Hybrid particle-ensemble Kalman filter for Lagrangian data assimilation

Amit Apte[1]

([1]International Centre for Theoretical Sciences – TIFR, Bengaluru 560089 India)
apte@icts.res.in (corresponding author)

Abstract

This talk discusses the recently proposed hybrid particle-ensemble Kalman filter to assimilate Lagrangian data into a non-linear, high-dimensional dynamical model of the system. The hybrid filter applies a particle filter to the highly nonlinear, low-dimensional Lagrangian instrument variables while applying an ensemble Kalman type update to the high-dimensional Eulerian flow field. We will discuss some of the results of application of this filter to various model problems, as well as possible shortcomings and directions for future research.

1 Introduction

Lagrangian instruments such as drifters and floats provide useful sources of observations of ocean state, for example, temperature, salinity field, etc. and are highly informative for estimating and forecasting of the ocean. But the trajectories of these Lagrangian instruments are highly nonlinear and chaotic, thus posing a great challenge to some of the commonly used data assimilation methods such as ensemble Kalman filter (EnKF). The velocity field that advects these instruments is modelled by high dimensional dynamics, which poses a challenge to the methods such as particle filters because of the problems commonly known as 'curse of dimensionality.' In this talk, I will present a hybrid particle–ensemble Kalman filter method that aims to overcome both of these challenges, in particular for Lagrangian data assimilation.

2 Hybrid particle-ensemble Kalman filter

The hybrid filter is based on particle filter with a modification that for each 'particle' or sample of the fluid flow, we use multiple 'particles' of Lagrangian drifters. The weight of each fluid particle is the sum of the weights of the Lagrangian drifters. Thus one way to think of the hybrid filter is to consider it as 'degenerate' sampling, in the sense that the some of the variables of a set of particles is identical. This is shown schematically in Figure (1).

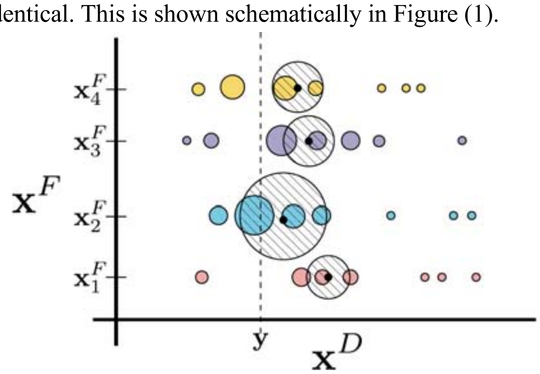

Fig.(1) The schematic of weighted particles used in the bybrid filter, where the area is proportional to weight.

The prediction step of this filter is identical to the usual particle filter or EnKF – thus each particle is propagated in time with the dynamical model. The main difference between the hybrid filters and other filters is the update step as explained below.

At the update step, the weights of each of the particles are changed by multiplying them by the likelihood of the observations at that time, conditioned on the state being equal to the position of the particle. If the weights of a significant number of particles becomes small, which is reflected in the effective sample size becoming small compared to the total number of particles, then the hybrid filter uses the following update, instead of the usual resampling of particles.

Recall that the flow field belongs to high dimensional state space. With this in mind, the hybrid filter uses the usual perturbed observation EnKF for the flow field, but using the difference between the Lagrangian observations and the mean of the Lagrangian drifter particles as the innovation vector. The drifter particles are then resampled either using the actual observations or the prior mean, and the flow field particles are resampled as well.

3 Numerical results

The hybrid filter above is applied to a simple linear shallow water equation [1] as well as to a quasigeostrophic model [2], and its efficacy is illustrated in Figure (2). This figure shows that the root mean square error for both the flow field and the position of the drifter is higher for the EnKF than for the other filters. The reference [1] also discusses a detailed mechanism for this phenomena, which relates to the nonlinearity of the drifter dynamics.

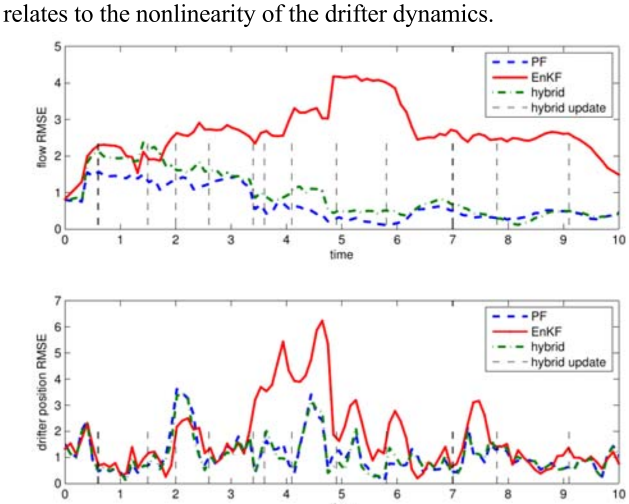

Fig.(2) The RMSE for the flow variables (top panel) and drifter variables (bottom panel), comparing the EnKF, particle filter, and the hybrid filter, for linear shallow water equation model.

The hybrid filter is also efficient at capturing the true posterior or filtering distribution, but at a much smaller computational cost. A comparison of posterior distributions of particle filter, ensemble Kalman filter, and hybrid filter (figure 3 in reference [1]) clearly shows that for highly nonlinear flows, the EnKF does not represent the posterior distribution well enough, in particular, when the prior and / or the posterior is bimodal, which leads to the failure of the implicit Gaussian assumption of EnKF. This shortcoming overcome by the hybrid filter. A more detailed analysis of this mechanism is still ongoing.

4 Discussion

The "curse of dimensionality" is one of the major challenges faced by any nonlinear filtering algorithms. Some of the commonly used methods such as the different variants of ensemble Kalman filter (EnKF) apparently work well for high dimensional system, but do not represent non-Gaussian distributions well enough. For Lagrangian data assimilation, we have demonstrated the efficacy of the proposed hybrid particle-Kalman filter in overcoming both the above difficulties. The proposed filter uses the natural splitting of the phase space into high dimensional flow field and relatively low dimensional, nonlinear Lagrangian coordinates. Generalization of this hybrid methodology to more general dynamical systems is an interesting direction of future research.

References

This talk is based mainly on the following two papers, which contain a detailed list of references that were used to develop the methods described above.

1. Slivinski, L., Spiller, E., Apte, A., & Sandstede, B. (2015). A hybrid particle–ensemble Kalman filter for Lagrangian data assimilation. Monthly Weather Review, 143(1), 195-211.
2. Slivinski, L., Spiller, E., & Apte, A. (2014, November). A hybrid particle-ensemble Kalman filter for high dimensional Lagrangian data assimilation. In International Conference on Dynamic Data-Driven Environmental Systems Science (pp. 263-273).