

Modèles à espace d'état en météorologie et en océanographie

Valerie Monbet

Université de Bretagne Sud

14 mai 2009

Plan

- 1 Introduction
 - Généralités
 - Exemples
 - Filtrage
- 2 Modèle à espace d'état
 - Définition
 - Inférence dans les modèles à espace d'état
 - Filtre de Kalman
- 3 Inférence pour les modèles non linéaires
 - Filtre de Kalman d'Ensemble
 - Filtre particulière
 - Assimilation variationnelle
- 4 Applications
 - Exemple 1 : fusion de données
 - Exemple 2 : estimation d'une vitesse latente
 - Exemple 3 : assimilation de données
- 5 Conclusion

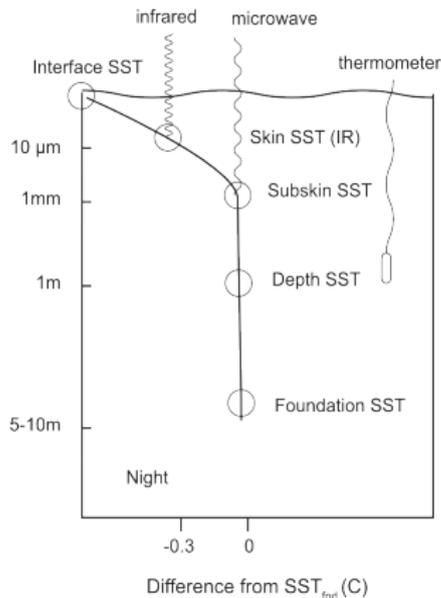
Introduction

- Modèles à chaîne de Markov cachée
 - Variable cachée : chaîne de Markov, à valeurs discrètes
 - Modélisation de régimes dans les séries temporelles
- Modèles à espace d'état
 - Variable cachée : processus markovien, à valeurs continues
 - Combinaison d'information : fusion de données, assimilation de données, extraction de signal sur bruit
 - Modélisation d'un phénomène latent (vitesse de déplacement)

Plan

- 1 Introduction
 - Généralités
 - **Exemples**
 - Filtrage
- 2 Modèle à espace d'état
 - Définition
 - Inférence dans les modèles à espace d'état
 - Filtre de Kalman
- 3 Inférence pour les modèles non linéaires
 - Filtre de Kalman d'Ensemble
 - Filtre particulière
 - Assimilation variationnelle
- 4 Applications
 - Exemple 1 : fusion de données
 - Exemple 2 : estimation d'une vitesse latente
 - Exemple 3 : assimilation de données
- 5 Conclusion

Exemple 1 : fusion de données Température de surface de la mer (Sea State Temperature : SST)



La SST est utilisée pour

- l'étude du réchauffement climatique
- le recalage des modèles de courant
- la biologie

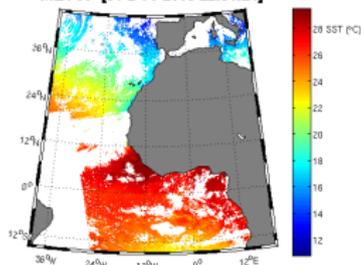
FIG.: Définition de la température de surface

Exemple 1 : fusion de données

Mesure de la SST

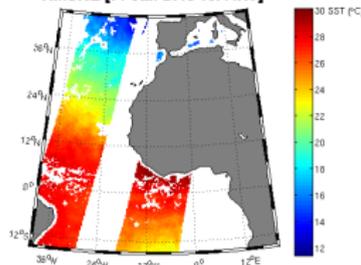
Nom	Description	Résolution spatiale	Résolution temporelle
(a) METOP	infrared	$\sim 0.05^\circ$	~ 12 h
(b) AMSRE	microwave	$\sim 0.25^\circ$	~ 24 h
(c) SEVIRI	infrared	$\sim 0.10^\circ$	~ 1 h

METOP [31-Dec-2008 22:31:21]



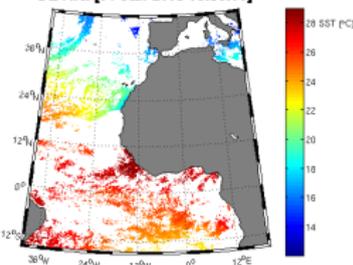
(a) 22h31

AMSRE [01-Jan-2009 03:44:07]



(b) 03h44

SEVIRI [01-Jan-2009 03:53:00]

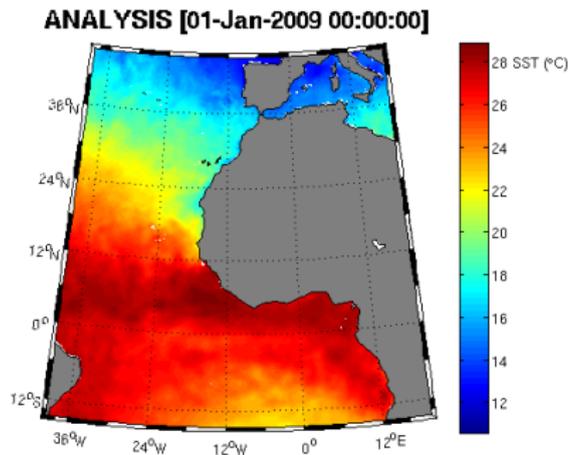


(c) 03h53

FIG.: Exemple de d'observations du champ de SST pour une nuit

Exemple 1 : fusion de données

Objectif



Générer des cartes avec ces caractéristiques :

- combiner toutes les observations
- pas de données éparées
- physiquement réaliste
- bonne résolution spatiale
- analyses en temps réel

Aujourd'hui : analyse à partir du meilleur capteur en chaque point

FIG.: Champ de SST interpolé

Exemple 2 : modélisation d'un phénomène latent

Suivi de structures météorologiques

• Vent

- Améliorer les prévisions locales : énergie éolienne, plaisanciers
- Générer des climatologies : estimation de rentabilité (parc éolien), dimensionnement (éoliennes, ponts, ...)

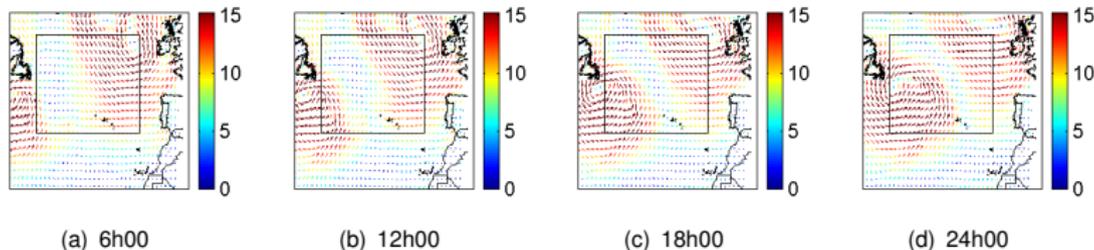


FIG.: Séquence de champs de vent - 1er déc. 2002

On veut estimer la vitesse de déplacement des structures (dépressions, fronts).

Exemple 2 : modélisation d'un phénomène latent

Suivi de structures météorologiques

- Hauteur significative des vagues : mesure d'énergie, donnée par $4\sqrt{\text{Var}(\eta_t)}$ avec η_t l'élévation de la surface libre en un point fixe.
 - Améliorer les prévisions : pêche, surf
 - Générer des climatologies : dimensionnement de structures (plate-forme, bateau), estimation de la fatigue (transport maritime)

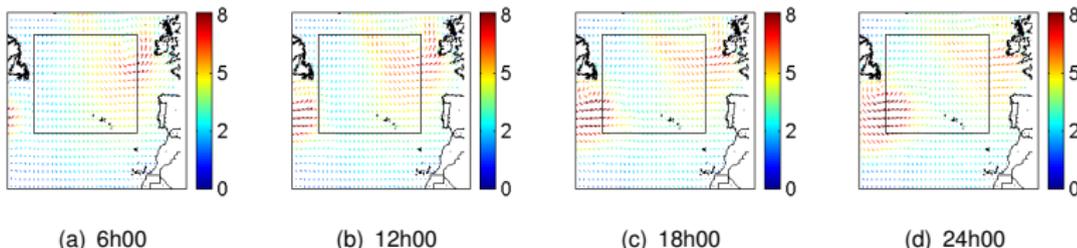


FIG.: Séquence de champs de hauteur significative des vagues - 1er déc. 2002

On veut estimer la vitesse de déplacement des structures (suivi de systèmes d'état de mer).

