The limits of present methods. Future perspectives. Particle filters
Part 6

Olivier Talagrand
School *Data Assimilation*
Nordic Institute for Theoretical Physics (NORDITA)
Stockholm, Sweden
30 April 2011

Exact bayesian estimation ?

Particle filters

Predicted ensemble at time t : $\{x^b{}_n, n = 1, \ldots, N\}$, each element with its own weight (probability) $P(x^b{}_n)$

Observation vector at same time : $y = Hx + \varepsilon$

Bayes' formula

$$P(x^b{}_n|y) \sim P(y|x^b{}_n)\, P(x^b{}_n)$$

Defines updating of weights

Bayes' formula

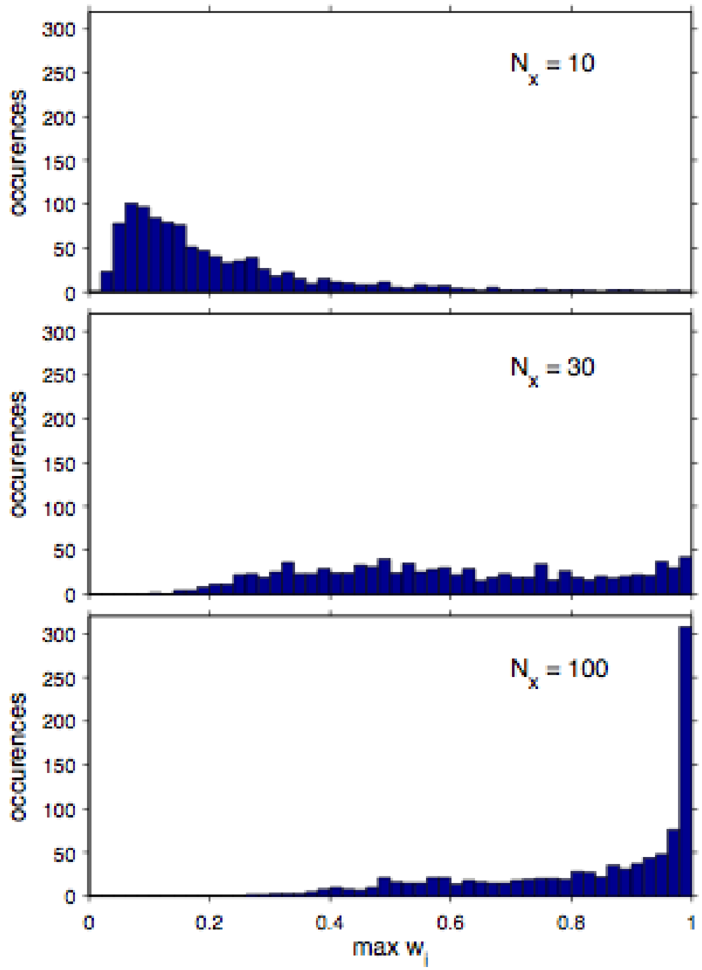
$$P(x^b{}_n|y) \sim P(y|x^b{}_n)\ P(x^b{}_n)$$

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 $= 5.5$

… $= 25$

… $= 127$

C. Snyder,
http://www.cawcr.gov.au/staff/pxs/wmoda5/Oral/Snyder.pdf

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.

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, 2009, *Mon. Wea. Rev.*, 4089-4114)

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_n are multiplied by ensemble dimension N. Then p copies of each particle n are taken, where p is the integer part of Nw_n. Remaining particles, if needed, are taken randomly from the resulting distribution.

Importance Sampling.

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). Independence between observations is then lost in the computation of likelihood $P(y|x)$.

In particular, *Guided Sequential Importance Sampling* (van Leeuwen, 2002). Idea : use observations performed at time k to resample ensemble at some timestep anterior to k, or 'nudge' integration between times k-1 and k towards observation at time k.

