

Calculs d'impacts des observations sur la réduction d'erreurs de prévision basés sur le dual (PSAS) du système d'analyse 4D-Var

Simon Pellerin

Stéphane Laroche, Josée Morneau, Monique Tanguay,
Jean-François Caron
MRB/MSC

Plan de la présentation

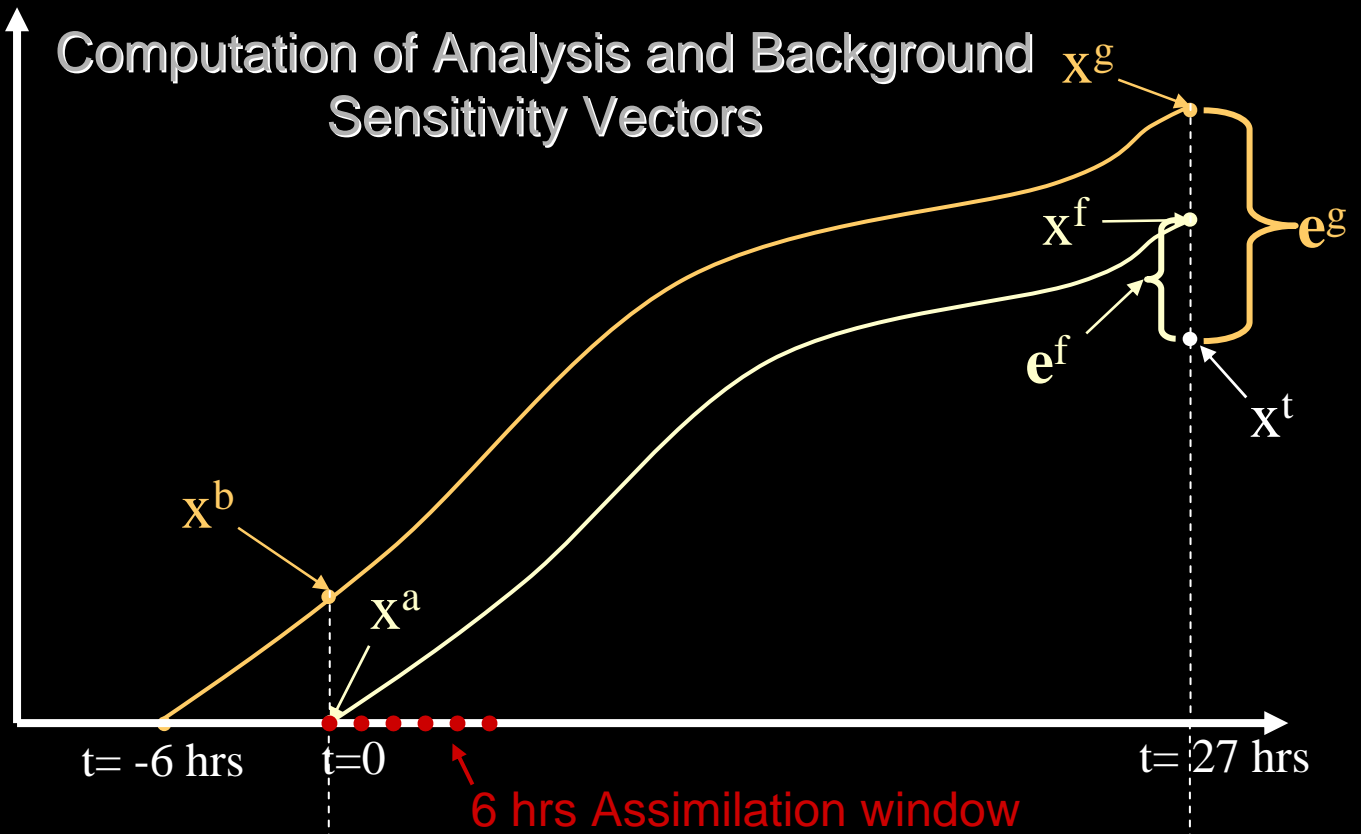
- Impacts des observations sur la réduction des erreurs de prévisions
 - Methodologie
 - Sensibilité dans l'espace des observations & Impact d'observation
 - 4D-Var Dual (PSAS) : Description & Validation
 - 4D-Var Observation System Experiments (OSEs)
 - Motivation
 - Description & resultats
 - Études comparatives d'impact d'observations des systèmes 4D-VAR, 3D-FGAT et 3D-Var
 - Description of experiments
 - Results and comparisons with OSEs
 - Conclusions & travaux futures
- Développement d'une librairie d'analyse objective : Plan de développement

Estimation of Observation Impact based on sensitivity calculation

- Sensitivity analysis system (key analysis) based on the TL and Adjoint of GEM (Global Environmental Multiscale) model (Laroche et al., 2002)
- Method described by Langland and Baker (2004) to calculate the impact of observations on forecast error
- Observation Impact (OI) is estimated using observation sensitivity vector and any subset of the innovation vector
- Unnecessary to add or remove observations (as in OSE)
- A Posteriori monitoring of observation quality control
- The method is used to assess the OI of the MSC operational 4D-Var

Computation of Analysis and Background Sensitivity Vectors

Extension of the adjoint-based technique described in Laroche *et al.*, 2002



$$J^f = \frac{1}{2} \langle (\mathbf{x}^f - \mathbf{x}^t), \mathbf{C}(\mathbf{x}^f - \mathbf{x}^t) \rangle$$

$$J^g = \frac{1}{2} \langle (\mathbf{x}^g - \mathbf{x}^t), \mathbf{C}(\mathbf{x}^g - \mathbf{x}^t) \rangle$$

$$\frac{\partial J^f}{\partial \mathbf{x}^a}$$

$$\frac{\partial J^g}{\partial \mathbf{x}^b}$$

$$\frac{\partial J^f}{\partial \mathbf{x}^f}$$

$$\frac{\partial J^g}{\partial \mathbf{x}^g}$$

\mathbf{L}^T

Estimé de la réduction d'erreur $\Delta e^{f,g}$

Langland and Baker, 2004

Errico, 2007 : re-dérivation

Procédure
d'analyse

Adjoint de la
procédure
d'analyse

Réduction
d'erreur

$$\mathbf{x}^a - \mathbf{x}^b = \mathbf{K} [\mathbf{y} - H(\mathbf{x}^b)]$$

Innovations

$$\Delta e^{f,g} = e^f - e^g \approx \left\langle \mathbf{K} [\mathbf{y} - H(\mathbf{x}^b)], \frac{\partial J^f}{\partial \mathbf{x}^a} + \frac{\partial J^g}{\partial \mathbf{x}^b} \right\rangle = \left\langle \mathbf{y} - H(\mathbf{x}^b), \mathbf{K}^T \left[\frac{\partial J^f}{\partial \mathbf{x}^a} + \frac{\partial J^g}{\partial \mathbf{x}^b} \right] \right\rangle$$

$$\mathbf{L}^T \frac{\partial J^f}{\partial \mathbf{x}^f} + \mathbf{L}^T \frac{\partial J^g}{\partial \mathbf{x}^g}$$

\mathbf{w} : Vecteur de sensibilité dans l'espace des observations (OSV)

Adjoint de l'analyse (K^T)

Différentes Approches Possibles

- Contenu en information de la Hessienne (vecteurs singuliers) : ECMWF 4D-Var (C. Cardinali)
- Adjoint ligne par ligne (analogie: adjoint d'un modèle numérique) : GMAO – 3D-FGAT (R. Gelaro)
- Propriétés self-adjointes (symétries) de B et ($HBH^T + R$) : NRL 3D-PSAS-FGAT (Langland & Baker), MRB/MSC 4D-Var (Pellerin et al.)

