Modélisation probabiliste des écoulements atmosphériques turbulents afin d'en filtrer la mesure par approche particulaire

Séminaire de Probabilités Toulouse

Christophe Baehr

Météo-France / IMT-LSP Univ. Paul Sabatier

christophe.baehr@math.ups-tlse.fr

Mardi 16 Septembre 2008

METEO FRANCE
Toujours un temps d'avance

Que va-t-on voir ?

1. Le problème posé
2. Un premier outil : le processus d'acquistion
3. Le second outil : le filtrage non-linéaire d'un processus à champ moyen
4. Le filtre particulaire pour les fluides turbulents
5. Applications aux observations de fluides réels uni et tridimensionnels
6. Vers de nouveaux travaux

Préambule

Représentation probabiliste d'équation aux dérivées partielles
Soit L l'opérateur elliptique donné par

$$L = \frac{1}{2}\sum_{i,j=1}^{d} a^{ij}(x)\frac{\partial^2}{\partial x^i \partial x^j} + \sum_{i=1}^{d} b^i(x)\frac{\partial}{\partial x^i}$$

Soit le problème de Cauchy défini par l'équation aux dérivées partielles (EDP):

$$\begin{cases} \frac{\partial}{\partial t}u(t,x) = Lu(t,x) + c(x)u(t,x) \\ u(0,x) = u_0(x) \end{cases}$$

Soit maintenant le processus de diffusion défini par l'équation différentielle stochastique (EDS) sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ avec $a = \sigma\sigma^*$

$$\begin{cases} dX_t^{x_0} = b(X_t^{x_0})dt + \sigma(X_t^{x_0})dW_t \\ X_0^{x_0} = x_0 \end{cases}$$

Préambule

$(X_t^{x_0}, \mathbb{P})$ est un processus de Feller-Markov continu. T_t donné pour une fonction $\mathcal{C}^\infty$ bornée par

$$T_t f(x_0) = \mathbb{E}[f(X_t^{x_0}) \ \exp(\int_0^t c(X_s^{x_0}) ds)] = \mathbb{E}_{x_0}[f(X_t) \ \exp(\int_0^t c(X_s) ds)]$$

est un semi-groupe de générateur infinitésimal L. On peut alors montrer la formule de Feynman-Kac

$$u(t,x) = T_t f(x) = \mathbb{E}_x[f(X_t) \ \exp(\int_0^t c(X_s) ds)]$$

Le processus X_t est la représentation probabiliste de la solution au problème de Cauchy posé.

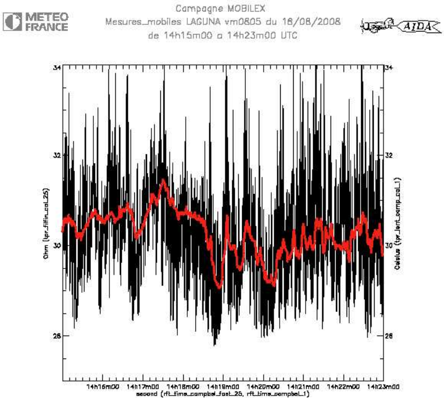
Le problème posé : traitement rapide de la mesure en micrométéorologie

Mesures atmosphériques

- Des mesures sur l'atmosphère pour étudier et pour être assimilées par les modèles de prévisions météorologiques.
- On a besoin de :
  - Améliorer les systèmes de mesures et de numérisation.
  - Mettre en place un traitement des signaux rapides.
  - Estimer les grandeurs caractéristiques de la turbulence.
- En météorologie expérimentale, les plates-formes de mesures sont multiples.

Mesures atmosphériques

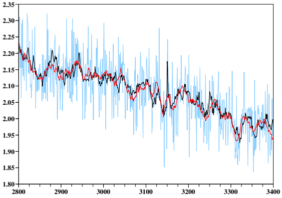
Mesures lentes / Mesures rapides

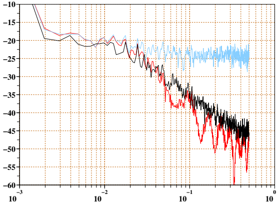


- Exemple de la température relevée le 18 Août 2008 à 14h15 UTC, en rouge à 1 Hz et en noir à 25 Hz.

Traitements linéaires en micro-météorologie

- Exemple de traitement actuel sur une série d'observation de vent bruitées

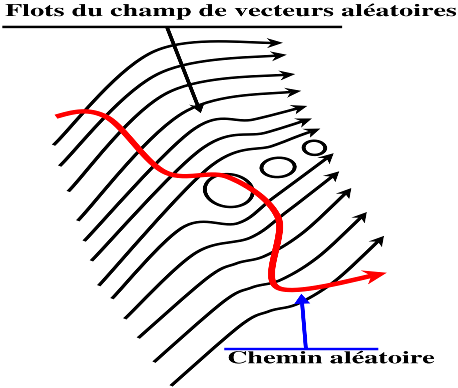




- L'amélioration passe par un traitement non-linéaire de l'observation.

Processus d'acquisition d'un champ de vecteur aléatoire le long d'un chemin aléatoire

Acquisition d'un champ de vecteur



Acquisition d'un champ de vecteur

⊙ Définition: Soient $(d, d') \in \mathbb{N}^* \times \mathbb{N}^*$, $(E \subset \mathbb{R}^d, \mathcal{E})$, espace des configurations ou espace des sites et $(E' \subset \mathbb{R}^{d'}, \mathcal{E}')$, espace des phases. On se donne enfin un espace de probabilité filtré complet $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$.
Soient $T < \infty$ un réel, $t \in [0, T[$, $x \in E$ un point de l'espace des sites, X_t une famille de variables aléatoires sur $(\Omega, \mathcal{F}, \mathcal{F}_t)$ à valeurs dans $(E, \mathcal{E})$ et $X'_{t,x}$ une famille de variables aléatoires sur $(\Omega, \mathcal{F}, \mathcal{F}_t)$ à valeurs dans $(E', \mathcal{E}')$.
Le couple d'applications $\mathcal{F}_t$-mesurables $(X_t, X'_{t,x})$ est appelé système d'acquisition du champ de vecteur aléatoire.
Le processus X_t est appelé le chemin du processus d'acquisition et la famille $X'_{t,x}$ est le champ d'acquisition.

