

# Filtrage stochastique non linéaire et approximation particulière pour des modèles météorologiques

LEFE-ASSIM

Christophe Baehr

Météo-France / IMT-LSP Univ. Paul Sabatier

Paris

Vendredi 4 Décembre 2009



**METEO FRANCE**  
Toujours un temps d'avance

## De va-t-on voir ?

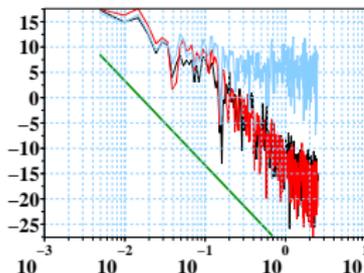
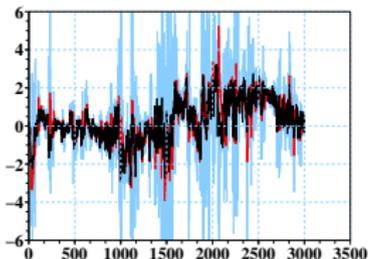
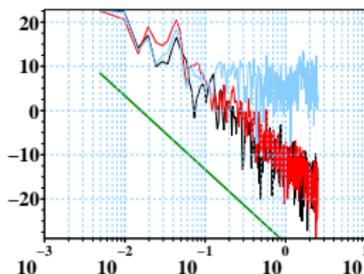
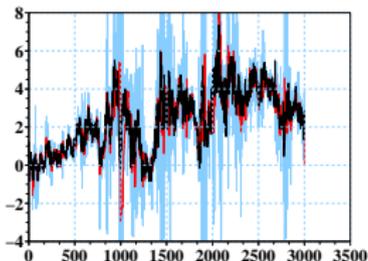
1. Une digression introductive.
2. Le problème filtrage non-linéaire.
3. Les problèmes liés à la sélection des états.
4. L'approximation particulière des estimateurs non-linéaires.
5. Asymptotiques et questions posées par les dimensions dans l'approximation par particules.
6. Etude sur Lorenz95.
7. Le couplage entre un PF local et un EnKF, une solution ?
8. Une application du PF à l'assimilation d'observations dans COBEL/ISBA.
9. Bilan et Perspectives.



## Une digression introductive

# Filtrage d'une observation ponctuelle sur un champ de vent turbulent

⊕ Série temporelle et DSP : Toulouse, le 18 Juillet 2006 entre 16h58 et 17h00 UTC.



## De quoi a-t-on besoin pour ça ?

1. Une représentation probabiliste du problème de filtrage.
2. Une représentation locale du milieu : un modèle Lagrangien.
3. Une modélisation du processus d'acquisition d'un champ le long d'un trajectoire.
4. Une technique de couplage du processus d'acquisition et du modèle Lagrangien.
5. Un estimateur non-linéaire trajectorien.
6. De nouveaux algorithmes d'approximations par particules, notamment pour les champs moyens (coll. avec F. Le Gland).
7. De réduire l'espace d'état qui est infinidimensionnel en conditionnant la dynamique aux observations.

## Le problème de filtrage non-linéaire

# Système d'état du problème de filtrage

⊖ Le système d'état fixe le problème de filtrage.

$$\begin{cases} X_{n+1} &= X_n + A_n(X_n)\Delta t + B_n(X_n)\Delta W_n \\ Y_n &= H_n(X_n) + \sigma_n V_n \end{cases}$$

où  $V_n$  et  $W_n$  sont des processus martingales indépendants.  $A_n$ ,  $B_n$  sont des fonctions à valeurs dans  $\mathbb{R}^d$  de  $X_n$  et  $H_n$  est une fonction à valeurs dans  $\mathbb{R}^{d'}$ ,  $\sigma_n$  une constante bornée positive.

⊖ Le problème de filtrage associé à ce système peut être, par exemple, la recherche de la  $\mathcal{L}oi(X_n|Y_n)$  ou si l'on parle de dynamiques trajectorielles de la  $\mathcal{L}oi(X_{[0,n]}|\mathcal{Y}_{[0,n]})$

⊖ On définit alors 2 lois, l'une pour la mise à jour du système, l'autre pour la prédiction :

$$\hat{\eta}_n(f) = \mathbb{E}[f(X_0, \dots, X_n) \mid Y_0 = y_0, \dots, Y_n = y_n]$$

$$\eta_n(f) = \mathbb{E}[f(X_0, \dots, X_n) \mid Y_0 = y_0, \dots, Y_{n-1} = y_{n-1}]$$

# Filtrage non-linéaire

⊕ A partir du système du filtrage précédent, on définit :

⊕ un noyau de transition Markovien  $M_{n+1}$  associé à

$$X_{n+1} = X_n + A_n(X_n)\Delta t + B_n(X_n)W_n$$

⊕ Associé à l'équation d'observation on définit une fonction potentiel non nulle :

$$G_n(x_n) \stackrel{\text{def}}{=} g_n(x_n, y_n) \text{ avec}$$

$$\mathbb{P}(H(x_n) + \sigma_n V_n \in dy_n \mid X_n = x_n) = g_n(x_n, y_n)\lambda_n(dy_n)$$

⊕ On passe d'une dynamique de Markov simple à un processus de Markov pondéré (MAMA).