OSV solution

$$\Delta e^{f,g} = \langle \mathbf{y} - H(\mathbf{x}^b), \mathbf{w} \rangle \quad \mathbf{w} = \mathbf{K}^T \left[\frac{\partial J^f}{\partial \mathbf{x}^a} + \frac{\partial J^g}{\partial \mathbf{x}^b} \right]$$

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

Using the Kalman gain matrix definition

$$(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})\mathbf{w} = \mathbf{H}\mathbf{B} \left(\frac{\partial J^f}{\partial \mathbf{x}^a} + \frac{\partial J^g}{\partial \mathbf{x}^b} \right)$$

Linear system equation for “w”

Solving PSAS linear system equation for “w”

$$\mathbf{u} \equiv \sqrt{\mathbf{R}} \mathbf{w}$$

$$F(\mathbf{u}) = \frac{1}{2} \mathbf{u}^T (\mathbf{I} + \sqrt{\mathbf{R}^{-1}} \mathbf{H} \mathbf{B} \mathbf{H}^T \sqrt{\mathbf{R}^{-1}}) \mathbf{u} - \mathbf{u}^T \sqrt{\mathbf{R}^{-1}} \mathbf{d}$$

$$\nabla_{\mathbf{u}} F = (\mathbf{I} + \sqrt{\mathbf{R}^{-1}} \mathbf{H} \mathbf{B} \mathbf{H}^T \sqrt{\mathbf{R}^{-1}}) \mathbf{u} - \sqrt{\mathbf{R}^{-1}} \mathbf{d}$$

Courtier, 1997

PSAS Preconditioned objective function

$$\mathbf{d} \equiv \mathbf{H} \mathbf{B} \left(\frac{\partial J^f}{\partial \mathbf{x}^a} + \frac{\partial J^g}{\partial \mathbf{x}^b} \right) \quad \Delta e^{f,g} = \langle \mathbf{y} - H(\mathbf{x}^b), \mathbf{w} \rangle$$

Observation impact

K and K^T consistency

4D-Var vs. 4D-PSAS DA

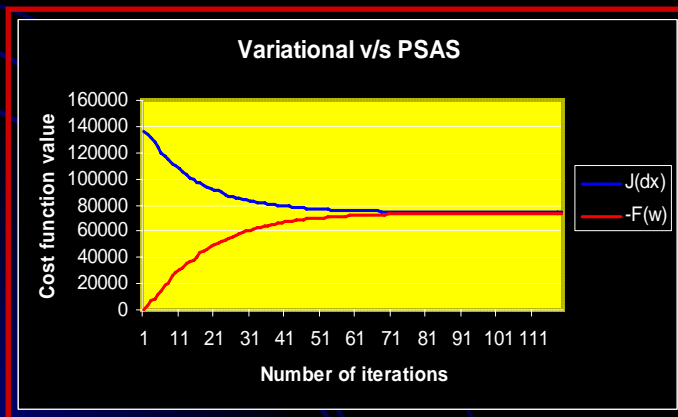
$$\Delta e^{f,g} = \left\langle \mathbf{K} [\mathbf{y} - H(\mathbf{x}^b)], \frac{\partial J^f}{\partial \mathbf{x}^a} + \frac{\partial J^g}{\partial \mathbf{x}^b} \right\rangle = \left\langle \mathbf{y} - H(\mathbf{x}^b), \mathbf{K}^T \left[\frac{\partial J^f}{\partial \mathbf{x}^a} + \frac{\partial J^g}{\partial \mathbf{x}^b} \right] \right\rangle$$

Operational 4D-Var DAS

4D-PSAS OSV Computation

Var vs. PSAS Equivalence demonstration

- Courtier, 1997
- Experimentally in 3D at MSC : El Akkraoui, 2004



PSAS equivalent to VAR : Verification done over 1 month of DA Cycle

Experimental differences in K and K^T elementary operator

Data Assimilation System (DAS)

- 2 different “ K ” definitions used in 2 Inner Loops (IL)
 - Linearization with respect to 2 outer loops (OL) trajectories
 - Simplified physics:
 - 1st IL : Vertical Diffusion (VD)
 - 2nd IL : VD, Subgrid-Scale Orographic Effects, Large Scale Precipitation, Deep Moist Convection

Sensitivity Assimilation System (SAS)

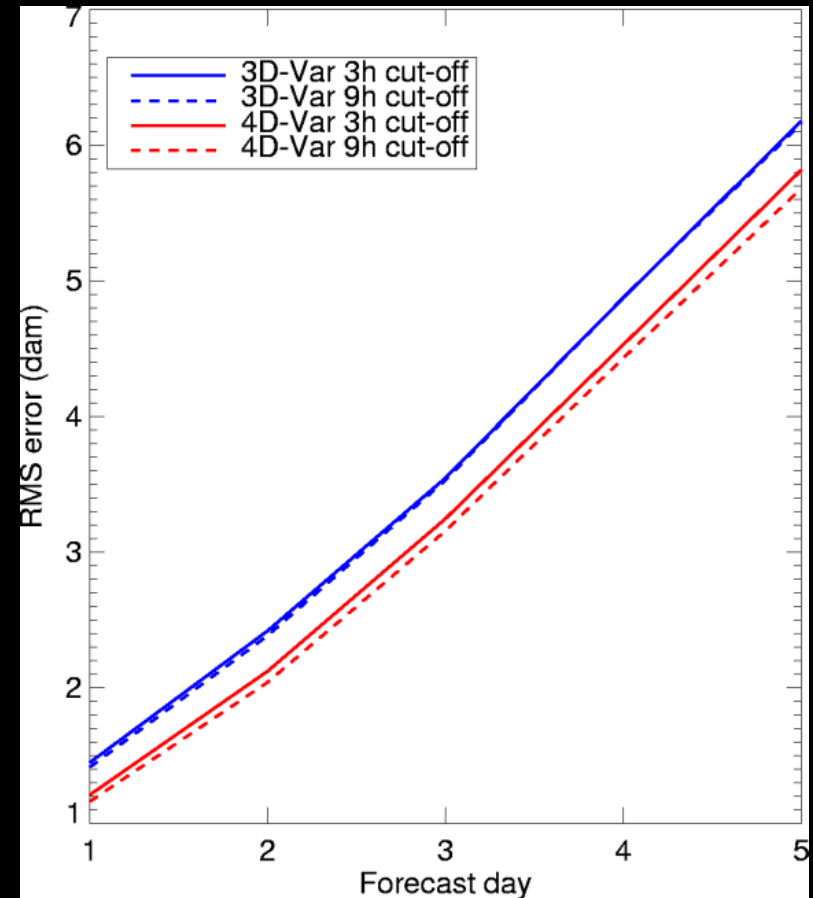
- Single “ K^T ” definition (1 minimization)
 - Linearization with respect to the 1st DAS OL trajectory
 - Simplified physics of the 2nd DAS IL

Observation Impact Estimation Technique : Motivations

- Current operational constraints for global and regional DAS (delivery schedule & computation cost) impose short cut-off time for the reception of observations
 - There is a need to understand the impact of observations distributed over the assimilation window
- Increasing analysis resolution / number of data assimilated / computational power require frequent examination of the cut-off time impact on forecast skill
 - We seek for a method that could be easily implemented in the operational DA suites

Observation Impact Estimation Technique : Motivations (cont.)

500 hPa Geopotential Height
Southern Hemisphere



“The impact of using a shorter cut-off time for the reception of observations,..., is more detrimental with 4D-Var [then it is with a 3D-Var]”