Acquisition d'un champ de vecteur

- Définition: Soit $(X_t, X'_{t,x})$ un système d'acquisition sur un espace l'espace de probabilité $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ à valeurs dans $((E, \mathcal{E}) \times (E', \mathcal{E}'))$. On définit pour tout temps $t \in [0, T]$ le processus d'acquisition A_t sur $(\Omega, \mathcal{F}, \mathcal{F}_t)$ à valeurs dans $(E', \mathcal{E}')$ par

$$A_t \stackrel{def}{=} X'_{t,X_t}$$

- Exemple : Soit $U_{t,x}$ un champ Eulérien donné par exemple une Equation Différentielle Stochastique. Soit $X_t^{x_0}$ le flot de l'EDS partant du point x_0. Alors la quantité $V_t = U_{t,X_t^{x_0}}$ est appelée acquisition Lagrangienne.

Acquisition d'un champ de vecteur

➲ On se donne un système d'acquisition (X_t, A_t) à valeurs $E \times E'$. On suppose que la variable aléatoire (X_t, A_t) possède une densité p^{X_t,A_t} par rapport à la mesure de Lebesgue et est markovienne. On cherche par exemple à calculer la moyenne :

$$\mathbb{E}(f(X_t, A_t)|X_t) = \int_{E\times E'} f(x,a)\ p^{A_t|X_t}(a|z)\ da\ dz$$

Remarque : Si f ne porte que sur l'acquisition il faut régulariser l'espérance :

$$\mathbb{E}(f(A_t)|X_t) = \lim_{\delta\to 0} \mathbb{E}^{\delta}(f(A_t)|X_t)$$

avec K^{δ} un noyau de régularisation faible de la mesure de Dirac et

$$\mathbb{E}^{\delta}(f(A_t)|X_t) = \int_{E\times E'} f(a)\ K^{\delta}(X_t, z)\ p^{A_t|X_t}(a|z)\ da\ dz$$

Acquisition d'un champ de vecteur

- Exemple 1: Soit $(X_t^{x_0}, V_t)$ le système d'acquisition Lagrangien précédent. Avec le noyau Gaussien $K^\delta(X_t, x) = e^{\frac{-\|X_t - x\|^2}{2\delta}}$ et $\mathbb{F}^\delta(f(V_t))(X_t) = \int_{E\times E'} f(v)\ K^\delta(X_t, x)\ p^{X_t, A_t}(x, v)\ dx\ dv$, on a :

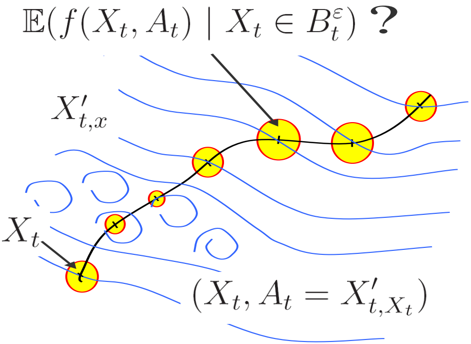
$$\mathbb{E}(f(V_t)|X_t) = \lim_{\delta\to 0} \frac{\mathbb{F}^\delta(f(V_t))(X_t)}{\mathbb{F}^\delta(1)(X_t)}$$

que l'on appellera Espérance Lagrangienne.

- Exemple 2: On prend un système d'acquisition général (X_t, A_t) dans un milieu localement homogène en loi, c'est à dire que pour tout $x \in E$ il existe une boule $B(x, \varepsilon)$ telle que pour tout élément y de $B(x, \varepsilon)$, on a $\mathbb{P}(A_t \in da \mid X_t = x) = \mathbb{P}(A_t \in da \mid X_t = y)$
Alors en se donnant pour chaque temps t une boule B_t^ε, on s'intéressera à l'espérance $\mathbb{E}(f(X_t, A_t)|X_t \in B_t^\varepsilon)$

- Cette nouvelle acquisition est le conditionnement d'un champ à une boule.

Acquisition d'un champ de vecteur



➲ En modélisation sur point de grille c'est la méthode à utiliser pour estimer les lois de probabilité d'un milieu localement homogène en loi sur le chemin stationnaire $X_t = x_{i,j}$, i et j étant les indices du points de grille.

Acquisition d'un champ de vecteurs décrits par ses flots

- Un champ de vecteur peut être décrit en chacun de ses points (Eulérien) ou par ses flots (Lagrangien).
- Dans ce dernier cas, l'acquisition est composée de :
  - un processus d'évolution selon le flot.
  - un processus de saut pour passer d'un flot à l'autre.
- Il faut alors mixer 2 systèmes d'acquisition (écrits en temps discret):
  - l'un étant l'acquisition Lagrangienne : $(X_n^{x_0}, X'_{n,X_n^{x_0}}) = (X_n, X'_n)$
  - l'autre étant l'acquisition du système Lagrangien sur un chemin indépendant du milieu Z_n : (Z_n, X'_n)

Acquisition discrète en milieu homogène

On travaille en temps discret et pour tout instant $n \in \mathbb{N}$, on dispose du chemin d'acquisition $(Z_i)_{0 \leq i \leq n \leq \infty}$ et d'un paramètre $\varepsilon_n > 0$.

- On se donne pour chaque instant n les boules $B_n^{\varepsilon}(Z_n) = \{x \in E \ : \ d(x, Z_n) \leq \varepsilon_n\}$
- On définit les mesures χ_n et $\hat{\chi}_n$ telle que pour toute fonction test f :

$$\begin{aligned} \hat{\chi}_n(f) &= \mathbb{E}(f(X_n, X_n') \mid X_0 \in B_0^{\varepsilon}(Z_0), \ldots, X_n \in B_n^{\varepsilon}(Z_n)) \\ \chi_n(f) &= \mathbb{E}(f(X_n, X_n') \mid X_0 \in B_0^{\varepsilon}(Z_0), \ldots, X_{n-1} \in B_{n-1}^{\varepsilon}(Z_{n-1})) \end{aligned}$$

- (X_n, X_n') est supposé Markovien de noyau de transition $M_{n+1,\pi_n}((x,x'), d(z,z')) = \mathbb{P}((X_{n+1}, X_{n+1}') \in d(z,z') | (X_n, X_n') = (x,x'))$ où π_n est la loi de (X_n, X_n').

