Convergence properties of the primal and dual forms od the strong and weak constraint variational data assimilation

Amal EL AKKRAOUI[1] and Pierre GAUTHIER[1,2]

[1]Department of Atmospheric and Oceanic Sciences, McGill University, Canada.
[2]Department of Earth and Atmospheric Sciences,
Université du Québec à Montréal (UQAM), Canada.

15 mai 2009

1. Introduction
2. Dual behavior
3. The minimization algorithms
4. Weak-constraint formulation
5. conclusion

Introduction

- **Primal** : 3D and 4D-Var — **Dual** : 3D and 4D-PSAS.
  PSAS : Physical-space Statistical Analysis System.

- Solving the same variational data assimilation problem in two different spaces : **model** space (primal) and **observation** space (dual).

Introduction

- **Primal** : 3D and 4D-Var — **Dual** : 3D and 4D-PSAS.
  PSAS : Physical-space Statistical Analysis System.

- Solving the same variational data assimilation problem in two different spaces : **model** space (primal) and **observation** space (dual).

Why the dual formualtion ?

- It is a smaller space compared to the model space.

- It is expected to be particularly interesting when the size of the control variable of the assimilation problem becomes very large :

  - Extended data assimilation window ;
  - Weak-Constraint formulation.

Why the dual formualtion ?

- It is a smaller space compared to the model space.
- It is expected to be particularly interesting when the size of the control variable of the assimilation problem becomes very large :

  - Extended data assimilation window ;
  - Weak-Constraint formulation.

3D-Var/3D-PSAS

- The objective functions of the primal and dual 3D form are respectively :

$$J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}(\mathbf{H}\delta\mathbf{x} - \mathbf{y}')^T\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{y}')$$

$$F(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)\mathbf{w} - \mathbf{w}^T\mathbf{y}'$$

- At convergence :

$$\underbrace{\delta\mathbf{x}_a}_{\substack{\text{dimension n} \\ \text{(model space)}}} = \underbrace{\mathbf{B}\mathbf{H}^T}_{\text{representer matrix}} \underbrace{\mathbf{w}_a}_{\substack{\text{dimension m} \\ \text{(observation space)}}}$$

- 3D-Var and 3D-PSAS are preconditioned with $B^{-\frac{1}{2}}$ and $R^{\frac{1}{2}}$ respectively. (Amodei, 1995)

3D-Var/3D-PSAS

- The objective functions of the primal and dual 3D form are respectively :

$$J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}(\mathbf{H}\delta\mathbf{x} - \mathbf{y}')^T\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{y}')$$

$$F(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)\mathbf{w} - \mathbf{w}^T\mathbf{y}'$$

- At convergence :

$$\underbrace{\delta\mathbf{x}_a}_{\substack{\text{dimension n}\\ \text{(model space)}}} = \underbrace{\mathbf{B}\mathbf{H}^T}_{\text{representer matrix}} \underbrace{\mathbf{w}_a}_{\substack{\text{dimension m}\\ \text{(observation space)}}}$$

- 3D-Var and 3D-PSAS are preconditioned with $B^{-\frac{1}{2}}$ and $R^{\frac{1}{2}}$ respectively. (Amodei, 1995)

3D-Var/3D-PSAS

- The objective functions of the primal and dual 3D form are respectively :

$$J(\delta \mathbf{x}) = \frac{1}{2}\delta \mathbf{x}^T \mathbf{B}^{-1}\delta \mathbf{x} + \frac{1}{2}(\mathbf{H}\delta \mathbf{x} - \mathbf{y}')^T \mathbf{R}^{-1}(\mathbf{H}\delta \mathbf{x} - \mathbf{y}')$$

$$F(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T(\mathbf{R} + \mathbf{HBH}^T)\mathbf{w} - \mathbf{w}^T \mathbf{y}'$$

- At convergence :

$$\underset{\substack{\uparrow \\ \text{dimension n} \\ \text{(model space)}}}{\delta \mathbf{x}_a} = \underset{\substack{\uparrow \\ \text{representer matrix}}}{\mathbf{BH}^T} \underset{\substack{\uparrow \\ \text{dimension m} \\ \text{(observation space)}}}{\mathbf{w}_a}$$

- 3D-Var and 3D-PSAS are preconditioned with $B^{-\frac{1}{2}}$ and $R^{\frac{1}{2}}$ respectively. (Amodei, 1995)

