

# SOME ISSUES AND RESULTS ON THE EnKF AND PARTICLE FILTERS FOR METEOROLOGICAL MODELS

Christophe Baehr<sup>1,2</sup> and Olivier Pannekoucke<sup>1</sup>

- <sup>1</sup> Météo-France / CNRS  
CNRM / GAME URA1357  
42 Avenue G. Coriolis, 31057 Toulouse Cedex 1, France  
(e-mail: [christophe.baehr@meteo.fr](mailto:christophe.baehr@meteo.fr), [olivier.pannekoucke@meteo.fr](mailto:olivier.pannekoucke@meteo.fr))
- <sup>2</sup> Université de Toulouse Paul Sabatier  
Institut de Mathématiques  
118 route de Narbonne, 31062 Toulouse Cedex 9, France

**Abstract.** In this paper we examine the links between Ensemble Kalman Filters (EnKF) and Particle Filters (PF). EnKF can be seen as a Mean-Field process with a PF approximation. We explore the problem of dimensionality on a toy model. To by-pass this difficulty, we suggest using Local Particle Filters (LPF) to catch nonlinearities and feed larger scale EnKF. To go one step forward we conclude with a real application and present the filtering of perturbed measurements of atmospheric wind in the domain of turbulence. This example is the cornerstone of the LPF for the assimilation of atmospheric turbulent wind. These local representation techniques will be used in further works to assimilate singular data of turbulence linked parameters in non-hydrostatic models.

**Keywords:** Ensemble Kalman Filter, Particle Filter, Data Assimilation, Mean-Field Process.

## 1 Introduction

The major problems in data assimilation for geophysical models come from nonlinearity of dynamics, non-gaussianity of perturbations and high dimensions of state space. Ensemble Kalman Filters (EnKF) was a first answer to these difficulties. For a few years some authors have tried to use Particle Filters (PF) roughly to propose an alternative strategy. But directly applied this new approach stumbles across the problem of the dimensionality. In this paper, we present the links between EnKF and PF, we also remind that the EnKF converges but tends to a particular process and we describe the dynamical system of the nonlinear filter distribution. In the case of the PF, with a modified selection step, we investigate the effect of an increasing state space dimension for a constant number of particles. Then we suggest to couple EnKF and Local Particle Filter (LPF) to propose solutions in the vicinity of strong uncertainties. The next step is the use of LPF with a stochastic representation of the medium and we present some results on the filtering

of real turbulent wind measurements. In the conclusion we see the expected consequences of this work for meteorological models.

## 2 EnKF as a mean-field process

The nonlinear filtering process has a complete description in terms of Feynman-Kac distribution (see [3]). For instance in discrete time, we consider the dynamical system for the state vector  $X_n \in \mathbb{R}^d$  partially observed by the process  $Y_n$ :

$$\begin{cases} X_n = A_n(X_{n-1}) + \sqrt{Q_n} \cdot W_n \\ Y_n = C_n(X_n) + \sqrt{R_n} \cdot V_n \end{cases} \quad \begin{matrix} W_n \sim \mathcal{N}(0, 1) \\ V_n \sim \mathcal{N}(0, 1) \end{matrix} \quad (1)$$

