

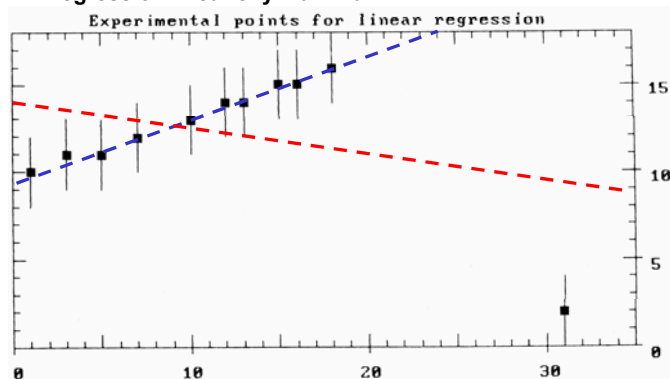
Contrôle de qualité

Introduction

- **Nature des données reçues par les centres opérationnels**
 - * Grande variété de données provenant de plusieurs sources différentes
 - * Plusieurs problèmes peuvent corrompre les données
- **Données incorrectes peuvent avoir un impact significatif sur l'analyse**
- **Acquisition de données et le contrôle de qualité**
 - * Réception des données
 - * Vérifie la qualité des données et rejet de celles ayant une forte probabilité d'être en erreur

Exemple: ajustement par une méthode des moindres carrés en présence d'une donnée en erreur grossière (tiré de Tarantola)

- Régression linéaire: $y = ax + b$

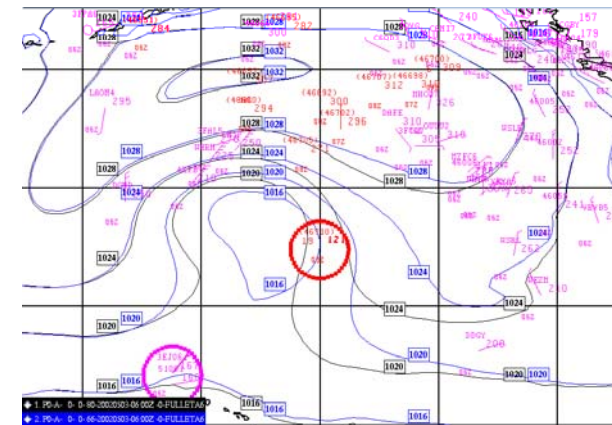


Impact sur l'analyse d'une seule donnée en erreur

Rapport d'une bouée dérivante: $p = 1012.1 \text{ hPa}$ (10 hPa trop basse)

Analyse avec QC en noir

Analyse sans QC en bleu



Contrôle de qualité

• **Sources d'erreurs :**

- * Erreurs de mesures associées à l'instrument
- * Erreur de représentativité
- * Instruments mal calibrés
- * Erreur d'inscription de l'observation
- * Erreurs dans la transmission des données

• **Objectif :**

- * Rejeter toutes les erreurs autres que celles associées à l'erreur de mesure
- * Identificateurs associés à chaque observation durant le processus d'assimilation

5

Contrôle de qualité: rejets grossiers basé sur les valeurs extrêmes climatologiques

Limit values for surface temperature								
Area	Winter				Summer			
	Min2	Min1	Max1	Max2	Min2	Min1	Max1	Max2
45°S - 45°N	-40°C	-30°C	+50°C	+55°C	-30°C	-20°C	+50°C	+60°C
45°N - 90°N	-90°C	-80°C	+35°C	+40°C	-40°C	-30°C	+40°C	+50°C
45°S - 90°S								

Limit values for surface dew-point temperature								
Area	Winter				Summer			
	Min2	Min1	Max1	Max2	Min2	Min1	Max1	Max2
45°S - 45°N	-45°C	-35°C	+35°C	+40°C	-35°C	-25°C	+35°C	+40°C
45°N - 90°N	-99°C	-85°C	+30°C	+35°C	-45°C	-35°C	+35°C	+40°C
45°S - 90°S								

6

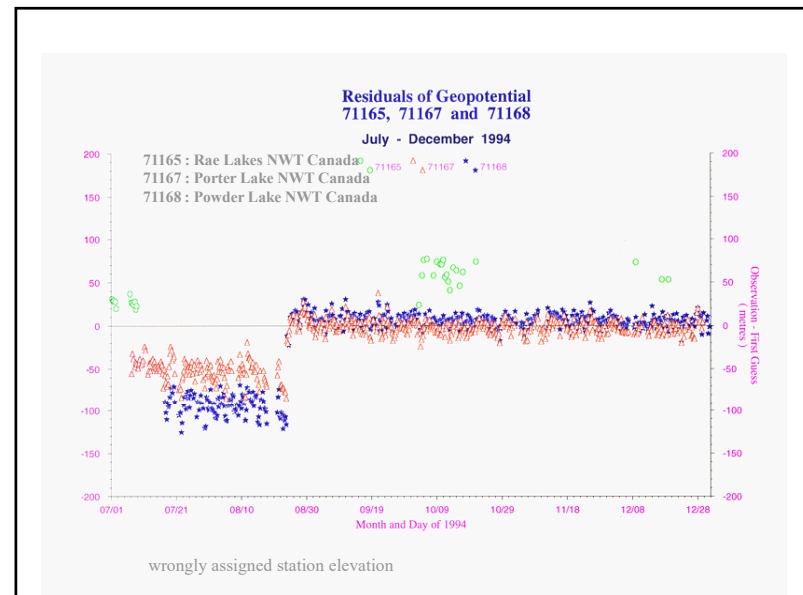
Information contenue dans les innovations