Exemple 3 : Assimilation de données dans un modèle hydrodynamique

Assimilation de données séquentielle de température de surface dans un modèle hydrodynamique côtier dans le golfe de Gascogne

- pour améliorer les prévisions en température et salinité dans le domaine côtier
- pour corriger les prévisions les courants de surface dans le domaine côtier

Modèle physique

- Modèle de type Blumberg & Mellor (variables du modèle : courants, température, salinité)
- Grille horizontale : 4 km x 4 km
- Résolution verticale : 30 niveaux sigma (résolution plus fine près de la surface)

Exemple 3 : Assimilation de données dans un modèle hydrodynamique

Les prévisions du modèle hydrodynamique comportent des erreurs.

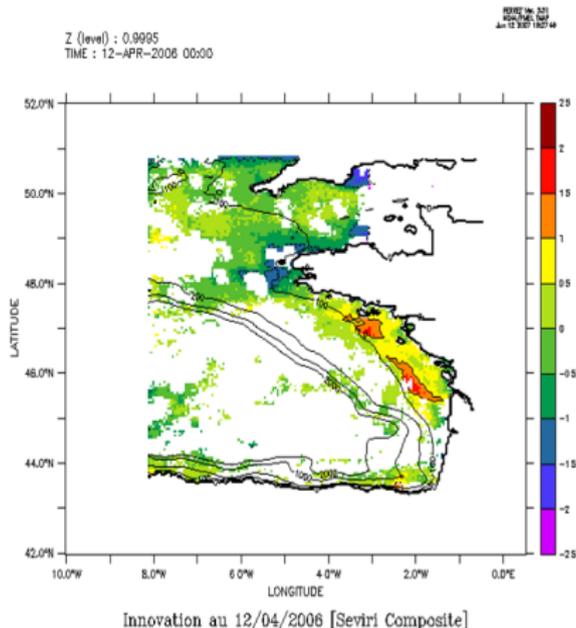


FIG.: Erreur entre la SST modèle et la SST SEVIRI - 12 avril 2006

Exemple 3 : Assimilation de données dans un modèle hydrodynamique

L'erreur du modèle est saisonnière et due à plusieurs sources dont la turbulence, le coefficient d'extinction de la lumière, les tensions de vent (modèle), la rugosité du fond,...

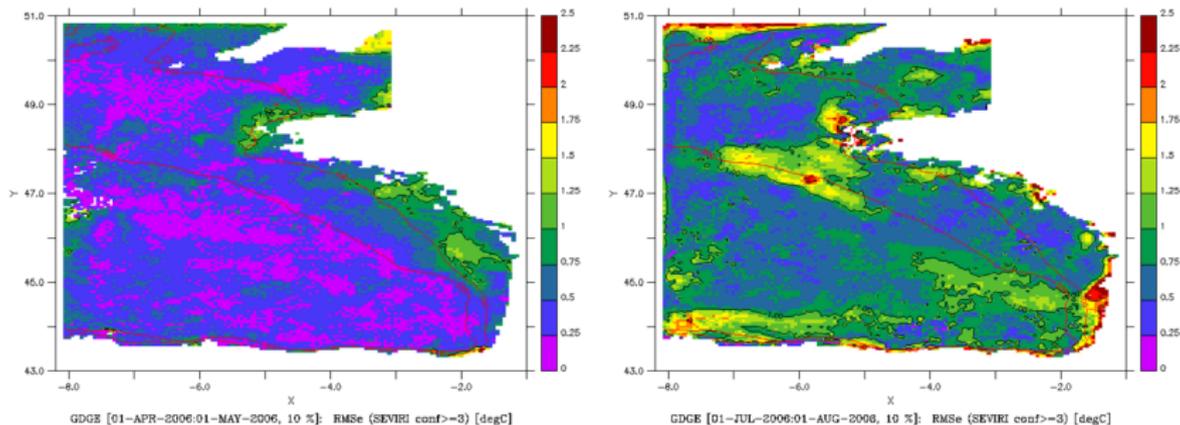


FIG.: Erreur en moyenne quadratique entre le modèle et SEVIRI (avril et juillet)

Exemple 3 : Assimilation de données dans un modèle hydrodynamique

- Objectif : assimiler des données de température de surface de la mer pour corriger l'énergie potentielle interne (température, salinité -> densité de l'eau)
- On suppose que l'énergie potentielle barotrope (niveau de la mer -> courants) s' "ajuste" dynamiquement lors de la prévision.
- Contraintes : méthode séquentielle, temps réel.

Plan

- 1 Introduction
 - Généralités
 - Exemples
 - Filtrage
- 2 Modèle à espace d'état
 - Définition
 - Inférence dans les modèles à espace d'état
 - Filtre de Kalman
- 3 Inférence pour les modèles non linéaires
 - Filtre de Kalman d'Ensemble
 - Filtre particulière
 - Assimilation variationnelle
- 4 Applications
 - Exemple 1 : fusion de données
 - Exemple 2 : estimation d'une vitesse latente
 - Exemple 3 : assimilation de données
- 5 Conclusion

Filtrage

Le filtrage consiste à estimer l'état S d'un système dynamique à partir d'observations partielles, généralement bruitées.

- Exemple 1 : S = SST
- Exemple 2 : S = vitesse de déplacement
- Exemple 3 : S = intensité et direction des courants, température, salinité

On dispose d'une suite d'observations $\{y_1, \dots, y_T\}$. Chaque observation est reliée à l'état par une relation probabiliste du type

$$P(Y_t \in dy | S_t = s) = p(y|s)dy$$

par exemple

$$Y_t = h(S_t) + W_t$$

avec W_t un bruit indépendant de S_t qui modélise l'erreur d'observation.

- Exemple 1 : h = opérateur linéaire (identité partielle)
- Exemple 2 : h = basé sur la comparaison de champs successifs (non linéaire)
- Exemple 3 : h = opérateur linéaire (identité partielle)

Remarque : ce problème, formulé ainsi, est mal posé dans le cas général.

Filtrage (suite)

- Considérons un exemple,
 - $S_t = s$, $\forall t$ et $s \in \mathbb{R}^m$ inconnu.
 - $Y_t = Hs + W_t$, H matrice $d \times m$
 - s^0 : vraie valeur du paramètre

- Problème d'optimisation

$$s^* = \arg \min_{s \in \mathbb{R}^m} \left\{ \frac{1}{2} \sum_{t=1}^T |Y_t - Hs|^2 \right\} \text{ d'où } Hs^* = \frac{1}{T} \sum_{t=1}^T Y_t$$

- Si $m = d$ et H inversible,

$$\begin{aligned} \hat{s}^* &= H^{-1} \frac{1}{T} \sum_{t=1}^T Y_t \\ &= s^0 + H^{-1} \frac{1}{T} \sum_{t=1}^T W_t \rightarrow s^0 \text{ quand } T \rightarrow \infty \end{aligned}$$

- Si $d < m$, on a une infinité de solutions avec $Hs^* \rightarrow Hs^0$ quand $T \rightarrow \infty$

Filtre et modèle à espace d'état

- On a besoin d'information supplémentaire sur S_t pour choisir un point particulier de l'espace affine $\{s \in \mathbb{R}^m | Hs = Hs^0\}$.
- On introduit en général
 - une équation d'évolution pour l'état qui va donner un *a priori* sur l'état (background),
 - des relations d'indépendance conditionnelles pour écrire les problèmes d'inférence.
- Exemples 1 et 2 : l'évolution de l'état est modélisée par un modèle stochastique inféré à partir des données (modèle autorégressif)
- Exemple 3 : le modèle hydrodynamique donne un *a priori* sur l'état

Plan

- 1 Introduction
 - Généralités
 - Exemples
 - Filtrage
- 2 **Modèle à espace d'état**
 - **Définition**
 - Inférence dans les modèles à espace d'état
 - Filtre de Kalman
- 3 **Inférence pour les modèles non linéaires**
 - Filtre de Kalman d'Ensemble
 - Filtre particulière
 - Assimilation variationnelle
- 4 **Applications**
 - Exemple 1 : fusion de données
 - Exemple 2 : estimation d'une vitesse latente
 - Exemple 3 : assimilation de données
- 5 **Conclusion**