where  $A_n$  and  $C_n$  are nonlinear functions of  $X_n$  ( $C_n$  is linear for EnKF),  $Q_n$  and  $R_n$  are covariances of the Gaussian noises. We denote  $M_n$  the Markov Kernel of the state vector  $X_n$ . The filtering problem consists in calculating the distributions  $\hat{\eta}_n = \mathcal{L}aw(X_n | Y_{[0,n]})$  and  $\eta_n = \mathcal{L}aw(X_n | Y_{[0,n-1]})$ , where  $Y_{[0,n]}$  is the collection  $\{Y_0, \dots, Y_n\}$ . The nonlinear filter is therefore a sequential algorithm that gives the solution of the dynamic system  $\eta_n = \eta_{n-1} K_{n-1, \eta_{n-1}}$ , where  $K_{n-1, \eta_{n-1}}$  is a (non-unique) Markov kernel representation of the filtering process. For genetic type filtering algorithm, the kernel  $K_{n-1, \eta_{n-1}}$  is  $K_{n-1, \eta_{n-1}} = S_{n-1, \eta_{n-1}} M_n$  with  $S_{n-1, \eta_{n-1}}$  a (also non-unique) selection of adapted states to be defined. We call mean-field process, a process for which the evolution depends on a probability law  $\pi_n$ , a priori or conditioned to the observations, for instance with the dynamical evolution  $X_{n+1} = F(X_n, \pi_n)$  where  $F$  is a nonlinear function. The filtering process is mean-field type. Now we will see the corresponding mean-field process of the EnKF. The EnKF is a clever technic to use an ensemble of state to approximate the covariance error matrix by an empirical matrix. The motivation of this approximation is the high dimensional size of the state vectors. The equations of the EnKF are described in [5], the filter has a prediction step and a corrective update. The convergence of the filter is proved in [6], but the limit process is not the filtering process. Denoting  $Z_n$  the update process, it is a Markov process following the nonlinear equation

$Z_n = A_n(Z_{n-1}) + \sqrt{Q_n} \cdot W_n + G_n \cdot [Y_n - C_n \cdot A_n(Z_{n-1}) - C_n \cdot \sqrt{Q_n} \cdot W_n + \sqrt{R_n} \cdot V_n]$  where  $G_n = P_n C_n^T [C_n P_n C_n^T + R_n]^{-1}$ ,  $R_n = \mathbb{E}([\sqrt{R_n} V_n] \sqrt{R_n} V_n^T)$  and  $P_n = \mathbb{E}([A_n(Z_{n-1}) + \sqrt{Q_n} W_n] [A_n(Z_{n-1}) + \sqrt{Q_n} W_n]^T)$ . This mean-field process could have a particle approximation, and with a  $N$  particles system  $(Z_n^i)_{1 \leq i \leq N}$  computing the empirical average  $\bar{Z}_n = \frac{1}{N} \sum_{i=1}^N Z_n^i$  we have  $\bar{Z}_n = \bar{A}_n(\bar{Z}_{n-1}) + G_n [Y_n - C_n \bar{A}_n(\bar{Z}_{n-1})]$ , and if  $A_n$  is a linear function, we get exactly the Kalman estimator. The estimation is exact if the pair  $(X_n, Y_n)$  is linear and Gaussian, and that is the best linear estimator of  $Z_n$  in the other cases.  $G_n$  is an mean-field operator according to  $\eta_n$  and the EnKF filtering process approaches the dynamical equation  $\eta_n = \eta_{n-1} \mathcal{C}_{n, Y_n, \eta_{n-1}} M_n$  where  $\mathcal{C}_{n, Y_n, \eta_{n-1}}$  is a correction kernel induced by  $G_n [Y_n - C_n F(Z_{n-1}, W_n) - V_n]$ .

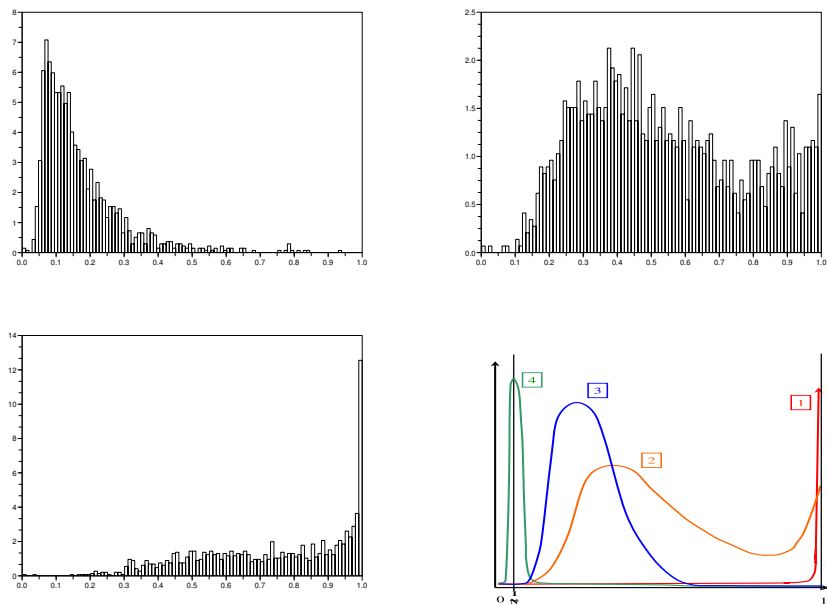
For Gaussian noises,  $\mathcal{C}_{n,Y_n,\eta_{n-1}}(x, dz) = \exp\left(-\frac{[z-x-G_n(Y_n-C_nx)]^2}{2G_n^2 R_n}\right) dz$ . The EnKF, as the number of elements goes to  $\infty$ , tends to a mean-field process  $Z_n$  different from the filtering process. With a small number of elements, the EnKF by its correction method is better than a PF, but when this number increases largely, only the PF converges to the optimal filter.

