

L'observation adaptative en météorologie : techniques mathématiques utilisées

A. Doerenbecher

CNRM, Météo-France, Toulouse



METEO FRANCE
Toujours un temps d'avance

Introduction

● Plan :

- 1 Présentation du concept d'observation adaptative ;
- 2 Un brin de formalisme - assimilation de données ;
- 3 Revue (partielle) de quelques techniques utilisées ;
- 4 Zoom sur la méthode KFS (*Kalman Filter Sensitivity*).

● Introduction :

- ▶ L'observation adaptative est une discipline qui se trouve à un carrefour de la prévision numérique du temps :
 - ▶ prévisibilité et contrôle,
 - ▶ assimilation des données,
 - ▶ gestion des réseaux d'observation.
- ▶ Cadre est celui du contrôle optimal.
- ▶ Les algorithmes utilisés actuellement ne parviennent pas à optimiser toutes les contraintes.

Observation adaptative

- L'observation adaptative consiste en une variation locale et temporaire d'un réseau d'observations, e.g. ajout d'observations (lieu, nombre, paramètre).
- L'objectif est d'améliorer la prévision qui découle de l'assimilation du jeu d'observations modifié, par rapport à la prévision qui découle du jeu d'observations par défaut.
- Des techniques numériques dites *objectives* permettent d'optimiser ces modifications de réseau d'observation.
- L'observation adaptative est aussi appelée *ciblage* des observations.

Adaptive observation

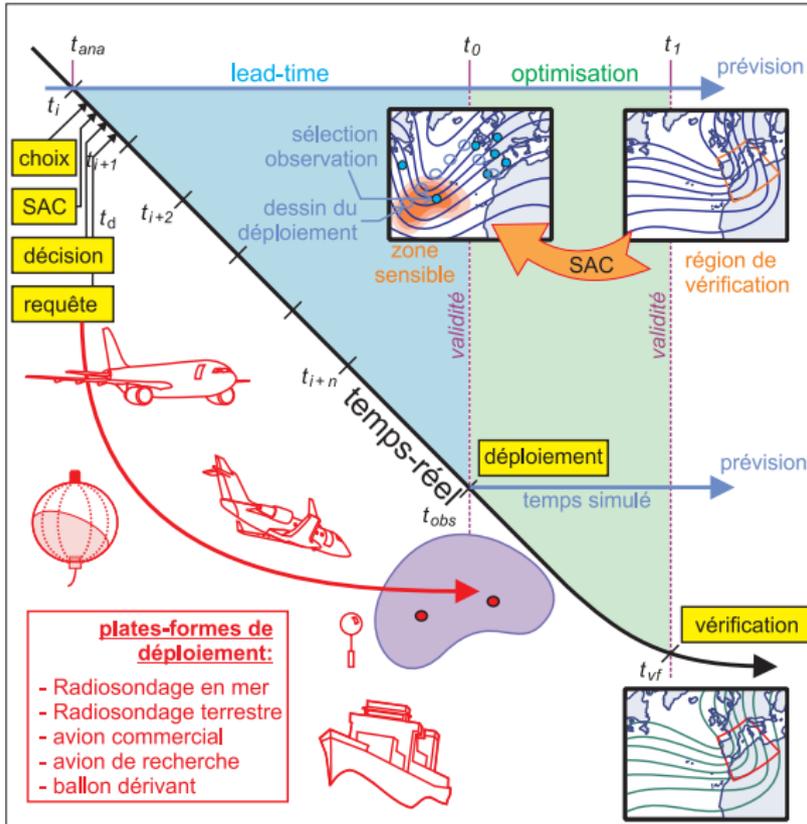


Schéma d'un opération d'observation adaptative en temps réel.

t_{ana} : instant (présent) d'analyse,

t_{obs} : instant (futur) d'observation,

t_{vf} : instant (futur) de vérification.

Camapagnes de mesure & observation adaptative

- Quelques exemples de campagnes de mesure :

1997 FASTEX (**FR**onts and **AT**lantic **S**torm **T**rack **E**xperiment)

1998 NORPEX (**N**orth **P**acific **E**xperiment)

2000-2009 WSRP (**W**inter **S**torm **R**econnnaissance **P**rogramme)

2003 NA-TReC (**N**orth **A**tlantic **T**HORPEX **R**egional **C**ampaign)

2006 AMMA (**A**frican **M**onsoon **M**ultidisciplinary **A**nalyses)

2008 Eurorisk-Preview DTS (**D**ata **T**argeting **S**ystem) et T-PARC

2009 CONCORDIASI (Antarctique)

- L'observation adaptative se pratique aussi bien sous les latitudes tempérées que autour des cyclones tropicaux (USA, DOTSTAR (Taiwan), SWICE (La Réunion)).