Modèle à espace d'état

Définition

Un modèle à espace d'état est un processus markovien à 2 composantes

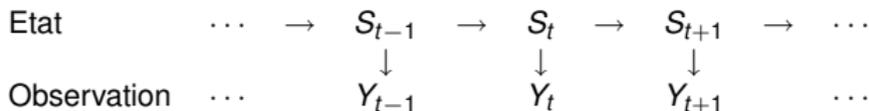
$$\{X_t\} = \{S_t, Y_t\} \in \{S \times \mathcal{Y}\} \subset \mathbb{R}^d \times \mathbb{R}^m$$

avec $\{S_t\}$ non observable et les propriétés markoviennes suivantes :

$$P(S_t | S_0 = s_0, \dots, S_{t-1} = s_{t-1}, Y_1 = y_1, \dots, Y_{t-1} = y_{t-1})$$

$$= P(S_t | S_{t-1} = s_{t-1})$$

$$P(Y_t | S_0 = s_0, \dots, S_t = s_t, Y_1 = y_1, \dots, Y_{t-1} = y_{t-1}) = P(Y_t | S_t = s_t)$$



Modèle à espace d'état

Deux écritures

- Probabiliste

- Loi initiale : $P(S_0 \in ds) = p(s)ds$
- Probabilité de transition : $P(S_t \in ds | S_{t-1} = s_{t-1}) = p(s | s_{t-1})ds$
- Probabilité d'émission : $P(Y_t \in dy | S_t = s_t) = p(y | s_t)dy$

- Système dynamique

- Condition initiale : $P(S_0 \in ds) = p(s)ds$
- Équation d'état : $S_t = f_t(S_{t-1}, V_t)$ avec V_t un bruit blanc
- Équation d'observation : $Y_t = h_t(S_t, W_t)$ avec W_t un bruit blanc indépendant de V_t et de S_0 .

- Cas particulier du modèle linéaire gaussien

- Condition initiale : $S_0 \sim \mathcal{N}(m_0, Q_0)$
- $S_t = F_t S_{t-1} + V_t$ avec V_t un bruit blanc gaussien de covariance Q
- $Y_t = H_t S_t + W_t$ avec W_t un bruit blanc gaussien de covariance R , indépendant de V_t , R inversible

Plan

- 1 Introduction
 - Généralités
 - Exemples
 - Filtrage
- 2 **Modèle à espace d'état**
 - Définition
 - **Inférence dans les modèles à espace d'état**
 - Filtre de Kalman
- 3 **Inférence pour les modèles non linéaires**
 - Filtre de Kalman d'Ensemble
 - Filtre particulière
 - Assimilation variationnelle
- 4 **Applications**
 - Exemple 1 : fusion de données
 - Exemple 2 : estimation d'une vitesse latente
 - Exemple 3 : assimilation de données
- 5 **Conclusion**

Inférence statistique

- Problème : pour tout temps t , estimer l'état S_t du système sachant une séquence d'observations $\{y_1, \dots, y_T\}$ et connaissant le modèle.
- Trois cas -
 - *Prédiction* : $T = t - 1$

$$p(s_t | y_1, \dots, y_{t-1}) = \int_{\mathbf{s}} p(s_t | s_{t-1}) p(s_{t-1} | y_1, \dots, y_{t-1}) ds_{t-1}$$

- *Filtrage* : $T = t$

$$p(s_t | y_1, \dots, y_t) \propto p(y_t | s_t) p(s_t | y_1, \dots, y_{t-1})$$

- *Lissage* : $T > t$

$$\begin{aligned} & p(s_t | y_1, \dots, y_T) \\ &= p(s_t | y_1, \dots, y_t) \int_{\mathbf{s}} \frac{p(s_{t+1} | s_t)}{p(s_{t+1} | y_1, \dots, y_t)} p(s_{t+1} | y_1, \dots, y_T) ds_{t+1} \end{aligned}$$

- En pratique, on utilise $\hat{s}_t = E[s_t | y_1, \dots, y_T]$ et $Var[s_t | y_1, \dots, y_T]$
- Remarque : on ne peut pas faire ces calculs de manière exacte sauf dans des cas particuliers
 - Chaîne de Markov cachée : algorithme Forward-Backward
 - Modèle linéaire Gaussien : filtre de Kalman

Plan

- 1 Introduction
 - Généralités
 - Exemples
 - Filtrage
- 2 **Modèle à espace d'état**
 - Définition
 - Inférence dans les modèles à espace d'état
 - **Filtre de Kalman**
- 3 **Inférence pour les modèles non linéaires**
 - Filtre de Kalman d'Ensemble
 - Filtre particulière
 - Assimilation variationnelle
- 4 **Applications**
 - Exemple 1 : fusion de données
 - Exemple 2 : estimation d'une vitesse latente
 - Exemple 3 : assimilation de données
- 5 **Conclusion**

Filtre de Kalman - approche contrôle

L'objectif est d'estimer S_t à partir de $\{y_1, \dots, y_t\}$ de manière optimale et récursive.

- Modèle à espace d'état

$$S_0 \sim \mathcal{N}(m_0, Q_0)$$

$$S_t = F_t S_{t-1} + V_t \text{ avec } V_t \text{ un bruit blanc gaussien de covariance } Q$$

$$Y_t = H_t S_t + W_t \text{ avec } W_t \text{ un bruit blanc gaussien de covariance inversible } R, \\ \text{indépendant de } V_t \text{ et de la condition initiale.}$$

- Par définition, le filtre de Kalman est le meilleur estimateur linéaire non biaisé de S_t (estimateur BLUE) sachant une suite d'observations $\{y_1, \dots, y_t\}$ d'où

$$\hat{S}_t = (I - KH)S_t^f + Ky_t \text{ avec l'état prédit } S_t^f = FS_{t-1}^f$$

- On a $\text{Var}(\hat{S}_t) = \hat{S}_t \hat{S}_t' = P_t - 2P_t H' K' + K(R + HP_t H') K'$
 avec P_t la covariance de S_t^f
 Or

$$K^* = \arg \min_{K \in \mathbb{R}^{d \times m}} (P_t - 2P_t H' K' + K(R + HP_t H') K')$$

conduit à

$$K^* = P_t H' (R + HP_t H')^{-1}$$

Filtre de Kalman

Filtre de Kalman

- Initialisation

$$s_0 \sim \mathcal{N}(0, R_0)$$

- Prédiction

$$s_t^f = F\hat{s}_{t-1}$$

$$P_t^f = FP_{t-1}F' + Q$$

- Correction

$$K_t = P_t^f H' (HP_t^f H' + R)^{-1}$$

$$\hat{s}_t = s_t^f + K_t(y_t - Hs_t^f)$$

$$P_t = (I - K_t H)P_t^f$$

Remarque : le gain de Kalman s'écrit comme un rapport de variances, qui mesurent la confiance accordée à chacune des approximations de l'état (prédiction du modèle, observations).

Ref : Kalman (1966)

Filtre de Kalman - approche probabiliste

Comme le cadre est gaussien, il suffit d'estimer

$$\hat{S}_t = E[S_t | Y_1, \dots, Y_t] \text{ et } P_t = E[(S_t - \hat{S}_t)(S_t - \hat{S}_t)' | Y_1, \dots, Y_t]$$

Ainsi P_t ne dépend pas des observations d'après la proposition suivante.

Proposition

Soit $Z = (S, Y) \sim \mathbf{N} \left(\begin{pmatrix} m_S \\ m_Y \end{pmatrix}, \begin{pmatrix} Q_S & Q_{SY} \\ Q_{YS} & Q_Y \end{pmatrix} \right)$. Si Q_Y est inversible, $p_{S|Y=y}(s)$ est une loi de Gauss de moyenne $S + Q_{SY}Q_Y^{-1}(y - m_Y)$ et de matrice de covariance $Q_S - Q_{SY}Q_Y^{-1}Q_{YS}$.