All stochastic nonlinear filter have two steps, one is the prediction according to the dynamic model, the other is an update through a selection process. At this present time, no correction process is available to ensure the convergence of the nonlinear filter. The exact filter laws are not analytically known ( except in the linear Gaussian case with the Kalman estimator ) and we have to use a particle approximation to learn these probability laws. To filter mean-field processes, there are various particle algorithms (see [1]). All are based on a mean-field Markovian model, a genetic selection rule and particle approximation for the filtering laws and for the mean-field laws. Now we turn the discussion onto the selection step with limited numerical resources which is the core of the problems and the success of any particle or ensemble filter.

### 3 Particle Filters regimes

Initially for nonlinear filters, the selection step was an Importance Sampling (IS). This kind of selection brings some difficulties with filter collapses. This is the motivation of the recent paper [8] relatively to the dimensionality. But since the late 90's, genetic selection have shown their efficiency to the filtering problems. In this selection, there is an acceptance/rejection of the state and only the rejected state are resampled. More precisely, the observational equation  $Y_n = C_n(X_n) + \sqrt{R_n} \cdot V_n$  leads to a potential function  $G_n$  ( see [3] ) which evaluates the adaptation of a state point  $X_n$  with respect to  $Y_n$ . For a parameter  $\varepsilon_n \geq 0$  such that  $\varepsilon_n G_n \in [0, 1]$ , the selection kernel is defined by  $S_{n,\eta_n}(x, dy) = \varepsilon_n G_n(x) \delta_x(dy) + [1 - \varepsilon_n G_n(x)] \Psi_n(\eta_n)(dy)$  where  $\Psi_n(\eta_n)(dy) = \frac{\eta_n(dy) G_n(y)}{\eta_n(G_n)}$  is the resampling law. In the case of the high dimensional state space we suggest to choose for the parameter  $\varepsilon_n = 1/\text{ess sup}(G_n)$ . A small noise is added on each particle to insure the exploration of the state space, and the potential  $G_n$  is corrected consequently. The use of genetic selection and this choice of the parameter  $\varepsilon_n$  provide a very different behavior in comparison with the IS selection, especially with limited computational resources. *Snyder et al.* suggest to examine the possibility of a PF collapse with  $W_n^{max}$ , the maximum of the weight  $W_n = \frac{G_n}{\eta_n(G_n)}$ . The filter is reputed to be collapsed if  $W_n^{max}$  is almost surely (a.s.) equal to 1. We conduct some numerical experiments using the dynamical model proposed by Lorenz in 1996 (see [7]). We used this chaotic model because we can easily increase the size  $d$  of its state space. We observe directly half of the state space and perturb the observation vector with a standard Gaussian noise. A PF using  $N = 1000$  particles with a genetic selection filters the signal during 1460

time steps. We examine here the histograms of maximum of weight  $W_n^{max}$ . The results for dimensions  $d = 200, 600$  and  $1500$ , presented figure 1, are quite unlike the diagrams proposed by *Snyder* for the IS. For a fixed number  $N = 1000$  of particles, the collapse is reached for a higher dimension. In fact the question seems not to be on the dimension but on the existence of a critical number of particles for a given model and a given selection rule. Theoretical works are still to be done, but numerically it seems that three regimes are possible. The first one is a collapsed filter where  $W_n^{max} = 1$  a.s., the filter is fully divergent. The second is a transitional regime and the filter may be locally divergent. The third is the optimal situation when the filter converges and  $\mathbb{P}(W_n^{max} \leq \alpha_n) = 1$  where  $\alpha_n < 1$ . A fourth regime occurs when all particles have a.s. the weight  $1/N$  which corresponds to an ill-adapted system. In the case of Lorenz-96 model, with the chosen selection rule, the critical number of particles seems to be  $\mathcal{O}(d)$ , while *Snyder et al* have shown for IS an exponential critical number. Keep in mind that for other models or other selection scheme, the result could be very different.