Formalisme

- Le cadre formel est celui de l'assimilation de données. L'article Ide *et al.*, 1997 constitue la référence pour les notations dans le domaine de l'assimilation de données en météorologie.

- Vecteur d'état, espace du **modèle** :

\mathbf{x}_j est le vecteur d'état du modèle, de dimension n , à l'instant t_j ;

\mathbf{x}_j^t est la réalité (inconnue) à l'instant t_j ;

\mathbf{x}_j^a est l'analyse à l'instant t_j ;

$\mathbf{x}_{j \rightarrow j+1}^f$ est la prévision issue de \mathbf{x}_j^a et valide à t_{j+1} : $\mathbf{x}_{j+1}^f = \mathcal{M}_{j \rightarrow j+1}(\mathbf{x}_j^a)$.

- Espace des **observations** :

\mathbf{y}_j^o est le vecteur des observations à l'instant t_j (dimension p) ;

\mathcal{H}_j est l'opérateur non linéaire permettant de relier les espaces des obs. et du modèle.

- Incertitudes** sur les états :

$\mathbf{P}_{j-1 \rightarrow j}^f$: matrice ($n \times n$) de var.-covar. des erreurs de prévision ($\varepsilon_j^f = \mathbf{x}_{j-1 \rightarrow j}^f - \mathbf{x}_j^t$) ;

\mathbf{P}_j^a : matrice de var.-covar. des erreurs d'analyse ($\varepsilon_j^a = \mathbf{x}_j^a - \mathbf{x}_j^t$) ;

\mathbf{R}_j : matrice de var.-covar. des erreurs d'observation ($\varepsilon_j^o = \mathbf{y}_j^o - \mathcal{H}(\mathbf{x}_j^t)$).

Formalisme (fin) & BLUE

• Modèles linéaires :

$\mathbf{M}_{j \rightarrow j+1}$ représente le modèle linéaire tangent.

$\mathbf{M}_{j \rightarrow j+1}^*$ représente l'adjoint de $\mathbf{M}_{j \rightarrow j+1}$.

\mathbf{H}_j représente la forme linéarisée pour \mathcal{H}_j .

• EKF : mise à jour (BLUE : *Best Linear Unbiased Estimator*)

$$\mathbf{K}_j = \mathbf{P}_{j-1 \rightarrow j}^f \mathbf{H}_j^T (\mathbf{H}_j \mathbf{P}_{j-1 \rightarrow j}^f \mathbf{H}_j^T + \mathbf{R}_j)^{-1}$$

$$\mathbf{x}_j^a = \mathbf{x}_{j-1 \rightarrow j}^f - \mathbf{K}_j (\mathbf{y}_j^o - \mathcal{H}_j \mathbf{x}_{j-1 \rightarrow j}^f)$$

$$\mathbf{P}_j^a = \mathbf{P}_{j-1 \rightarrow j}^f - \mathbf{K}_j \mathbf{H}_j \mathbf{P}_{j-1 \rightarrow j}^f$$

• EKF : propagation

$$\mathbf{x}_{j+1}^f = \mathcal{M}_{j \rightarrow j+1} (\mathbf{x}_j^a)$$

$$\mathbf{P}_{j \rightarrow j+1}^f = \mathbf{M}_{j \rightarrow j+1} \mathbf{P}_j^a \mathbf{M}_{j \rightarrow j+1}^*$$

Simplification des notations

- Simplifions la notation en nous plaçant à t_{obs} , par exemple :

$$\begin{aligned} P_{obs}^a &= A \\ P_{obs-1 \rightarrow obs}^f &= B \\ P_{obs \rightarrow vf}^f &= P_{vf} \end{aligned}$$

L'assimilation variationnelle

- L'analyse \mathbf{x}^a vérifie le BLUE et est obtenue en minimisant une fonction quadratique d'écart de \mathbf{x} à l'ébauche \mathbf{x}^b et aux observations (fonction optimale du variationnel) :

$$J(\mathbf{x}) = \frac{1}{2} (\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}^b) + \frac{1}{2} (\mathcal{H}(\mathbf{x}^b) - \mathbf{y}^o)^T \mathbf{R}^{-1} (\mathcal{H}(\mathbf{x}^b) - \mathbf{y}^o).$$

- En pratique, on effectue le changement de variable $\chi = \mathbf{L}^{-1} (\mathbf{x} - \mathbf{x}^b)$ pour minimiser la fonction :

$$J(\chi) = \frac{1}{2} \chi^T \chi + \frac{1}{2} (\mathbf{H}\mathbf{L}\chi - \mathbf{y}^o)^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{L}\chi - \mathbf{y}^o).$$

- Dans le variationnel, la matrice \mathbf{A} n'est pas calculée explicitement. Cependant, la Hessienne J , ou dérivée seconde, donne une estimation de l'inverse de \mathbf{A} :

$$\nabla \nabla J(\mathbf{x}) = \mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \simeq \mathbf{A}^{-1}.$$

- Dans l'espace de χ la Hessienne devient $\nabla \nabla J(\chi) = \mathbf{I} + \mathbf{L}^T \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}\mathbf{L}$.