Filtre de Kalman

- *Etape de prédiction* - $p(s_t | y_1, \dots, y_{t-1})$?
- *Etape de correction* - On considère l'innovation $l_t = Y_t - E[Y_t | Y_1, \dots, Y_{t-1}]$ qui est un processus gaussien centré de covariance $Q_t' = HP_tH' + Q$ et $E[(S_t - F\hat{S}_{t-1})l_t'] = P_tH'$

- Loi initiale - D'après la proposition

$$S_0 | Y_0 = y_0 \sim \mathcal{N} \left(m_0 + Q_0 H' [H Q_0 H' + R]^{-1} [y_0 - H m_0], Q_0 - Q_0 H' [H Q_0 H' + R]^{-1} H Q_0 \right)$$

- Prédiction - D'après l'équation d'état

$$S_t^f | y_0, \dots, y_{t-1} \sim \mathcal{N} \left(F \hat{S}_{t-1}, F P_{t-1} F' + R \right)$$

- Correction - D'après la proposition

$$S_t | y_0, \dots, y_{t-1} \sim \mathcal{N} \left(\hat{S}_t, P_t \right)$$

avec $\hat{S}_t = E[S_t | y_0, \dots, y_t] = S_t^f + E[S_t - S_t^f | I_t]$
 et $P_t = E[((S_t - S_t^f) - E[S_t - S_t^f | I_t])((S_t - S_t^f) - E[S_t - S_t^f | I_t])]$.
 D'où, par la proposition,

$$S_t | y_0, \dots, y_t \sim \mathcal{N} \left(S_t^f + K_t (y_t - H S_t^f), (I - K_t H)(F P_{t-1} F' + R) \right)$$

Filtre de Kalman

Dans les modèles linéaires gaussiens, on utilise le filtre de Kalman pour

- Maximiser la vraisemblance du modèle et ainsi estimer ses paramètres

$$\begin{aligned}
 p_{\theta}(Y_1 = y_1, \dots, Y_t = y_T) &= \int_{S^{T+1}} p_{\theta}(s_0) \prod_{t=1}^T p_{\theta}(s_t | s_{t-1}) p_{\theta}(y_t | s_t) ds_0 ds_1 \dots ds_T \\
 &= \prod_{t=1}^T \int_S p_{\theta}(y_t | s_t) p_{\theta}(s_t | y_1, \dots, y_{t-1}) ds_t
 \end{aligned}$$

Dans le cas du modèle linéaire gaussien, $p_{\theta}(y_t | s_t)$ et $p_{\theta}(s_t | y_1, \dots, y_{t-1})$ sont des lois de Gauss et on peut calculer la vraisemblance.

- Estimer la densité de lissage $p(s_t | y_1, \dots, y_T)$
→ lissage de Kalman

Filtre de Kalman

Exemple jouet

Modèle

- $S_0 \sim \mathcal{N}(10, 1)$
- $S_t = S_{t-1} + R$ avec $R \sim \mathcal{N}(0, 0.2)$
- $Y_t = S_t + Q$ avec $Q \sim \mathcal{N}(0, 0.5)$

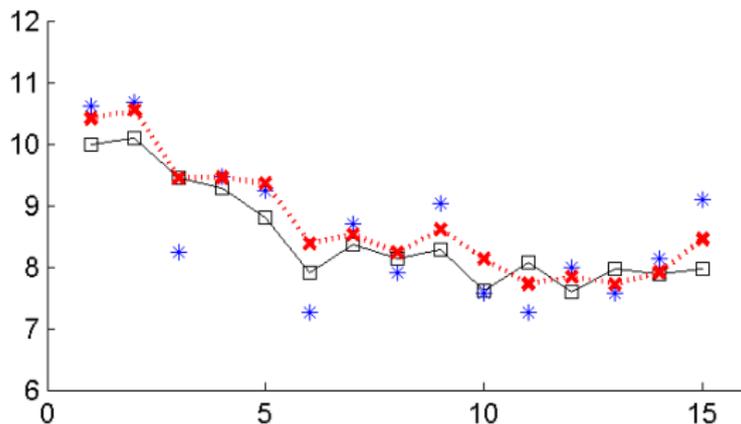


FIG.: Exemple d'assimilation par le filtre de Kalman - *Etat vrai* (noir), *Observations* (bleu), *Etat estimé* (rouge)

Filtre de Kalman

Avantages et inconvénients

- Les plus
 - Solution exacte
 - Coût de calcul faible en petite dimension
- Les moins
 - Modèles linéaires gaussiens (alternatives : filtre de Kalman linéarisé, filtre de Kalman étendu)
 - Coûts de calcul (produit et inversion de matrices) et de stockage en grande dimension
Exemple de pb météo : grille sur l'emprise du modèle (ex ARPEGE : Europe du Nord, MARS : Golfe de Gascogne) + différentes couches d'altitude (ou de profondeur)
- Les principales extensions du filtre de Kalman
 - Filtre de Kalman d'ensemble
 - Filtre particulière

Plan

- 1 Introduction
 - Généralités
 - Exemples
 - Filtrage
- 2 Modèle à espace d'état
 - Définition
 - Inférence dans les modèles à espace d'état
 - Filtre de Kalman
- 3 **Inférence pour les modèles non linéaires**
 - **Filtre de Kalman d'Ensemble**
 - Filtre particulière
 - Assimilation variationnelle
- 4 Applications
 - Exemple 1 : fusion de données
 - Exemple 2 : estimation d'une vitesse latente
 - Exemple 3 : assimilation de données
- 5 Conclusion

Filtre Kalman d'Ensemble (EnKF)

- Objectif historique : résoudre le problème du calcul du filtre de Kalman en grande dimension.
- Modèle
 - $P(S_0 \in ds) = p(s)ds$
 - $S_t = f(S_{t-1}) + V_t$: Équation d'état non linéaire
 - $Y_t = HS_t + W_t$: Équation d'observation linéaire
 - V_t, W_t : Bruits indépendants et de loi de Gauss de matrices de covariance R et Q , indépendants de S_0
- Idée : écrire une approximation de Monte Carlo du filtre de Kalman
 - Générer des éléments d'ensemble s_t^i selon l'équation d'état
 - Calculer l'estimateur empirique \hat{P}_t^f de la covariance P_t^f et en déduire le gain de Kalman
 - Réaliser l'étape de correction comme dans le filtre de Kalman
- Très utilisé dans les applications météorologiques et océanographiques.

Algorithme du filtre de Kalman d'ensemble

Filtre de Kalman

- Initialisation
 $s_0 \sim \mathcal{N}(m_0, Q_0)$
- Prédiction
 $s_t^f = F\hat{s}_{t-1}$
 $P_t^f = FP_t^a F' + Q$
- Correction
 $K_t = P_t^f H' (HP_t^f H' + R)^{-1}$
 $\hat{s}_t = s_t^f + K_t(y_t - Hs_t^f)$
 $P_t = (I - K_t H)P_t^f$

Filtre de Kalman d'Ensemble

- Initialisation
 $s_0^i \sim p(s_0), i = 1, \dots, n$
- Prédiction
 $v^i \sim \mathcal{N}(0, Q)$
 $s_t^{f,i} = f(\hat{s}_{t-1}^i) + v^i, i = 1, \dots, n$
 $\hat{P}_t^f = \frac{1}{n} \sum_{i=1}^n (s_t^{f,i} - \bar{s}_t^{f,i})(s_t^{f,i} - \bar{s}_t^{f,i})'$
- Correction
 $w^i \sim \mathcal{N}(0, R), i = 1, \dots, n$
 $\hat{\gamma} = \frac{1}{n} \sum_{i=1}^n w^i (w^i)'$
 $K_t = \hat{P}_t^f H' (H\hat{P}_t^f H' + \hat{R})^{-1}$
 $\hat{s}_t^i = s_t^{f,i} + K_t(y_t - Hs_t^{f,i} + w^i)$

Réf : Evensen (1994)

Filtre de Kalman d'ensemble (suite)

En pratique, pour réduire les coûts de calcul, on utilise

- les factorisations

$$\hat{P}_t^f H' = \frac{1}{n} \sum_{i=1}^n (S_t^{f,i} - \bar{S}_t^{f,i})(h_t^i)'$$

$$H \hat{P}_t^f H' = \frac{1}{n} \sum_{i=1}^n h_t^i (h_t^i)'$$

avec $h_t^i = H_t(S_t^{f,i} - \bar{S}_t^{f,i})$

- la formule de Sherman-Morrison-Woodbury qui est avantageuse si $m \gg d$

$$(R + UV')^{-1} = R^{-1} - R^{-1}U(I + V'R^{-1})^{-1}V'R^{-1}$$

avec $U = \frac{1}{n}h$ et $V = h$

- Dans les applications, $n \simeq 50$.