3D-Var/3D-PSAS

- In a compact form using $\mathbf{L} = \mathbf{R}^{-\frac{1}{2}}\mathbf{H}\mathbf{B}^{\frac{1}{2}}$:

$$J(\mathbf{v}) = \frac{1}{2}\mathbf{v}^T(\mathbf{I}_n + \mathbf{L}^T\mathbf{L})\mathbf{v} - \mathbf{v}^T\mathbf{L}^T\tilde{\mathbf{y}} + \frac{1}{2}\tilde{\mathbf{y}}^T\tilde{\mathbf{y}},$$

$$F(\mathbf{u}) = \frac{1}{2}\mathbf{u}^T(\mathbf{I}_m + \mathbf{L}\mathbf{L}^T)\mathbf{u} - \mathbf{u}^T\tilde{\mathbf{y}},$$

- Equivalence only valid at convergence + $\mathbf{H}$ is linear.
- The Hessians have the same condition number, and both methods should give the same results and converge at similar convergence rates (*Courtier, 1997*).
- The equivalence is extended to the SV of the Hessians (*El Akkraoui et al., 2008*).

3D-Var/3D-PSAS

- In a compact form using $\mathbf{L} = \mathbf{R}^{-\frac{1}{2}}\mathbf{H}\mathbf{B}^{\frac{1}{2}}$:

$$J(\mathbf{v}) = \frac{1}{2}\mathbf{v}^T(\mathbf{I}_n + \mathbf{L}^T\mathbf{L})\mathbf{v} - \mathbf{v}^T\mathbf{L}^T\tilde{\mathbf{y}} + \frac{1}{2}\tilde{\mathbf{y}}^T\tilde{\mathbf{y}},$$

$$F(\mathbf{u}) = \frac{1}{2}\mathbf{u}^T(\mathbf{I}_m + \mathbf{L}\mathbf{L}^T)\mathbf{u} - \mathbf{u}^T\tilde{\mathbf{y}},$$

- Equivalence only valid at convergence + **H** is linear.
- The Hessians have the same condition number, and both methods should give the same results and converge at similar convergence rates (*Courtier, 1997*).
- The equivalence is extended to the SV of the Hessians (*El Akkraoui et al., 2008*).

3D-Var/3D-PSAS

- In a compact form using $\mathbf{L} = \mathbf{R}^{-\frac{1}{2}}\mathbf{H}\mathbf{B}^{\frac{1}{2}}$:

$$J(\mathbf{v}) = \frac{1}{2}\mathbf{v}^T(\mathbf{I}_n + \mathbf{L}^T\mathbf{L})\mathbf{v} - \mathbf{v}^T\mathbf{L}^T\tilde{\mathbf{y}} + \frac{1}{2}\tilde{\mathbf{y}}^T\tilde{\mathbf{y}},$$

$$F(\mathbf{u}) = \frac{1}{2}\mathbf{u}^T(\mathbf{I}_m + \mathbf{L}\mathbf{L}^T)\mathbf{u} - \mathbf{u}^T\tilde{\mathbf{y}},$$

- Equivalence only valid at convergence + **H** is linear.
- The Hessians have the same condition number, and both methods should give the same results and converge at similar convergence rates (*Courtier, 1997*).
- The equivalence is extended to the SV of the Hessians (*El Akkraoui et al., 2008*).

3D-Var/3D-PSAS

- In a compact form using $\mathbf{L} = \mathbf{R}^{-\frac{1}{2}}\mathbf{H}\mathbf{B}^{\frac{1}{2}}$:

$$J(\mathbf{v}) = \frac{1}{2}\mathbf{v}^T(\mathbf{I}_n + \mathbf{L}^T\mathbf{L})\mathbf{v} - \mathbf{v}^T\mathbf{L}^T\tilde{\mathbf{y}} + \frac{1}{2}\tilde{\mathbf{y}}^T\tilde{\mathbf{y}},$$

$$F(\mathbf{u}) = \frac{1}{2}\mathbf{u}^T(\mathbf{I}_m + \mathbf{L}\mathbf{L}^T)\mathbf{u} - \mathbf{u}^T\tilde{\mathbf{y}},$$

- Equivalence only valid at convergence + $\mathbf{H}$ is linear.
- The Hessians have the same condition number, and both methods should give the same results and converge at similar convergence rates (*Courtier, 1997*).
- The equivalence is extended to the SV of the Hessians (*El Akkraoui et al., 2008*).

That is the theory...

...The practice is full of surprises.

That is the theory...

...The practice is full of surprises.