Quelle formulation du problème ?

- On introduit une distinction entre le réseau d'observation de routine \mathbf{y}_r^o et le réseau d'observation étendu \mathbf{y}_s^o .
- Modifier \mathbf{y}_{obs}^o (nouvelle dimension $p_r + p_s$) :

$$\mathbf{y}_{rs}^o = \begin{pmatrix} \mathbf{y}_r^o \\ \mathbf{y}_s^o \end{pmatrix}$$
- Minimiser/maximiser une fonction sur \mathbf{P}_{vf} .

Il faut donc évaluer \mathbf{P}_{vf} , **sous une forme ou une autre**, pour les 2 réseaux \mathbf{y}_r^o et \mathbf{y}_{rs}^o .

On remarque que l'estimation de

$$\mathbf{A}_{rs} = \mathbf{B}_r - \mathbf{K}_{rs} \mathbf{H}_{rs} \mathbf{B}_r$$

et de

$$\mathbf{K}_{rs} = \mathbf{B}_r \mathbf{H}_{rs}^T (\mathbf{H}_{rs} \mathbf{B}_r \mathbf{H}_{rs}^T + \mathbf{R}_{rs})^{-1}$$

ne requièrent pas la connaissance de \mathbf{y}_{rs}^o , mais seulement de \mathbf{H}_{rs} et de \mathbf{R}_{rs} .

Les points difficiles

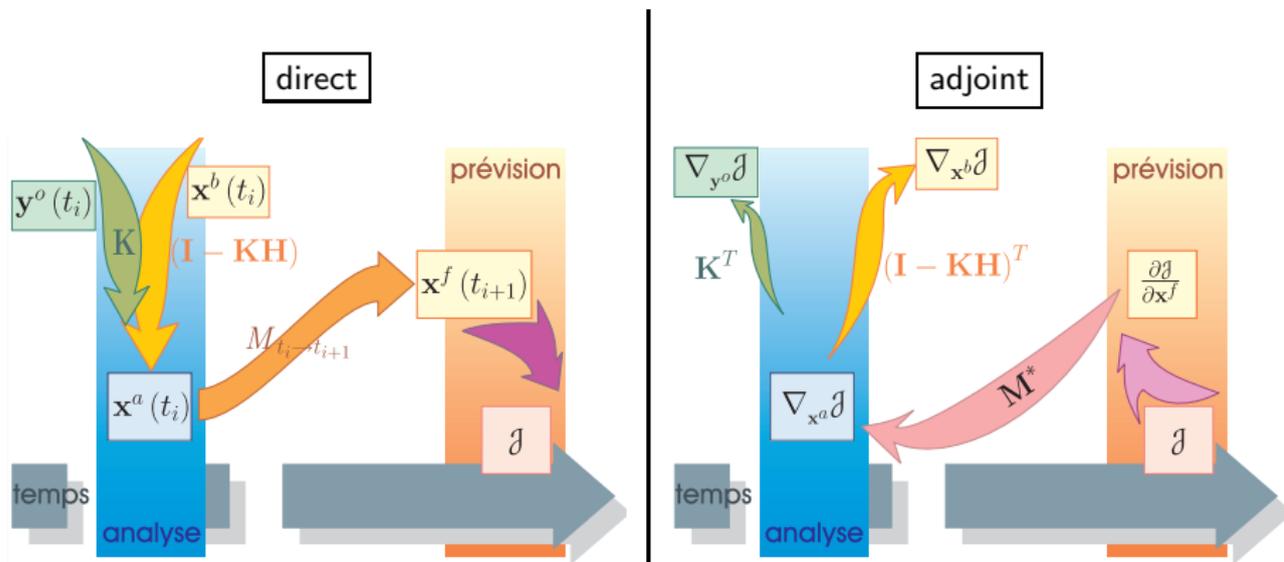
- L'inconnue \mathbf{y}_s^o joue sur la dimension p_{rs} .
 - On introduit une contrainte sur le réseau supplémentaire pour que p_{rs} soit connu.
-
- Les calculs se font entre t_{ana} et t_{obs} : \mathbf{y}_r^o est encore inconnu, lui aussi.
 - Peut-on faire une prévision raisonnable du réseau de routine \mathbf{H}_r et de \mathbf{R}_r ?
-
- Les dimensions des matrices correspondant aux systèmes opérationnels sont trop grandes pour être calculées explicitement : $n \simeq \mathcal{O}(10^7)$
 - On cherchera à travailler dans un sous-espace "sélectionné" pour diminuer la dimension du problème.