Filtre de Kalman d'Ensemble Exemple jouet

Modèle

- $S_0 \sim \mathcal{N}(10, 1)$
- $S_t = S_{t-1} + R$ avec $R \sim \mathcal{N}(0, 0.2)$
- $Y_t = S_t + Q$ avec $Q \sim \mathcal{N}(0, 0.5)$

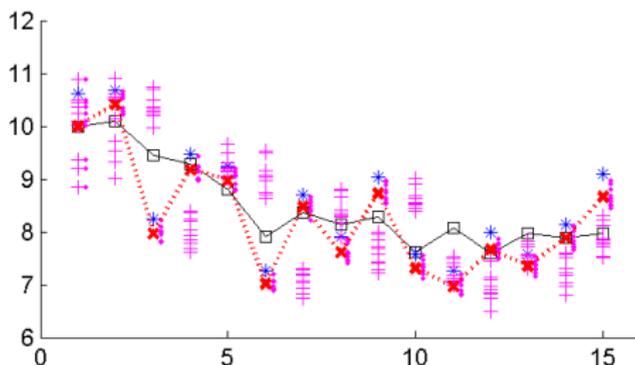


FIG.: Exemple d'assimilation par le filtre de Kalman d'ensemble - $n = 10$ - Etat vrai (noir), Observations (bleu), Etat estimé (rouge), Forecast (+ violets), Corrigé (points)

Filtre de Kalman d'Ensemble Exemple jouet

Modèle

- $S_0 \sim \mathcal{N}(10, 1)$
- $S_t = S_{t-1} + R$ avec $R \sim \mathcal{N}(0, 0.2)$
- $Y_t = S_t + Q$ avec $Q \sim \mathcal{N}(0, 0.5)$

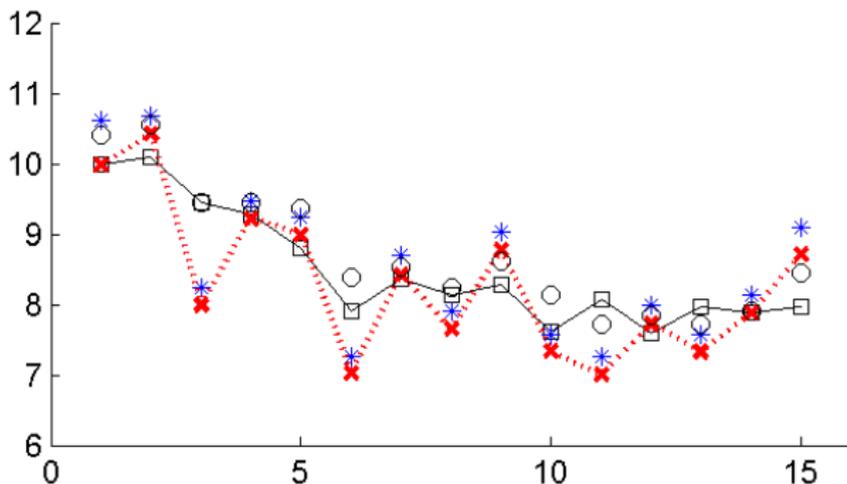


FIG.: Exemple d'assimilation par le filtre de Kalman d'ensemble - $n = 10000$ - *Etat vrai (noir)*,

Filtre de Kalman d'ensemble - loi limite

Considérons le modèle à un pas de temps suivant :

$$\begin{aligned} S_1 &= f(S_0) + V_1 \\ Y_1 &= H S_1 + W_1, \end{aligned}$$

où les bruits additifs V_1 et W_1 sont bruits gaussien de matrices de covariances Q et R respectivement (R étant inversible), et où l'état initial S_0 peut prendre un nombre fini de valeurs, i.e. sa distribution

$$\mu_0 = \sum_{i \in I} p_0^i \delta_{x_0^i} \quad \text{avec des poids positifs, et} \quad \sum_{i \in I} p_0^i = 1,$$

est un mélange fini de masses de Dirac.

On peut calculer les lois a posteriori

- du filtre de Kalman
- du filtre de Kalman d'ensemble
- du filtre bayésien

Filtre de Kalman d'ensemble - loi limite

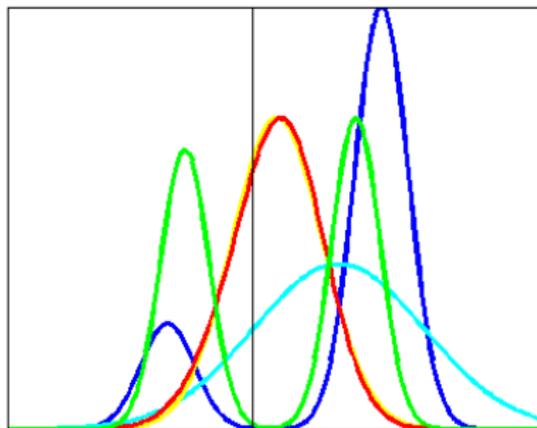


FIG.: Dist. a priori (bleu foncé), Dist. gauss. a priori (bleu clair), Dist. gauss. a posteriori (jaune), Dist. EnKF (rouge), Dist. Bayésienne (vert), Observation (trait vertical)

EnKF - Avantages et inconvénients

- Les plus

- Coût de calcul relativement faible : en pratique c'est souvent l'étape de prédiction qui coûte le plus cher
- Implémentation aisée
- Converge en $1/\sqrt{n}$
- "Robustesse" : donne des résultats satisfaisants même pour un faible nombre d'éléments quand la dynamique n'est pas trop complexe.

- Les moins

- Ne permet pas de reproduire la densité de filtrage si la dynamique est complexe (ex : bimodalités)
- La loi limite $\hat{p}(s_t|y_1, \dots, y_t)$ des éléments d'ensemble quand n tend vers l'infini n'est pas la loi du filtre bayésien.

$$p(s_t|y_1, \dots, y_t) \propto p(y_t|s_t)p(s_t|y_1, \dots, y_{t-1})$$

- Alternative : filtre particulière.

Plan

- 1 Introduction
 - Généralités
 - Exemples
 - Filtrage
- 2 Modèle à espace d'état
 - Définition
 - Inférence dans les modèles à espace d'état
 - Filtre de Kalman
- 3 **Inférence pour les modèles non linéaires**
 - Filtre de Kalman d'Ensemble
 - **Filtre particulière**
 - Assimilation variationnelle
- 4 Applications
 - Exemple 1 : fusion de données
 - Exemple 2 : estimation d'une vitesse latente
 - Exemple 3 : assimilation de données
- 5 Conclusion

Filtre particulaire

- Objectif : approcher le filtre bayésien.
- Modèle :
 - $P(S_0 \in ds) = p(s)ds$
 - $S_t = f(S_{t-1}) + V_t$ ou $P(S_t \in ds | S_{t-1} = s_{t-1}) = p(s | s_{t-1})ds$: Équation d'état non linéaire
 - $Y_t = h(S_t) + W_t$ ou $P(Y_t \in dy | S_t = s_t) = p(y | s_t)dy$: Équation d'observation non linéaire
 - V_t, W_t : Bruits blancs indépendants et indépendants des conditions initiales
- Idée - Le filtre particulaire est une approximation de Monte Carlo du filtre bayésien optimal : l'idée du filtre particulaire est d'approcher la distribution de filtrage $p(s | y_1, \dots, y_t)ds$ par la distribution empirique μ_t^n d'un ensemble de particules $\{s_t^1, \dots, s_t^n\}$ muni de poids d'importance $\{\omega_1, \dots, \omega_n\}$:

$$\mu_t^n(\cdot) = \sum_{i=1}^n \omega_i \delta_{s_t^i}(\cdot)$$

- Très utilisé en signal et en automatique : extraction de signal dans du bruit, suivi de cible, commande, ... très peu en météo.

Filtre particulaire

Filtre particulaire

- Loi initiale
 $s_0^i \sim p(s_0)$, $i = 1, \dots, n$
 $\omega_0^i = 1/n$
- Prédiction
 $s_t^i \sim p(s_t | s_{t-1}^i)$
- Correction
 $\omega_t^i \propto \omega_{t-1}^i p(y_t | s_t^i)$

Remarques :

- En pratique, dans l'étape de correction on peut redistribuer les particules.
- On peut choisir une *meilleure* loi d'importance $p(s_t | s_{t-1}, y_t)$ qui permet de tenir compte de l'observation dans l'étape de prédiction.