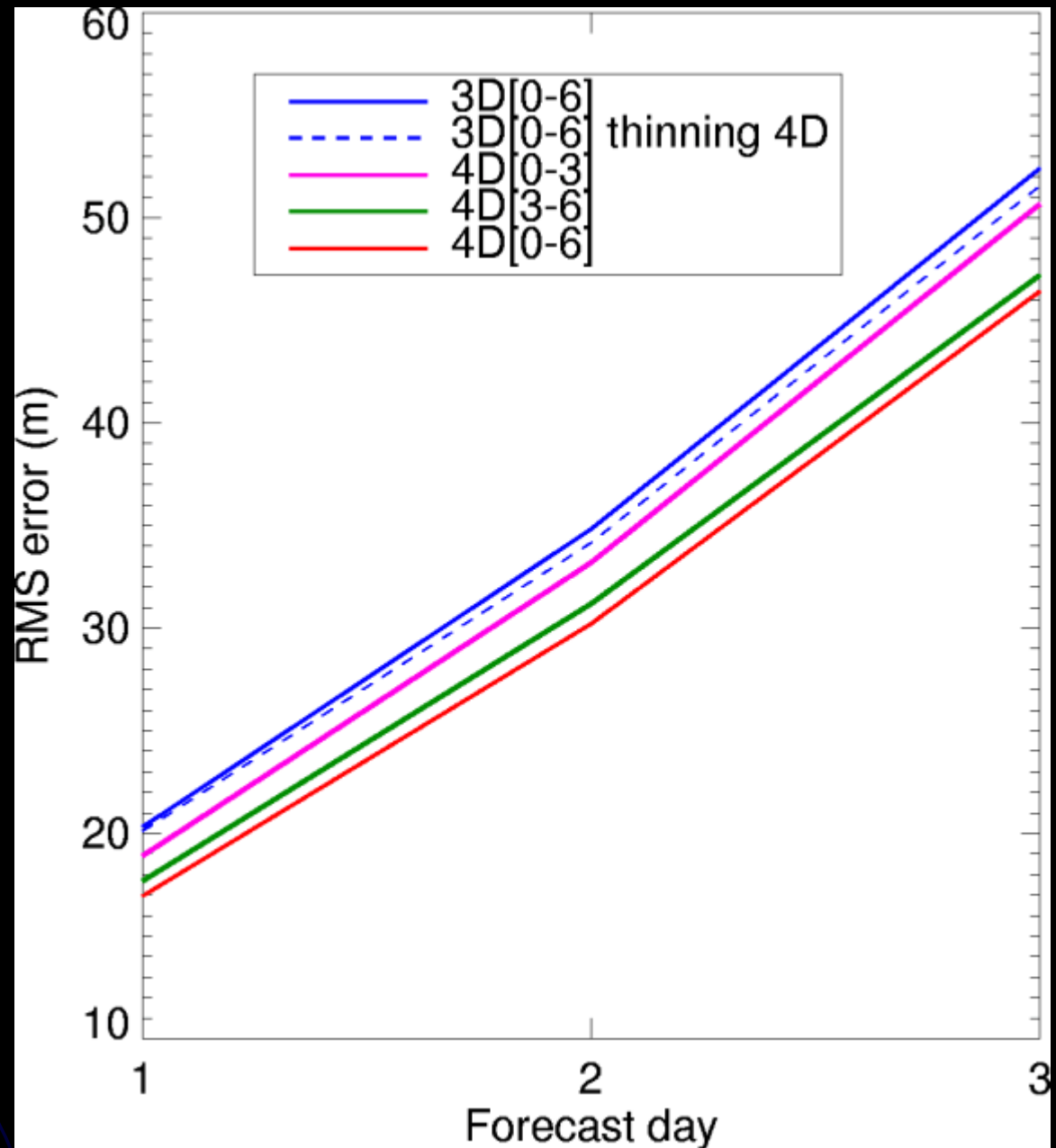
Laroche *et al.*, 2007

Observing System Experiments (OSE) Description

- 6 DA cycles from August 1st to August 31, 2004
 - 3D[0-6] Thinning 3D : 3D-Var using partial data set over the AW
 - 3D[0-6] Thinning 4D : 3D-Var using all data over the AW
 - 3F[0-6] Thinning 4D : 3D-FGAT using all data over the AW
 - 4D[0-6] : 4D-Var using all data over the AW
 - 4D[0-3] : 4D-Var using data over the 1st 3h of AW
 - 4D[3-6] : 4D-Var using data over the last 3h of AW
- In all experiments, data valid in the central assimilation « bin » ([2:37:30 – 3:22:30 [) are assimilated

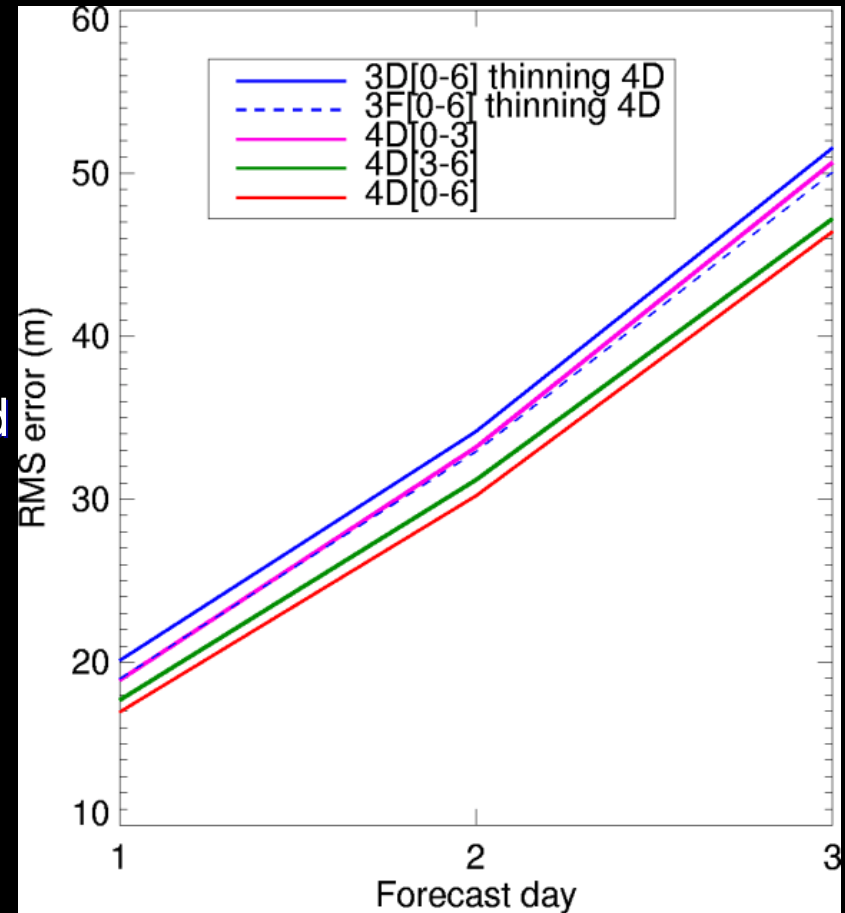
Impact du thinning des observation sur le 3D-Var (non FGAT)

Importance de développer des outils diagnostiques flexibles afin de ré-évaluer les stratégies de thinning des observations lors des changements apportés au système d'assimilation



OSEs : Results

- RMS 3 day forecast Errors over Southern Hemisphere (SH).
 - 3F[0-6] has similar (slightly better) performance than a 4D-Var using 1st half of the data AW
 - 4D-Var performance is more sensitive to data at the end of AW.
- Same conclusions could be inferred from verifications against radiosondes
- To conclude on time sensitivity of the different types of observation, more OSEs would be required



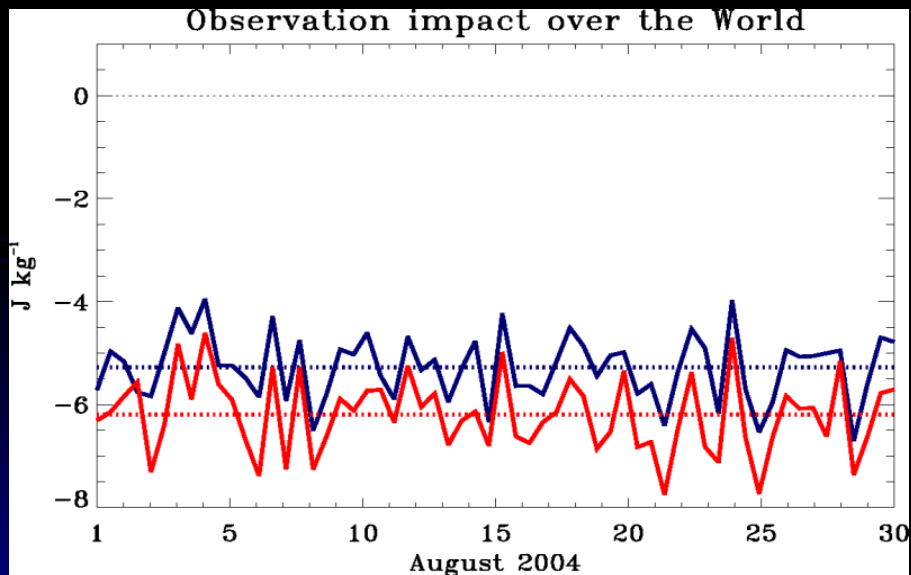
Observation Impact using OSV

- Only experiment “4D[0-6]” is required
 - 1) Computation of model space sensitivity vectors “ $\delta J^f / \delta x^a$ ” and “ $\delta J^g / \delta x^b$ ”
 - 2) Computation of observation sensitivity vector “w”
 - 3) Computation of Observation Impact of any subset of observations from a corresponding pair of innovation and observation sensitivity vectors : $\langle y - H(x^b), w \rangle$