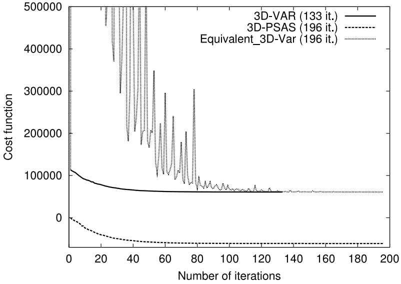
- **The good news :** At convergence, the dual method gives the same results as the primal one...as expected.

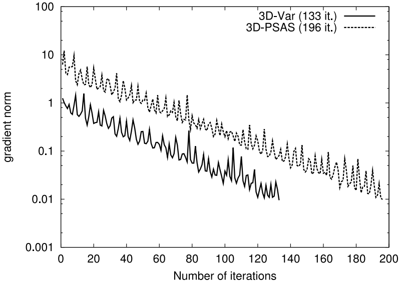
- **The problem :** During the minimization, the dual algorithm exhibits a spurious behavior, source of a serious concern. (From El Akkraoui *et al.*, 2008)

- **The good news :** At convergence, the dual method gives the same results as the primal one...as expected.

- **The problem :** During the minimization, the dual algorithm exhibits a spurious behavior, source of a serious concern. (From El Akkraoui *et al.*, 2008)

All roads lead to Rome...but some are stranger than others





At each PSAS iteration k, the iterate $\mathbf{u}_k$ is brought to the model space through the operator $\mathbf{L}^T$ and the 3D-Var objective function is calculated for $\mathbf{v}_k = \mathbf{L}^T\mathbf{u}_k$. That is, $J(\mathbf{L}^T\mathbf{u}_k)$

So, in the dual case, we note

- A big increase of the norm of the first gradient.
- The dual assimilation may give an analysis state worst than the background when using a finite number of iterations.
- As long as the problem is not solved, the dual method cannot be used in operational applications, nor is it reliable for a weak-constraint implementation.

- A closer look at the term of the primal function evaluated at the dual iterates shows that

$$\boxed{J(\mathbf{L}^T\mathbf{u}_k) = \frac{1}{2}\|\nabla F(\mathbf{u}_k)\|^2 - F(\mathbf{u}_k)}$$

- While F is being reduced gradually by the minimization algorithm (the CG), no constraint is imposed on its gradient.

- At the first iteration, $F(\mathbf{u}_1) = 0$...The gradient norm may be the dominant term in this formula.

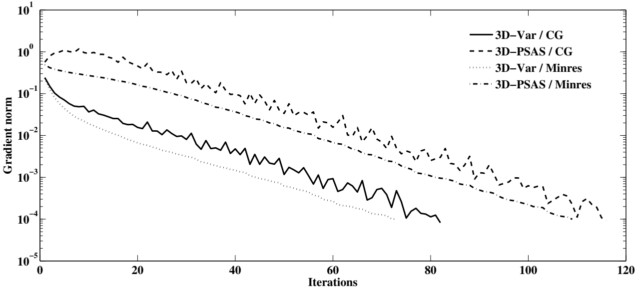
- Need a constraint on the gradient norm....change the minimization algorithm.

Minres Vs the Conjugate-Gradient

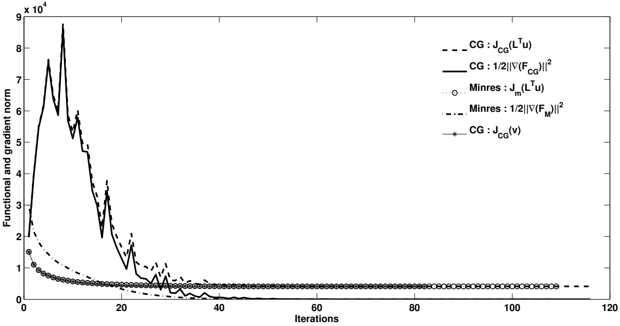
- Iterative methods for solving the linear system : $\mathbf{Ax} = \mathbf{b}$.
- Here, $\mathbf{A}$ is symmetric and positive definite and corresponds to the Hessians $\mathbf{J}''$ and $\mathbf{F}''$, and $\mathbf{b}$ to the terms $\mathbf{L}^T\tilde{\mathbf{y}}$ and $\tilde{\mathbf{y}}$ respectively.
- The gradients correspond to the residuals : $\mathbf{r} = \mathbf{Ax} - \mathbf{b}$.