Filtre particulaire Exemple jouet

Modèle

- $S_0 \sim \mathcal{N}(10, 1)$
- $S_t = S_{t-1} + R$ avec $R \sim \mathcal{N}(0, 0.2)$
- $Y_t = S_t + Q$ avec $Q \sim \mathcal{N}(0, 0.5)$

Filtre particulaire avec loi d'importance optimale et redistribution (n=10)

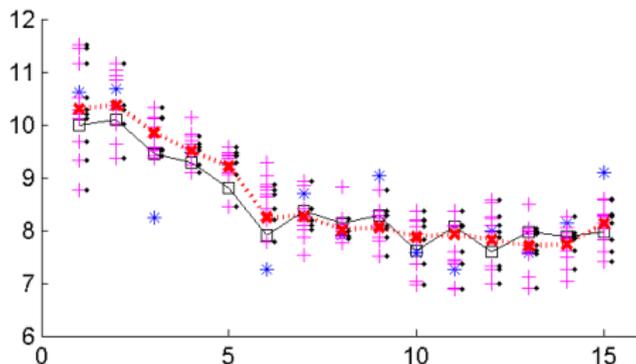


FIG.: Exemple d'assimilation par le filtre particulaire - *Etat vrai* (noir), *Observations* (bleu), *Etat estimé* (rouge), *Forecast* (+ violets), *Corrigé* (points noirs)

Filtre particulaire Exemple jouet

Modèle

- $S_0 \sim \mathcal{N}(10, 1)$
- $S_t = S_{t-1} + R$ avec $R \sim \mathcal{N}(0, 0.2)$
- $Y_t = S_t + Q$ avec $Q \sim \mathcal{N}(0, 0.5)$

Filtre particulaire avec loi d'importance optimale et redistribution (n=10000)

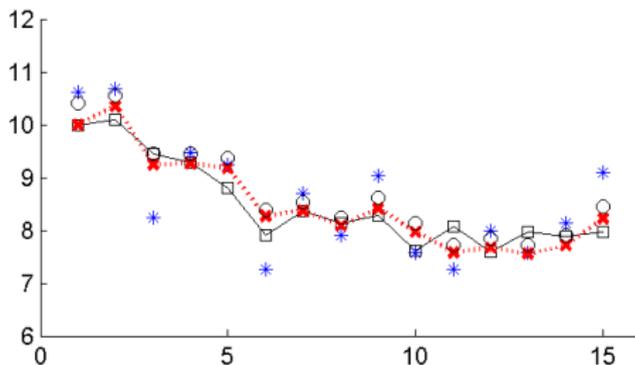


FIG.: Exemple d'assimilation par le filtre particulaire - *Etat vrai (noir)*, *Observations (bleu)*, *Etat estimé (rouge)*, *Kalman (ronds noirs)*

FP - Avantages et inconvénients

- Les plus
 - Facile à mettre en oeuvre
 - Peu d'hypothèses sur le modèle
 - Converge vers le filtre bayésien optimal à la vitesse $1/\sqrt{n}$ quand le nombre de particules n tend vers l'infini.
- Les moins
 - En pratique, nécessite un assez grand nombre de particules.

Plan

- 1 Introduction
 - Généralités
 - Exemples
 - Filtrage
- 2 Modèle à espace d'état
 - Définition
 - Inférence dans les modèles à espace d'état
 - Filtre de Kalman
- 3 **Inférence pour les modèles non linéaires**
 - Filtre de Kalman d'Ensemble
 - Filtre particulière
 - **Assimilation variationnelle**
- 4 Applications
 - Exemple 1 : fusion de données
 - Exemple 2 : estimation d'une vitesse latente
 - Exemple 3 : assimilation de données
- 5 Conclusion

Assimilation variationnelle de données

- Contexte : Les méthodes d'assimilation variationnelle de données, proposées par Sasaki dans les années 50. Ces méthodes relèvent de la théorie de l'optimisation.
- Modèle :
 - S_0 une variable aléatoire de covariance Q_0
 - $S_t = f(S_{t-1}) + V_t$: Équation d'état non linéaire
 - $Y_t = HS_t + W_t$: Équation d'observation linéaire
 - V_t, W_t : Bruits indépendants de matrices de covariance Q et R , indépendants de S_0 , R inversible
- Idée : trouver l'état qui minimise un coût aux moindres carrés pondérés prenant en compte l'écart entre l'état et l'*a priori* donné par le modèle ainsi que l'écart entre l'état et les observations.
- Ces méthodes sont très utilisées dans les grands centres de prévisions météorologiques.

Assimilation variationnelle (4D Var)

La solution de l'assimilation variationnelle sachant le modèle et une suite d'observations $\{y_1, \dots, y_T\}$ est la trajectoire déterministe x^* donné par

$$x^* = \arg \min_{x \in S^T} \sum_{t=1}^T \left((x_t^f - s_t)' (P_t^f)^{-1} (x_t^f - s_t) + (y_t - Hx_t)' R^{-1} (y_t - Hx_t)' \right) + (s_0 - x_0)' Q_0^{-1} (s_0 - x_0)$$

- En pratique, on résout ce problème par un algorithme de type Forward-Backward.
- La principale difficulté dans les méthodes d'assimilation variationnelles réside dans le calcul gradient de la fonction de coût. Quand l'équation d'état est non linéaire on peut
 - approcher le gradient par différence finie, mais ça pose des problèmes numériques ou
 - écrire un problème dual, c'est cette solution qui est en générale utilisée.
- Dans le cas d'un modèle linéaire gaussien :
 - Le résultat \hat{s}_t du filtre de Kalman sachant $\{y_1, \dots, y_t\}$ est identique à x_t .
 - Le lissage de Kalman est identique à l'assimilation 4D Var.

Assimilation variationnelle

Avantages et inconvénients

- Les plus

- Le coût de calcul de la solution d'une assimilation variationnelle est très faible par rapport au coût de calcul d'un filtre séquentiel.
- Peu d'hypothèses sur le modèle. L'hypothèse de linéarité de l'équation d'observation peut facilement être relaxée.

- Les moins

- L'écriture du problème dual est souvent difficile et unique pour chaque modèle.
- On remarque que l'assimilation variationnelle (4D Var) n'est pas une méthode séquentielle contrairement aux filtres introduits précédemment.
- Elle conduit à une solution déterministe et ne permet pas d'estimer les distributions de filtrage.

Plan

- 1 Introduction
 - Généralités
 - Exemples
 - Filtrage
- 2 Modèle à espace d'état
 - Définition
 - Inférence dans les modèles à espace d'état
 - Filtre de Kalman
- 3 Inférence pour les modèles non linéaires
 - Filtre de Kalman d'Ensemble
 - Filtre particulière
 - Assimilation variationnelle
- 4 Applications
 - **Exemple 1 : fusion de données**
 - Exemple 2 : estimation d'une vitesse latente
 - Exemple 3 : assimilation de données
- 5 Conclusion

Modèle à espace d'état

Exemple 1 : fusion de données

Modèle à espace d'état linéaire gaussien

Modèle :

$$\begin{cases} S(t_i) = F_{\Delta t} S(t_{i-1}) + V_{\Delta t}(t_i) & \rightarrow \text{Équation d'état} \\ y^{o(k)}(t_i) = H^{(k)} S^t(t_i) + W^{o(k)}(t_i) & \rightarrow k \text{ équations d'observation} \end{cases}$$

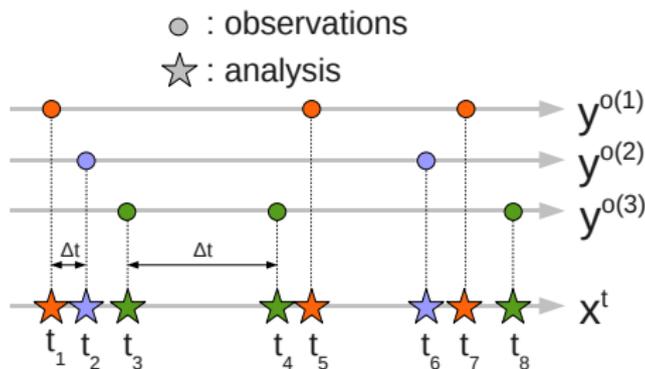


FIG.: Schéma d'analyse (ou de filtrage) en temps réel

Exemple 1 Inférence

Modèle :

$$\begin{cases} S(t_i) = F_{\Delta t} S^t(t_{i-1}) + V_{\Delta t}(t_i) & \rightarrow \text{Équation d'état} \\ Y^{o(k)}(t_i) = H^{(k)} S^t(t_i) + W^{o(k)}(t_i) & \rightarrow k \text{ Équations d'observation} \end{cases}$$

- Filtre de Kalman :

$$\hat{s}(t_i) = E(S(t_i) / y^o(t_0), \dots, y^o(t_i))$$

$$\hat{P}(t_i) = \text{Var}(S(t_i) / y^o(t_0), \dots, y^o(t_i))$$

- L'algorithme EM permet d'estimer F
- Les variances $\text{Var}(V)$, $\text{Var}(W^{o(k)})$ sont fixées d'après des considérations physiques.