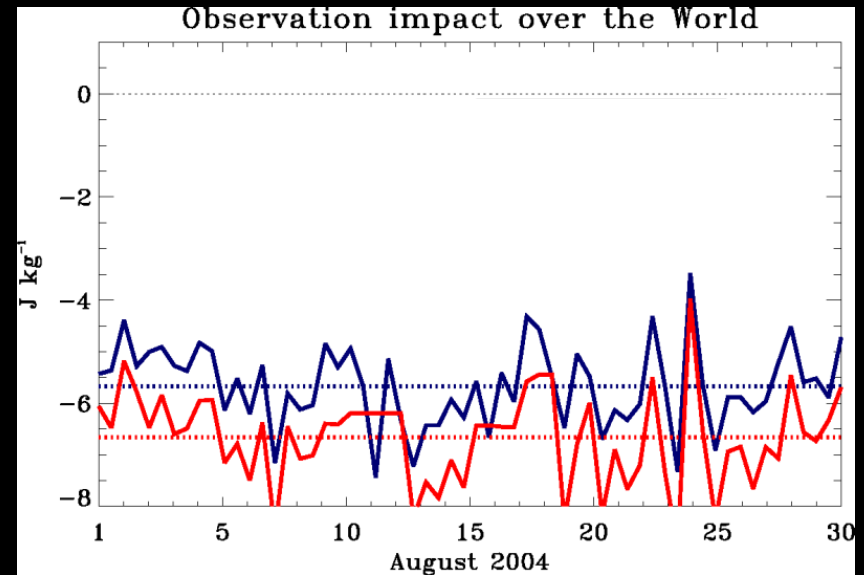
Time series of global OI

- OSV estimations (blue) explains ~85% of the non-linear error differences (red)
- 3D-Var OI (slightly) > 3D-FGAT OI > 4D-VAR OI

$$e^f - e^g \approx \langle \mathbf{y} - H(\mathbf{x}^b), \mathbf{w} \rangle$$



4D-Var



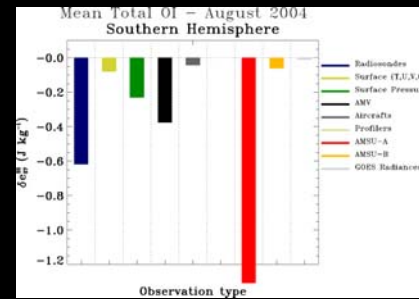
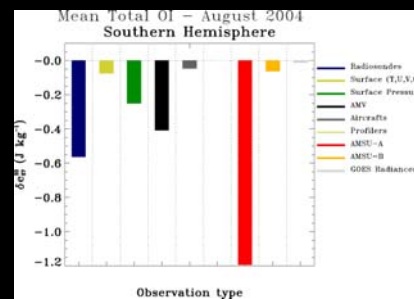
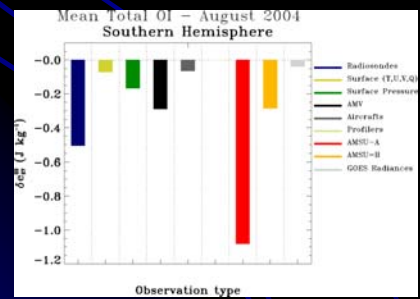
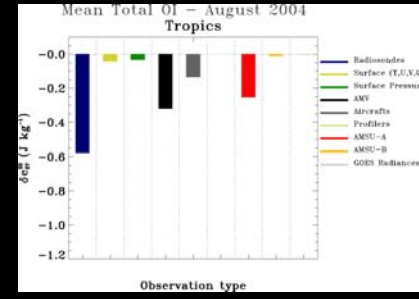
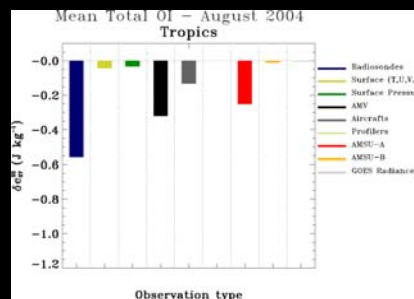
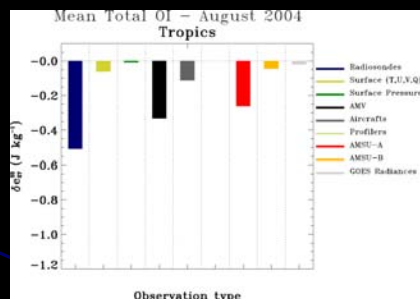
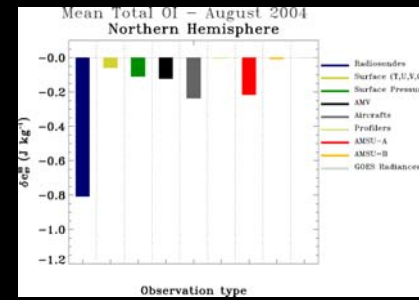
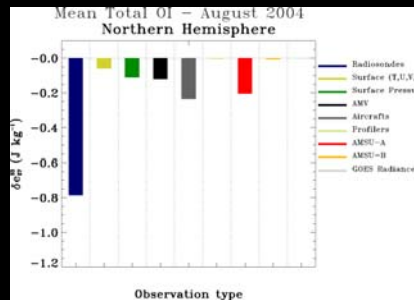
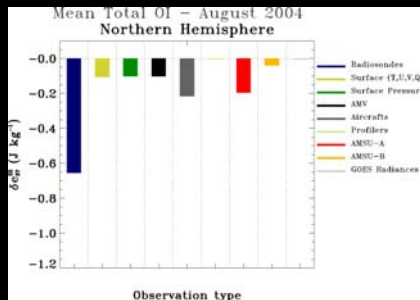
3D-Var

