Department of Meteorology

Exploring synchronisation in nonlinear data assimilation

Flavia Pinheiro | Peter Jan van Leeuwen

University of Reading

CIÊNCIA SEM FRONTEIRAS

National Centre for Earth Observation — Natural Environment Research Council

Introduction

Issue: Data assimilation methods based on linearizations. Serious problems in strongly nonlinear cases such as convection.

Solution: A full nonlinear method: Particle filter!

Limitations: Standard Particle Filter does not work well in high-dimensional cases.

How can we overcome this limitation?

- By using **synchronisation** on each particle of a particle filter, exploring the **proposal density** freedom.

How can we understand synchronisation?

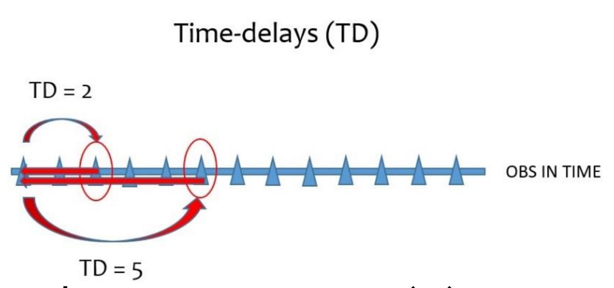
- To better investigate the method, we will compare **synchronisation** with the **Kalman Smoother** (KS), both of them **smoothing** methods.

Synchronisation – An innovative smoother

Main goal: to couple the model with the true evolution of the system via the observations, generating better estimates. With good **estimates** of the unobserved states, it is possible to make a more accurate **prediction** of them.

The basic framework used here follows **Rey et al (2014a,b)**:

- Use of time-delays (TDs) to bring additional information from measurements ahead, back to present time (assimilation window).



- A state vector (S) and an observation vector (Y) are constructed along these TDs.

$$S = (H(x(t)), H(x(t+\tau)), ..., H(x(t+(TD-1)\tau)))^T$$

$$Y = (y(t), y(t+\tau), ..., y(t+(TD-1)\tau))^T$$

- A Jacobian is calculated.

$$\frac{\partial S}{\partial x}(t) = \begin{pmatrix} J_{11}(t) & \dots & J_{1D}(t) \\ J_{11}(t+\tau) & \dots & J_{1D}(t+\tau) \\ \vdots & \ddots & \vdots \\ J_{11}(t+(TD-1)\tau) & \dots & J_{1D}(t+(TD-1)\tau) \end{pmatrix}$$

where

$$\frac{\mathrm{d}}{\mathrm{d}t} J_{ab}(t) = \sum_{c=1}^{D} \frac{\partial F_a}{\partial x_c}(t) J_{cb}(t)$$

The **variable estimates** are then calculated using:

$$\frac{\mathrm{d}x_a(t)}{\mathrm{d}t} = F_a(x(t)) + g_n \delta(t-t_n)(\partial S/\partial x)^{-1}(t)(Y(t)-S(t))$$

$(\partial S/\partial x)^{-1}$ is the pseudoinverse and it is responsible for spreading information from the time-delay data to all dynamical variables of the model (observed and unobserved).

Experiments and Results

We performed twin experiments using: Lorenz96 model, Δt = 0.01, TD = 5, observation window of 9,900 time steps and σ = 0.1 (observation noise).

The system has 20 variables to be predicted and only 4 variables have observations at each time step (20% of the system is observed).

RMSE

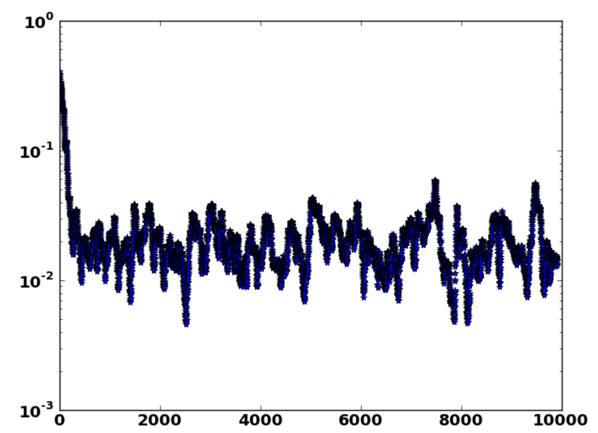


Figure 1 – Root Mean Square Error (RMSE) over all variables.

Estimates & Predictions

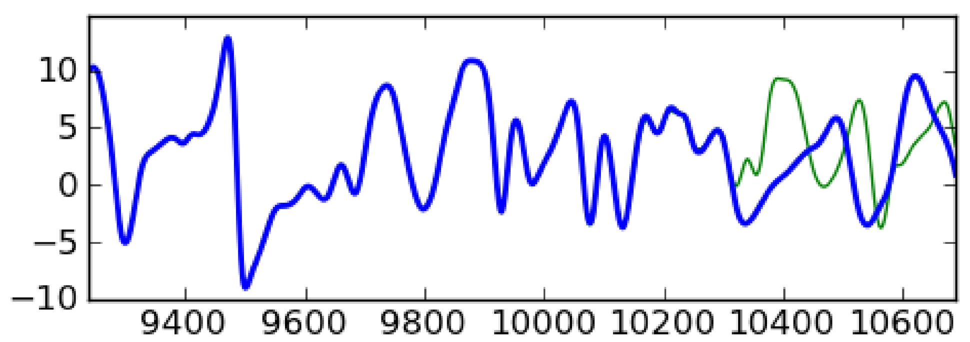


Figure 2 - Evolution of an **unobserved** variable. Blue lines: truth / Green lines: our estimates.