Algorithmes

- Il existe 2 approches :
 - ▶ les méthodes ensemblistes : ET-KF (**E**nsemble-**T**ransform **K**alman-**F**ilter);
 - ▶ les méthodes de l'adjoint :
 - ▶ la KFS (**K**alman **F**ilter **S**ensitivity);
 - ▶ la HRRE (**H**essian **R**educed **R**ank **E**stimate).

Sensibilité & KFS (1/2)

- La KFS est très proche de la sensibilité (de la prévision) aux valeurs observées (FSO) : elle utilise l'adjoint de l'opérateur d'assimilation.
- La KFS pourrait ainsi être appelée sensibilité (de l'incertitude de la prévision) au réseau d'observation.



Sensibilité & KFS (2/2)

- La sensibilité aux conditions initiales de la prévision :

$$\nabla_{\mathbf{x}} \mathcal{J} = \mathbf{M}^* \frac{\partial \mathcal{J}}{\partial \mathbf{x}^f}$$

$$\begin{aligned} \delta \mathcal{J} &= \left\langle \frac{\partial \mathcal{J}}{\partial \mathbf{x}^f}; \delta \mathbf{x}^f \right\rangle_{E_{x_1}} \\ &= \left\langle \frac{\partial \mathcal{J}}{\partial \mathbf{x}^f}; \mathbf{M} \delta \mathbf{x} \right\rangle_{E_{x_1}} \\ &= \left\langle \mathbf{M}^* \frac{\partial \mathcal{J}}{\partial \mathbf{x}^f}; \delta \mathbf{x} \right\rangle_{E_{x_0}} \end{aligned}$$

- La sensibilité aux observations (**R** et **A** sont auto-adjoints) :

$$\begin{aligned} \frac{\partial \mathcal{J}}{\partial \mathbf{y}^o} &= \mathbf{K}^* \mathbf{M}^T \frac{\partial \mathcal{J}}{\partial \mathbf{x}^f} \\ &= \mathbf{R}^{-1} \mathbf{H} \mathbf{A} \nabla_{\mathbf{x}} \mathcal{J} \end{aligned}$$

$$\begin{aligned} \delta \mathcal{J} &= \left\langle \frac{\partial \mathcal{J}}{\partial \mathbf{x}^f}; \mathbf{M} \mathbf{A} \mathbf{H}^T \mathbf{R}^{-1} \delta \mathbf{y}^o \right\rangle_{E_{y_0}} \\ &= \left\langle \mathbf{R}^{-1} \mathbf{H} \mathbf{A} \mathbf{M}^* \frac{\partial \mathcal{J}}{\partial \mathbf{x}^f}; \delta \mathbf{y}^o \right\rangle_{E_{y_0}} \end{aligned}$$

Définition d'une erreur de prévision

- La KFS (Bergot and Doerenbecher, 2002) mesure une variation (réduction) de la variance de l'erreur sur la prévision d'une fonction qui est un score théorique.
- Le score \mathcal{S} est l'écart entre la valeur d'une fonction (\mathcal{J}) de la prévision (\mathbf{x}^f) et la valeur non connue qu'elle aurait dans la réalité (\mathbf{x}^t).

la fonction : $\mathcal{J}(\mathbf{x}^f)$

le score : $\mathcal{S} = \mathcal{J}(\mathbf{x}^f) - \mathcal{J}(\mathbf{x}^t)$

- Pour l'observation adaptative, on s'intéresse à la variation de la variance de ce score : $\Delta \text{Var}(\mathcal{S})$.
- Hypothèses nécessaires :
 - ▶ modèle parfait,
 - ▶ analyse non biaisée,
 - ▶ hypothèse linéaire-tangente.

La variance des erreurs

- On peut montrer que la variance du score \mathcal{S} s'écrit :

$$\begin{aligned}
 \text{Var}(\mathcal{S}) &= \nabla_x \mathcal{J}^T \mathbf{A} \nabla_x \mathcal{J} \\
 &= \frac{\partial \mathcal{J}}{\partial \mathbf{x}^f}{}^T \mathbf{MAM}^* \frac{\partial \mathcal{J}}{\partial \mathbf{x}^f} \\
 &= \frac{\partial \mathcal{J}}{\partial \mathbf{x}^f}{}^T \mathbf{P}^f \frac{\partial \mathcal{J}}{\partial \mathbf{x}^f}
 \end{aligned}$$