Acquisition discrète en milieu homogène

- $\hat{\chi}_n(f)$ s'écrit simplement

$$\hat{\chi}_n(f) = \frac{\mathbb{E}(f(X_n, X'_n)\ \prod_{p=0}^{n} \mathbb{1}_{B^{\varepsilon}_{\boldsymbol{p}}(Z_{\boldsymbol{p}})}(X_p))}{\mathbb{E}(\prod_{p=0}^{n} \mathbb{1}_{B^{\varepsilon}_{\boldsymbol{p}}(Z_{\boldsymbol{p}})}(X_p))}$$

- ce qui permet de donner un noyau de transition et une fonction potentiel qui vont permettre de définir des mesures de Feynman-Kac :
  - $M_{n+1,\pi_{\boldsymbol{n}}}((x,x'), d(z,z'))$ le noyau Markovien.
  - pour tout $p > 0$, $G_p((X_p, X'_p)) = \mathbb{1}_{B^{\varepsilon}_{\boldsymbol{p}}(Z_{\boldsymbol{p}})\times E'}((X_p, X'_p))$
- Associé à ce potentiel on définit le noyau de sélection :

$$\begin{aligned} & S^{Z}_{n+1,\chi_{\boldsymbol{n+1}}}((x,x'), d(y,y')) \\ = \quad & \mathbb{1}_{B^{\varepsilon}_{\boldsymbol{n+1}}(Z_{\boldsymbol{n+1}})\times E'}(x,x')\delta_{(x,x')}(d(y,y')) \\ + \quad & \left[1 - \mathbb{1}_{B^{\varepsilon}_{\boldsymbol{n+1}}(Z_{\boldsymbol{n+1}})\times E'}(x,x')\right] \frac{\mathbb{1}_{B^{\varepsilon}_{\boldsymbol{n+1}}(Z_{\boldsymbol{n+1}})\times E'}(y,y')\chi_{n+1}(d(y,y'))}{\chi_{n+1}(B^{\varepsilon}_{n+1}(Z_{n+1})\times E')} \end{aligned}$$

Acquisition discrète en milieu homogène

On a besoin de rendre les transitions Markoviennes, et on obtient les flots de Feynman-Kac du processus d'acquisition. Pour se faire, on utilise une normalisation:

$$\tilde{M}_{n+1,\pi_{\boldsymbol{n}}}((x,x'),d(y,y')) = \frac{M_{n+1,\pi_{\boldsymbol{n}}}((x,x'),d(y,y'))\,\mathbb{1}_{B^{\varepsilon}_{\boldsymbol{n+1}}(Z_{\boldsymbol{n+1}})\times E'}(y,y')}{M_{n+1,\pi_{\boldsymbol{n}}}(\mathbb{1}_{B^{\varepsilon}_{\boldsymbol{n+1}}(Z_{\boldsymbol{n+1}})\times E'})(x,x')}$$

et

$$\tilde{G}_{n+1}(x,x') = M_{n+1,\pi_{\boldsymbol{n}}}(\mathbb{1}_{B^{\varepsilon}_{\boldsymbol{n+1}}(Z_{\boldsymbol{n+1}})\times E'})(x,x')$$

- $\tilde{M}_{n+1,\pi_{\boldsymbol{n}}}((x,x'),d(y,y'))$ est la transition locale restreinte au domaine $B^{\varepsilon}_{n+1}(Z_{n+1})$ et $\tilde{G}_{n+1}(x,x')$ donne les chances que l'acquisition Lagrangienne suive le chemin $Z_0,\ldots,Z_{n+1}$.
- On note χ^B_n et $\hat{\chi}^B_n$ les lois du processus restreint aux cylindres $B^{\varepsilon}_{n+1}(Z_{n+1})\times E'$.
- Ce sont des processus à champ moyen en interaction solutions de l'équation non-linéaire $\chi^B_{n+1} = \chi^B_n K^B_{n+1,\chi^{\boldsymbol{B}}_{\boldsymbol{n}},\pi_{\boldsymbol{n}}}$

Acquisition discrète en milieu homogène

⊕ On obtient le schéma des évolutions restreintes au chemin d'acquisition:

$$\hat{\chi}_n^B \xrightarrow{\tilde{M}_{\boldsymbol{n+1},\pi_{\boldsymbol{n}}}} \chi_{n+1}^B \xrightarrow{\tilde{S}_{\boldsymbol{n+1},\chi^{\boldsymbol{B}}_{\boldsymbol{n+1}}}} \hat{\chi}_{n+1}^B$$

et pour les états :

$$\underbrace{(X_n^B, X_n'^B)}_{\in\ (B_{\boldsymbol{n}}^{\varepsilon}(Z_{\boldsymbol{n}})\times E')} \xrightarrow{\text{Sélection}} \underbrace{(\hat{X}_n^B, \hat{X}_n'^B)}_{\in\ (B_{\boldsymbol{n}}^{\varepsilon}(Z_{\boldsymbol{n}})\times E')} \xrightarrow{\text{Mutation restreinte}} \underbrace{(X_{n+1}^B, X_{n+1}'^B)}_{\in\ (B_{\boldsymbol{n+1}}^{\varepsilon}(Z_{\boldsymbol{n+1}})\times E')}$$

⊕ Ces intégrales n'ont de solutions qu'approchées que l'on estime par un algorithme particulaire.

Approximation particulaire d'une acquisition discrète

- Initialisation : $N > 0$ particules $(\xi_0^i, \xi_0'^i)_{1 \le i \le N} \in B_0^\varepsilon(Z_0) \times E'$ *i.i.d.* selon χ_0^B et on note $\chi_0^{B,N} = \frac{1}{N}\sum_{i=1}^N \delta_{\xi_0^i, \xi_0'^i}$
- Au pas de temps n, $(\hat{\xi}_n^i, \hat{\xi}_n'^i)_{1 \le i \le N}$ dans $B_n^\varepsilon(Z_n) \times E'$ évoluent selon

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

- Avec une hypothèse de régularité classique on peut obtenir

<u>Théorème</u> Pour tout $n \ge 0$, pour tout $p \ge 1$, il existe des constantes C_n et C_n' telles que

$$\mathbb{E}(\|\chi_n^{B,N} - \chi_n^B\|_{\mathcal{H}}^p)^{\frac{1}{p}} \le [\frac{C_n}{\sqrt{N}} + \frac{C_n'}{\sqrt{d}}] I(\mathcal{H})$$

avec la seminorme pour 2 mesures η et η' définie par
$\|\eta - \eta'\|_{\mathcal{H}} = \sup\{|\eta(h) - \eta'(h)| \text{ tel que } h \in \mathcal{H}\}$

- La difficulté vient de l'annulation possible du potentiel $G_n(x, x')$.