Average Total OI : Observation type

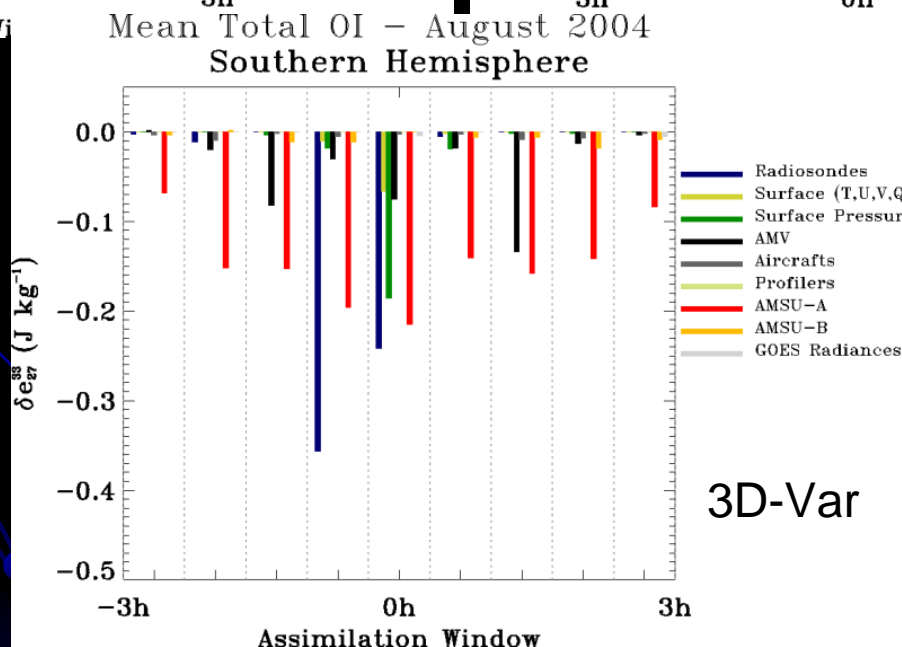
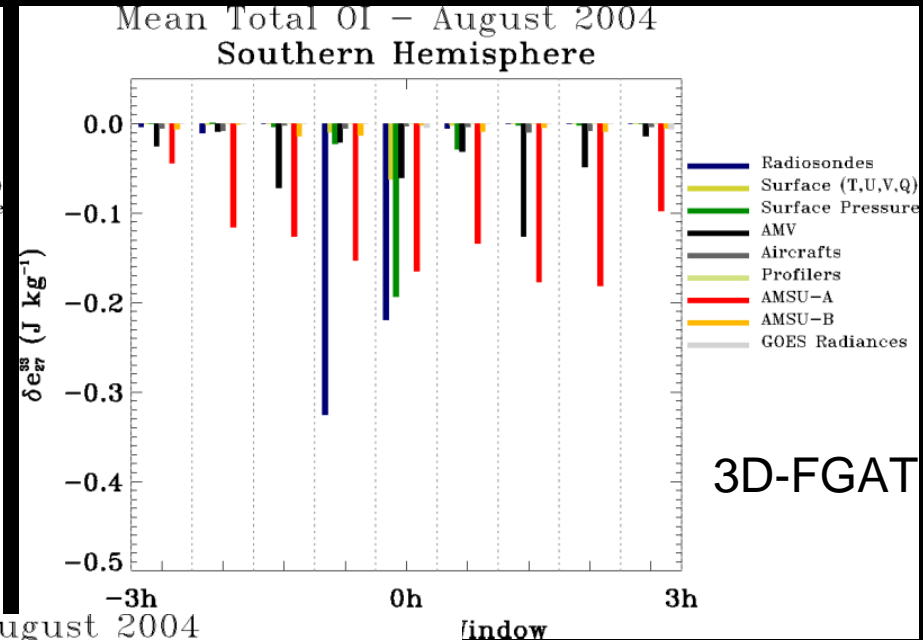
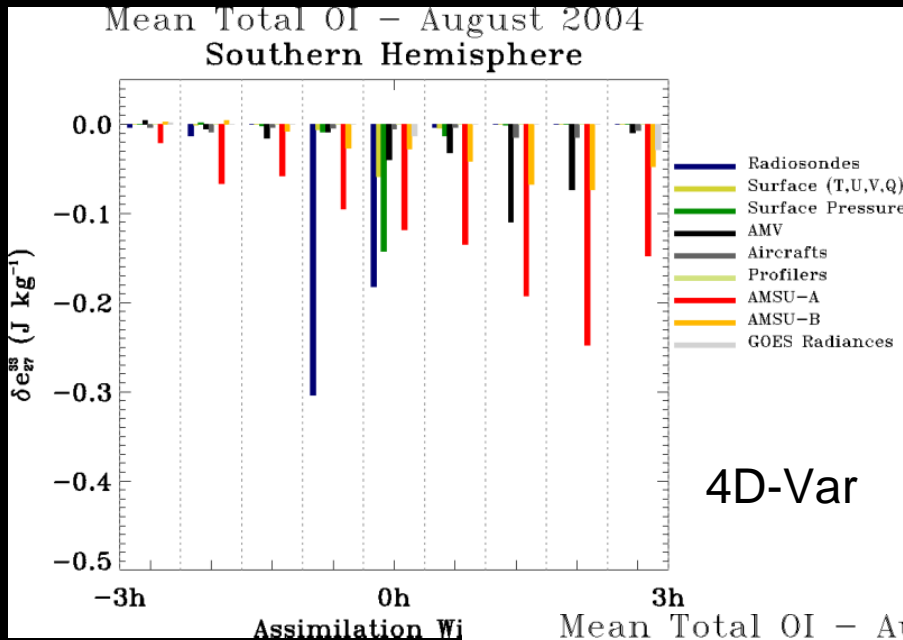
4D-var

3D-FGAT

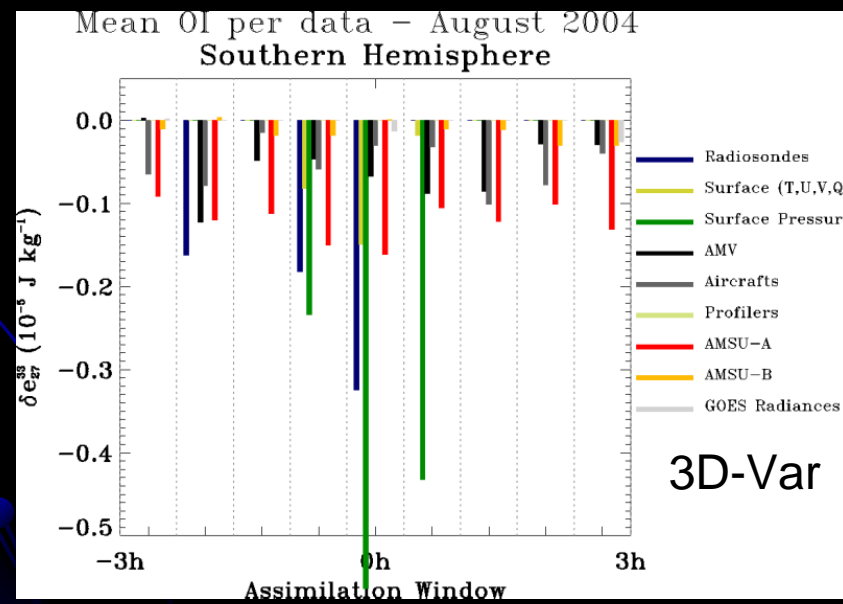
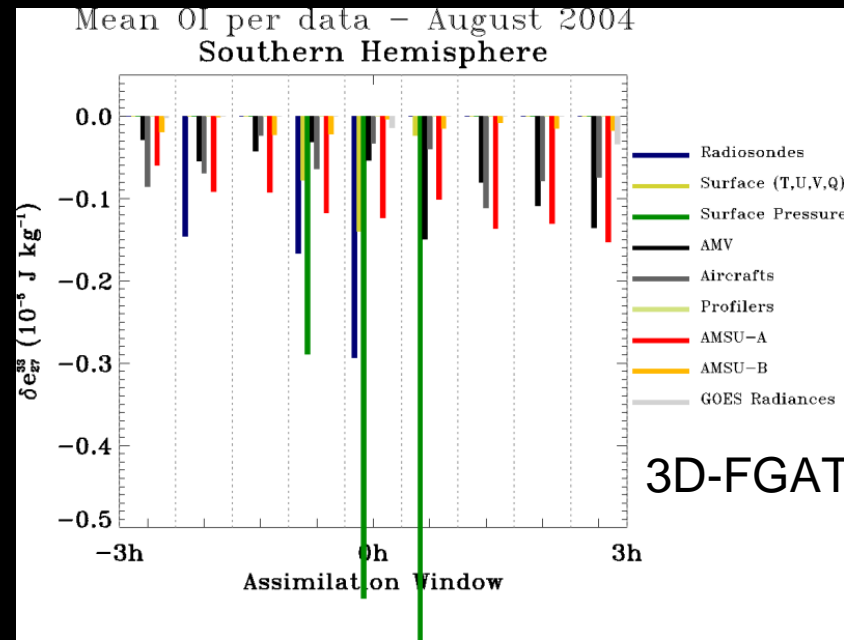
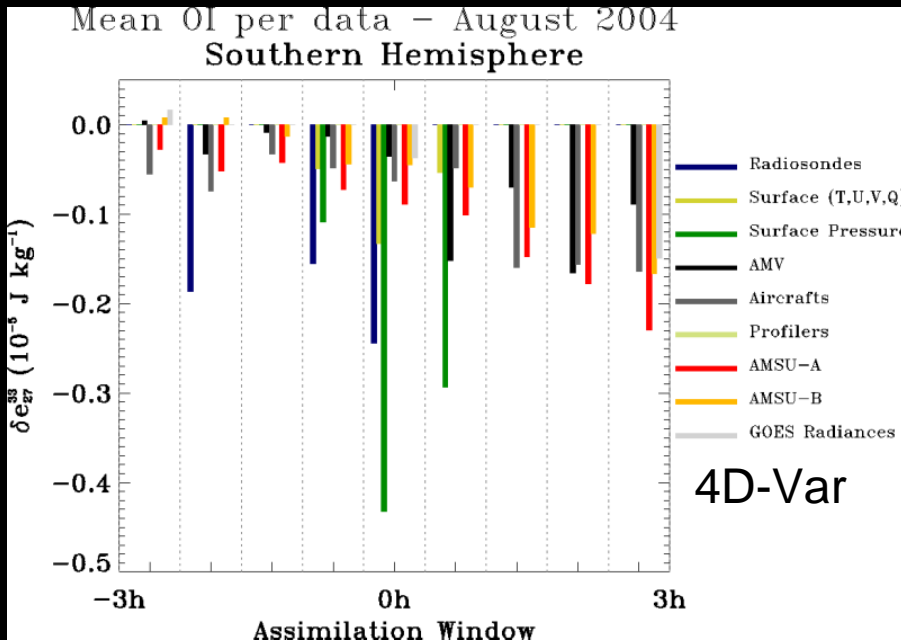
3D-var