⊕ Cette représentation amène un algorithme séquentiel en 2 étapes :

$$\eta_n \xrightarrow{S_{n,\eta_n}} \hat{\eta}_n = \eta_n \xrightarrow{M_{n+1}} \eta_{n+1} = \hat{\eta}_n M_{n+1}$$

où, en anticipant un peu,  $S_{n,\eta_n}$  sera un noyau Markovien de mise à jour.

# Filtrage non-linéaire

- ⊕ Les objets calculés séquentiellement sont les lois  $\hat{\eta}_n$  et  $\eta_{n+1}$ .
- ⊕ On peut montrer que l'on a

$$\hat{\eta}_n(f) = \mathbb{E}(f(X_n) \mid Y_0, \dots, Y_n) = \frac{\mathbb{E} \left[ f(X_n) \prod_{p=0}^n G_p(X_p) \right]}{\mathbb{E} \left[ \prod_{p=0}^n G_p(X_p) \right]}$$

$$\eta_{n+1}(f) = \mathbb{E}(f(X_{n+1}) \mid Y_0, \dots, Y_n) = \frac{\mathbb{E} \left[ f(X_{n+1}) \prod_{p=0}^n G_p(X_p) \right]}{\mathbb{E} \left[ \prod_{p=0}^n G_p(X_p) \right]}$$

- ⊕ et séquentiellement on montre que l'on a

$$\begin{aligned} \hat{\eta}_n(f) &= \frac{\eta_n(G_n \cdot f)}{\eta_n(G_n)} = \psi_n(\eta_n)(f) \\ \eta_{n+1}(f) &= \hat{\eta}_n M_{n+1}(f) \end{aligned}$$

## Le problème de sélection des états

## Le problème de la sélection

- ⊕ L'équation  $\hat{\eta}_n(f) = \psi_n(\eta_n)(f)$  suggère que la mise à jour des états se fait par une transformation de loi.
- ⊕ En général on préfère pour des processus Markoviens des transports de loi. On cherche donc une représentation de cette mise à jour comme un transport.
- ⊕ C'est la représentation de McKean. Mais McKean nous a montré que cette représentation n'est pas unique.
- ⊕ Tout noyau  $S_n$  respectant la relation de compatibilité  $\eta_n S_n(f) = \psi_n(\eta_n)(f)$  sera admissible.

## Exemples de noyaux de sélection

- ⊕ Le plus classique, et le premier historiquement décrit :

$$S_{n,\eta_n}^1(x_n, dx) = \frac{G_n(x)}{\eta_n(G_n)} \eta_n(dx)$$

- ⊕ Le noyau de sélection génétique :

$$S_{n,\eta_n}^2(x_n, dx) = G_n(x_n) \delta_{x_n}(dx) + [1 - G_n(x_n)] \frac{G_n(x)}{\eta_n(G_n)} \eta_n(dx)$$

- ⊕ Le noyau de sélection génétique à paramètre  $\varepsilon_n$  :

$$S_{n,\eta_n}^3(x_n, dx) = \varepsilon_n \cdot G_n(x_n) \delta_{x_n}(dx) + [1 - \varepsilon_n \cdot G_n(x_n)] \frac{G_n(x)}{\eta_n(G_n)} \eta_n(dx)$$

où  $\varepsilon_n \cdot G_n(x_n) \in [0, 1]$  pour tout  $x_n \in E$

- ⊕ Dans les 3 cas, l'équation de compatibilité est vérifiée et on a :