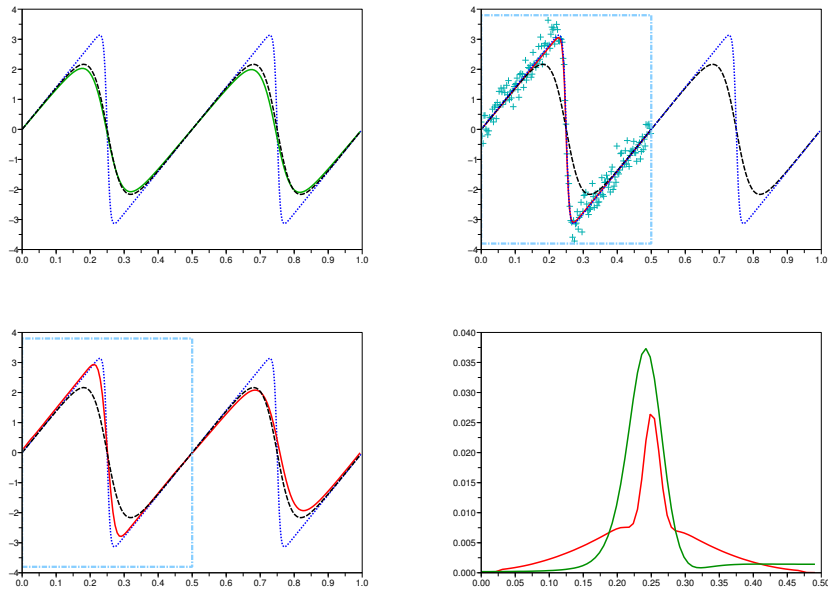
**Fig. 1.** Empirical histograms of maximum of weight  $W_n^{max}$  for a PF with 1000 particles and a genetic selection for the Lorenz-96 model with dimension  $d = 200$  (top left),  $d = 600$  (top right) and  $1500$  (bottom left). On bottom right, we summarize the 3 regimes w.r.t. to the dimension (1- collapsed filter, 2- transition regime, 3- convergent filter) and a 4th regime when the system is ill-adapted.

#### 4 Couple together a EnKF and a Local Particle Filter

To filter or assimilate data for meteorological system, operational forecast centers begin to use the EnKF. Joined to ensemble forecast systems, this way is very promising. But ensemble assimilation do not face rapid and local evolutions brought by nonlinear phenomena. A PF may be more efficient to catch these nonlinearities, and with an equal number of elements, PF are cheaper than EnKF. To bypass the problem of computational costs induced by PF in high dimension problems, it could be interesting to couple locally some Ensemble Filter and Local Particle Filter (LPF). This section is the presentation of a numerical example using a 1D Burgers equation. Discretized on the interval  $[0, 1]$  with a spectral computation scheme, this equation could be seen as the propagation of a wave along a latitude circle with the generation of a front. The Burgers equation has been chosen for its ability to generate nonlinearities. As a reference, we consider a fine scale model all over the domain, with 361 points (blue dot on each part of fig. 2 ) and generate with it perturbed observations. Then we use an EnKF (canadian type with 2 ensembles) with 100 elements, a larger grid with 161 points and we assimilate the observations every 10 time steps. On fig. 2 top left, we examine the results after a cycle of 37 assimilations. Due to the nonlinearities, the EnKF (green curve) shows some spatial shift in its reaction, and does not follow the discontinuities, rejecting the observations too close to the front. The black dash line is the same model without assimilation. These areas are therefore where the covariance prediction error matrix has its bigger values (see the dispersion of the ensemble on fig. 2 bottom right, the green curve). Then we place a smaller domain centered on the first front (light blue longdash rectangle on top right) with a refined grid, the state vector has 161 dimensions on the interval  $[0, \frac{1}{2}]$  and a Limited-Area Model (LAM). There, we use a LPF with 100 particles to filter the same set of observations (green crosses). The LAM has an adapted set of parameters : time, diffusion coefficient, etc. On fig. 2, top right, we see with the red curve that the LPF fits correctly the front and perform the best assimilation in the LAM area for an equivalent cost than EnKF in this case. Then a coupling of the LFP with large scale model is performed. To feedback the information of the LAM particles to the EnKF elements, for each element of the ensemble we randomize a particle according to the a posteriori law. The result is shown on fig. 2 bottom left and we see clearly the contribution of this coupling EnKF-LPF technique. On the fig. 2 bottom right the variance error of the LPF coupled with the EnKF is largely reduced in the LAM area.

