Filtering and smoothing with non parametric state space models

V. Monbet [1] joint work with T. T. Trang Chau[1], P. Ailliot[2]
Collaboration with P. Tandeo [3], A. Cuzol [4]
(part of inter-labex SEACS project)

[1]Université de Rennes 1/IRMAR & INRIA/Aspi,
[2]UBO/LMBA,
[3]IMT-Atlantique (Brest),
[4]UBS/LMBA

Ideas/Motivations

- **State Space Models & Data Assimilation**, general context

$$\begin{cases} x(t) = \mathcal{M}(x(t-1), \epsilon(t)) & \text{hidden} \\ y(t) = \mathcal{H}(x(t), \eta(t)) & \text{observed} \end{cases}$$

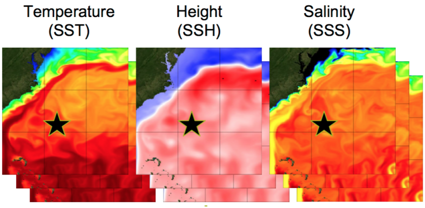
  Problem: reconstruction of **x** given observations **y**
  $\rightarrow$ data fusion, forecasting, reanalysis

- **Application in meteorology or oceanology**

Models $\mathcal{M}$
- are biased
- need forcings, boundary conditions, etc.
- need high computational cost

Observations
- Satellite (ex SST 40 years)
- In situ (ex rain 100 years)
(- Numerical weather models output.)
$\rightarrow$ hudge databases

- **Idea**: replace the numerical model $\mathcal{M}$ by a data driven (Machine Learning) model $\widehat{\mathcal{M}}_t$ learned on the available databases.



Particle filter (smoother)

- At each time t, estimation of $p(x_t|y_1,\cdots,y_t)$

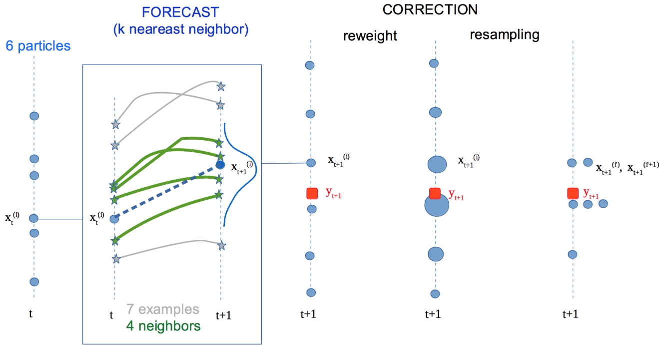
  Forecasting step

  A sample of particles is generated according to the model

$$x^{(i,f)}(k) = \widehat{\mathcal{M}}_t\left(x^{(i,a)}(k-1),\epsilon^{(i)}(k)\right),\quad i=1,\cdots,N$$

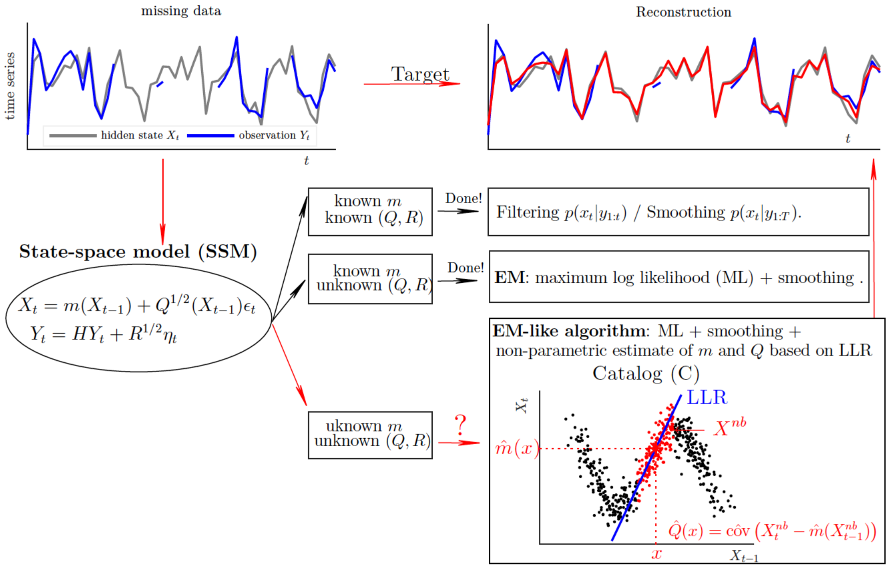
  Correction step

  The particle set $\{x^{(i,f)}(t)\}_{i=1,\cdots,N}$ is resampled according to the likelihood to match the current observation $y^{(f)}(t)$



- Estimation of $p(x_t|y_1,\cdots,y_T)$ (smoothing) needs a backward pass (CPF AS-BSS).

General problem, fundamental ingredients

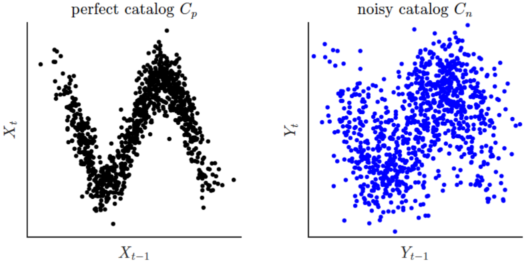


Toy example

- **Model**

$$\begin{cases} X_t = \sin(3X_{t-1}) + Q^{1/2}\epsilon_t, & \epsilon_t \sim \mathcal{N}(0,1) \\ Y_t = X_t + R^{1/2}\eta_t, & \eta_t \sim \mathcal{N}(0,1) \end{cases}$$

- Joint distribution of two successive variables in 1000-catalogs, (Q=R=0.1).