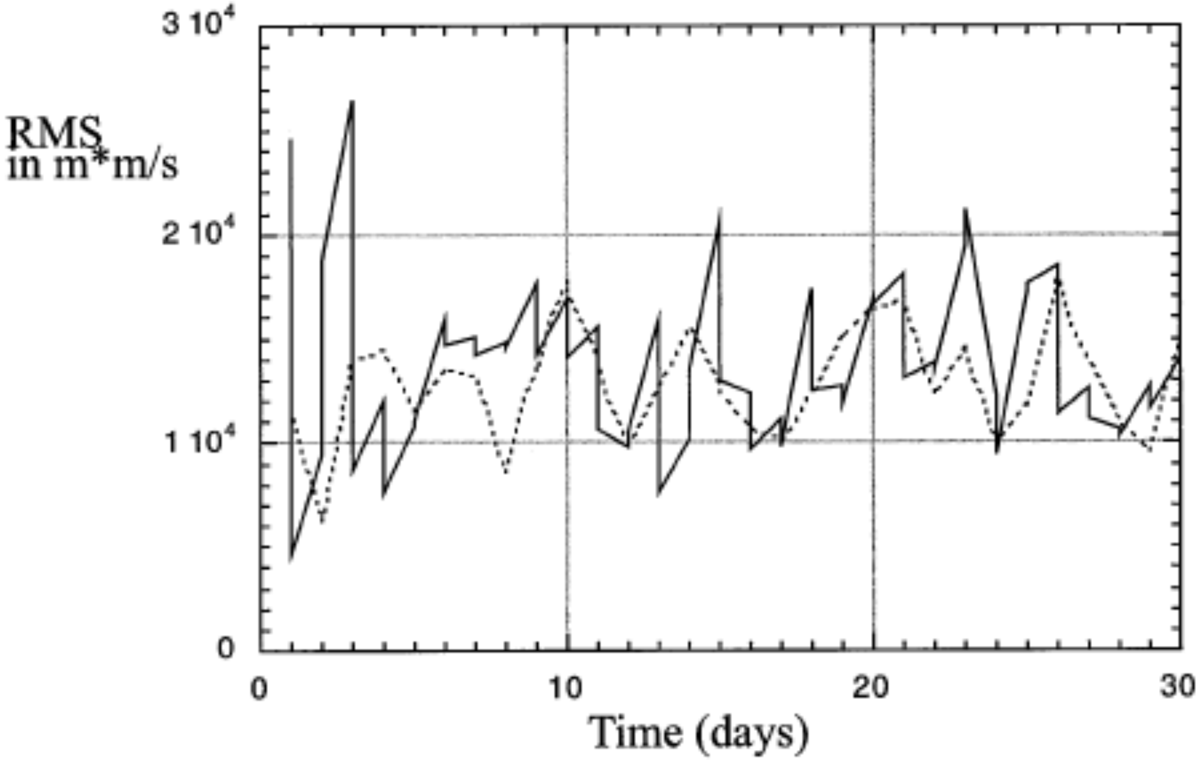


FIG. 12. Comparison of rms error ($m^2 s^{-1}$) between ensemble mean and independent observations (dotted line) and the std dev in the ensemble (solid line). The excellent agreement shows that the SIRF is working correctly.

van Leeuwen, 2003, *Mon. Wea. Rev.*, **131**, 2071-2084

Buehner (2008)

For the same numerical cost, and in meteorologically realistic situations, Ensemble Kalman Filter and Variational Assimilation produce results of similar quality.

Conclusion on Sequential Assimilation

$$x^a_k = x^b_k + P^b_k H_k^{\mathrm{T}} [H_k P^b_k H_k^{\mathrm{T}} + R_k]^{-1} (y_k - H_k x^b_k)$$

Pros

'Natural', and well adapted to many practical situations

Provides, at least relatively easily, explicit estimate of estimation error, especially in its 'ensemble' form.

Cons

Carries information only forward in time (of no importance if one is interested only in doing forecast)

In present form, optimality is possible only if errors are independent in time

Conclusion on Variational Assimilation

Pros

Carries information both forward and backward in time (important for reassimilation of past data).

Can easily take into account temporal statistical dependence (Järvinen *et al.*)

Does not require explicit computation of temporal evolution of estimation error

Very well adapted to some specific problems (*e. g.*, identification of tracer sources)

Cons

Does not readily provide estimate of estimation error

Requires development and maintenance of adjoint codes. But the latter can have other uses (sensitivity studies).

- Dual approach seems most promising. But still needs further development for application in non exactly linear cases.

- Is ensemble variational assimilation possible ? Probably yes. But also needs development.

Conclusions

- The main two classes of algorithms that presently exist for operational assimilation of observations in geophysical applications are Ensemble Kalman Filter (EnKF) and Variational Assimilation (4D-Var). They both are more or less empirical extensions to mildly nonlinear and nongaussian situations of algorithms which achieve Bayesian estimation in linear and gaussian situations.

- These two classes of algorithms produce useful (and of comparable quality) results.

- They are far fom optimality. There is no obvious way for improving on them, but ensemble methods, meant to produce a sample of the sought-for conditional probability distribution, seem most promising. Fully bayesian particle filters are the subject of active research.

Conclusions (continued)

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

- Oceanography
- 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)
- …

Assimilation is related to

- Estimation theory
- Probability theory
- Atmospheric and oceanic dynamics
- Atmospheric and oceanic predictability
- Instrumental physics
- Optimisation theory
- Control theory
- Algorithmics and computer science
- …

A few of the (many) remaining problems :

- Observability (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
- …