CG	Minres
symmetric positive definite	symmetric and indefinite
minimize the functional	minimize the residual (gradient)

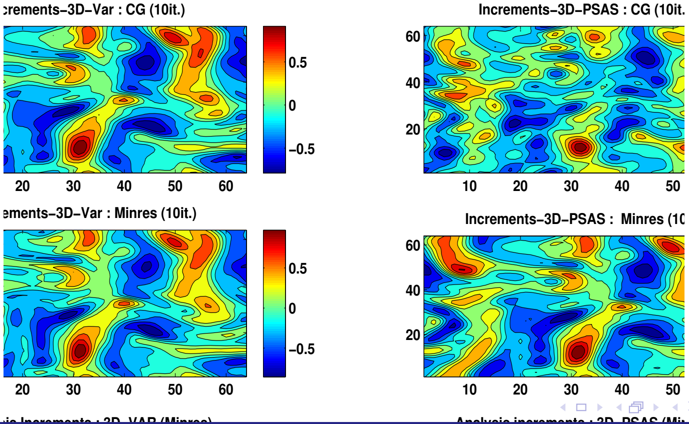
$$\frac{||\mathbf{r}^m{}_k||^2}{||\mathbf{r}^c{}_k||^2} = 1 - \frac{||\mathbf{r}^m{}_k||^2}{||\mathbf{r}^m{}_{k-1}||^2}$$

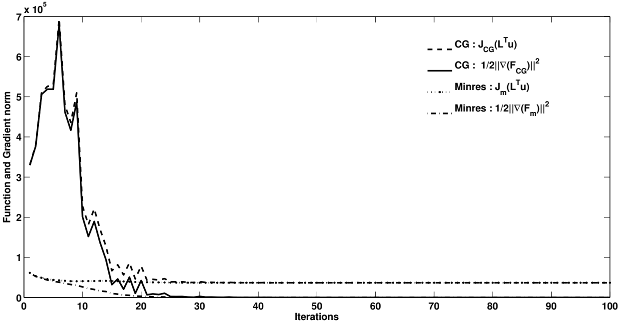


Minres and the CG residual norms of 3D-Var (solid and dotted lines), and 3D-PSAS (dashed and dash-dotted lines).



- The primal functional estimated for the dual iterates using the formula $J(\mathbf{L}^T\mathbf{u}_k) = \frac{1}{2}\|\nabla F(\mathbf{u}_k)\|^2 - F(\mathbf{u}_k)$ for the CG (dashed line) and Minres (dotted-line with the circle marker). Also the term $\frac{1}{2}\|\nabla(F)\|^2$ is plotted for the CG (solid line), and Minres (dashed-dotted line), and finally, the original primal function calculated with the CG (solid line with the star marker) is plotted for comparison.


crements–3D–Var : CG (10it.)
Increments–3D–PSAS : CG (10it.
ements–3D–Var : Minres (10it.)
Increments–3D–PSAS : Minres (10
0.5
0
–0.5
20 30 40 50 60
10 20 30 40 50
60
40
20



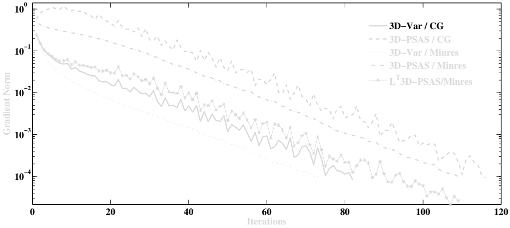
- Four dimensional case

The stopping criterion

- 3D/4D-PSAS needs more iterations to converge to the same stopping criterion as the 3D/4D-Var : $\frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_0\|} \leq \epsilon$.
- Recall

$$\mathbf{v}_k \equiv \mathbf{L}^T \mathbf{u}_k, \quad \text{and,} \quad \mathbf{r}_k^{primal} \equiv \mathbf{L}^T \mathbf{r}_k^{dual}$$

- The comparison needs to be in the same space.



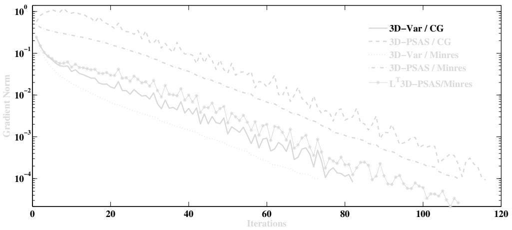
- Same as before. The star line representing the norm of the dual residuals in the model space.