- La réduction de variance correspond à la mise à jour de \mathbf{A} lorsqu'on ajoute des observations, comme dans un schéma d'assimilation séquentiel.

$$\Delta \text{Var}(\mathcal{S}) = \nabla_x \mathcal{J}^T \mathbf{A}_r \mathbf{H}_s^T (\mathbf{R}_s + \mathbf{H}_s \mathbf{A}_r \mathbf{H}_s^T)^{-1} \mathbf{H}_s \mathbf{A}_r \nabla_x \mathcal{J}$$

Les étapes du calcul de la K_{FS}

- Calcul de la variation de $Var(S)$:

- ▶ Sensibilité aux conditions initiales : $\nabla_x \mathcal{J} = \mathbf{M}^* \frac{\partial \mathcal{J}}{\partial \mathbf{x}^T}$;
- ▶ *Filtrage* de $\nabla_x \mathcal{J}$ par une estimation de \mathbf{A} : $\mathbf{A} \nabla_x \mathcal{J}$ (comme dans la sensibilité aux observations) ;
- ▶ Produit scalaire avec $\nabla_x \mathcal{J}$;
- ▶ Différence entre les produits scalaires : $\Delta Var(S) = \nabla_x \mathcal{J}^T \mathbf{A}_r \nabla_x \mathcal{J} - \nabla_x \mathcal{J}^T \mathbf{A}_{rs} \nabla_x \mathcal{J}$.

- Le calcul de $\mathbf{z} = \mathbf{A} \nabla_x \mathcal{J}$ se fait dans le cadre d'une assimilation variationnelle 4D.
- On utilise un algorithme de minimisation par gradient conjugué pour estimer l'inverse de la Hessienne $\nabla \nabla J^{-1}$.
- La routine CONGRAD du code IFS/ARPEGE calcule \mathbf{z} tel que

$$(\mathbf{I} + \mathbf{L}^T \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{L}) \mathbf{z} = \mathbf{L} \mathbf{z}_a$$

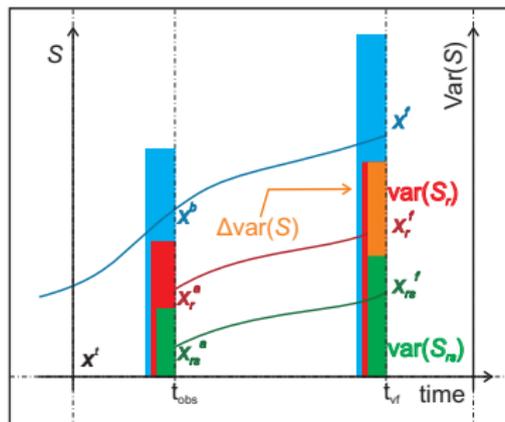
où $\mathbf{z}_a = \nabla_x \mathcal{J}$ (Cardinali, 2009).

Implémentation

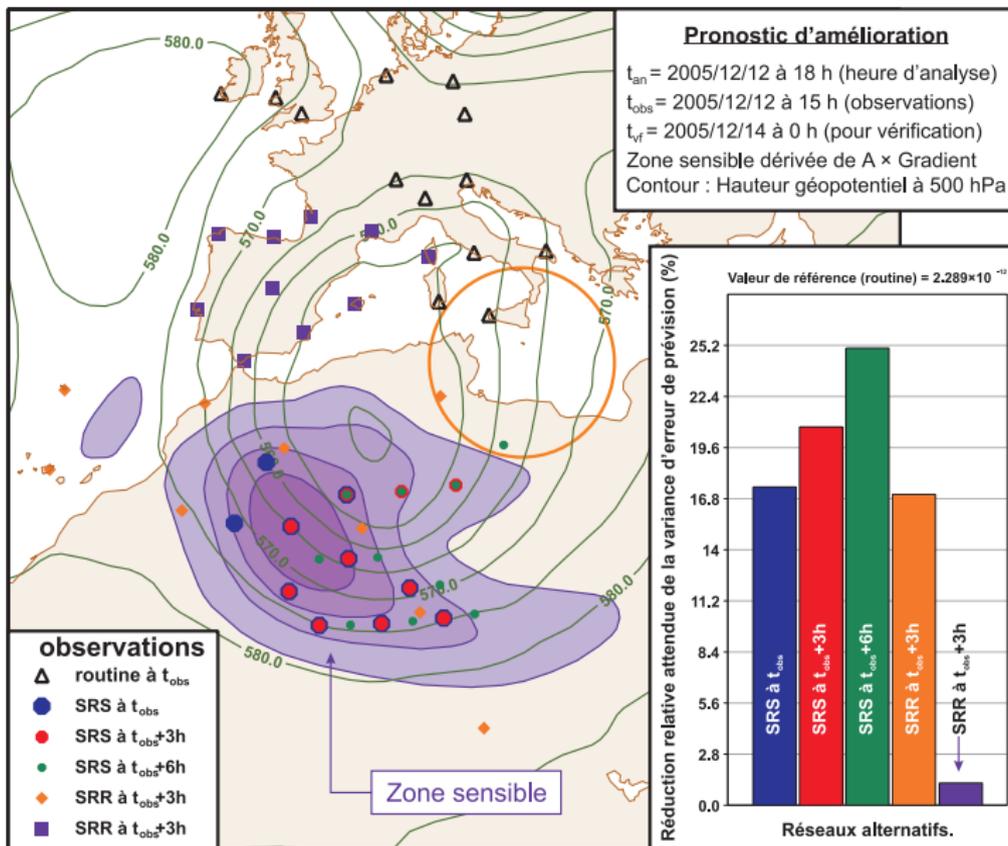
• Dans la pratique, on fait 2 assimilations :