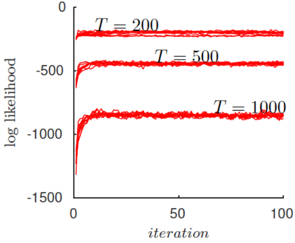


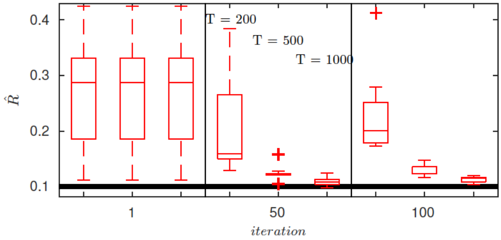
- **Problem**

  Estimation of Q, R and joint distribution of (X_{t-1}, X_t) (black scatter plot) given only $\{y_t\}_{t=1,\cdots,T}$ (blue scatter plot)
  Remark: $\mathcal{M}(x) = \sin(3x)$ is unknown.

Results

- 10 repetitions of EM-like algorithm for T = 200, T = 500, T = 1000.
- CPF-BSS is run with N_f = 20 filtering particles and N_s = 2 smoothing realizations.

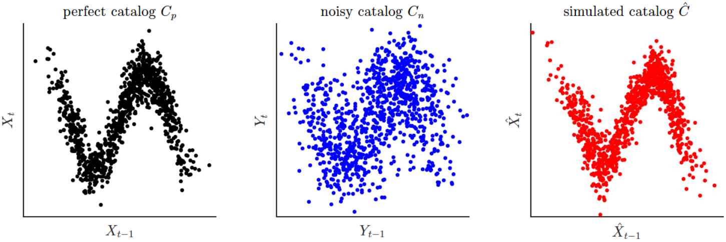




- EM-like algorithm converges to the true value if the catalog is large enough.

Results

- Joint distribution of of (X_{t-1}, X_t) learned in 1000-catalogs after 100 iterations of EM-like algorithm.



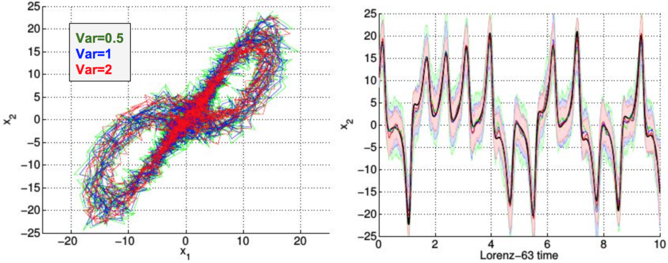
- MSE between true state and estimated state

MSE	$(\mathcal{M}, R)$	(C_p, R)	(C_n, R)	$(C_p, \hat{R})$	$(C_n, \hat{R})$	$(\hat{C}, \hat{R})$
smoothing	0.0475	0.0477	0.0604	0.0519	0.0589	0.0503

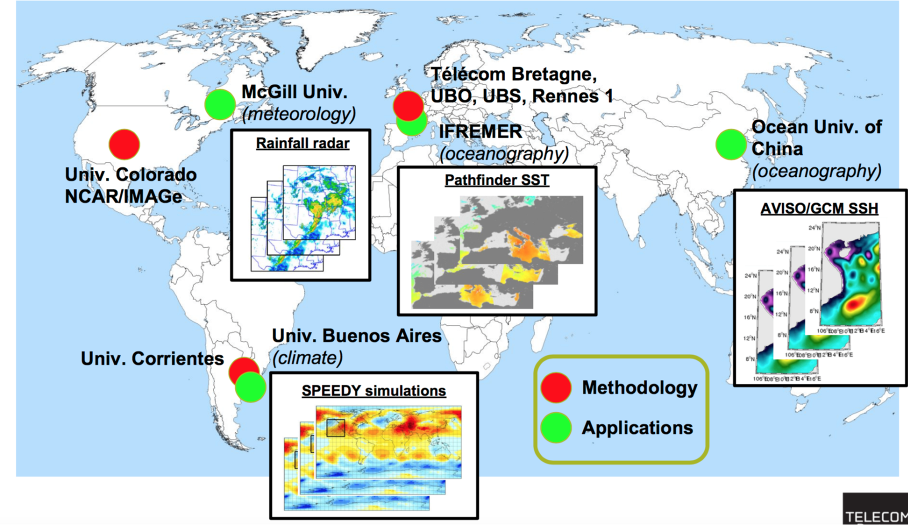
- The true model can be replaced by local linear regression.
- The EM like algorithm coupled to a smoothing algorithm allows to estimate the hidden state given noisy observations.

Concluding remarks

- **Smoothing+Machine learning**: the developped method allows to reconstruct the hidden state given noisy observations (for additive noise).
- Other studies have being done for **higher dimensional** data (without estimation of noise) [P. Tandeo et al.].



- Some **libraries** have been developed in Matlab and Python.
- **Next steps**
  - Mathematical properties of the smoother/estimators
  - Spatio-temporal dynamic kernels
  - Real data



Ref: P. Tandeo (IMT-Atlantique, Brest)