The stopping criterion

- 3D/4D-PSAS needs more iterations to converge to the same stopping criterion as the 3D/4D-Var : $\frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_0\|} \leq \epsilon$.
- Recall

$$\mathbf{v}_k \equiv \mathbf{L}^T \mathbf{u}_k, \quad \text{and}, \quad \mathbf{r}_k^{primal} \equiv \mathbf{L}^T \mathbf{r}_k^{dual}$$

- The comparison needs to be in the same space.



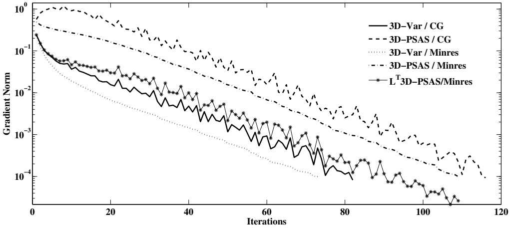
- Same as before. The star line representing the norm of the dual residuals in the model space.

The stopping criterion

- 3D/4D-PSAS needs more iterations to converge to the same stopping criterion as the 3D/4D-Var : $\frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_0\|} \leq \epsilon$.
- Recall

$$\mathbf{v}_k \equiv \mathbf{L}^T \mathbf{u}_k, \quad \text{and}, \quad \mathbf{r}_k^{primal} \equiv \mathbf{L}^T \mathbf{r}_k^{dual}$$

- The comparison needs to be in the same space.



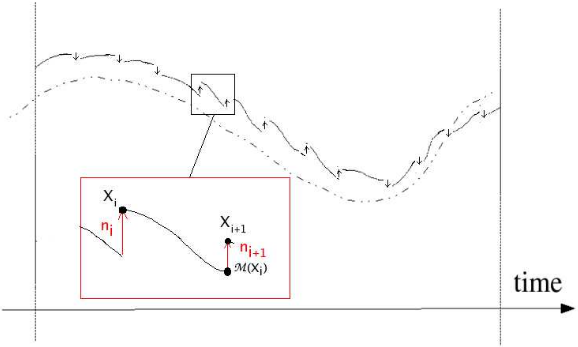
- Same as before. The star line representing the norm of the dual residuals in the model space.

The stopping criterion

- R_1 : stop the primal minimization when $\frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_0\|} \leq \epsilon$.
- R_2 : stop the dual minimization when $\frac{\|\mathbf{L}^T\mathbf{r}_k\|}{\|\mathbf{L}^T\mathbf{r}_0\|} \leq \epsilon$.

Weak-constraint formulation : Accounting for model errors in DA

$$\mathbf{x}_i = \mathcal{M}_{i-1,i}(\mathbf{x}_{i-1}) + \eta_i$$



- Strong-Constraint 4D-Var (incremental)

$$J(\delta\mathbf{x}_o) = \frac{1}{2}\delta\mathbf{x}_o^T\mathbf{B}^{-1}\delta\mathbf{x}_o + \frac{1}{2}(\mathbf{G}\delta\mathbf{x}_o - \mathbf{y}')^T\mathbf{R}^{-1}(\mathbf{G}\delta\mathbf{x}_o - \mathbf{y}')$$

where $\mathbf{G} = (..., \mathcal{H}_i\mathcal{M}_{0,i}, ...)$, (Courtier's notations)

- Weak-Constraint 4D-Var

$$J(\delta\mathbf{z}) = \frac{1}{2}\delta\mathbf{z}^T\mathbf{D}^{-1}\delta\mathbf{z} + \frac{1}{2}(\mathbf{S}\delta\mathbf{z} - \mathbf{y}')^T\mathbf{R}^{-1}(\mathbf{S}\delta\mathbf{z} - \mathbf{y}')$$

where $\delta\mathbf{z} = (\delta\mathbf{x}_o, ..., \eta_i, ...)$,

$$\mathbf{D} = \begin{pmatrix} B & 0 & \cdots & 0 \\ 0 & Q_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & Q_q \end{pmatrix}, \mathbf{S} = \begin{pmatrix} \vdots \\ \mathbf{G}_i \\ \vdots \end{pmatrix}$$

- Strong-Constraint 4D-Var (incremental)

$$J(\delta\mathbf{x}_o) = \frac{1}{2}\delta\mathbf{x}_o^T\mathbf{B}^{-1}\delta\mathbf{x}_o + \frac{1}{2}(\mathbf{G}\delta\mathbf{x}_o - \mathbf{y}')^T\mathbf{R}^{-1}(\mathbf{G}\delta\mathbf{x}_o - \mathbf{y}')$$

where $\mathbf{G} = (..., \mathcal{H}_i\mathcal{M}_{0,i}, ...)$, (Courtier's notations)