• **Vecteur d'innovation: $y - Hx_b$**

- * Etat d'ébauche obtenue d'une prévision de courte échéance (6-h) contient l'information obtenue l'ensemble des observations assimilées précédemment
- * Comparaison des observations à l'ébauche qui résume notre connaissance *a priori* de l'état de l'atmosphère
- * L'ébauche fournit un état atmosphérique de référence contre lequel on peut comparer toutes les observations

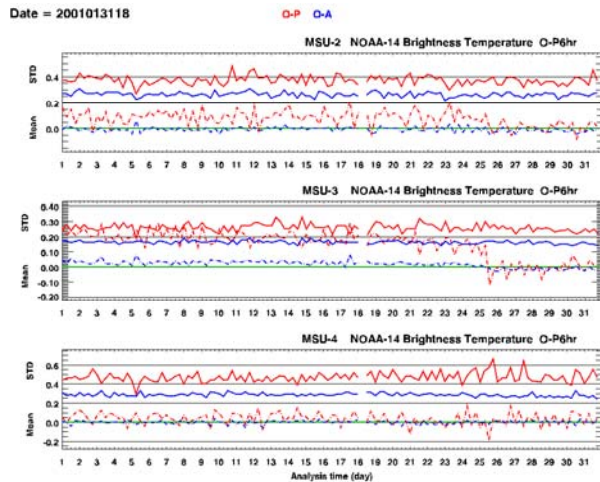
• **Monitoring (surveillance) des observations**

- * innovations sont représentées par types et moyennées sur un grand volume de données, après avoir été regroupées en différentes
- * Permet de détecter si un instrument particulier est défectueux



Monitoring et contrôle de qualité

Statistiques basées sur les innovations ($y - H(x_b)$): exemple extrait des radiances TOVS



Correction de biais (Stéphane Laroche)

Le processus de correction de biais des observations est une étape très importante pour extraire le maximum d'information des observations. En effet, une hypothèse importante de l'interpolation statistique est que les erreurs d'observation et de l'étanche sont sans biais, c'est-à-dire: $\langle \epsilon_o \rangle = 0$ et $\langle \epsilon_b \rangle = 0$. Ceci implique que $\langle y - H(x_b) \rangle = 0$. Le biais peut donc être estimé à partir des innovations obtenus d'un processus de monitoring des observations. Pour mettre en œuvre un schéma de correction de biais, il faut établir des liens entre le biais et un certains nombres de prédicteurs.

Soit b_j le biais d'un type d'observation y_j , l'observation corrigé est donnée par: $y_{jc} = y_j - b_j$

b_j est calculé à partir de prédicteurs p_i et de coefficients A_{ji} comme suit: $b_j = \sum_{i=1}^n A_{ji} p_i + c_j$

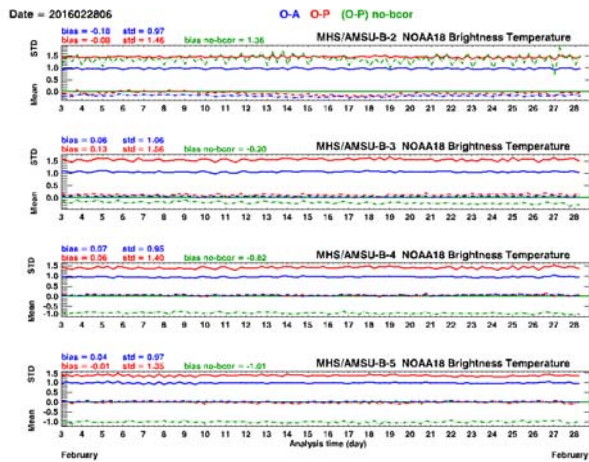
Par exemple, pour les observations de radiances satellitaires, les prédicteurs sont des épaisseurs géopotentielle décrivant différentes parties de l'atmosphère (e.g. 1000-300 hPa, 200-100 hPa, etc.). Il existe plusieurs façon de déterminer les coefficients. À Environment Canada, ces coefficients sont déterminés par régression multilinéaire:

$$A_{ji} = \sum_{k=1}^K [P^T P]_{ki}^{-1} P_k^T (y_j - H_j(x_b))$$

Source: Harris and Kelly, 2001: A satellite radiance-bias correction scheme for data assimilation. *Quart. J. Royal Met. Soc.*, 1453-1468.