- ▶ Une assimilation avec les observations de routine **uniquement**, durant laquelle on calcule la variance de S_r : $Var(S_r) = \nabla_x \mathcal{J}^T \mathbf{A}_r \nabla_x \mathcal{J}$
- ▶ Une seconde assimilation avec le réseau modifié (routine + adaptatif) pour lequel on calcule la variance de S_{rs} : $Var(S_{rs}) = \nabla_x \mathcal{J}^T \mathbf{A}_{rs} \nabla_x \mathcal{J}$,

On déduit aisément $\Delta Var(S)$.



Exemple de résultat.



Conclusions & remarques (1/2)

● Pour toutes les techniques :

1 Aucune technique ne calcule explicitement les matrices de var.-covariance des erreurs ; toutes utilisent un sous espace pour effectuer leurs calculs.

2 Ce sont des approches statistiques.

- ▶ Réduction de la variance d'erreur de prévision pour éviter les grandes erreurs en moyenne.
- ▶ Pas de garantie d'amélioration de la prévision à tous les coups.
- ▶ Le jeu d'observation supplémentaire est local, l'optimisation reste globale.

3 Quelle certitude sur les incertitudes ?

- ▶ Les techniques font l'impasse sur le processus d'assimilation entre t_{ana} et t_{obs} .
- ▶ Quels effets sur la trajectoire $\mathbf{x}_{ana \rightarrow vf}^f$?
- ▶ Quels effets sur l'incertitude $\mathbf{P}_{ana}^f \rightarrow \mathbf{P}_{vf}^f$?

4 Comment prendre davantage d'inconnues/contraintes en compte, e.g. nombre d'observation p_s ?

Conclusions & remarques (2/2)

● Remarques :

- ▶ Système (assimilation/prévision/cas) dépendant ;
- ▶ Coûts numériques variés (ensemble, vecteurs singuliers) ;
- ▶ Coûts d'implémentation (avions, RS, satellites) ;
- ▶ Comparaison et preuve d'efficacité difficiles en pratique ; les techniques sont pourtant apparentées en théorie.

● Limitations :

- ▶ Taille de l'ensemble,
- ▶ Limites de l'hypothèse linéaire tangente (cf.4D-Var incrémental).
- ▶ ...

● Perspectives :

- ▶ L'observation adaptative comme aide à l'assimilation adaptative (gestion du flux des données)
- ▶ ...

Références

Bergot T. and A. Doerenbecher, 2002.

A study on the optimization of the deployment of targeted observations using adjoint-based methods.

Quart. J. Roy. Meteor. Soc., **128**, 1689–1712.

Bishop, C.H., B.J. Etherton and S.J. Majumdar, 2001.

Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I : Theoretical Aspects.
Mon. Wea. Rev., **129**, 420 – 436.

Cardinali C., 2009.

Monitoring the observation impact on the short range forecast.

Quart. J. Roy. Meteor. Soc., **135**, 239–250.

Leutbecher M., 2003.

A reduced rank estimate of forecast error changes due to intermittent modifications of the observing network.

J. Atmos. Sci., **60**, Vol.5, 729–742.

Majumdar S.J., C.H. Bishop, B.J. Etherton, I. Szunyogh and Z. Toth, 2001.

Can an ensemble transform Kalman filter predict the reduction in forecast-error variance produced by targeted observations ?

Quart. J. Roy. Meteor. Soc., **127**, 2803 – 2820.

Majumdar S.J., C.H. Bishop and B.J. Etherton, 2002.

Adaptive Sampling with the Ensemble Transform Kalman Filter. Part II : Field Program Implementation.

Mon. Wea. Rev., **130**, 1356 – 1369.