- Weak-Constraint 4D-Var

$$J(\delta\mathbf{z}) = \frac{1}{2}\delta\mathbf{z}^T\mathbf{D}^{-1}\delta\mathbf{z} + \frac{1}{2}(\mathbf{S}\delta\mathbf{z} - \mathbf{y}')^T\mathbf{R}^{-1}(\mathbf{S}\delta\mathbf{z} - \mathbf{y}')$$

where $\delta\mathbf{z} = (\delta\mathbf{x}_o, ..., \eta_i, ...)$,

$$\mathbf{D} = \begin{pmatrix} B & 0 & \cdots & 0 \\ 0 & Q_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & Q_q \end{pmatrix}, \mathbf{S} = \begin{pmatrix} \vdots \\ \mathbf{G}_i \\ \vdots \end{pmatrix}$$

- Strong-Constraint 4D-PSAS

$$F(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T(\mathbf{R} + \mathbf{GBG}^T)\mathbf{w} - \mathbf{w}^T\mathbf{y}'$$

- Weak-Constraint 4D-PSAS

$$F(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T(\mathbf{R} + \mathbf{SDS}^T)\mathbf{w} - \mathbf{w}^T\mathbf{y}'$$

- The control variable is still defined in the observation space (does not change).
- Preconditioning ($\mathbf{u} = \mathbf{R}^{\frac{1}{2}}\mathbf{w}$)

$$F(\mathbf{u}) = \frac{1}{2}\mathbf{u}^T(\mathbf{I}_m + \mathbf{R}^{-\frac{1}{2}}\mathbf{SDS}^T\mathbf{R}^{-\frac{1}{2}})\mathbf{u} - \mathbf{u}^T\mathbf{R}^{-\frac{1}{2}}\mathbf{y}'$$

- Strong-Constraint 4D-PSAS

$$F(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T(\mathbf{R} + \mathbf{GBG}^T)\mathbf{w} - \mathbf{w}^T\mathbf{y}'$$

- Weak-Constraint 4D-PSAS

$$F(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T(\mathbf{R} + \mathbf{SDS}^T)\mathbf{w} - \mathbf{w}^T\mathbf{y}'$$

- The control variable is still defined in the observation space (does not change).
- Preconditioning ($\mathbf{u} = \mathbf{R}^{\frac{1}{2}}\mathbf{w}$)

$$F(\mathbf{u}) = \frac{1}{2}\mathbf{u}^T(\mathbf{I}_m + \mathbf{R}^{-\frac{1}{2}}\mathbf{SDS}^T\mathbf{R}^{-\frac{1}{2}})\mathbf{u} - \mathbf{u}^T\mathbf{R}^{-\frac{1}{2}}\mathbf{y}'$$

- Strong-Constraint 4D-PSAS

$$F(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T(\mathbf{R} + \mathbf{GBG}^T)\mathbf{w} - \mathbf{w}^T\mathbf{y}'$$

- Weak-Constraint 4D-PSAS

$$F(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T(\mathbf{R} + \mathbf{SDS}^T)\mathbf{w} - \mathbf{w}^T\mathbf{y}'$$

- The control variable is still defined in the observation space (does not change).
- Preconditioning ($\mathbf{u} = \mathbf{R}^{\frac{1}{2}}\mathbf{w}$)

$$F(\mathbf{u}) = \frac{1}{2}\mathbf{u}^T(\mathbf{I}_m + \mathbf{R}^{-\frac{1}{2}}\mathbf{SDS}^T\mathbf{R}^{-\frac{1}{2}})\mathbf{u} - \mathbf{u}^T\mathbf{R}^{-\frac{1}{2}}\mathbf{y}'$$

- Strong-Constraint 4D-PSAS

$$F(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T(\mathbf{R} + \mathbf{GBG}^T)\mathbf{w} - \mathbf{w}^T\mathbf{y}'$$

- Weak-Constraint 4D-PSAS

$$F(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T(\mathbf{R} + \mathbf{SDS}^T)\mathbf{w} - \mathbf{w}^T\mathbf{y}'$$