$$\eta_{n+1}(f) = \eta_n S_{n,\eta_n} M_{n+1}(f) = \eta_n K_{n+1,\eta_n}(f)$$

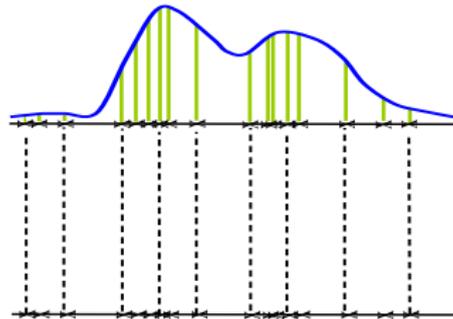
## Adaptation au cas trajectoriel

- ⊕ Tout ce que l'on vient de voir s'écrit de manière identique si la dynamique est trajectorielle (c'est à dire que  $X_n = (X'_0, \dots, X'_n)$ ).
- ⊕ On utilise le modèle de dynamique sur  $X'_n$  pour étendre la trajectoire  $X_n$ .
- ⊕ On définit  $G_n(X_n) = G_n(X'_0, \dots, X'_n) = G'_n(X'_n)$
- ⊕ On sélectionne sur la tête de la trajectoire mais c'est tout une trajectoire qui est retenue ou redistribuée sur une autre.
- ⊕ C'est ici que l'on peut placer des corrélations temporelles dans les dynamiques ou les observations (voir l'exemple introductif).

## L'approximation particulière des estimateurs non-linéaires

## Approximation particulière

- ⊕ Les équations intégrales que l'on a vu, ne sont pas calculables analytiquement hors du cas linéaire Gaussien. Nous devons utiliser une méthode d'approximation pour résoudre ces équations.
- ⊕ Nous utilisons une représentation des lois de probabilités par particules. On peut montrer que cette approche particulière résout le problème de filtrage.
- ⊕ Ce filtrage particulier est apparenté méthodes de Monte-Carlo.



## Approximation particulière

⊖ A chaque pas de temps  $n \geq 0$  on dispose d'un ensemble de particules  $(\hat{\xi}_n^{i,N})_{1 \leq i \leq N}$  (on ne regarde que la loi de mise à jour).

⊖  $(\hat{\xi}_n^{1,N}, \dots, \hat{\xi}_n^{N,N})$  sont distribuées selon

$$\lim_{N \rightarrow \infty} \sum_{i=1}^N \delta_{\hat{\xi}_n^{i,N}} = \mathcal{L}oi(X_n \mid Y_0, \dots, Y_n)$$

la  $n$ ième marginale de  $\mathcal{L}oi(X_0, \dots, X_n \mid Y_0, \dots, Y_n)$

⊖ La population de particules bouge selon l'algorithme de sélection/prédiction que l'on a vu :

- ⊕ La prédiction utilise le modèle de dynamique.
- ⊕ Dans le cas génétique, l'étape de sélection garde en vie les états acceptés et redistribue les particules rejetées.

## Approximation particulière

- ⊕ L'algorithme de filtrage est alors assez simple.
- ⊕ Initialisation :  $N > 0$  particules  $(\hat{\xi}_0^i)_{1 \leq i \leq N} \in E$  *i.i.d.* selon  $\eta_0$  et on note  $\hat{\eta}_0^N = \frac{1}{N} \sum_{i=1}^N \delta_{\hat{\xi}_0^i}$
- ⊕ Au pas de temps  $n$ ,  $(\hat{\xi}_n^i)_{1 \leq i \leq N}$  dans  $E$  évoluent selon

$$\hat{\xi}_n^i \xrightarrow{\text{Mutation}} \xi_{n+1}^i \xrightarrow{\text{Sélection}} \hat{\xi}_{n+1}^i$$

- ⊕ Les particules donnent les lois approchées :  $\hat{\eta}_n^N = \frac{1}{N} \sum_{i=1}^N \delta_{\hat{\xi}_n^i}$  et  $\eta_n^N = \frac{1}{N} \sum_{i=1}^N \delta_{\xi_n^i}$
- ⊕ Tout est dans la méthode de sélection utilisée et dans la représentation ponctuelle ou trajectorielle du problème de filtrage.

Asymptotiques et problèmes posés  
par les grandes dimensions.

Convergence en norme  $L^p$ 

- ⊖ L'algorithme que l'on a présenté converge vers la solution du problème de filtrage.
- ⊖ On cherche alors à quantifier l'erreur commise en utilisant  $\hat{\eta}_n^N = \frac{1}{N} \sum_{i=1}^N \delta_{\xi_n^i}$  et  $\eta_n^N = \frac{1}{N} \sum_{i=1}^N \delta_{\xi_n^i}$  au lieu des lois exactes.
- ⊖ Avec une hypothèse de régularité classique on peut obtenir Théorème Pour tout  $n \geq 0$ , pour tout  $p \geq 1$ , il existe des constantes  $C_n(p)$  telles que

$$\mathbb{E}(\|\eta_n^N - \eta_n\|_{\mathcal{H}}^p)^{\frac{1}{p}} \leq \frac{C_n(p)}{\sqrt{N}} I(\mathcal{H})$$

avec la seminorme pour 2 mesures  $\eta$  et  $\eta'$  définie par  $\|\eta - \eta'\|_{\mathcal{H}} = \sup\{|\eta(h) - \eta'(h)| \text{ tel que } h \in \mathcal{H}\}$  avec  $\mathcal{H} = \{h : E \rightarrow \mathbb{R} \text{ mesurable telle que } \|h\| \leq 1\}$