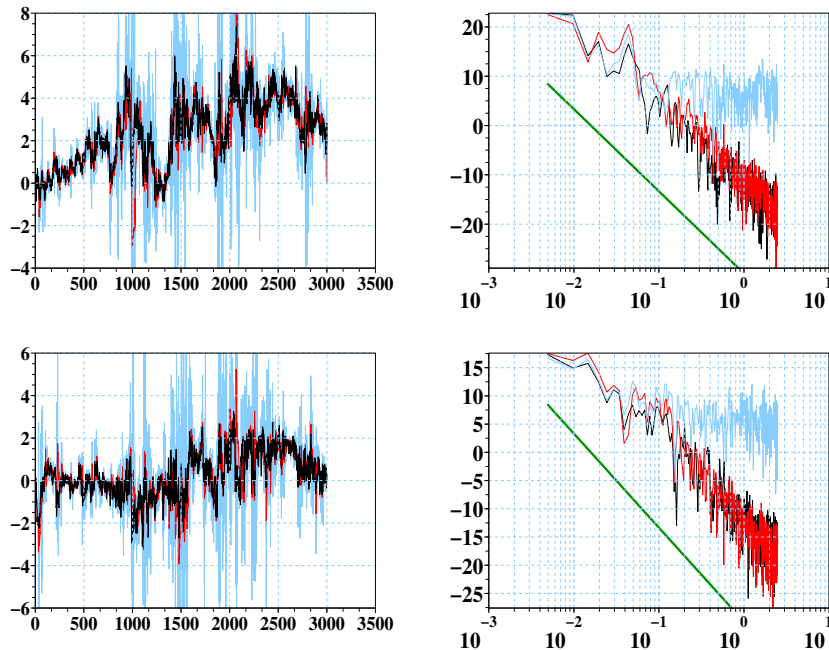
#### 5 The Filtering of Atmospheric turbulent Wind with Local Particle Filter

Regionalization of PF seems to be a response to the problem of dimensionality. It is possible to go one step forward with pointwise PF (one PF per grid-



**Fig. 2.** Observed 1d Burgers system, on each part the true state is in blue dots and black dash lines are the EnKF without assimilation. The observations (green crosses) are perturbed by white noise. On top left, the green line is the EnKF estimation with 100 elements. On top right, the red line is the LPF estimation in the LAM area with 100 particles and a genetic selection. Bottom left part, in red line is LPF coupled with large scale EnKF. Bottom right part is the dispersion of the EnKF in green and the red line is the coupled LPF-EnKF in the LAM area.

point). This technique requires two ingredients. The first one is a stochastic representation of the medium, for the atmospheric wind this is a Stochastic Lagrangian Model (SLM), and then a conditioning process to a ball centered on each gridpoint. Large scale components may be learned from the ensemble assimilation system, each PF estimates subgrid quantities and uploads the information to the larger scale model or learned from observations. The PF for turbulent wind and the conditioning process are described in [1]. Here we present the result of real data numerically perturbed and filtered by a LPF with a SLM for 3D stratified turbulent flows inspired by [2]. The model has 7 dimensions (3 for the location, 3 for wind components and one for temperature). The figure shows series of horizontal wind with the perturbed signal in blue, in black the real signal and in red the denoised signal with LPF using 300 particles. On the right part of the picture we examine the energetic structure with Power Spectral Density (PSD) and see that the corrections are very impressive even if the noises are strong.