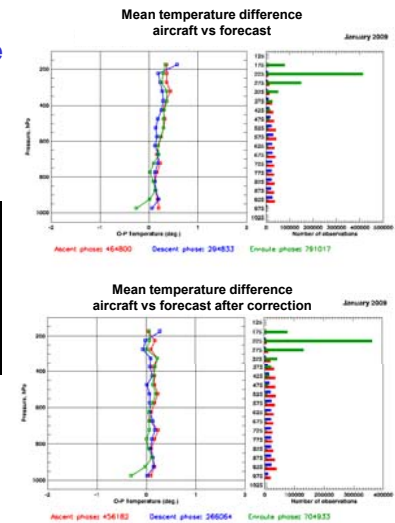
Monitoring et contrôle de qualité

Statistiques basées sur les innovations ($y - H(x_b)$): exemple extrait des radiances ATOVS



Correction de biais des observations de température des rapports d'avion

Layer	Correction
100 hPa < P < 300 hPa	0.5 C
300 hPa < P < 400 hPa	0.3 C
400 hPa < P < 500 hPa	0.2 C
500 hPa < P < 700 hPa	0.1 C



Vérification contre l'ébauche

- Ecart entre l'observation et X_b : $\mathbf{y} - \mathbf{H}(\mathbf{x}_b)$

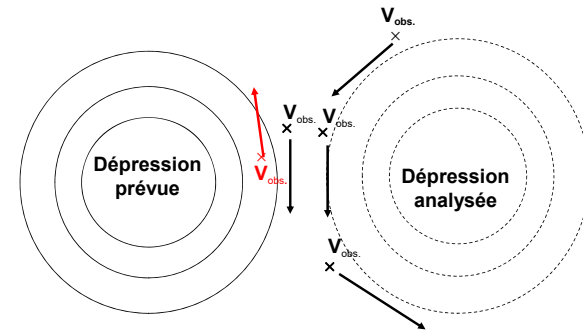
$$\mathbf{y} - \mathbf{H}(\mathbf{X}_b) \cong \varepsilon_o - \mathbf{H}'\varepsilon_b$$

$$\langle (\varepsilon_o - \mathbf{H}'\varepsilon_b)(\varepsilon_o - \mathbf{H}'\varepsilon_b)^T \rangle = \mathbf{R} + \mathbf{H}'\mathbf{B}\mathbf{H}'^T$$

- Pour une seule observation: $(y - H(\mathbf{X}_b))^2 \cong \sigma_o^2 + \sigma_b^2$
 - $\mathbf{H}'\mathbf{B}\mathbf{H}'^T$ est généralement estimé à partir d'une méthode de perturbations
- Observation est rejetée si $\hat{y} = y - \mathbf{H}\mathbf{x}_b \geq \lambda(\sigma_o^2 + \sigma_b^2)^{1/2}$

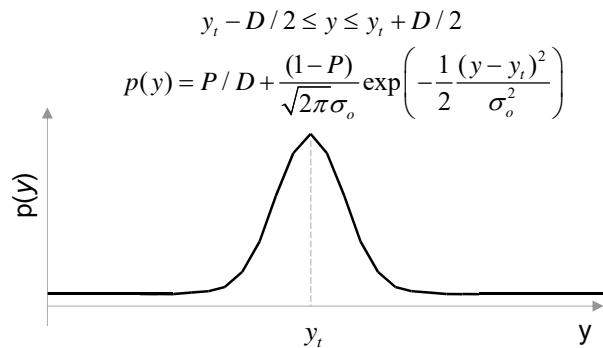
En prenant λ suffisamment grand.

Difficulté associée avec la procédure du "background check"



Contrôle de qualité variationnel (QC-Var)

- Dharssi *et al.* (1992), Ingleby et Lorenc (1993), Andersson et Järvinen (1999)
- Probabilité d'avoir une erreur grossière
 - * Considère que



Distribution conditionnelle de y

- Probabilité de \mathbf{y} si $\mathbf{x} = \mathbf{x}_i$: cas gaussien

$$p(\mathbf{y} | \mathbf{x}) = \frac{1}{C_3} \exp\left\{-\frac{1}{2}(\mathbf{y} - \mathbf{H}(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}(\mathbf{x}))\right\}$$

- Probabilité a posteriori

$$p(\mathbf{x} | \mathbf{y}) = C \frac{\exp\left\{-\frac{1}{2}(\mathbf{y} - \mathbf{H}(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}(\mathbf{x}))\right\} \exp\left\{-\frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b)\right\}}{\exp\left\{-\frac{1}{2}(\mathbf{y} - \mathbf{y}_o)^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{y}_o)\right\}}$$

- Probabilité a posteriori

$$J(\mathbf{x}) = -\ln p(\mathbf{x} | \mathbf{y})$$

$$= \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(\mathbf{H}(\mathbf{x}) - \mathbf{y})^T \mathbf{R}^{-1}(\mathbf{H}(\mathbf{x}) - \mathbf{y}) + C$$