Estimates: until 9,900 time steps, our estimates are so **perfect** that we can't see the green line!

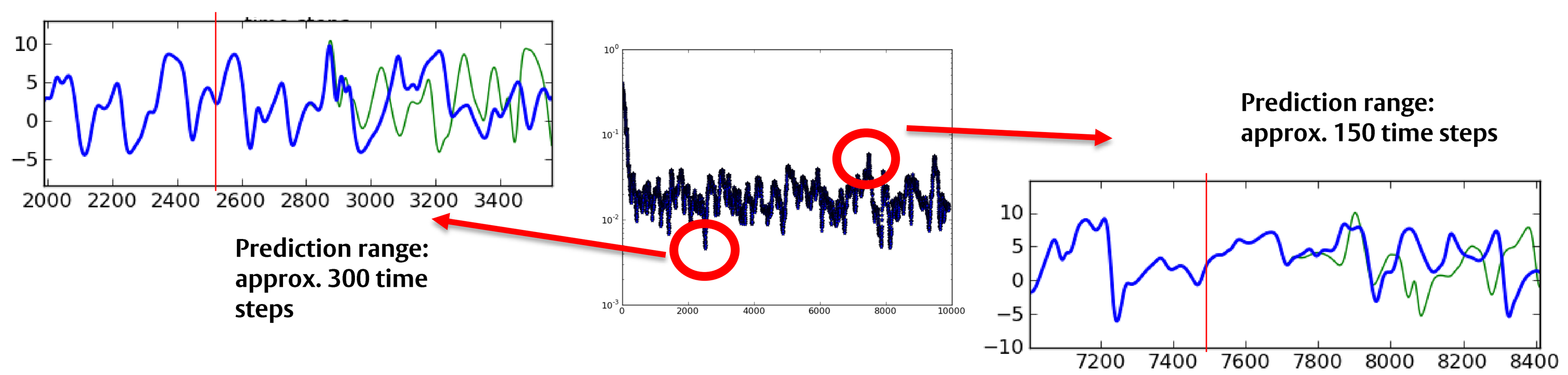
Predictions: after 9,900 time steps, we let the system run freely, with **no** data assimilation. We can get **very precise** predictions for around **400** time steps.

Step forward!

In this work the synchronisation method includes **errors** in the observations, which has not been done before in this technique.

Prediction ranges in a synchronised system

Our runs don't need to go until 9,900 time steps. When the length of our observation window coincides with the moment our system is more well synchronised, it is possible to generate longer forecasts.



A connection between synchronisation and Kalman Smoother (KS)

Kalman Smoother: Using the covariance form of the KS we obtained bad results, compared to synchronisation. Estimates struggled to match the truth until not far than 300 time steps and there were many rank issues in these calculations.

So, we decided to use the information form of the smoother:

$$x_a = x_b + (P^{-1} + H^T R^{-1} H)^{-1} H^T R^{-1} (Y - S)$$

In this case, *H* contains model evolution over time: $H = (\mathrm{H}, \mathrm{HM}, \mathrm{HM}^2, ..., \mathrm{HM}^{(TD-1)})$, which is equivalent to the synchronisation Jacobian, leading us to the connection between synchronisation and the KS framework:

$$x_a = x_b + (\sigma^2 P^{-1} + (\partial S/\partial x)^T (\partial S/\partial x))^{-1} (\partial S/\partial x)^T (Y - S)$$

where σ^2 is the variance in the observation error covariance matrix *R*.

- ✓ **Results** for this adapted KS formulation: practically equal to synchronisation!
- ✓ A KS problem is the propagation of the full P, which can be avoided in the information form we are using.
- ✓ **Advantage** of this KS formulation: the P covariance matrix can be chosen to be an identity matrix, much simpler than the usual complex P (composed of many submatrices). Note the connection with regularisation.

Conclusions

- Results indicate that synchronisation is a promising tool to steer model states to the true evolution of the system.
- A connection was found between synchronisation and the Kalman Smoother, with a KS version that does not need to propagate the P matrix.
- The next step is to use ensembles to construct $\frac{\partial S}{\partial x}$ to expand to bigger systems.
- If this new method is successful in high-dimensional systems, synchronisation will be a valuable tool in particle filtering.

References

1. Rey, D., M. Eldridge, M. Kostuk, H. Abarbanel, J. Schummann-Bischoff, and U. Parlitz, 2014a: Accurate state and parameter estimation in nonlinear systems with sparse observations. Physics Letters A, 378, 869873, doi:10.1016/j.physleta.2014.01.027.
2. Rey, D., M. Eldridge, U. Morone, H. Abarbanel, U. Parlitz, and J. Schummann-Bischoff, 2014b: Using waveform information in nonlinear data assimilation. Physical Review E, 90, 062916, doi:10.1103/PhysRevE.90.062916.

Acknowledgements

- This work is supported by CNPq/Brazil (National Council for Scientific and Technological Development) and the National Centre for Earth Observation (NCEO).

Contact information

- Department of Meteorology, University of Reading, Whiteknights, RG6 6BB
- Email: **f.rodriguespinheiro@pgr.reading.ac.uk**