Exemple 1 Résultats pour une série temporelle

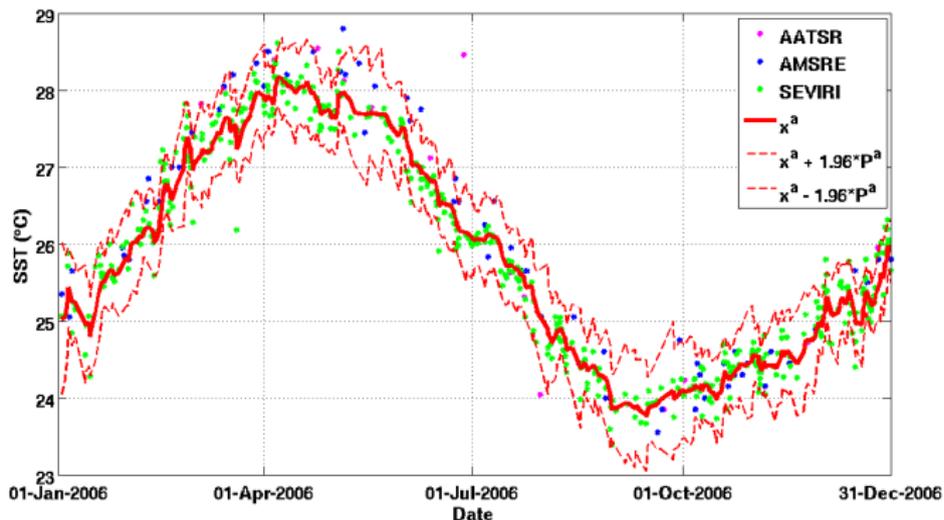


FIG.: Résultat du filtrage de Kalman (année 2006 & 3 observations)

Plan

- 1 Introduction
 - Généralités
 - Exemples
 - Filtrage
- 2 Modèle à espace d'état
 - Définition
 - Inférence dans les modèles à espace d'état
 - Filtre de Kalman
- 3 Inférence pour les modèles non linéaires
 - Filtre de Kalman d'Ensemble
 - Filtre particulière
 - Assimilation variationnelle
- 4 Applications
 - Exemple 1 : fusion de données
 - **Exemple 2 : estimation d'une vitesse latente**
 - Exemple 3 : assimilation de données
- 5 Conclusion

Champs de vitesse estimés

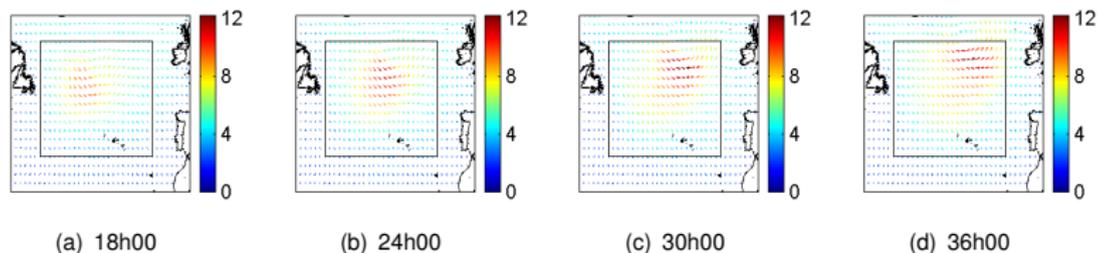
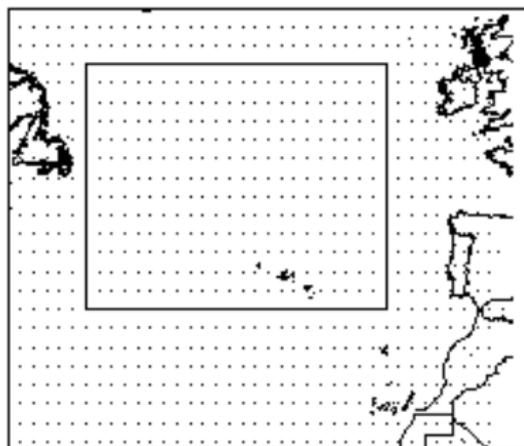


FIG.: Séquence de champs de hauteur significative des vagues - 15 déc. 2002

Estimation de la vitesse de déplacement (H_s)

Données de réanalyse (ERA-Interim)

- Echantillonnage : $\Delta lat = \Delta lon = 1.5^\circ$, $\Delta t = 6h$
- Période 1989-2007
- Paramètres : H_s , T_m , Θ_m , u , v ,...



Estimation des vitesses , données ERA-Interim

La vitesse V n'est pas directement observable : modèle à espace d'état

- Hypothèse : conservation de l'énergie (équation d'observation non linéaire)

$$H_s(p, t) = H_s(p - V(p, t)\Delta t, t - \Delta t) + W(p, t)$$

avec W un bruit (blanc en temps mais corrélé en espace)

- Dynamique

- Modèle physique : relation de dispersion (houle)
Vitesse de groupe : $C_g = \frac{g}{4\pi} T_m$ dans la direction : Θ_m

- Modèle statistique :

$$V(:, t) = AV(:, t - \Delta t) + E(:, t)$$

avec A un opérateur de lissage spatial (noyau gaussien)

- En pratique : on mélange les deux dynamiques.

Filtrage séquentiel

- Filtre particulaire

- Prédiction

- $v^{(i)}(.; t) = C_g + E(., t)$, pour $i = 1, \dots, n/3$

- $v^{(i)}(.; t) = Av^{(i)}(., t - \Delta t) + E(., t)$ pour $i = n/3 + 1, \dots, n$

- Correction

$$\omega_t^i \propto \omega_{t-1}^i \exp\left(-\|H_s(., t) - H_s(., v^{(i)}(., t)\Delta t, t - \Delta t)\|_{\Sigma_{obs}}^{\nu}\right)$$

- Estimation de $E[V(., t)|H_s(., 0), \dots, H_s(., t)] \sim \sum_{i=1}^N \omega_i v^{(i)}(.; t)$

- 10000 particules

Champs de vitesse estimés

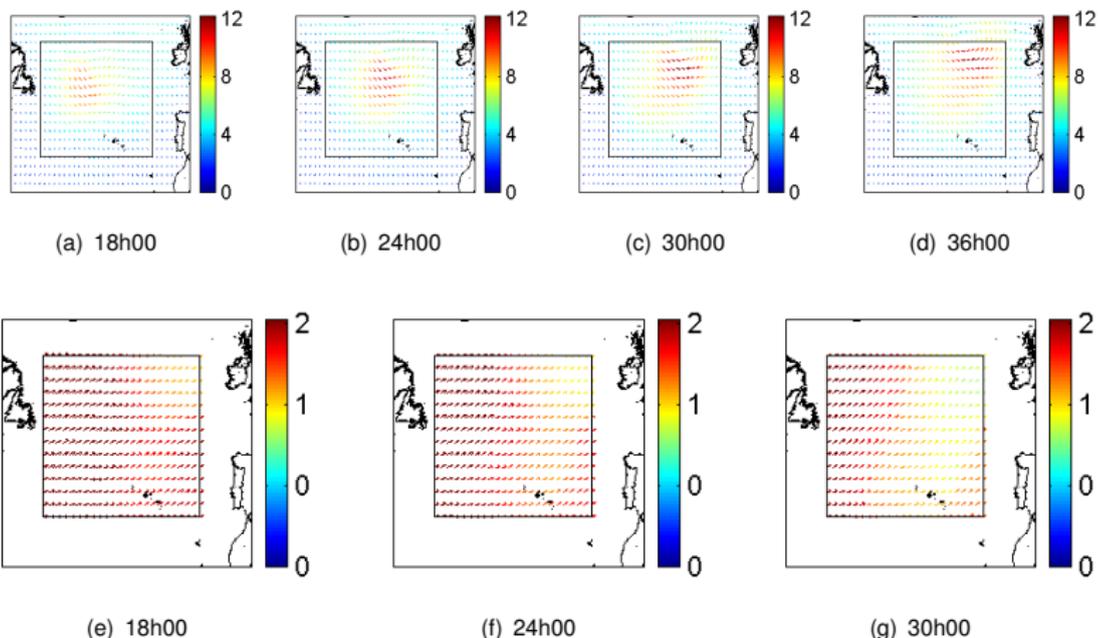


FIG.: Séquence de champs de hauteur significative des vagues et vitesses- 15 déc. 2002