Distribution conditionnelle de y

- Probabilité de y si $\mathbf{x} = \mathbf{x}_t$: cas non-gaussien

$$p(y|x) = P/D + \frac{(1-P)}{\sqrt{2\pi}\sigma_o} \exp\left(-\frac{1}{2} \frac{(y-H(x))^2}{\sigma_o^2}\right)$$

$$-\ln(p(y|x)) = -\ln\left(P/D + \frac{(1-P)}{\sqrt{2\pi}\sigma_o} \exp\left(-\frac{1}{2} \frac{(y-H(x))^2}{\sigma_o^2}\right)\right)$$

$$= -\ln\left(P/D + \frac{(1-P)}{\sqrt{2\pi}\sigma_o} \exp(-J_{ov}(x))\right)$$

QC-Var

Définition de la fonction objective

$$J_o^{QC}(\mathbf{x}) \equiv J_o^{QC}(\hat{y}(\mathbf{x})) = -\ln p(y_o | H(\mathbf{x}))$$

$$= -\ln\left(P/D + C \exp(-J^N(\hat{y}))\right)$$

où

$$J^N(\hat{y}(\mathbf{x})) = \frac{1}{2} \hat{y}^T \mathbf{R}^{-1} \hat{y} \equiv \frac{1}{2} \frac{(H(\mathbf{x}) - y_o)^2}{\sigma_o^2}$$

Gradient de la fonction coût du QC-Var

$$\nabla_{\hat{y}} J_o(\mathbf{x}) = \frac{\exp(-J_o^N)}{\gamma + \exp(-J_o^N)} \nabla_{\hat{y}} J_o^N(\hat{y}) \equiv W_{QC} \nabla_{\hat{y}} J_o^N(\hat{y})$$

$$= W_{QC} \frac{(H(\mathbf{x}) - y_o)}{\sigma_o^2}$$

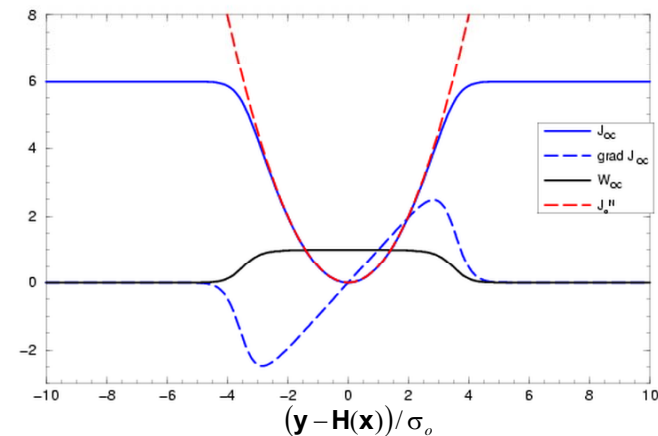
$$\nabla_{\mathbf{x}} J_o(\mathbf{x}) = \left[\frac{\partial H}{\partial \mathbf{x}}(\mathbf{x}) \right]^T W_{QC} \frac{(H(\mathbf{x}) - y_o)}{\sigma_o^2}$$

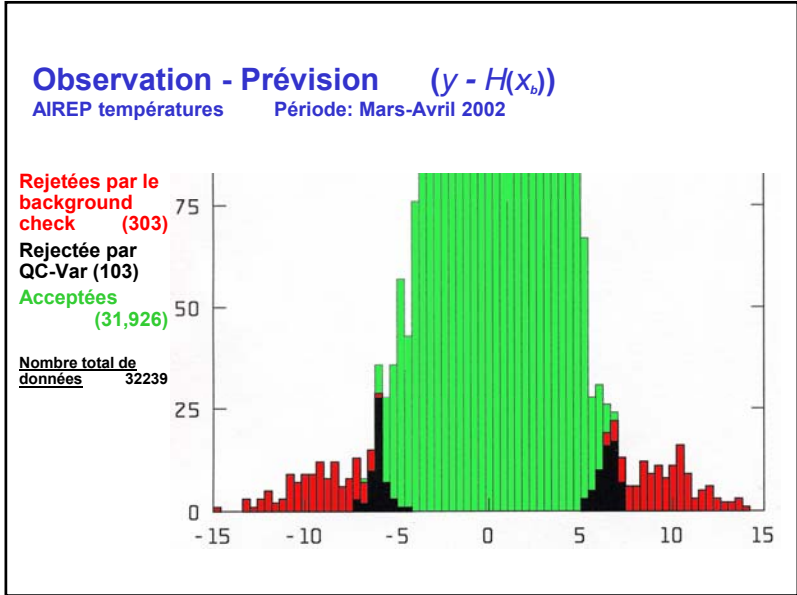
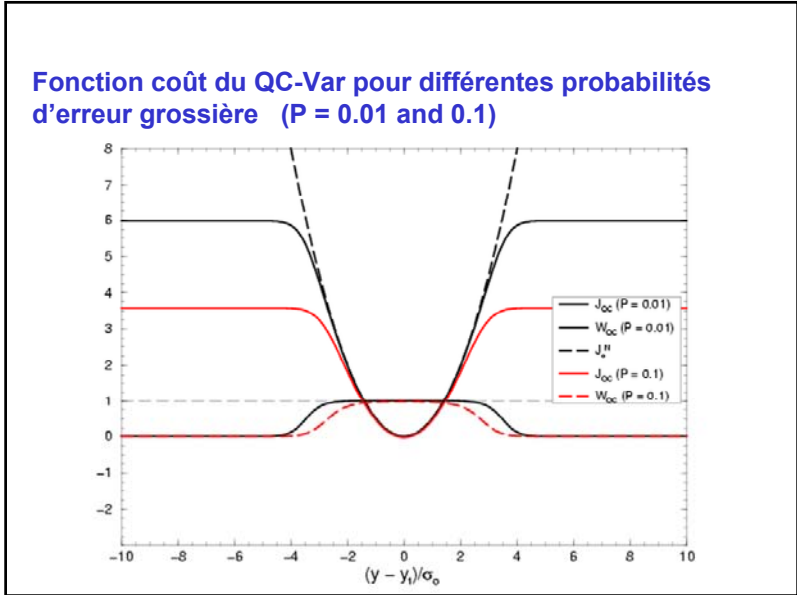
où $\gamma = (P\sigma_o\sqrt{2\pi})/(1-P)D$

W_{QC} dépend de l'état courant de l'atmosphère.