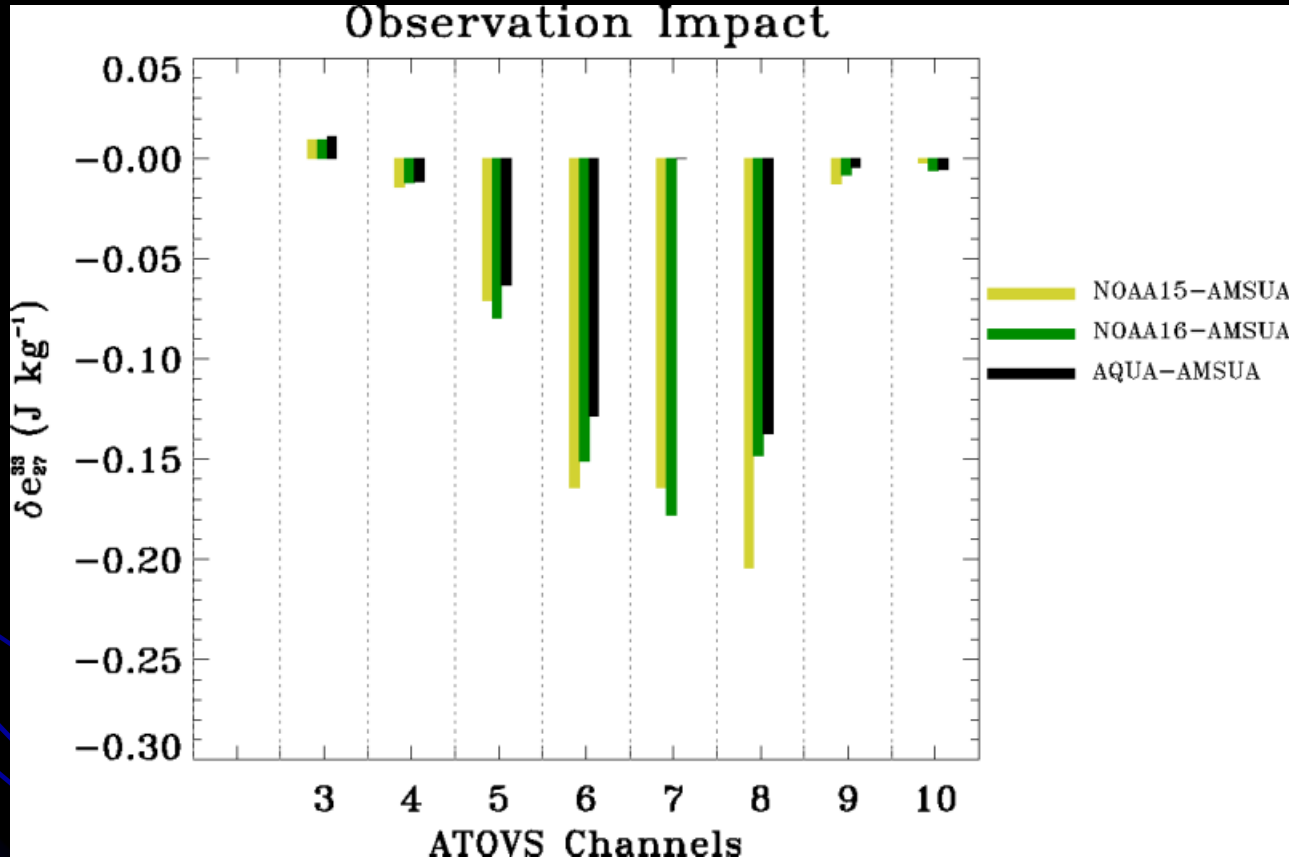
Average Total OI: Obs. types over AW



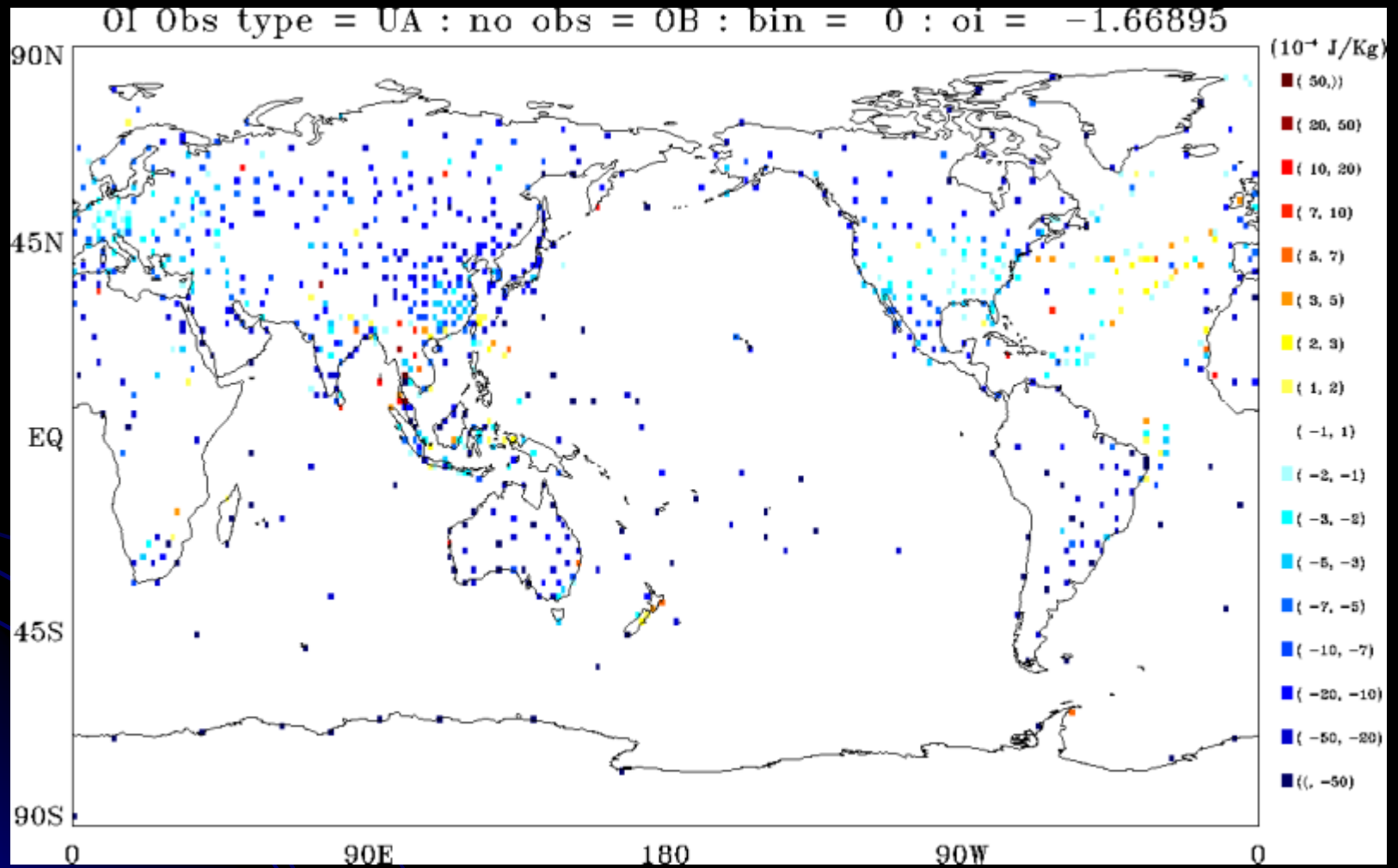
Average OI per data: Southern Hemisphere