- The control variable is still defined in the observation space (does not change).
- Preconditioning ($\mathbf{u} = \mathbf{R}^{\frac{1}{2}}\mathbf{w}$)

$$F(\mathbf{u}) = \frac{1}{2}\mathbf{u}^T(\mathbf{I}_m + \mathbf{R}^{-\frac{1}{2}}\mathbf{SDS}^T\mathbf{R}^{-\frac{1}{2}})\mathbf{u} - \mathbf{u}^T\mathbf{R}^{-\frac{1}{2}}\mathbf{y}'$$

The adjoint variables

- The primal case

$$\nabla_{\delta\mathbf{z}} J = \mathbf{D}^{-1}\delta\mathbf{z} + \mathbf{S}^T\mathbf{R}^{-1}(\mathbf{S}\delta\mathbf{z} - \mathbf{y}')$$

the adjoint variable is $\delta\mathbf{x}^*{}_i = \mathbf{M}_{i+1}^T\delta\mathbf{x}^*{}_{i+1} - \mathbf{H}_i{}^T\mathbf{R}_i{}^{-1}\mathbf{y}'{}_i$, with $\mathbf{H}_n{}^T\mathbf{R}_n{}^{-1}\mathbf{y}'{}_n$. (Trémolet, 2007)

- The dual case

$$\nabla_{\mathbf{w}} F = (\mathbf{R} + \mathbf{S}\mathbf{D}\mathbf{S}^T)\mathbf{w} - \mathbf{y}'$$

the adjoint variable is $\mathbf{w}_i{}^* = \mathbf{M}_{i+1}^T\mathbf{w}_{i+1}^* + \mathbf{H}_i{}^T\mathbf{w}_i$, with $\mathbf{w}_n^* = \mathbf{H}_n{}^T\mathbf{w}_n$

- The gradient can still be calculated with one backward integration + one forward integration.

The adjoint variables

- The primal case

$$\nabla_{\delta \mathbf{z}} J = \mathbf{D}^{-1}\delta\mathbf{z} + \mathbf{S}^T\mathbf{R}^{-1}(\mathbf{S}\delta\mathbf{z} - \mathbf{y}')$$

  the adjoint variable is $\delta\mathbf{x}^*{}_i = \mathbf{M}_{i+1}^T\delta\mathbf{x}^*{}_{i+1} - \mathbf{H}_i{}^T\mathbf{R}_i{}^{-1}\mathbf{y}'{}_i$, with $\mathbf{H}_n{}^T\mathbf{R}_n{}^{-1}\mathbf{y}'{}_n$. (Trémolet, 2007)

- The dual case

$$\nabla_{\mathbf{w}} F = (\mathbf{R} + \mathbf{SDS}^T)\mathbf{w} - \mathbf{y}'$$

  the adjoint variable is $\mathbf{w}_i{}^* = \mathbf{M}_{i+1}^T\mathbf{w}_{i+1}^* + \mathbf{H}_i{}^T\mathbf{w}_i$, with $\mathbf{w}_n^* = \mathbf{H}_n{}^T\mathbf{w}_n$

- The gradient can still be calculated with one backward integration + one forward integration.

The adjoint variables

- The primal case

$$\nabla_{\delta\mathbf{z}} J = \mathbf{D}^{-1}\delta\mathbf{z} + \mathbf{S}^T\mathbf{R}^{-1}(\mathbf{S}\delta\mathbf{z} - \mathbf{y}')$$

  the adjoint variable is $\delta\mathbf{x}^*{}_i = \mathbf{M}_{i+1}^T\delta\mathbf{x}^*{}_{i+1} - \mathbf{H}_i{}^T\mathbf{R}_i{}^{-1}\mathbf{y}'{}_i$, with $\mathbf{H}_n{}^T\mathbf{R}_n{}^{-1}\mathbf{y}'{}_n$. (Trémolet, 2007)

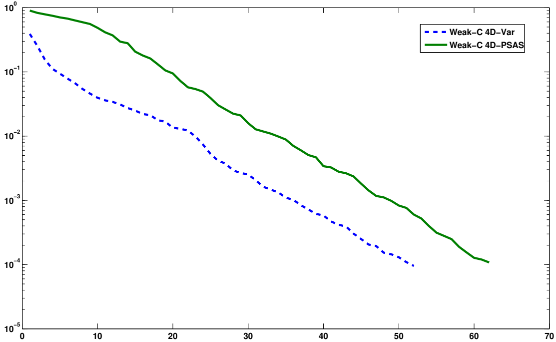
- The dual case

$$\nabla_{\mathbf{w}} F = (\mathbf{R} + \mathbf{SDS}^T)\mathbf{w} - \mathbf{y}'$$