- Poids *a posteriori* sont donc basés sur l'écart par rapport à l'analyse

Représentation de la fonction coût du QC-Var ($P = 0.01$)





Évaluation de l'impact des observations sur une analyse: le contenu en information

UQAM Département des sciences de la terre et de l'atmosphère
 Université du Québec à Montréal

Diagnostic *a posteriori* pour tester l'optimalité

- Talagrand (1998)
 - * Valeur de la fonction objective au minimum $J(x_a) = m/2$ où m = nombre d'observations
 - $$J(x_a) = \frac{1}{2} (x_a - x_b)^T B^{-1} (x_a - x_b) + \frac{1}{2} (Hx_a - y)^T R^{-1} (Hx_a - y)$$

$$= \frac{1}{2} \text{tr} (B^{-1} (x_a - x_b)(x_a - x_b)^T) + \frac{1}{2} \text{tr} (R^{-1} (Hx_a - y)(Hx_a - y)^T)$$
- On a $x_a - x_b = K(y - Hx_b) = Kd$
 $Hx_a - y = (HK - I)d$ Trace d'une matrice
- Les covariances d'erreur *a posteriori* sont alors $D = \langle (\varepsilon_a - H\varepsilon_b)(\varepsilon_a - H\varepsilon_b)^T \rangle = \langle dd^T \rangle = R + HBH^T$
- Les statistiques *a priori* sont celles utilisées par l'assimilation c'est-à-dire R et B. On note alors $D = R + HBH^T$

Diagnostic de J_b

$$J_b(\mathbf{x}) = \frac{1}{2} \text{tr}(\mathbf{B}^{-1}(\mathbf{x}_a - \mathbf{x}_b)(\mathbf{x}_a - \mathbf{x}_b)^T) = \frac{1}{2} \text{tr}(\mathbf{H}^T \mathbf{D}^{-1} \mathbf{D} \mathbf{H} \mathbf{B}) = \frac{1}{2} \text{tr}(\mathbf{D}^{-1} \mathbf{D} \mathbf{D}^{-1} \mathbf{H} \mathbf{B} \mathbf{H}^T)$$

- On obtient ceci en utilisant l'expression de \mathbf{K} et les relations précédentes pour montrer que

$$(\mathbf{x}_a - \mathbf{x}_b)(\mathbf{x}_a - \mathbf{x}_b)^T = \mathbf{K} \mathbf{D} \mathbf{K}^T = \mathbf{B} \mathbf{H}^T \mathbf{D}^{-1} \mathbf{D} \mathbf{D}^{-1} \mathbf{H} \mathbf{B}$$

Diagnostic de J_o

$$J_o(\mathbf{x}) = \frac{1}{2} \text{tr}(\mathbf{R}^{-1}(\mathbf{H}\mathbf{x}_a - \mathbf{y})(\mathbf{H}\mathbf{x}_a - \mathbf{y})^T) = \frac{1}{2} \text{tr}(\mathbf{D}^{-1} \mathbf{D} \mathbf{D}^{-1} \mathbf{R})$$

- La définition de \mathbf{K} permet d'obtenir

$$(\mathbf{H}\mathbf{x}_a - \mathbf{y}) = (\mathbf{H}\mathbf{K} - \mathbf{I})\mathbf{d} = (\mathbf{H}\mathbf{B}\mathbf{H}^T \mathbf{D}^{-1} - \mathbf{I})\mathbf{d} = (\mathbf{H}\mathbf{B}\mathbf{H}^T - \mathbf{R} - \mathbf{H}\mathbf{B}\mathbf{H}^T) \mathbf{D}^{-1} \mathbf{d} = -\mathbf{R} \mathbf{D}^{-1} \mathbf{d}$$

- Donc

$$\text{tr}(\mathbf{R}^{-1}(\mathbf{H}\mathbf{x}_a - \mathbf{y})(\mathbf{H}\mathbf{x}_a - \mathbf{y})^T) = \text{tr}(\mathbf{R}^{-1} \mathbf{R} \mathbf{D}^{-1} \mathbf{D} \mathbf{D}^{-1} \mathbf{R}) = \text{tr}(\mathbf{D}^{-1} \mathbf{D} \mathbf{D}^{-1} \mathbf{R})$$

Et en combinant ces deux résultats

$$J(\mathbf{x}) = \frac{1}{2} \text{tr}(\mathbf{D}^{-1} \mathbf{D} \mathbf{D}^{-1} \mathbf{H} \mathbf{B} \mathbf{H}^T) + \frac{1}{2} \text{tr}(\mathbf{D}^{-1} \mathbf{D} \mathbf{D}^{-1} \mathbf{R}) = \frac{1}{2} \text{tr}(\mathbf{D}^{-1} \mathbf{D} \mathbf{D}^{-1} (\mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R})) = \frac{1}{2} \text{tr}(\mathbf{D}^{-1} \mathbf{D})$$

Si les statistiques *a priori* et *a posteriori* sont cohérentes entre elles alors $\mathbf{D} = \mathbf{D}$ et conséquemment,

$$J(\mathbf{x}) = \frac{M}{2}$$

Si $J(\mathbf{x}_a) < m/2$, ceci implique que les statistiques de l'erreur totale sont surestimées.