Filtrage non-linéaire de processus à champ moyen

Filtrage d'un processus à champs moyen à temps discret

- Nous allons décrire le filtrage à temps discret des observations bruitées de processus à champ moyen pour diverses lois du champ moyen.
- On dispose d'états $X_0, \dots, X_n$ et d'observations $Y_0 = y_0, \dots, Y_n = y_n$ avec pour tout $n \geq 0$, $Y_n = H(X_n, V_n)$.
- Le problème de filtrage non-linéaire revient à chercher les 2 mesures:

$$\begin{aligned} \eta_n &= Loi(X_n | Y_0 \dots Y_{n-1}) \\ \hat{\eta}_n &= Loi(X_n | Y_0 \dots Y_n) \end{aligned}$$

- On suppose que Y_n et $h(X_n) + \sigma_n^Y . W_n^Y$ sont absolument continues de densité g_n et on définit la fonction potentiel $G_n(X_n) = g_n(X_n, Y_n)$
- On suppose également que les transitions de X_n sont données par $M_{n+1,\pi_n}(x_n, dx_{n+1}) = \mathbb{P}(X_{n+1} \in dx_{n+1} | X_n = x_n)$

Filtrage d'un processus à champs moyen à temps discret

- Le système lié au problème de filtrage peut alors s'écrire

$$\left\{\begin{array}{rcl} X_{n+1} & = & X_n + b(X_n, \pi_n)\Delta t + \sigma_n^X.\Delta W_n^X \\ Y_n & = & h(X_n) + \sigma_n^Y.W_n^Y \\ X_0 & \sim & \eta_0 \\ Y_0 & = & y_0 \end{array}\right.$$

où $h : X_n \rightarrow Y_n$ est une fonction bornée, W_n^Y est un $\mathcal{F}_n$-mouvement Brownien, et σ_n^Y une fonction positive et Δt le pas de temps.

- Pour le noyau b il faut se donner une forme et quelques hypothèses et on prendra

$$b(X_n, \pi_n) = A(X_n) + \int B(X_n, x)\pi_n(dx)$$

où A est bornée, la fonction B est supposée bornée en sa première variable et Lipschitzienne de paramètre L^B en sa seconde variable.

Filtrage d'un processus à champs moyen à temps discret

- On cherche pour toute fonction mesurable bornée f, les mesures de Feynman-Kac, $\eta_{\boldsymbol{n}}$ et $\hat{\eta}_{\boldsymbol{n}}$.

- On définit le noyau de sélection

$$S_{n,\eta_{\boldsymbol{n}}}(x, dy) = G_n(x)\delta_x(dy) + (1 - G_n(x))\Psi_n(\eta_n)(dy)$$

avec la loi de redistribution de Boltzmann-Gibbs

$$\Psi_n(\eta_n)(dy) = \frac{G_n(y).\eta_n(dy)}{\eta_n(G_n)}$$

- Il existe une représentation de McKean de ces mesures sous forme d'un noyau de transport, les mesures sont alors solutions du système dynamique : $\eta_{n+1}(f) = \eta_n K_{n+1,\eta_{\boldsymbol{n}},\pi_{\boldsymbol{n}}}(f) = \eta_n \; S_{n,\eta_{\boldsymbol{n}}} \; M_{n+1,\pi_{\boldsymbol{n}}}(f)$

Filtrage d'un processus à champs moyen à temps discret

- Ce système non-linéaire doit être résolu par approximation particulaire. Dans notre cas il faut représenter 2 types de loi:
  - les 2 Feynman-Kac du filtrage η_n et $\hat{\eta}_n$.
  - la loi de champ moyen π_n.
- Si π_n est une loi a priori, il faut se doter de 2 systèmes de particules, l'un apprenant les Feynman-Kac, l'autre le champ moyen.
- Le système auxiliaire ne verra pas les observations, l'évolution est a priori et il n'y a pas d'étape de sélection.

Filtrage d'un processus à champs moyen à temps discret

- Il faut approcher π_n par une loi empirique π_n^d.

- Pour tout pas de temps $n \geq 0$ les particules Z_n^i évoluent selon $M_{n+1,\pi_n^d}(Z_n^i, .)$ avec $\pi_n^d = \frac{1}{d}\sum_{j=1}^{d} \delta_{Z_n^j}$

- Le système de particules évolue selon la dynamique : $Z_{n+1}^{j,d} = Z_n^{j,d} + b(Z_n^{j,d}, \pi_n^d)\Delta t + \sigma_n^Z \Delta W_n^{Z,j,d}$

- On est dans un cas classique d'approximation de champ moyen, on peut disposer des résultats de convergences asymptotiques.

Filtrage d'un processus à champs moyen à temps discret

- Maintenant on peut faire du filtrage avec un système de particules $(X_n^i)_{i=1}^N$ et on note $\eta_n^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_n^i}$
- Les particules évoluent selon $X_{n+1}^{i,N} = X_n^{i,N} + b(X_n^{i,N}, \pi_n^d)\Delta t + \sigma_n^X \Delta W_n^{X,i,N}$
- avec $\eta_{n+1}^N(f) \sim \eta_n^N \; S_{n,\eta_n^N} \; M_{n,\pi_n^d}(f)$.
- On va chercher à contrôler l'écart $|\eta_{n+1}^N(f) - \eta_{n+1}(f)|$

Filtrage d'un processus à champs moyen à temps discret

- Théorème Avec les hypothèses générales, pour tout $n \leq 0$ et tout $p \leq 1$, il existe des constantes finies $C_n(p) > 0$ et $C'_n(p) > 0$ telles que

$$\mathbb{E}(\|\eta_n^N - \eta_n\|_{\mathcal{H}}^p)^{\frac{1}{p}} \leq [\frac{C_n(p)}{\sqrt{N}} + \frac{C'_n(p)}{\sqrt{d}}] I(\mathcal{H})$$