...vos questions sont bienvenues

Approche ensembliste (1/4)

- Soit un ensemble de N_{ens} prévisions (ou membres \mathbf{x}_k).
- \mathbf{X} est la matrice ($N_{ens} \times n$) dont les colonnes sont les perturbations $\delta\mathbf{x}_k$ de l'ensemble par rapport à sa moyenne $\bar{\mathbf{x}}$ ($\delta\mathbf{x}_k = \mathbf{x}_k - \bar{\mathbf{x}}$).
- L'ensemble, initialisé à t_{ana} , constitue un échantillon de la fonction de densité de probabilité (PDF) des erreurs de prévisions.
À l'instant t_{vf} on a : $\mathbf{P}_{vf} = \mathbf{X}\mathbf{X}^T$.

➤ Il faut tenir compte de l'assimilation des données à t_{obs} .

- L'ET-KF détermine une transformée \mathbf{T} , associée à un réseau d'observation \mathbf{H}_r , tel que $\mathbf{A}_r = \mathbf{X}\mathbf{T}_r\mathbf{T}_r^T\mathbf{X}^T = \mathbf{Z}_r\mathbf{Z}_r^T$.
- On utilise une hypothèse d'assimilation séquentielle des réseaux de routine et supplémentaire (observations non corrélées entre elles).
- La variance du signal \mathbf{S} , lié à l'ajout des observations \mathbf{H}_s , s'écrit $\mathbf{S}(t_{obs}) = \mathbf{A}_r - \mathbf{A}_{rs}$

Bishop *et al.*, 2001

Majumdar *et al.*, 2001 et 2002

Approche ensembliste (2/4)

- La variance du signal \mathbf{S} s'écrit :

$$\mathbf{S}(t_{obs}) = \mathbf{Z}_r \mathbf{C}_s \mathbf{\Gamma}_s (\mathbf{\Gamma}_s + \mathbf{I})^{-1} \mathbf{C}_s^T \mathbf{Z}_r^T$$
 où $\mathbf{Z}_r = \mathbf{X} \mathbf{T}_r$ et \mathbf{C}_s et $\mathbf{\Gamma}_s$ sont les valeurs et vecteurs propres de $\mathbf{Z}_r \mathbf{H}_s^T \mathbf{R}_s \mathbf{H}_s \mathbf{Z}_r^T$.
- La propriété

$$\mathbf{M}_{obs \rightarrow vf} \delta \mathbf{x}_k(t_{obs}) = \mathcal{M}_{obs \rightarrow vf}(\mathbf{x}_k(t_{obs})) - \bar{\mathbf{x}}(t_{vf})$$
 dispense de l'usage de \mathbf{M} .
 La variance du signal \mathbf{S} peut être calculée à peu de frais pour tout instant $t \geq t_{obs}$.
- La variance du signal \mathbf{S} est une grande matrice de dimension $(n \times n)$, on calcule sa diagonale *diag*(\mathbf{S}) ou sa trace.
- *diag*(\mathbf{S}) permet de tracer des cartes de signal (voir figure).
- $\sigma_s = Tr[\mathbf{P}_{\Sigma_{vf}} \mathbf{S}(t_{vf})]$ permet de prévoir la réduction de la variance des erreurs de prévision à l'instant t_{vf} dans le domaine Σ_{vf} obtenue par l'ajout d'observations à t_{obs} (voir figure).