Diagnostic de l'information statistique à partir des résultats de l'analyse

- Desroziers (2005), Lupu et al. (2011)

$$\mathbf{d} = \mathbf{y} - \mathbf{H}\mathbf{x}_b \quad \mathbf{a} = \mathbf{y} - \mathbf{H}\mathbf{x}_a \quad \mathbf{d}_a^b = \mathbf{H}(\mathbf{x}_a - \mathbf{x}_b) = \mathbf{H}\mathbf{K}\mathbf{d}$$

alors

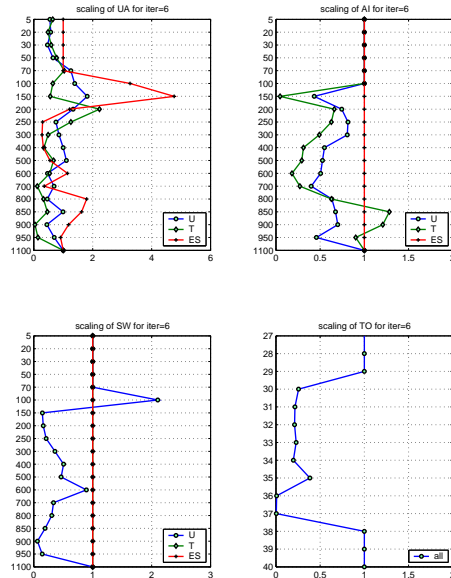
$$\langle \mathbf{d} \mathbf{d}^T \rangle \equiv \tilde{\mathbf{D}} = (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T) \equiv \mathbf{D}$$

$$\langle \mathbf{a} \mathbf{d}^T \rangle \equiv \tilde{\mathbf{R}} = \mathbf{R}(\mathbf{D}^{-1} \tilde{\mathbf{D}})$$

$$\langle \mathbf{d}_a^b \mathbf{d}^T \rangle \equiv \mathbf{H} \tilde{\mathbf{B}} \mathbf{H}^T = \mathbf{H} \mathbf{B} \mathbf{H}^T (\mathbf{D}^{-1} \tilde{\mathbf{D}})$$

$$\langle \mathbf{d}_a^b \mathbf{a}^T \rangle = \mathbf{H} \tilde{\mathbf{P}}_a \mathbf{H}^T = \mathbf{H} \mathbf{B} \mathbf{H}^T \mathbf{D}^{-1} \tilde{\mathbf{D}} \mathbf{D}^{-1} \mathbf{R}$$

Facteur d'ajustement des erreurs d'observation obtenue avec la méthode de Desroziers et Ivanov (Liu *et al.*, 2005)



Approches pour mesurer l'impact des observations assimilées

Contenu en information

- * Basé sur la précision relative des observations par rapport à celle de l'ébauche

Expériences sur les systèmes d'observations (ESO)

(ou *Observing System Experiments, OSE*)

- * Retrait de types d'observations globalement ou sur de grandes régions
- * Donne une vue globale de l'impact des observations sur l'analyse et sur les prévisions qui en résultent

Impact des observations sur la qualité des prévisions

- * Sensibilités des prévisions par rapport aux observations (Baker and Daley, 2000; Langland and Baker, 2003; Gelaro *et al.*, 2010)
- * Méthodes basées sur le filtre de Kalman

Contenu en information

- Rapport entre de la covariance d'erreur d'analyse et B

$$\text{tr}(\mathbf{P}_a \mathbf{B}^{-1}) = \text{tr}(\mathbf{I}) - \text{tr}(\mathbf{K}\mathbf{H}) = N - \text{tr}(\mathbf{K}\mathbf{H})$$

L'information apportée en assimilant un jeu d'observations particulier est associée au deuxième terme et N est la dimension de l'état-modèle

- ... et dans l'espace des observations

$$\mathbf{P}_a \rightarrow \mathbf{H}\mathbf{P}_a\mathbf{H}^T \quad \mathbf{B} \rightarrow \mathbf{H}\mathbf{B}\mathbf{H}^T$$

$$\text{tr}\left(\left(\mathbf{H}\mathbf{P}_a\mathbf{H}^T\right)\left(\mathbf{H}\mathbf{B}\mathbf{H}^T\right)^{-1}\right) = M - \text{tr}(\mathbf{H}\mathbf{K})$$

où M est le nombre d'observations

DFS = Degrees of Freedom per signal

Diagnostics de l'information statistiques à partir des résultats de l'analyse

- Desroziers (2005)