**Fig. 3.** Left, series of horizontal wind (velocity ( $m.s^{-1}$ ) vs time step number) and right, PSD (with a log-log scale, power ( $dB$ ) vs frequency ( $Hz$ )). The series on the top is the  $U$  component and the bottom the  $V$  component. Time series are taken the 18th of July 2006 between 16h58 and 17h00 UTC. In light blue the perturbed signal to be denoised, in black the reference signal and in red the filtered with 300 particles.

This estimation technique provides not only unperturbed states but also the estimation of quantities used is the dynamical model. In our case, it could be able to produce series of turbulent dissipation rate, kinetic energy, buoyancy coefficient, etc. With these estimations it is possible to close the large scale model not with empirical closures but with the observations. For a model with fully decorrelated dimensions, pointwise PF is a cheaper technique than global PF even if a pointwise filter is computed for each gridpoint.

## 6 Outcomes and further developments

In this short paper we have seen that EnKF converges to a mean-field process which is not the filter process, while the PF converges to the optimal filter. Particle Filter is a generic name, everything takes place in the stage of selection. We have seen that it is worthwhile to use a genetic selection instead of IS for high dimensional problems. But anyway the PF requires a lot of particles as the dimension of the state space goes to infinity. We have developed

a strategy to couple together EnKF and LPF, and test it on a discretized toy model. We have pursued our investigations with Pointwise PF and seen that they are efficient for filter measurements in the domain of turbulence. Our next step will be the use of LPF, with stochastic representation or not, to assimilate data with strong nonlinearities for a barotropic model and carry information to an ensemble assimilation system. In this work we have seen that nonlinear filters have a non-unique representation with a kernel  $K_{n,\eta_n}$ . Correction process with the kernel  $\mathcal{C}_{n,Y_n,\eta_n}$  and genetic selection  $S_{n,\eta_n}$  are known answers. It may have other responses, with for instance a combination of the 2 previous kernels, but also with adaptive resampling procedure (see [4]). There are also some strategies of piloting/tuning of the selection parameters. Find new kernels and more efficient selection rules will be the next challenges for atmospheric data assimilation.

## Acknowledgments

The authors thank P. Del Moral for the helpful suggestions on EnKF mean-field process interpretation and particle filter selection strategies and also thank N. Oger, F. Suzat and R. Locatelli for their collaboration in the simulation part of this work.

## References

- 1.C. Baehr. *Modélisation probabiliste des écoulements atmosphériques turbulents afin d'en filtrer la mesure par approche particulière*. PhD thesis, Ecole Doctorale Mathématiques Appliquées Université Toulouse III - Paul Sabatier, 2008.
- 2.S.K. Das and P.A. Durbin. A Lagrangian stochastic model for dispersion in stratified turbulence. *Phys. of Fluids*, 17:025109.1–025109.10, 2005.
- 3.P. Del Moral. *Feynman–Kac Formulae, Genealogical and Interacting Particle Systems with Applications*. Springer-Verlag, 2004.
- 4.P. Del Moral, A. Doucet, and A. Jasra. On adaptive resampling procedures for sequential monte carlo methods. Technical report, INRIA, 2008.
- 5.G. Evensen. *Data Assimilation: The Ensemble Kalman Filter*. Springer-Verlag, 2006.
- 6.François Le Gland, Valérie Monbet, and Vu-Duc Tran. Large sample asymptotics for the ensemble Kalman filter. In Dan O. Crisan and Boris L. Rozovskii, editors, *Handbook on Nonlinear Filtering*. Oxford University Press, to appear.
- 7.E. N. Lorenz. Predictability: A problem partially solved (1996). In *Proceedings Seminar on Predictability ECMWF*, 1996.
- 8.C. Snyder, T. Bengtsson, P. Bickel, and J. Anderson. Obstacles to high-dimensional particle filtering. *Mon. Wea. Rev.*, 136:4629–4640, 2008.