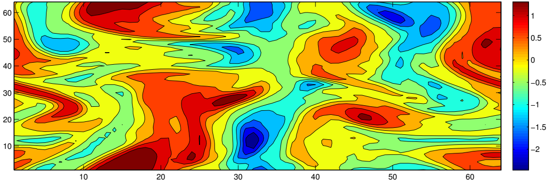
  the adjoint variable is $\mathbf{w}_i{}^* = \mathbf{M}_{i+1}^T\mathbf{w}_{i+1}^* + \mathbf{H}_i{}^T\mathbf{w}_i$, with $\mathbf{w}_n^* = \mathbf{H}_n{}^T\mathbf{w}_n$

- The gradient can still be calculated with one backward integration + one forward integration.

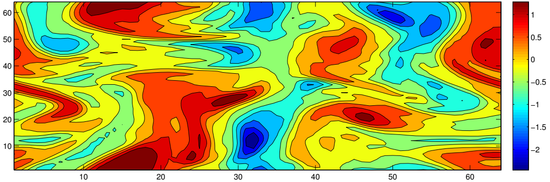
- Experiments : 2D-turbulent model solving for the barotropic vorticity on the β-plane.
- The model error : $\beta_{control} = 0.4$, and $\beta = 0.5$.
- Model error covariance matrices : $\mathbf{Q}_i = \alpha \mathbf{B}$.



- gradient norms of the Weak-C 4D-Var and 4D-PSAS with iterations



Analysis Increments (Strong−Constraint 4D−Var)



Analysis Increments (Weak−Constraint 4D−Var)

Questions currently examined (or about to be) in this context :

- The fit to the observations in the assimilation window and the total error in a weak-constraint assimilation compared to the S-C case...The impact of J_Q. (longer assimilation windows)

- Need to make sure the TLM validity holds.

- $\mathbf{Q} = \alpha\mathbf{B}$ is not the way to go. (the analysis increments are at best as "good" as the SC increments).

Conclusion

- The dual formulation of the variational data assimilation is a intresting scheme :
  $\hookrightarrow$ Equivalence of the results at convergence for the primal and dual cases (3D and 4D).
  $\hookrightarrow$ Both methods have similar convergence rates (Courtier, 1997), and the SV of their Hessians are equivalent (useful in preconditioning and cycling process).
  $\hookrightarrow$ With appropriate termination criterion, both methods converge with similar number of iterations.

Conclusion

- The biggest concern for the dual method has been fully explained.
- Using Minres as a minimization algorithm insead of the CG solves this problem.
- 3D/4D-PSAS can be used with confidence in operational implementations and in a weak-constraint framework.
- The implementation of a weak-constraint scheme (primal and dual) was made "relatively" easier with the modularity of the operators.
- The work on the model errors is still ongoing.

Conclusion

- The biggest concern for the dual method has been fully explained.
- Using Minres as a minimization algorithm insead of the CG solves this problem.
- 3D/4D-PSAS can be used with confidence in operational implementations and in a weak-constraint framework.
- The implementation of a weak-constraint scheme (primal and dual) was made "relatively" easier with the modularity of the operators.
- The work on the model errors is still ongoing.

Conclusion

- The biggest concern for the dual method has been fully explained.
- Using Minres as a minimization algorithm insead of the CG solves this problem.
- 3D/4D-PSAS can be used with confidence in operational implementations and in a weak-constraint framework.
- The implementation of a weak-constraint scheme (primal and dual) was made "relatively" easier with the modularity of the operators.
- The work on the model errors is still ongoing.

Conclusion

- The biggest concern for the dual method has been fully explained.
- Using Minres as a minimization algorithm insead of the CG solves this problem.
- 3D/4D-PSAS can be used with confidence in operational implementations and in a weak-constraint framework.
- The implementation of a weak-constraint scheme (primal and dual) was made "relatively" easier with the modularity of the operators.
- The work on the model errors is still ongoing.

Conclusion

- The biggest concern for the dual method has been fully explained.
- Using Minres as a minimization algorithm insead of the CG solves this problem.
- 3D/4D-PSAS can be used with confidence in operational implementations and in a weak-constraint framework.
- The implementation of a weak-constraint scheme (primal and dual) was made "relatively" easier with the modularity of the operators.
- The work on the model errors is still ongoing.

Thank you