- * Emploie les résultats de l'assimilation pour estimer les covariances d'erreur d'observation, d'ébauche et d'analyse dans l'espace des observations

$$\mathbf{d} = \mathbf{y} - \mathbf{H}\mathbf{x}_b \quad \mathbf{a} = \mathbf{y} - \mathbf{H}\mathbf{x}_a \quad \mathbf{d}_a^b = \mathbf{H}(\mathbf{x}_a - \mathbf{x}_b) = \mathbf{H}\mathbf{K}\mathbf{d}$$

$$\text{et on montre alors,} \quad \langle \mathbf{d}\mathbf{d}^T \rangle \equiv \tilde{\mathbf{D}} = (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T) \equiv \mathbf{D}$$

$$\langle \mathbf{a}\mathbf{d}^T \rangle \equiv \tilde{\mathbf{R}} = \mathbf{R}(\mathbf{D}^{-1}\tilde{\mathbf{D}})$$

$$\langle \mathbf{d}_a^b \mathbf{d}_a^b{}^T \rangle \equiv \mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^T = \mathbf{H}\mathbf{B}\mathbf{H}^T (\mathbf{D}^{-1}\tilde{\mathbf{D}})$$

$$\langle \mathbf{d}_a^b \mathbf{a}^T \rangle = \mathbf{H}\tilde{\mathbf{P}}_a\mathbf{H}^T = \mathbf{H}\mathbf{B}\mathbf{H}^T \mathbf{D}^{-1} \tilde{\mathbf{D}} \mathbf{D}^{-1} \mathbf{R}$$

Estimation du contenu en information (ou Degrees of Freedom per signal)

- En notant que $\tilde{\mathbf{R}} \triangleq \langle \mathbf{a} \mathbf{d}^T \rangle = \mathbf{R}(\mathbf{D}^{-1} \tilde{\mathbf{D}})$ et $\tilde{\mathbf{H}} \mathbf{B} \mathbf{H}^T \triangleq \langle \mathbf{d}'_i \mathbf{d}'_i^T \rangle = \mathbf{H} \mathbf{B} \mathbf{H}^T (\mathbf{D}^{-1} \tilde{\mathbf{D}})$
- Si les statistiques *a priori* et *a posteriori* sont parfaitement cohérentes, alors $\tilde{\mathbf{D}} = \mathbf{D}$ et conséquemment,

$$\tilde{\mathbf{R}} = \mathbf{R} \quad \tilde{\mathbf{H}} \mathbf{B} \mathbf{H}^T = \mathbf{H} \mathbf{B} \mathbf{H}^T$$

- Estimation du DFS

$$DFS = \text{tr}(\mathbf{H} \mathbf{K}) = \text{tr}(\mathbf{H} \mathbf{B} \mathbf{H}^T (\mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H}^T)^{-1}) = \text{tr}(\mathbf{H} \mathbf{B} \mathbf{H}^T \mathbf{D}^{-1})$$

$$\text{tr}(\mathbf{H} \tilde{\mathbf{K}}) = \text{tr}(\tilde{\mathbf{H}} \tilde{\mathbf{B}} \tilde{\mathbf{H}}^T \tilde{\mathbf{D}}^{-1}) = \text{tr}(\mathbf{H} \mathbf{B} \mathbf{H}^T \mathbf{D}^{-1} \tilde{\mathbf{D}} \tilde{\mathbf{D}}^{-1}) = DFS$$

Ceci conduit à la même estimation du contenu en information que celui obtenu à partir des statistiques *a priori*

Estimation du contenu en information

$$D\tilde{F}S_{APOST}^{(1)} = \text{tr}(\tilde{\mathbf{H}} \tilde{\mathbf{B}} \tilde{\mathbf{H}}^T \tilde{\mathbf{D}}^{-1})$$

$$D\tilde{F}S_{APOST}^{(2)} = \text{tr}(\tilde{\mathbf{R}}^{-1} (\tilde{\mathbf{H}} \tilde{\mathbf{P}}_a \tilde{\mathbf{H}}^T))$$

$D\tilde{F}S_{DIAG}$: seulement les termes diagonaux de la deuxième méthode sont retenus

L (km)	DFS_{THEOR}	DFS_{GIRARD}	$D\tilde{F}S_{APOST}^{(1)}$	$D\tilde{F}S_{APOST}^{(2)}$	$D\tilde{F}S_{DIAG}$
300	11.03	10.88	10.81	10.80	10.70
500	9.50	9.37	9.21	9.20	9.07
1000	7.34	7.08	6.79	6.79	6.75

- DFS_{GIRARD} : estimation obtenue d'analyses perturbées
- DFS_{THEOR} : estimation des valeurs exactes utilisées pour générer les observations et les ébauches

Contenu en information dans les analyses 3D-Var et 4D-Var issues du système d'Environnement Canada

- Résultats d'expérience d'assimilation de Laroche and Sarrazin (2010 a,b) sur la période du 21 Décembre 2006 au 28 Février 2007

* Les 11 premiers jours sont exclus (spin-up du cycle d'assimilation)