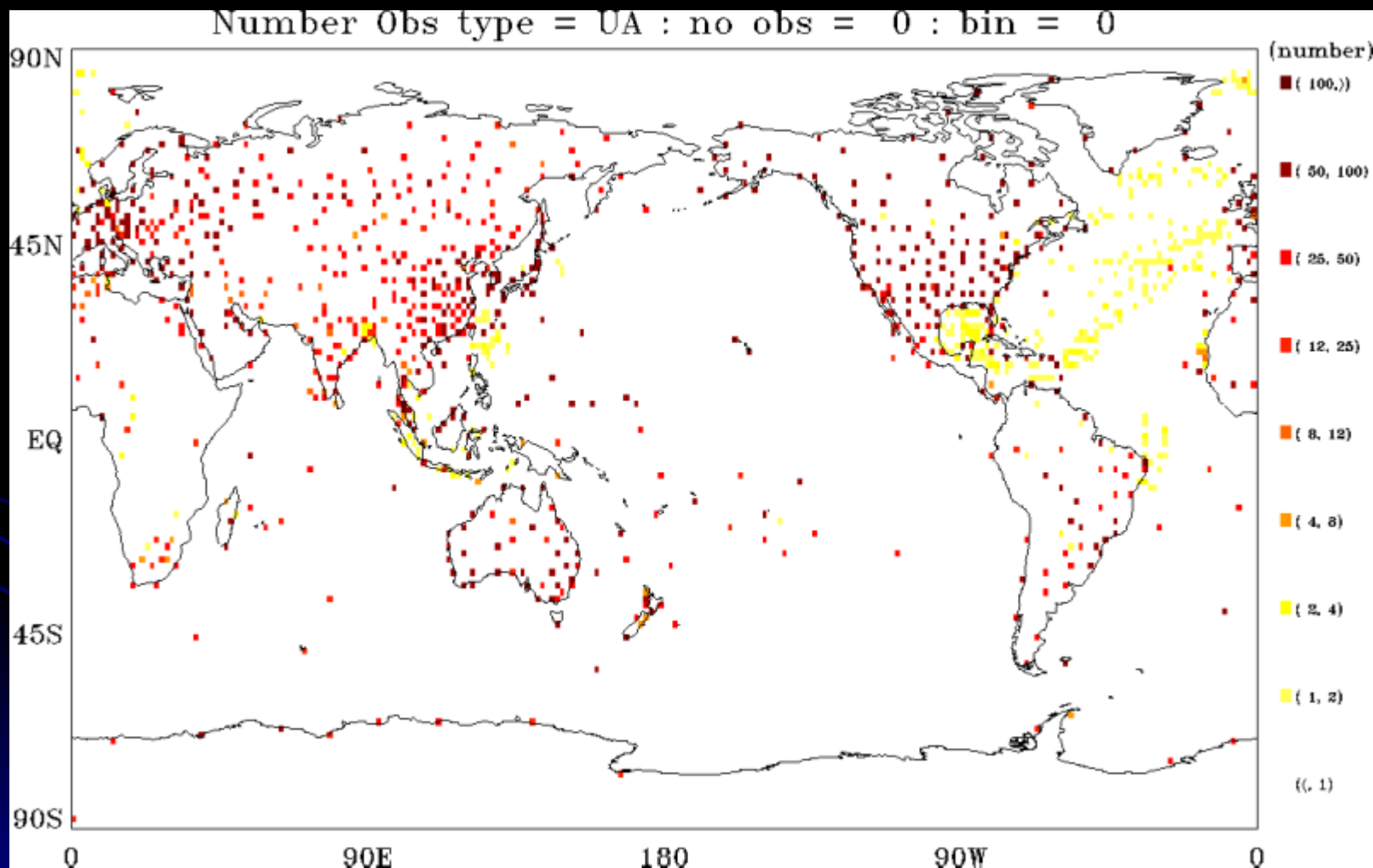
Average Total Global OI



http://iweb.cmc.ec.gc.ca/~afsdjmo/SENSIB/OB_IMPACT/fed_4d_vs_fge04t5/obimaps.htm
http://iweb.cmc.ec.gc.ca/~afsdjmo/SENSIB/OB_IMPACT/fed_4d_vs_fge04t5/obimaps_tovs.htm
http://iweb.cmc.ec.gc.ca/~afsdjmo/SENSIB/OB_IMPACT/fed_4d_vs_fge04t5/obihisto.htm



Nombre de données (rés. 240x120)



Conclusions

- MSC modular implementation of basic transformation operators involved in DA algorithms enabled the development of both model and observation space formulations of the variational problem
- Using the adjoint of GEM and the dual of our operational 4D-Var DAS, we successfully developed an OSV computation system that allows the estimation of the impact of arbitrary partitioning of observations on the skill of short term forecast
- To address the problem of strong sensitivity of forecast errors to the operational cut-off time of observations assimilated in the MSC 4D-var, a sensitivity study of short-term forecast to time distribution and type of observations over the assimilation window have been shown
- OSVs computed from a single experiment enables an *A posteriori* detailed partitioning of OI
- Results from OSVs are in good agreement with OSEs experiments
- OSV based diagnostic allows us to examine the 4D-Var “temporal signature” and clearly show the increasing impact of observations valid later in the AW
- The temporal “signature” is even stronger in the Southern Hemisphere where the assimilation of evenly distributed AMSU-A dominates the RS OI

Le future...

- Norme énergie humide
- Adjoint 'humide' pour le calcul des sensibilités (key analysis)
- Projet d'inter comparaison NRL-GMAO-MRB/MSC
 - En préparation à la campagne hivernale du Projet THORPEX (Hiver 2008-2009)
 - Mêmes types d'observations
 - Même approche diagnostique (Langland & Baker)
 - Période : 1 mois hiver 2006
 - Système OSV mis au niveau de la 'passe observations'
- Usage d'RDB pour les diagnostics
- Étude sur la balance des sensibilités adjointes et son influence sur la mesure de OI (Jean-François Caron)

Librairie d'Analyse Objective ODA (Objective Data Analysis)

Plan de développement

Simon Pellerin

Section de l'Assimilation de Données et de Météorologie Satellitaire

Objectifs du projet

- Repenser le traitement des observations dans nos systèmes d'analyses 3D/4D-Var/EnKF
 - La représentation des données en mémoire dans la programmathèque actuelle du 3D-Var ne convient plus à la grande diversité fonctionnelle reliée aux traitements des observations
 - d'introduire plus facilement de nouvelles données satellitaires tel AIRS, IASI, SSM/I, GPS, QuikScat, COSMIC et HIRS
- Mettre en commun une collection d'opérateurs élémentaires destinés aux algorithmes d'assimilation de données ou à leurs diagnostics
 - 3D/4D-Var/PSAS, EnKF
 - Diagnostique : Calcul des sensibilités des erreurs de prévisions dans l'espace des obs → Impact des observations

Qu'est-ce qu'ODA?