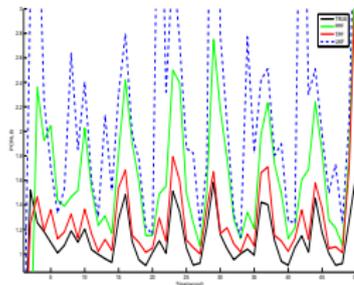
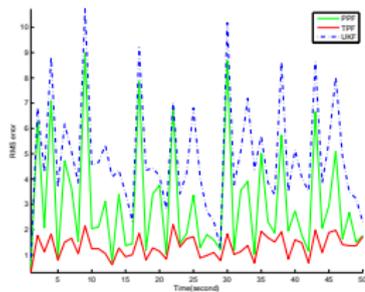
## Bornes de Cramer-Rao

(travail conjoint avec M. Lei et P. Del Moral)

⊕ On teste le modèle classique de la littérature

$$\begin{cases} x_n = 0.5x_{n-1} + \frac{25x_{n-1}}{1 + x_{n-1}^2} + 8 \cos[1.2(n - 1)] + w_n, \\ y_n = \frac{x_n^2}{20} + v_n, \quad n = 1, 2, \dots, T \end{cases}$$

⊕ La borne de Cramer-Rao est une borne inférieure sur la variance d'un estimateur.



## La question de la dimension

- ⊖ Dans le problème de filtrage non-linéaire la question de la dimension de l'espace d'état n'apparaît pas.
- ⊖ C'est dans l'approximation particulaire que se cache la difficulté.
- ⊖ Pas partout, mais seulement dans la partie redistribution de l'étape de sélection.
- ⊖ Depuis 1998 on sait que le bootstrap a des difficultés en grandes dimensions.

Etude sur Lorenz95  
(travail conjoint avec O. Pannekoucke)

## Filtre à Particules et grandes dimensions

⊖ C. Snyder et al. ont proposé d'utiliser le modèle de Lorenz-96 et ont filtré par un algorithme avec une sélection classique.

⊖ On rappelle brièvement la forme du modèle de Lorenz-96 :

$$\frac{dx_\alpha}{dt} = -x_{\alpha-2} \cdot x_{\alpha-1} + x_{\alpha-1} \cdot x_{\alpha+1} - x_\alpha + F$$

où  $x_\alpha$  est l' $\alpha$ ème composante du vecteur d'état

$X = (x_1, \dots, x_\alpha, \dots, x_d)$  and  $1 \leq \alpha \leq d$ .

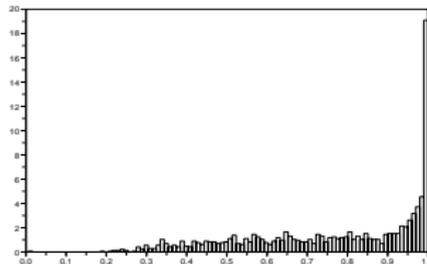
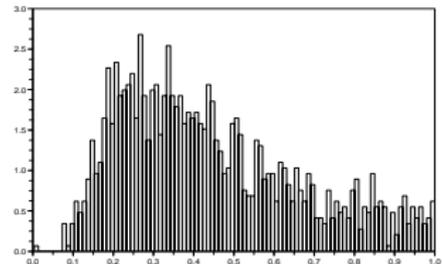
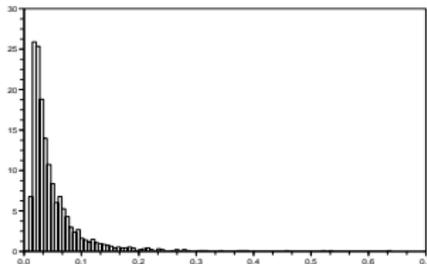
⊖ On propose une expérience où seulement la moitié des dimensions sont observées, les autres sont cachées.