Plan

- 1 Introduction
 - Généralités
 - Exemples
 - Filtrage
- 2 Modèle à espace d'état
 - Définition
 - Inférence dans les modèles à espace d'état
 - Filtre de Kalman
- 3 Inférence pour les modèles non linéaires
 - Filtre de Kalman d'Ensemble
 - Filtre particulière
 - Assimilation variationnelle
- 4 Applications
 - Exemple 1 : fusion de données
 - Exemple 2 : estimation d'une vitesse latente
 - **Exemple 3 : assimilation de données**
- 5 Conclusion

Exemple 3 : assimilation de données

- Équation d'état
 - modèle hydrodynamique, fortement non linéaire
 - dimension de l'espace d'état (1.910^6 points de grille)
- Bruits
 - erreurs sur les forçages (vent)
 - erreurs sur certains paramètres physiques (coefficient d'extinction)
 - un bruit additif
 - les bruits sont supposés gaussiens
- Équation d'observation
 - observation partielle de la température (une seule variable)
 - pas de temps irrégulier
 - bruit gaussien à covariance diagonale
 - rayon d'influence

Exemple 3 : assimilation de données Observations

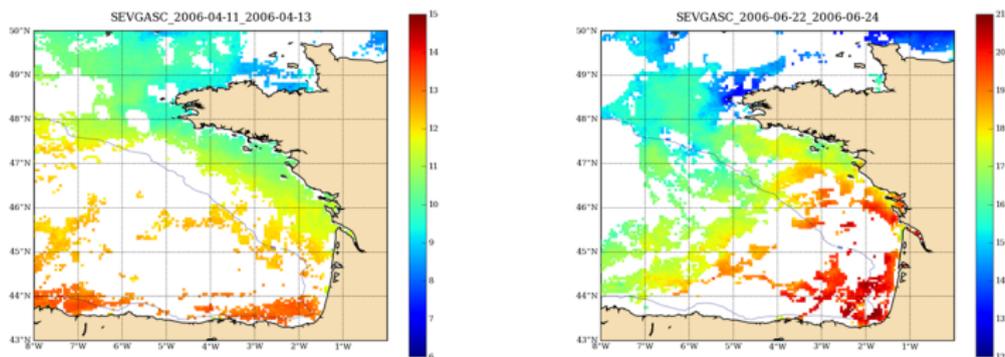


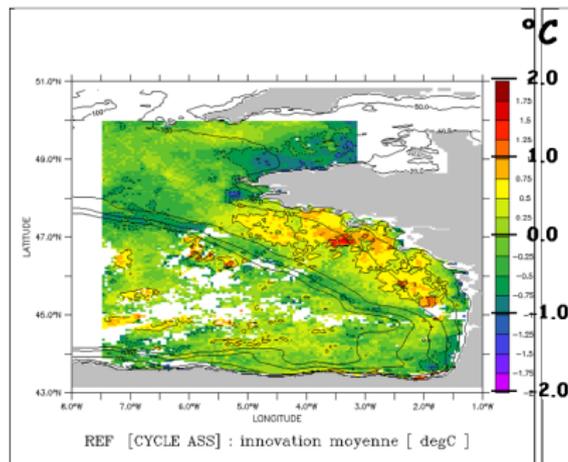
FIG.: Exemple d'observations SEVIRI 12 avril et 26 juin 2006

Exemple 3 : assimilation de données

Assimilation

EnKF, 30 éléments d'ensemble, une assimilation tous les 3 jours

Control Run



Ensemble Mean Forecast

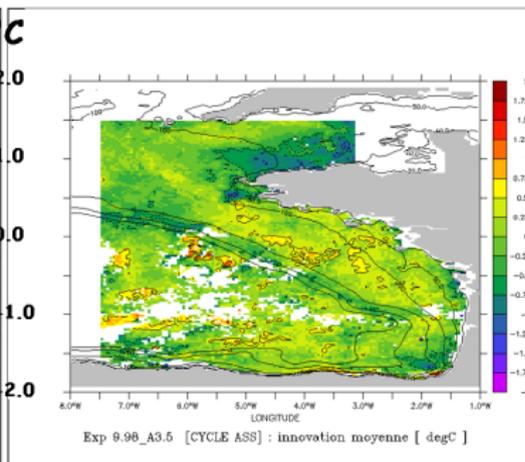
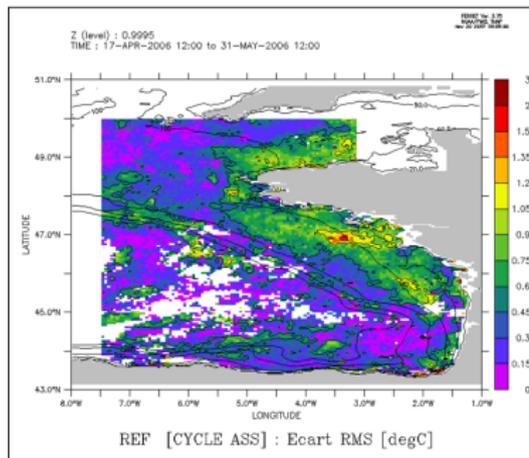


FIG.: Biais entre la prédiction et l'observation, sans assimilation (gauche) et avec assimilation (droite)

Exemple 3 : assimilation de données

Assimilation

Control Run



Ensemble Mean Forecast

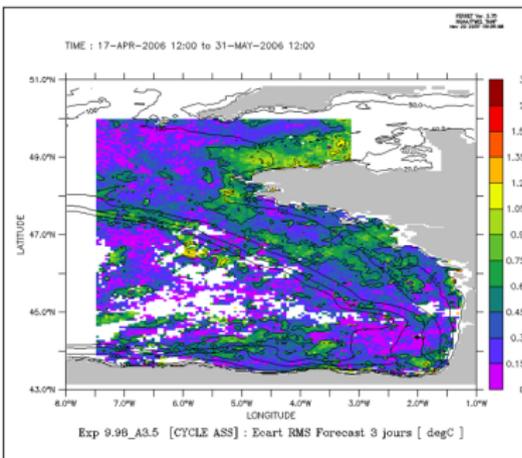
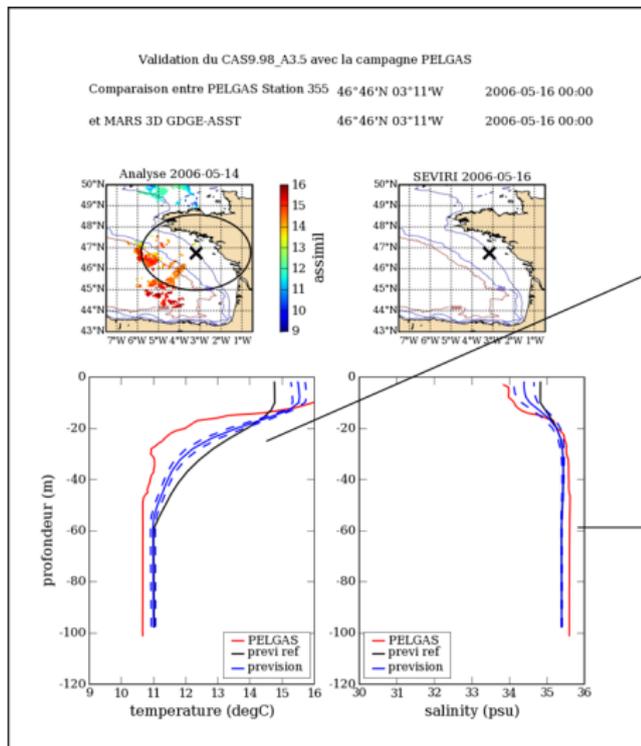


FIG.: EQM entre la prédiction et l'observation, sans assimilation (gauche) et avec assimilation (droite)

Exemple 3 : assimilation de données

Validation sur la colonne d'eau



Conclusion

- Modèles à espace d'état
 - Extraction de signal sur bruit
 - Fusion de données
 - Modélisation de processus latents
- Inférence
 - Modèle linéaire gaussien : solution exacte par Kalman
 - Modèles non linéaires : approximations par le filtre Kalman d'ensemble ou le filtre particulaire.
 - En météorologie et océanographie, on utilise surtout le filtre de Kalman d'ensemble.
- Autres domaines d'application
 - Contrôle, Reconnaissance de la parole, Suivi de cible (dans des vidéos, des images radar, ...), Finance, Gestion de réseaux télécom, ...
- Les difficultés des méthodes de filtrage séquentiel stochastique
 - Coûts de calcul/Nombre d'éléments d'ensemble
 - Modélisation des erreurs de modèle, d'observation
 - Estimation des paramètres dans les modèles non linéaires