Key Analysis

Prétraitement des observations

- Collection d'opérateurs élémentaires destinés au problème général de l'assimilation de données

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}[\mathbf{y}^c - H(\mathbf{x}^b)]$$

$$\mathbf{K} = \mathbf{P}\mathbf{H}^T (\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R})^{-1}$$

3D → $\mathbf{P} = \mathbf{B}$ $\mathbf{P} = \mathbf{L}\mathbf{B}\mathbf{L}^T$ → 4D

$$\mathbf{y}^c = D(\mathbf{x}^b, \mathbf{y})$$

Sélection de données

- Contrôle de qualité
- Consistance Topographique
- Contrôle de qualité variationnel
- Thinning

3D-Var

3D-PSAS

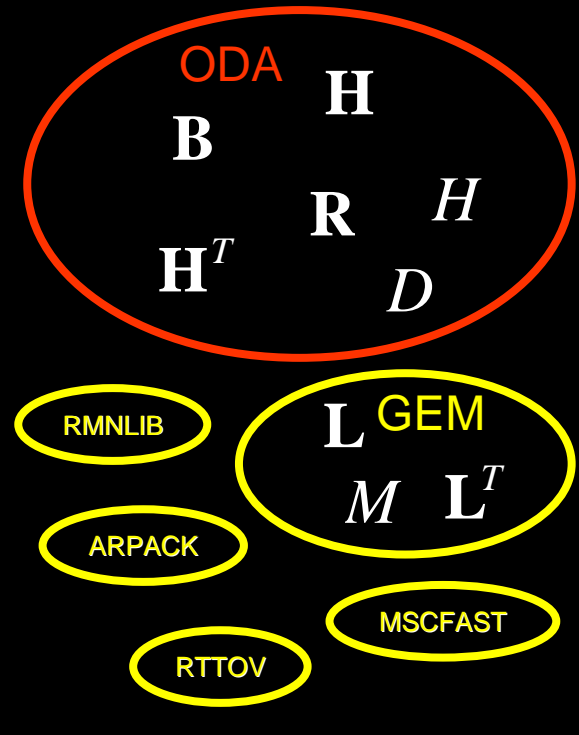
EnKF

4D-Var

4D-PSAS

OSV

Librairies de service



Méthodologie

- Association de types dérivés Fortran90 à des 'familles' d'observations partageant le même traitement fonctionnel
 - Chaque donnée traitée est encapsulé dans un 'objet' dont le type est associé à un ensemble spécifique d'opérateurs : filtres, opérateurs d'observation, covariances d'erreurs d'observation
- À chaque type dérivé est associé deux ensembles d'attributs
 - Attributs communs à l'ensemble des familles d'observations : latitude et longitude de l'observation, valeur observée, etc.
 - Attributs spécifiques au type : numéro du canal pour les données de radiance, hauteur de l'observation pour les radio-sondages ou les données de surface, etc.
- Remarque: La possibilité d'ajouter des attributs spécifiques facilitent l'introduction de nouvelles observations en sécurisant le développement des nouveaux opérateurs

Encapsulation

- Chaque donnée traitée est encapsulé dans un 'objet' dont le type est associé à un ensemble spécifique d'opérateurs

Objet global 'observation' → type (obs)

1 Objet global d'attributs
communs

```
type (comobs) :: tcom
```

```
type (pprof)  :: tua
```

```
type (pprof)  :: tai
```

```
type (pprof)  :: tsw
```

Partagent les
mêmes
opérateurs

3 Objets 'profiles pression'

1 Objet 'profiles hauteur'

```
type (zprof)  :: tpr
```

1 Objet 'radiance rtov'

```
type (tovs)   :: tto
```

1 Objet 'radiance MSCFAST'

```
type (goes)   :: tgo
```

1 Objet 'surface'

```
type (surface) :: tsf
```

- Filtre
- Opérateurs d'observations
- Covariances d'erreurs
- Etc...

Opérations de haut niveau

Exemple : Simulateur Variationnel
appelé par le minimiseur

- Opérations Globales
- Boucles implicites

```
subroutine simvar(na_indic,na_dim,da_v,da_J,da_gradJ)
  use oda_pub_int, only : oda_HT, oda_H
  use mododa, only : tg_y, dg_vbar
  implicit none
```

```
! Dummies
integer :: na_dim
integer :: na_indic
real*8 :: da_J
real*8, dimension(na_dim) :: da_gradJ, da_v
```

Vecteur d'état

```
! Local declaration
integer :: nl_ilev, nl_err
real*8, dimension(na_dim) :: dl_v
```

dl_v = da_v + dg_vbar ! dg_vbar = sum(v) of previous outer-loops

```
! Computation of vTv/2 term
da_J = dot_product(dl_v,dl_v)/2.d0
```

```
call oda_sqrtB(dl_v,na_dim)
call oda_L
```

```
! Computation of normalized residuals [ R-1/2 (HdX - d) ]
tg_y%dres = (oda_H(tg_y) - tg_y%dino)/tg_y%rstd
```

```
da_J = da_J + sum(tg_y%dres * tg_y%dres / 2.d0)
```

```
tg_y%ddelw = oda_HT(tg_y,tg_y%dres/tg_y%rstd)
```

```
call oda_LT
```

```
da_gradJ = 0.d0
call oda_sqrtBT(da_gradJ,na_dim)
```

```
da_gradJ = dl_v + da_gradJ
end subroutine simvar
```

$$J(\mathbf{v}) = \frac{1}{2} \mathbf{v}^T \mathbf{v} + \frac{1}{2} (\mathbf{HL}\sqrt{\mathbf{B}}\mathbf{v} - \mathbf{d})^T \mathbf{R}^{-1} (\mathbf{HL}\sqrt{\mathbf{B}}\mathbf{v} - \mathbf{d})$$

$$\nabla_{\mathbf{v}} J = \mathbf{v} + \sqrt{\mathbf{B}}\mathbf{L}^T \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{HL}\sqrt{\mathbf{B}}\mathbf{v} - \mathbf{d})$$

Historique et état d'ODA

- Il existe deux branches de développement d'ODA
 1. Prototype et cadre informatique pour la recherche et le développement : outils diagnostiques, recherche PSAS/Var
 1. supporte 3 modes d'exécutions : 'VAR', 'PSAS' et 'OSV'
 2. ODA n'est 'actif' que lors de la minimisation
 3. BURP → CMA → ODA → Minimisation → ODA → CMA → BURP
 4. 5 opérateurs d'observation (linéaires) et leurs adjoints

Historique et état d'ODA (suite)

2. ODA et le prétraitement (pré-solveur) des données
 1. Généralisation de l'accès aux données via une interface indépendante du format d'archive des données observées
 2. Les données ne transitent plus par une matrice de type 'CMA'
 3. Les structures de données d'ODA ont subi des modifications majeures par rapport à celles utilisées par le prototype
 4. Les structures des données d'observations et d'états modèles sont stabilisées
 5. Lecture des données via SQLite est fonctionnelle

Portée et livrables du projet

- Programmathèque de 'service' constituée d'une collection d'opérateurs impliqués dans les algorithmes d'assimilation de données
 - Interface aux données d'observations : 2 livraisons
 - 1^{ère} livraison
 - BRP→CMA→ODA→Prétraitement→Kd→Post-traitement→ODA→CMA→BRP
 - 2^{ème} livraison
 - BDRA→ODA→Prétraitement→Kd→Post-traitement→ODA→BDRA
 - Opérateurs d'initialisation et de prétraitement
 - Lecture/écriture et de stockage mémoire des états modèles
 - Opérateur de contrôle de qualité variationnel
 - Interpolateurs horizontaux des états modèles
 - Opération de consistance entre le domaine de l'analyse et les coordonnées des observations
 - Un ensemble d'opérateurs 'élémentaires'
 - H, H', H^T
 - R : corrélations temporelles et spatiales des observations, corrélations inter canaux
 - B : MPI (Bin He), Global, LA, Var, EnKF
 - Opérations d'échange d'objets avec TLM-ADJ
 - Interface aux données d'observations

Portée et livrables du projet (suite)

- Documentation
 - Formations des programmeurs sur la structure de données d'ODA
 - Guide de l'utilisateur de la programmation
 - description des APIs
 - documentation technique et scientifique des opérateurs de la programmation
 - Atelier de formation des usagers
- Bancs d'essai pour les applications principales
 - l'analyse variationnelle et le PSAS 3D et 4D
 - le filtre de Kalman d'ensemble
 - calcul des sensibilités aux erreurs de prévisions dans l'espace des observations
- Conforme à un usage distribué de type MPI et OpenMP ('thread safe'. Code parallèle de haut niveau de type MPI sera maintenu par les applications porteuses.

Merci de votre
attention