- Idées de la preuve qui s'effectue par induction. Au rang initial,

$$\mathbb{E}(\|\eta_0^N - \eta_0\|_{\mathcal{H}}^p)^{\frac{1}{p}} \leq \frac{c_p(0)}{\sqrt{N}} I(\mathcal{H})$$

- On suppose alors qu'elle est réalisée au rang n. On utilise la décomposition :

$$\begin{array}{rcll} \eta_{n+1}^N(f) - \eta_{n+1}(f) & = & \eta_{n+1}^N(f) - \eta_n^N K_{n+1,\eta_n^N,\pi_n^d}(f) & (*) \\ & + & \eta_n^N K_{n+1,\eta_n^N,\pi_n^d}(f) - \eta_n^N K_{n+1,\eta_n,\pi_n}(f) & (**) \\ & + & \eta_n^N K_{n+1,\eta_n,\pi_n}(f) - \eta_n K_{n+1,\eta_n,\pi_n}(f) & (***) \end{array}$$

Filtrage d'un processus à champs moyen à temps discret

- Le terme $(*)$ est un calcul sur les fluctuations

$\mathbb{E}((*)^2 \mid \eta_n^N) \leq \frac{1}{N}\|f\|^2$

$$\mathbb{E}[(*)^2 \mid \eta_n^N] = \frac{1}{N^2}\sum_{i=1}^{N} K_{n+1,\eta_n^N,\vartheta_n}([f - K_{n+1,\eta_n^N,\vartheta_n}(f)(\xi_n^i)]^2)(\xi_n^i)$$

via l'inégalité de Marcinkiewicz-Zygmund on obtient

$$\mathbb{E}(|\eta_{n+1}^N(f) - \eta_n^N K_{n+1,\eta_n^N,\vartheta_n}(f)|^p \;/\; \eta_n^N)^{\frac{1}{p}} \leq \frac{C_p}{\sqrt{N}}\|f\|$$

Filtrage d'un processus à champs moyen à temps discret

➲ Le membre $(***)$ est décomposé en faisant apparaitre des écarts sur certaines fonctions

$$
\begin{aligned}
& \eta_n^N K_{n+1,\eta_n^N,\vartheta_n}(f) - \eta_n K_{n+1,\eta_n,\vartheta_n}(f) \\
= \; & \eta_n^N K_{n+1,\eta_n^N,\vartheta_n}(f) - \eta_n^N K_{n+1,\eta_n,\vartheta_n}(f) \\
+ \; & \eta_n^N K_{n+1,\eta_n,\vartheta_n}(f) - \eta_n K_{n+1,\eta_n,\vartheta_n}(f)
\end{aligned}
$$

On utilise alors des lemmes techniques classiques sur la convergence des processus empiriques qui amène l'hypothèse de récurrence et donne une inégalité en $\frac{C_n(p)}{\sqrt{N}}$

Filtrage d'un processus à champs moyen à temps discret

- Le dernier terme $(**)$ fait travailler le modèle Markovien et va faire apparaître l'approximation de champ moyen du système de particules auxilliaire.

- La différence des noyaux de McKean contre une fonction test f mesurable bornée s'écrit:

$$\eta_n^N K_{n+1,\eta_n^N,\pi_n^d}(f) - \eta_n^N K_{n+1,\eta_n^N,\pi_n}(f)$$
$$= \int f(y)\eta_n^N(dx) S_{n,\eta_n^N}(x,dz)[M_{n+1,\pi_n^d} - M_{n+1,\pi_n}](z,dy)$$

- Ce sont alors les hypothèses sur le noyau b (Lipschitz + borné) qui amènent au travers de majoration de Gaussienne d'obtenir l'existence de constante et une majoration par $\frac{C_n'(p)}{\sqrt{d}}$ ce qui termine la preuve

Filtrage d'un processus à champs moyen à temps discret

- On peut modifier la loi de champ moyen pour la conditionner aux observations.
- C'est à dire que l'on peut considérer:

$X_{n+1} = X_n + b(X_n, \eta_n)\Delta t + \sigma_n^X \Delta W_n^X$

ou

$X_{n+1} = X_n + b(X_n, \hat{\eta}_n)\Delta t + \sigma_n^X \Delta W_n^X$

- Dans ce cas, un seul système de particules est nécessaire.
- Les estimations d'erreur utilisent les mêmes décompositions que précédemment et se font par récurrence.

Filtrage d'un processus à champs moyen à temps discret

- Cas d'un modèle possèdant une commande aléatoire X_n^1 pour un vecteur d'état X_n^2 d'évolution $X_{n+1}^2 = X_n^2 + b(X_n^1, X_n^2, \theta_n)\Delta t + \sigma_n^X \Delta W_n^X$ où θ_n est la $Loi(X_n^2 \mid X_0^1 \dots X_n^1)$.

- Si le problème de filtrage fournit une équation Markovienne pour X_n^1 il faut faire une augmentation d'état et régulariser. On a alors un algorithme déjà décrit et des estimées controlées.

- Sinon on paie le manque d'information en devant échantillonner la loi de X_n^1 et pour chaque tirage de $(X_n^{1,i})_{1 \leq i \leq M}$ il faut construire un nouveau système de particules qui nous ramène au cas a priori.

- On peut retrouver par ce procédé un algorithme de filtrage et des estimées comparables, mais c'est trés lourd avec la nécessité de générer $M + 1$ systèmes de particules.

Filtrage d'un processus à champs moyen à temps discret

Lois de champ moyen	Nombre de systèmes de particules	Nombres de particules	Majoration de l'Erreur d'Approximation
$Loi(X_n)$	2	$N+d$	$\frac{1}{\sqrt{N}}+\frac{1}{\sqrt{d}}$
$Loi(X_n \mid Y_0 \ldots Y_{n-1})$	1	N	$\frac{1}{\sqrt{N}}$
$Loi(X_n \mid Y_0 \ldots Y_n)$	1	N	$\frac{1}{\sqrt{N}}$
$Loi(X_n^2 \mid X_0^1 \ldots X_n^1)$ avec modèle markovien pour X_n^1	2	$N+d$	$\frac{1}{\sqrt{N}}+\frac{1}{\sqrt{d}}$
$Loi(X_n^2 \mid X_0^1 \ldots X_n^1)$ sans modèle markovien	$M+1$	$M\times d+N$	$\frac{1}{\sqrt{N}}+\frac{1}{\sqrt{d}}+\frac{1}{\sqrt{M}}$

Filtrage particulaire pour les mesures sur un fluide turbulent

Filtrage d'un fluide turbulent

- On va utiliser les outils que l'on a développé avec un modèle Lagrangien stochastique pour la turbulence homogène isotrope.