⊖ Comme suggéré par Snyder et al, le comportement du PF est évalué à chaque instant par  $W_n$  le poids maximal qui est porté par l'une des particules :

$$W_n = \max_{i \in [1, N]} w_n^i = \max_{i \in [1, N]} \frac{G_n^i}{\sum_{j=1}^N G_n^j}$$

## Filtre à Particules et grandes dimensions

⊕ Pour  $N = 1000$  particules ils obtiennent des histogrammes de ce poids maximum quand la dimension est de  $d = 10, 30$  et  $80$



## Filtre génétique face aux dimensions

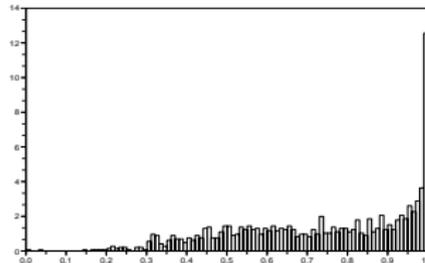
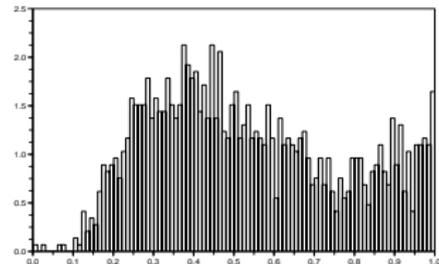
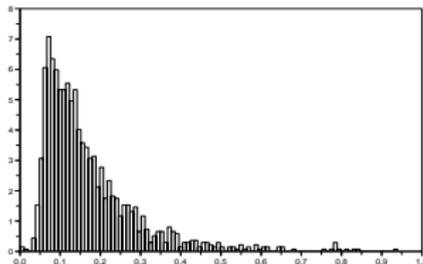
- ⊖ Donc la conclusion de l'expérience Snyder est que les filtres à particules ont des soucis quand la dimension augmente.
- ⊖ On propose d'utiliser un filtre trajectoriel à sélection génétique avec le paramètre  $\varepsilon_n = 1/\text{sup}(G_n)$  :

$$S_{n,\eta_n}(x_n, dx) = \varepsilon_n \cdot G_n(x_n) \delta_{x_n}(dx) + [1 - \varepsilon_n \cdot G_n(x_n)] \frac{G_n(x)}{\eta_n(G_n)} \eta_n(dx)$$

- ⊖ On ajoute un second ingrédient, en perturbant chaque particule avec un petit bruit Gaussien et en prenant en compte cet ajout dans la fonction de potentiel  $G_n$ .
- ⊖ Cela change radicalement le comportement de l'approximation particulaire !
- ⊖ Nous avons un filtre convergent pour des dimensions plus grandes que 400 avec toujours 1000 particules !

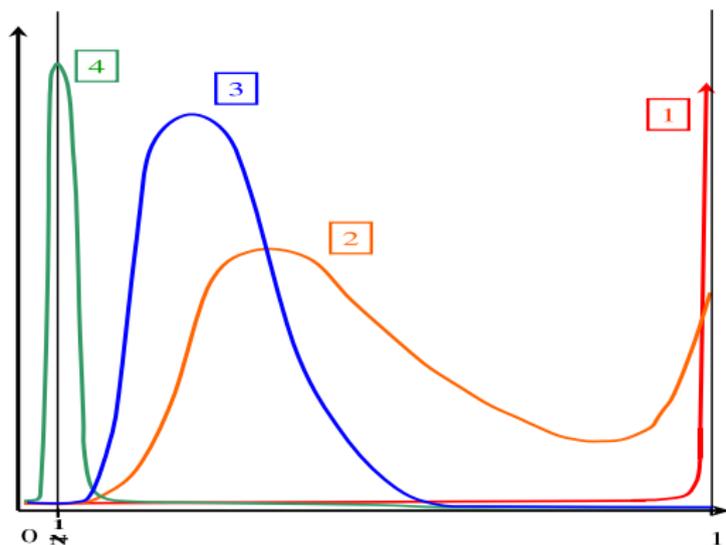
# Filtre génétique face aux dimensions

⊖ Pour  $N = 1000$  particules et des dimensions  $d = 200, 600$  et  $1500$  nous obtenons les histogrammes :



# Filtre génétique face aux dimensions

- ⊖ Le nombre critique de particules semble être de  $\mathcal{O}(d)$ , alors que Snyder *et al* ont trouvé un nombre exponentiel.



# Filtre trajectoriel génétique face aux dimensions

⊖ **Existerait-il ce théorème ?**

Théorème On se donne un espace d'état  $E$  de dimension  $d$ , un potentiel de sélection  $G_n$ , un noyau de mutation  $M_n$  ( et des hypothèses qui nous manquent ). Il existe une taille critique  $N^*(d)$  de particules dépendant des objets précédents telle que, quand  $N$  est plus grand que  $N^*(d)$ , « l'algorithme génétique ne diverge pas ».

