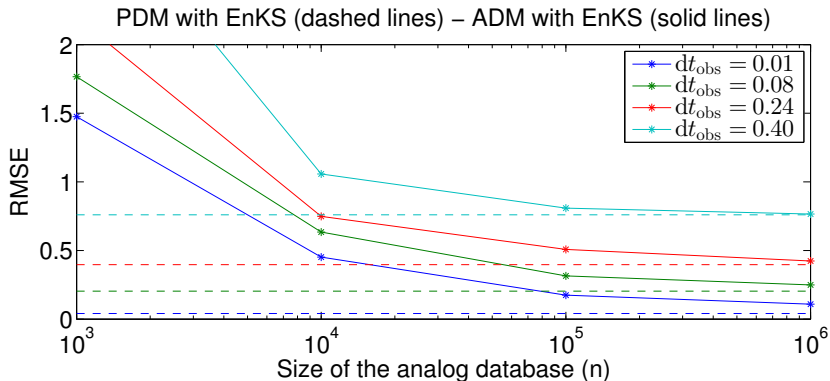
Application sur Lorenz-63

Résultats (modèle statistique autorégressif paramétrique VS non paramétrique)



Application sur Lorenz-63

Résultats (assimilation classique VS assimilation non paramétrique)



Tandeo et al. (2015), à paraître aux éditions Springer,
Machine Learning and Data Mining Approaches to Climate Science

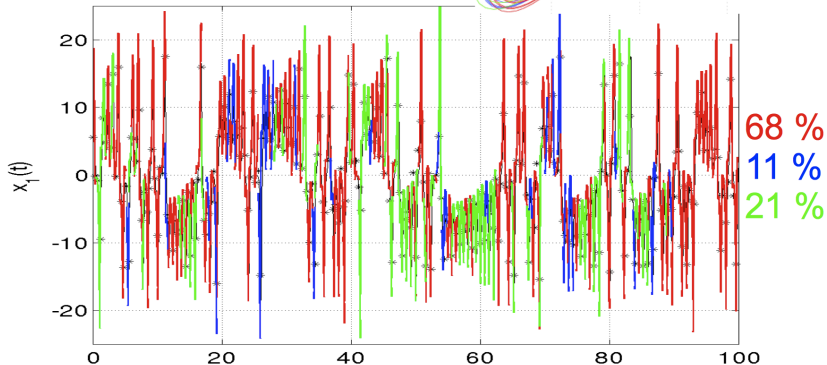
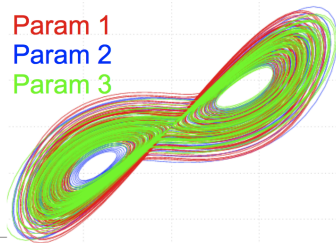
Application sur Lorenz-63

Résultats (estimation de paramètres)

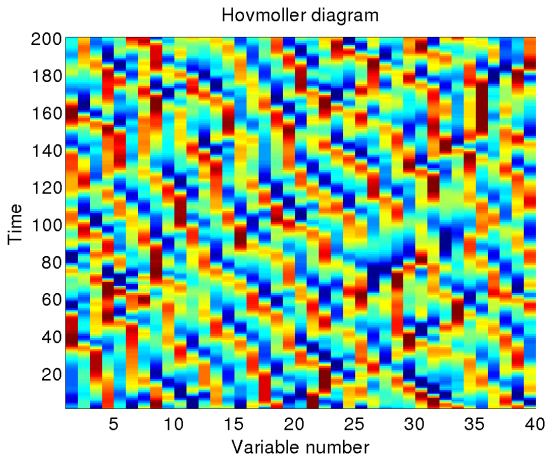
$$\frac{dx(t)}{dt} = \sigma (y(t) - x(t))$$

$$\frac{dy(t)}{dt} = x(t) (\rho - z(t)) - y(t)$$

$$\frac{dz(t)}{dt} = x(t) y(t) - \beta z(t)$$



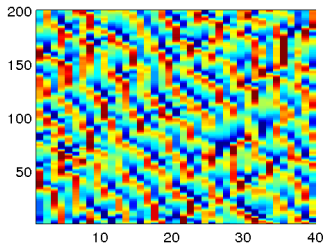
$$\frac{dx_i(t)}{dt} = (-x_{i-2}(t) + x_{i+1}(t)) x_{i-1}(t) - x_i(t) + F$$



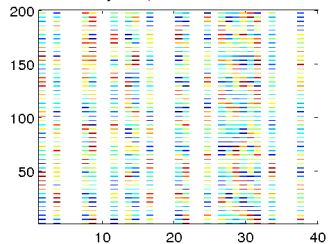
Application sur Lorenz-96

Résultats

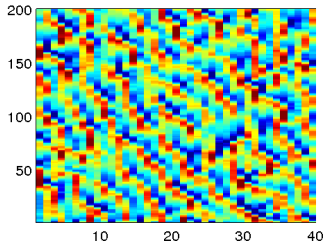
True state



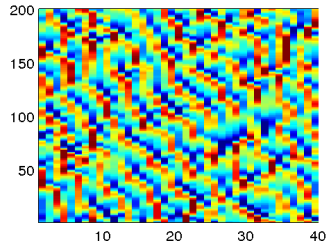
Noisy and partial observations



Nonparametric data assimilation (with analogs)



Classical data assimilation (with L-96 model)



- Conclusions :
 - méthode basée observations
 - peu couteuse
 - autres applications (estimation de paramètres)
- Perspectives :
 - fouille de données efficace
 - analogues locaux ou globaux
 - application à des archives satellitaires
 - données altimétriques (20 ans)

- [Lorenz \(1963\)](#) Deterministic nonperiodic flow. *Journal of the atmospheric sciences*.
- [Lorenz \(1996\)](#) Predictability: A problem partly solved. *Proc. Seminar on predictability*.
- [Evensen \(2009\)](#) Data assimilation: the ensemble Kalman filter. *Springer Science & Business Media*.
- [Tandeo, Ailliot, Ruiz, Hannart, Chapron, Cuzol, Monbet, Easton, Fablet \(2015\)](#) Combining analog method and ensemble data assimilation: application to the Lorenz-63 chaotic system. *Machine Learning and Data Mining Approaches to Climate Science*.