- Le modèle de Pope à temps continu s'écrit

$$\left\{ \begin{array}{lcl} dX_t & = & V_t \; dt \\ dV_t & = & -\nabla_x < p > dt + (\frac{1}{2} + \frac{3}{4} C_0) \frac{\varepsilon_t}{k_t} [V_t - < v >] dt + \sqrt{C_0 . \varepsilon_t} dW_t \end{array} \right.$$

- On va supposer l'existence des solutions et on discrétise par un schéma d'Euler explicite.

- $< v >$ est remplacé par $\mathbb{E}^{\delta}(V_t | X_t = x, X_0) \stackrel{def}{=} \Pi_t^{\delta}(V_t)(X_t)$

Filtrage d'un fluide turbulent

- Dans le modèle de Pope $-\nabla_x < p > dt$ et ε_t sont 2 commandes externes.

- Pour ne pas pénaliser le filtre en terme de coût de calcul, nous allons leur donner des modélisations qui vont fermer le système par l'observation.

- Par une remarque sur la structure du modèle, on va considérer que $-\nabla_x < p > dt$ s'approche par l'espérance des incréments de vitesse : $\mathbb{E}(dV_t) \stackrel{def}{=} \mathbb{E}(\alpha_t)$.

- De la même façon on approche le taux de dissipation turbulente par $\mathbb{E}(dV_t.dV_t) \stackrel{def}{=} \mathbb{E}(\varepsilon_t)$.

Filtrage d'un fluide turbulent

- En temps discret le système dynamique pour la turbulence s'écrit:

$$\begin{cases} X_{n+1} = X_n + V_n\,\Delta t + \sigma_n^X \Delta W_n^X \\ V_{n+1} = V_n + \mathbb{E}(\alpha_n) + C_1 \frac{\mathbb{E}(\varepsilon_n)}{k_n}[V_n - \Pi_n^\delta(V_n)(X_n)]\Delta t \\ \qquad + \sqrt{C_0.\mathbb{E}(\varepsilon_n)}\Delta W_n^V \end{cases}$$

avec $k_n = \frac{1}{2}\Pi_n^\delta[(V_n - \Pi_n^\delta(V_n)(X_n))^2](X_n)$

- Par conditionnement nous allons coupler ce système à l'observation et le localiser à son chemin d'acquisition.

- C'est un modèle à champ moyen conditionnel, le problème de filtrage associé, retrouver $Loi(V_n|Y_0, \ldots, Y_n)$, est résolu par un système de particules unique.

- Pour la partie pratique, nous allons utiliser des estimations trajectorielles, la convergence se faisant sur les lignes ancestrales.

Algorithme pour un fluide turbulent

- Initialisation : on distribue N particules selon η_0 pour le couple (X_0, V_0).

- A l'étape n on dispose de N couples (X_n^i, V_n^i), $1 \leq i \leq N$ situés dans la boule B_n approchant empiriquement la loi $\eta_n^{B_n}$.

- L'étape de sélection $(X_n^i, V_n^i) \xrightarrow{\text{Sélection}} (\hat{X}_n^i, \hat{V}_n^i)$ utilise le noyau $G_n(X_n^i, V_n^i) = e^{-\frac{1}{2\sigma_n^Y}(Y_n - H(V_n^i))^2}$ avec une sélection trajectorielle de type génétique.
- Les particules sélectionnées représentent au mieux le milieu, on calcule les paramètres locaux déduits des moyennes sur les incréments :

$$\hat{A}_n^N = \frac{\Delta t}{N} \sum_{i=1}^{N} [\hat{V}_n^i - \hat{V}_{n-1}^i]$$

$$\hat{E}_n^N = \frac{\Delta t}{N.C_0} \sum_{i=1}^{N} [\hat{V}_n^i - \hat{V}_{n-1}^i].[\hat{V}_n^i - \hat{V}_{n-1}^i]$$

Algorithme pour un fluide turbulent

- Pour chaque particule on calcule:

$$\Pi_n^{\delta,i,N}(\hat{V}_n^i)(\hat{X}_n^i) = \frac{\sum_{j=1}^N \hat{V}_n^j G^\delta(\hat{X}_n^i - \hat{X}_n^j)}{\sum_{k=1}^N G^\delta(\hat{X}_n^i - \hat{X}_n^k)}$$

et

$$\hat{k}_n^i = \frac{1}{2} \frac{\sum_{j=1}^N [\hat{V}_n^j - \Pi_n^{\delta,N}(\hat{V}_n^i)(\hat{X}_n^i)]^2 G^\delta(\hat{X}_n^i - \hat{X}_n^j)}{\sum_{k=1}^N G^\delta(\hat{X}_n^i - \hat{X}_n^k)}$$

- L'étape de mutation utilise le modèle de signal fluide en champ moyen conditionné aux observations et à l'acquisition,

$$(\hat{X}_n^i, \hat{V}_n^i) \xrightarrow{\text{selon } M_{\boldsymbol{n+1},\hat{\eta}_{\boldsymbol{n}}^{\boldsymbol{B_n},\boldsymbol{N}}}} (\tilde{X}_{n+1}^i, \tilde{V}_{n+1}^i) :$$

$$\left\{ \begin{array}{lcl} \tilde{X}_{n+1}^i & = & \hat{X}_n^i + \hat{V}_n^i \Delta t + \sigma_n^X \Delta W_n^{X,i} \\ \tilde{V}_{n+1}^i & = & \hat{V}_n^i + \hat{A}_n^N - C_1 \frac{\hat{E}_{\boldsymbol{n}}^{\boldsymbol{N}}}{\hat{k}_{\boldsymbol{n}}^{\boldsymbol{i}}} \, [\hat{V}_n^i - \Pi_n^{\delta,i,N}(\hat{V}_n^i)(\hat{X}_n^i)] \, \Delta t \\ & + & \sqrt{C_0 \hat{E}_n^N} \, \Delta W_n^{V,i} \end{array} \right.$$

Algorithme pour un fluide turbulent

- On termine en localisant autour de la position suivante sur le chemin d'acquisition.