⊖ **plus précisément en séparant les cas**

⊖ Cas 1 :  $\exists N_1^{crit}$  s.t.  $\forall N < N_1^{crit}, \mathbb{P}(W_{max} = 1) = 1$

⊖ Cas 3 :  $\exists N_2^{crit}$  s.t.  $\forall N > N_2^{crit}, \mathbb{P}(W_{max} \leq \beta) \geq \alpha$

⊖ Cas 4 : Quand  $\mathbb{P}(|W_{max} - \frac{1}{N}| \leq \beta) = 1$ , l'algorithme est mal adapté.

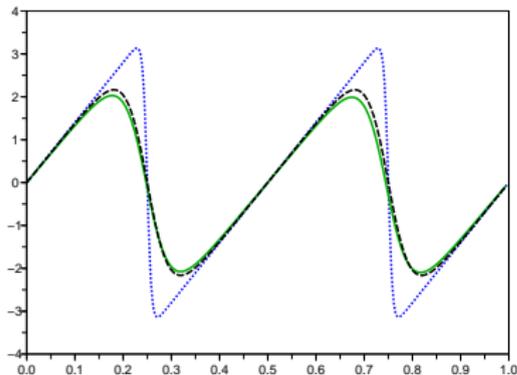
Couplage entre un Filtre Particulaire et un EnKF  
(travail conjoint avec O. Pannekoucke)

## Coupling of an EnKF and PF

- ⊕ Even if the EnKF converges to a process which is not the filtering process, it requires few elements to work in high dimensions.
- ⊕ Even if the PF converges to the right filtering process, it requires a number of particles in  $\mathcal{O}(d)$  for a genetic selection.
- ⊕ We propose now to couple together a EnKF and a Local Particle Filter. (I pass over Rao-Blackwellized Particle filters in silence ... may be later for questions if you wish ...)
- ⊕ For the numerical experiment, we use a discretized 1D-Burgers equation on  $[0, 1]$  with 361 points for an estimation of the true reality and 161 points for the model used by the EnKF.
- ⊕ Then we use a local model on the first front (161 dimensions on the intervalle  $[0, \frac{1}{2}]$ ) with a Particle Filter to assimilate the observations.

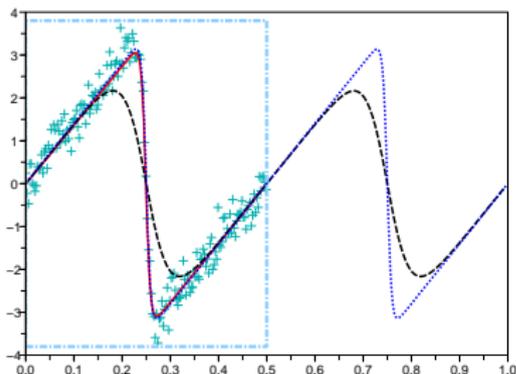
# Coupling Local PF and EnKF

EnKF result with 100 elements and after a cycle of 37 assimilations



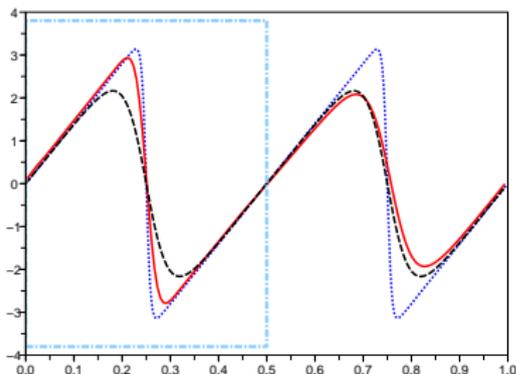
# Coupling Local PF and EnKF

Limited-Area Model, with a Local PF using 100 particles after 37 cycles of assimilation.



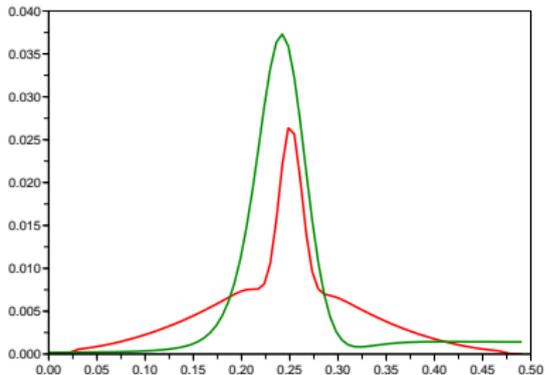
## Coupling Local PF and EnKF

Feedback of the particles to the EnKF elements in the LAM domain by randomization.



## Coupling Local PF and EnKF

Variance error of the EnKF and the LPF coupled with the EnKF.