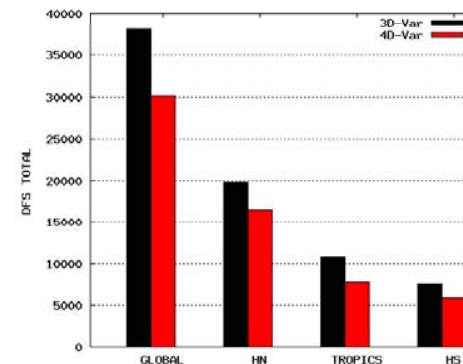
- Observations utilisées

- * Radiosondes, avions, données de surface et de bateaux, profileurs de vent
- * AMV des satellites géostationnaires
- * Radiances d'orbiteurs polaires (AMSU-a,b) et de satellites géostationnaires (GOES-Est et Ouest)

- Diagnostic de la cohérence statistique: $\chi^2 / M \sim 1$

- * Pour le 3D-Var et le 4D-Var, on a trouvé que $\chi^2 / M = 0.56$
- * Statistiques d'erreur du système sont surestimées
- * Desroziers et Ivanov (2001) et Chapnik et al. (2004) utilisent cette information pour recalibrer les statistiques d'erreur
- * Ce n'était pas l'objet de cette étude.

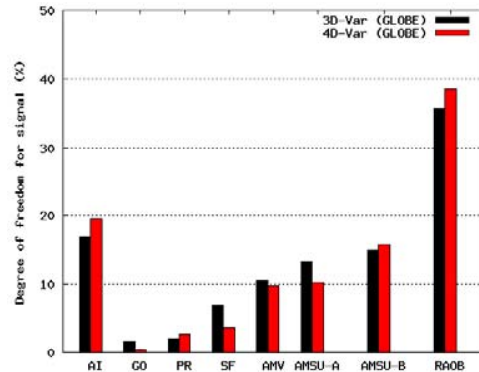
DFS total estimé sur différentes régions pour le 3D-Var et le 4D-Var (Janvier-Février 2007)



Calcul du DFS par type d'observations pour le 3D-Var et le 4D-Var

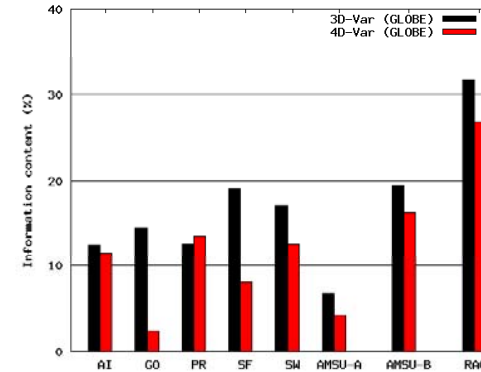
$$DFS_{Obs_type}^{Region} (\%) = 100 \cdot \frac{DFS_{Obs_type}^{Region}}{DFS_{Total_obs}^{Region}}$$

Région : Globe
 Types Obs.: AI, GO, PR, SF, SW, AMSU-A, AMSU-B, RAOB



Lupu et al. (2010)

Impact des observations par observation dans chaque région



$$IC(\%) = 100 \cdot \frac{DFS_k^{Region}}{P_k}$$

Lupu et al. (2010)

Expériences sur les systèmes d'observations

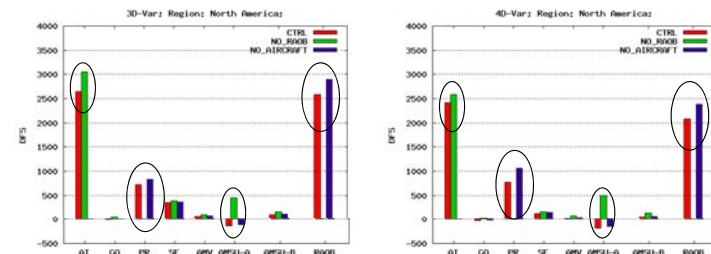
(Observing System Experiments, OSEs)

- **Expériences présentées dans Laroche et Sarrazin (2010-a,b)**
- **Evaluation de l'impact des observations par retrait de types de données**
 - * Analyse comprenant la totalité des observations est prise comme référence
 - * Un type d'observation est retiré et on mesure la dégradation
 - * Modification de la couverture en observations affect l'importance relative des observations
- **Comparaison du contenu en information pour ces expériences donne une image détaillée de la complémentarité entre les types d'observations**

Expériences OSEs : 3D-Var et 4D-Var, Amérique du Nord

$$\frac{DFS_k^{NA}}{P_k}$$

Valeurs du DFS par type d'observations normalisé par le nombre d'observations



NO_RAOB: DFS par observation augmente, particulièrement pour AMSU-A et GO;

NO_AIRCRAFT: DFS par observation augmente, particulièrement pour les RAOB et les PR; Augmentation du DFS par obs. augmente légèrement pour les autres types (GO, SW and AMSU-B) DFS per obs also increases slightly.

Résumé

- **Contenu en information peut être évalué en diagnostiquant les résultats de l'assimilation**
- **Fournit une image détaillée de l'impact des observations dans le cadre où elles sont utilisées**
- **Application aux résultats obtenus d'expérience de type OSEs montre comment l'impact des observations sur les analyses dépend de la couverture et de la nature du réseau d'observations**
- **OSEs par contre donne une mesure de l'impact des observations sur les prévisions qui en résultent**