- Cette localisation est une sélection où l'on garde les particules déjà présentes dans B_{n+1}, de rayon R_{n+1} et on redistribue selon $\hat{\eta}_{n+1}^{B_n,N}$,

$\tilde{\mathcal{X}}_{n+1}^i = (\tilde{X}_{n+1}^i, \tilde{V}_{n+1}^i) \xrightarrow{\text{selon } \mathcal{S}^{B_{n+1}}_{n+1,\tilde{\eta}^{B_n,N}_{n+1}}} \mathcal{X}_{n+1}^i = (X_{n+1}^i, V_{n+1}^i)$, et l'ensemble des particules (X_{n+1}^i, V_{n+1}^i) se trouvent dans B_{n+1}.

Filtrage d'un fluide turbulent atmosphérique

- Pour traiter dans l'atmosphère la turbulence en dimension 3 il faut changer de modèle.

- On s'est inspiré du modèle de dispersion pour la turbulence stratifiée proposé par Das & Durbin.

$$\left\{ \begin{array}{rcl} d\tilde{v}_t^h & = & -\frac{C_1}{2}\frac{\varepsilon_t}{k_t}\tilde{v}_t^h\ dt + (C_{2\theta}-1)\ \tilde{w}_t\ \frac{d<V^h>}{dz}\ dt + (C_0\varepsilon_t)^{\frac{1}{2}}\ dB_t^h \\ d\tilde{w} & = & -\frac{C_1}{2}\frac{\varepsilon_t}{k_t}\tilde{w}_t\ dt + (1-C_{5\theta})\ \beta_t\ g\ \tilde{\theta}_t\ dt + (C_0\varepsilon_t)^{\frac{1}{2}} dB_t^W \\ d\tilde{\theta}_t & = & -\left(C_{1\theta}-\frac{C_1}{2}\right)\frac{\varepsilon_t}{k_t}\tilde{\theta}_t\ dt - \tilde{w}_t\ \frac{d<\Theta>}{dz}\ dt + (C_\theta)^{\frac{1}{2}} dB_t^\theta \end{array} \right.$$

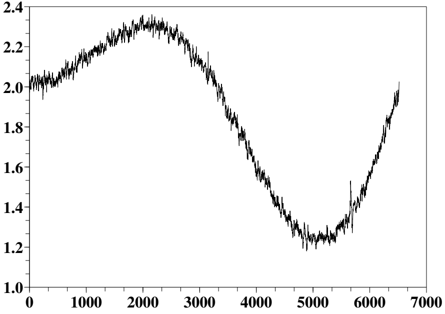
où $< V^h >=< U^h >_{t,X_t^{x_0}}$, $< \Theta >=< \theta >_{t,X_t^{x_0}}$. Les constantes ont été ajustées expérimentalement.

- Avec les mêmes méthodes, on a obtenu un algorithme pour filtrer les observations de la turbulence en atmosphère sèche.

Applications au filtrage des observations de fluides turbulents atmosphériques

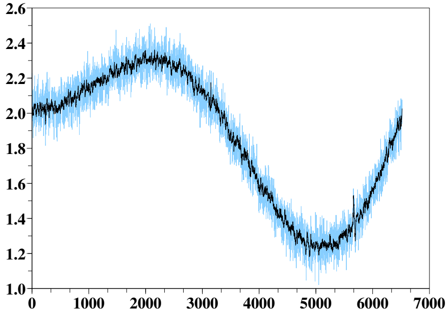
Filtrage d'un vent 1D simulé

- Avec le matériel dont on dispose, on est en mesure de simuler un vent 1D en choisissant des commandes.



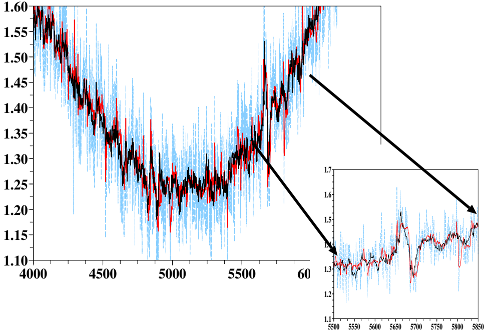
Filtrage d'un vent 1D simulé

➲ On bruite notre signal de référence par un bruit, ici blanc et gaussien.



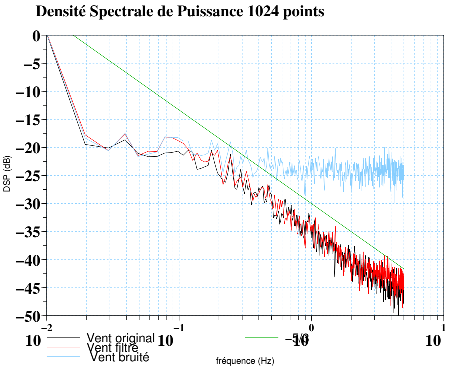
Filtrage d'un vent 1D simulé

- On débruite avec 300 particules le signal perturbé.



Filtrage d'un vent 1D simulé

- On complète l'analyse en examinant les spectres de puissance.



Filtrage d'un vent 1D simulé

- Par cet exemple on vient de tester l'algorithme.

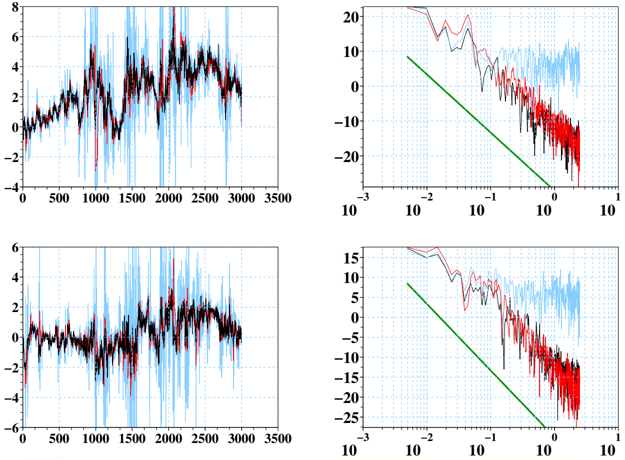
- On peut alors passer directement à des filtrages de mesures de vent 1D ou 2D.