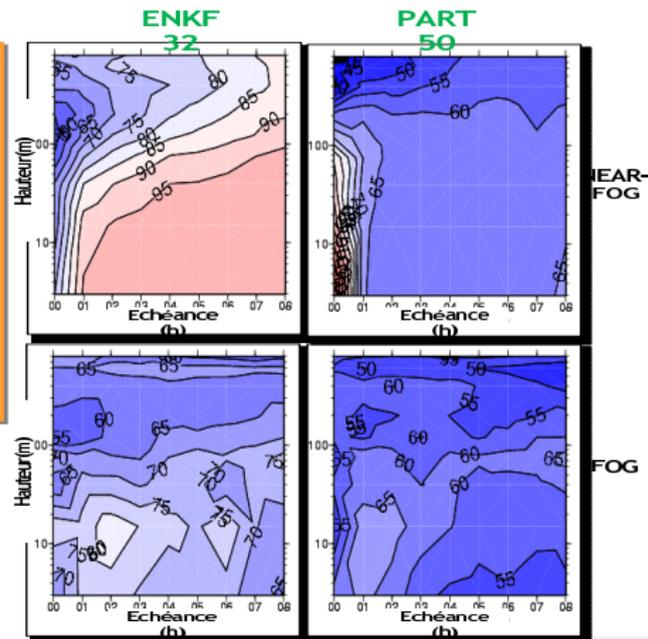
Une application plus météorologique  
Utilisation du filtre trajectorien génétique  
pour l'assimilation d'observations dans COBEL/ISBA  
(travail conjoint avec S. Remy, T. Bergot et O.  
Pannekoucke)

## Résultats avec des observations simulées

### Impact sur l'humidité initiale et prévue

- Comparaison avec l'expérience REF : système d'assimilation opérationnel
- ENKF32 : grande différence entre NEAR-FOG et FOG
- PART50 : moins de différences entre les deux situations
- Prévission de meilleure qualité pour PART50 par rapport à ENKF32

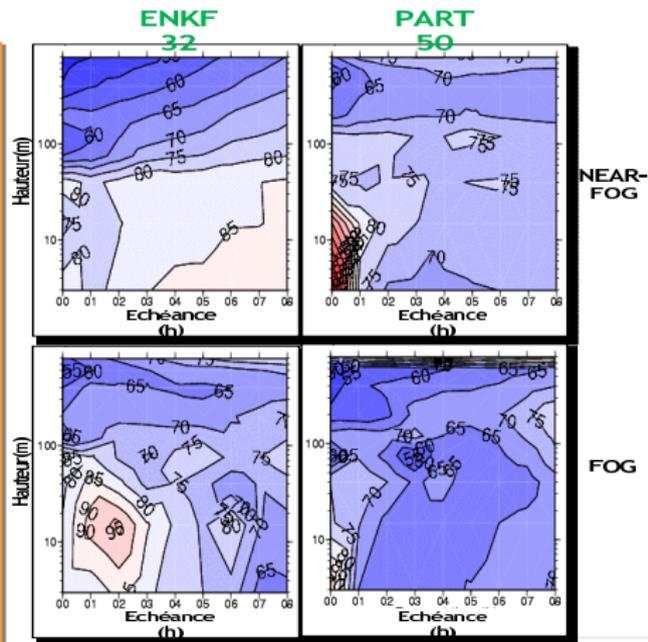
Erreur Quadratique Moyenne de la température, en pourcentage de l'EQM avec REF



## Résultats avec des observations simulées

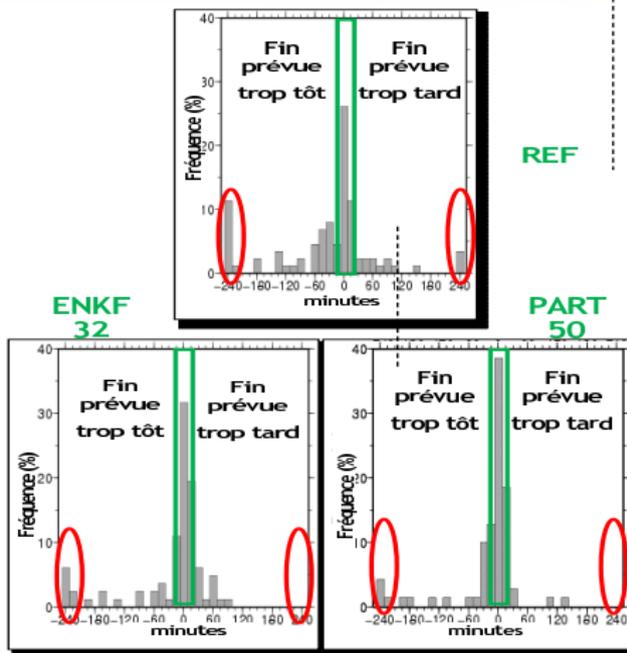
### Impact sur la température initiale et prévu