ETKF (3/4) : réduction de variance

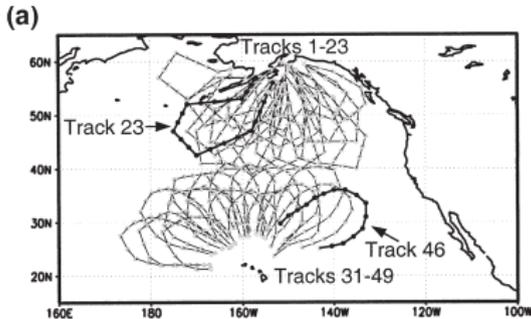
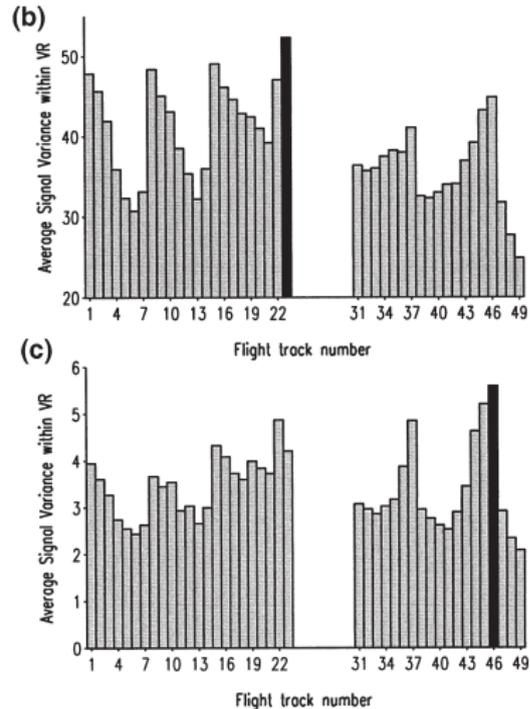
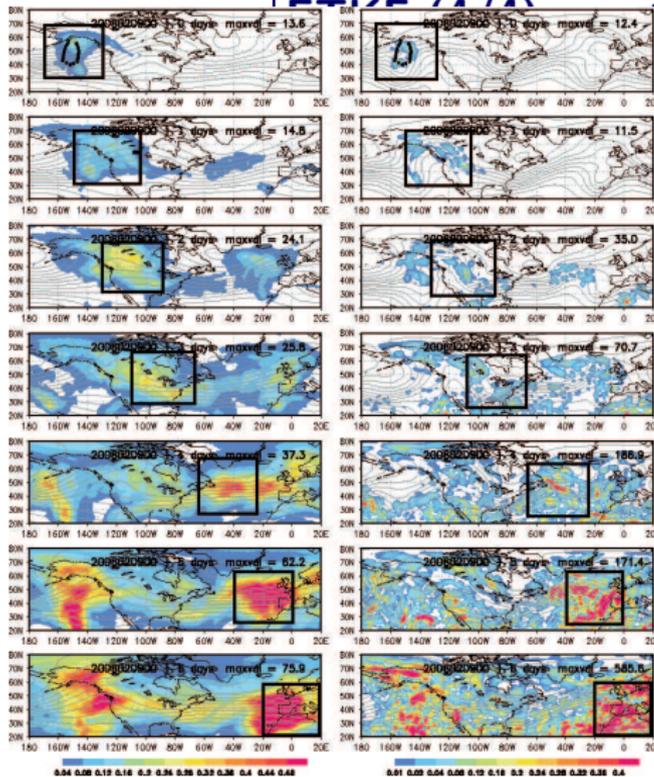


FIG. 6. (a) Map of all predesigned flight paths used during the 2000 WSR program. Noflights were officially assigned numbers 24–30. (b) Bar chart of σ^v for all preselected flight tracks, targeting time 0000 UTC 23 Jan 2000, and verification time 0000 UTC 26 Jan 2000. Verification region of radius 1000 km is centered at 30°N, 90°W. (c) Bar chart of σ^v for the same set of flight tracks, having already assimilated observations from flight track 23 at the targeting time.



Étapes de la variance du signal



Sellwood *et al.*, 2008

Figure 1. ETKF predicted signal variance (left) and squared GFS signal (right) of 0–144 hour forecasts, for dropwindsonde data assimilated on 9 February 2006. The dropwindsonde locations are marked by the black dots in the topmost panels. The black squares in each panel depict the 40° by 40° verification regions. This figure is available in colour online at www.interscience.wiley.com/journal/qj

La Hessienne de rang réduit (1/2)

- La technique HRRE utilise un jeu de vecteurs pour décrire une version de rang réduit de la matrice \mathbf{A}_r^{-1} . Cette technique fonctionnent avec un 4D-Var quasi-opérationnel.
- La HRRE est dite *Hessian* car elle utilise la Hessienne de la fonction $\nabla\nabla J$ du variationnel pour définir la norme initiale de vecteurs singuliers (optimisés sur $[t_{obs}; t_{vf}]$).

$$\mathbf{M}^T \boldsymbol{\varepsilon} \mathbf{M} \mathbf{V} = \boldsymbol{\Sigma} \mathbf{A}_r^{-1} \mathbf{V}$$

- À t_{vf} , les SV de la Hessienne sont les vecteurs propres de la matrice de variance-covariance des erreurs de prévisions : \mathbf{P}_{vf} . Les valeurs singulières sont les valeurs propres de \mathbf{P}_{vf} .

$$\mathbf{M} = \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^T \mathbf{A}_r^{-1}$$

- On ne calcule que les k ($k \simeq 10$) premiers vecteurs singuliers. La trace de la matrice diagonale $\boldsymbol{\Lambda}_r$ donne une estimation de variance d'erreur de prévision.

La Hessienne de rang réduit (2/2)

- Les directions des vecteurs singuliers fixent le sous-espace de travail.
- L'ajout des observations supplémentaires constitue une transformation Γ_s dans ce sous-espace. Les valeurs singulières sont réduites : $\Lambda_{rs} = \Gamma_s^T \Lambda_r \Gamma_s$.
- La réduction de variance d'erreur de prévision est finalement la différence des 2 traces.

Leutbecher M., 2003.