- Il est possible de tester différents algorithmes (a priori vs conditionnel) ou de proposer des bruits corrélés avec la turbulence.

- Nous avons également mené des tests de pertinence du modèle et d'intérêt du champ moyen.

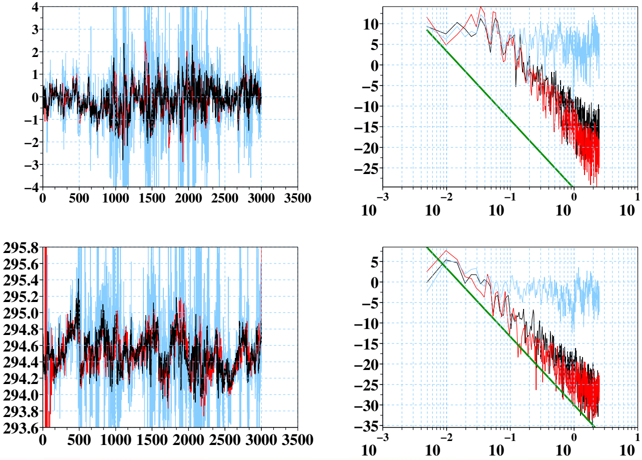
Filtrage d'un vent 3D réel

- Séries temporelles et DSP des composantes U et V de référence, bruitées et filtrés. Bruit corrélé à la variance locale du signal, estimateur à 800 particules.



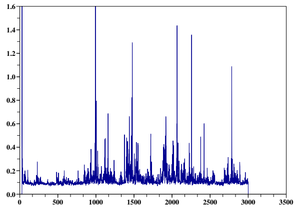
Filtrage d'un vent 3D réel

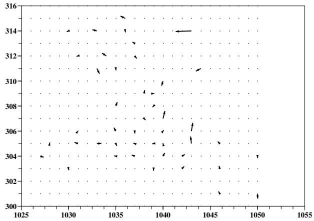
- Séries temporelles et DSP des composantes W et T de référence, bruitées et filtrés. Bruit corrélé à la variance locale du signal, estimateur à 800 particules.



Filtrage d'un vent 3D réel

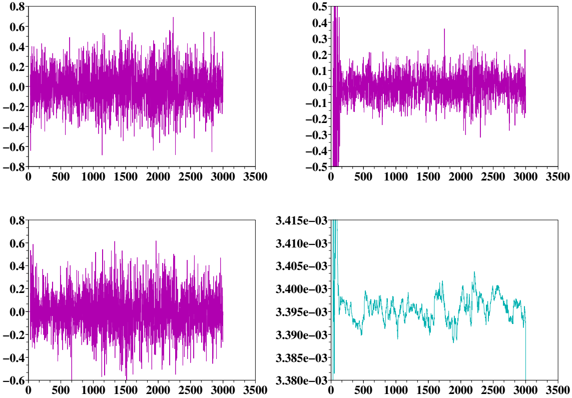
- Séries temporelles des paramètres du milieu turbulent, estimateur à 800 particules.





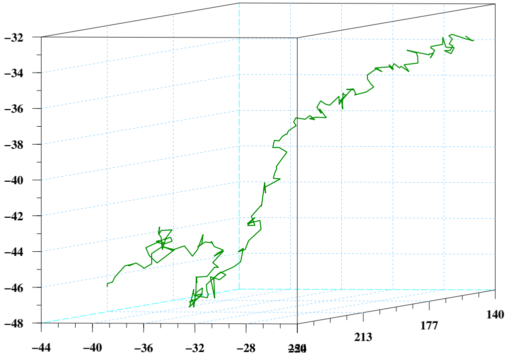
Filtrage d'un vent 3D réel

- Séries temporelles des paramètres du milieu turbulent, estimateur à 800 particules.



Filtrage d'un vent 3D réel

- Pseudo-trajectoires des 800 particules.

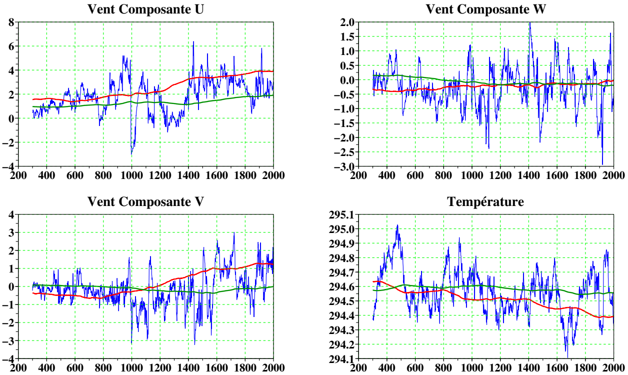


Quelques Perspectives ....

- Adaptation du modèle géophysique à la turbulence atmosphérique.
- Travailler sur l'estimation par cette technique de paramètres turbulents compatibles avec les modèles météos.
- Le problème de filtrage est formulé pour des mesures pouvant être mobiles. Il faut vérifier l'acuité de la solution sur données simulées et réelles.
- Il y a des réflexions à mener sur l'étape de sélection du processus d'acquisition. Y-a-t-il des solutions plus douces que le simple conditionnement à une boule ?
- Concernant la remarque sur le flot Lagrangien, il semble que l'on puisse faire une description de la dynamique du fluide par un processus de branchement. Il faut compléter cette approche et faire le lien avec des grandeurs physique, par exemple, avec les fréquences locales de la turbulence, l'échelle de Taylor, le coefficient de Hausdorff de l'écoulement, etc...

Application du Filtrage Particulaire

Cas d'un vent 3D réel bruité artificiellement



w'T' = 0.0125464

w'T' méthode classique = 0.0124804

tau ~ 0.2204135

tau méthode classique ~ 0.2105395

Calculs effectués sur l'ensemble de la période pour estimer :
$\mathbb{E}(w'T') \sim < w'T' >$ et $\tau \sim \sqrt{< u'w' >^2 + < v'w' >^2}$