- Comparaison avec l'expérience REF : système d'assimilation opérationnel
- Filtre de Kalman d'ensemble avec 32 membres : expérience ENKF32
- Filtre particulaire avec 50 particules : expérience PART50
- Amélioration des conditions initiales plus importante pour ENKF32
- Amélioration de la prévision plus importante pour PART50
- Pas d'amélioration pour le filtre particulaire corrigé par EnKF (non montré) par rapport à PART50



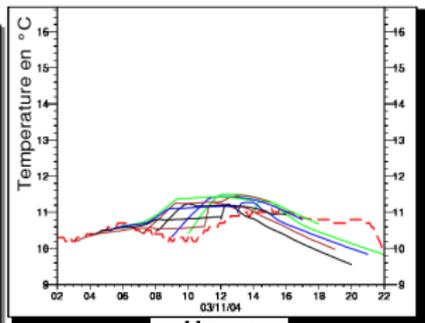
- Statistiques calculées uniquement sur FOG
- Biais quasi nul pour toutes les expériences
- Moins de grandes erreurs (>240mn) avec ENKF32 et PART50
- Erreurs inférieures à 30mn :
  - 40 % des cas pour REF
  - 60% des cas pour ENKF32
  - 70% des cas pour PART50

Distribution des fréquences d'erreur sur l'heure de fin prévue des événements LVP

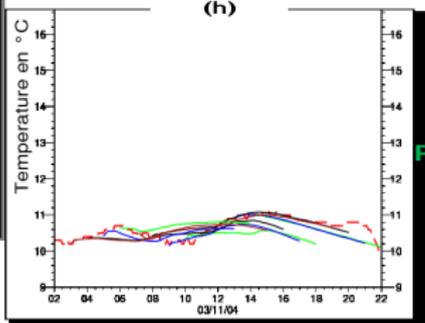


- *Faire fonctionner le filtre à particules avec des observations réelles*
- *En observations réelles, les ébauches ne représentent pas bien l'état de l'atmosphère : mauvais comportement de PART50*
- *On est donc obligé de corriger chaque particule avec un filtre de Kalman d'ensemble, pour la rapprocher des observations*
- *Résultats préliminaires avec un filtre à particules corrigé par un filtre de Kalman d'ensemble : expérience PART50\_COR*

Température à 2m observée (courbe rouge en tireté) et simulées (courbes de couleurs) le 3/11/2004 (de 4 UTC à 14 UTC)



REF

PART50\_C  
OR

**METEO FRANCE**  
Toujours un temps d'avance

## Conclusion

## Bilan et perspectives

- ⊕ Les estimateurs génétiques trajectoriels présentés sont prometteurs. Mais ne sont pas des boites à outils sur étagères.
- ⊕ Le problème de la dimension reste largement ouvert. On en est encore à des tests numériques.
- ⊕ Des algorithmes ont été décrits pour apprendre les erreurs de modèles. Ils semblent fonctionner tant que l'erreur a une dynamique lente.
- ⊕ Le couplage PF/EnKF pose surtout des questions techniques relatives à l'emboîtement.
- ⊕ Il existe des essais de PF locaux, sans aucun cadre théorique satisfaisant.
- ⊕ Une piste serait à explorer sur l'utilisation des filtres ponctuels couplés avec un système d'assimilation de grande échelle.

## Quelques références

- ⊕ C. Baehr. *Nonlinear Filtering for Observations on a Random Vector Field along a Random Path. Application to Atmospheric Turbulent Velocities*, Submitted to M2AN - ESAIM, 2009
- ⊕ C. Baehr et O. Pannekoucke. *Some issues and results on the EnKF and particle filters for meteorological models*, chapter in Chaotic Systems : Theory and Applications. C. H. Skiadas and I. Dimotikalis Ed. World Scientific, 2010.
- ⊕ C. Snyder, T. Bengtsson, P. Bickel, and J. Anderson. *Obstacles to high-dimensional particle filtering*, Mon. Wea. Rev., 136 :4629-4640, 2008.
- ⊕ P. Del Moral. *Feynman-Kac Formulae, Genealogical and Interacting Particle Systems with Applications*. Springer-Verlag, 2004.
- ⊕ M. Lei, P. Del Moral and C. Baehr *Analysis of Approximated PCRLBs for Nonlinear Dynamics Using Different Moments of State Estimate*, Submitted to IEEE transaction on Control and Automation, 2009.