Automatic Differentiation: A Tool For Data Assimilation and Sensitivity Analysis in Oceanography

Moulay HICHAM TBER*

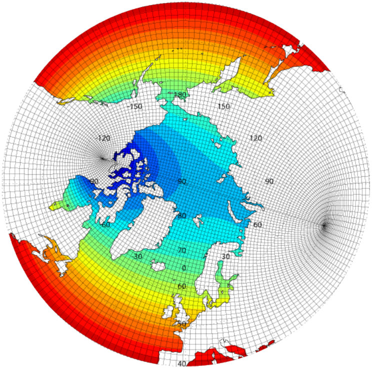
Projet TROPICS, INRIA, 2004 route des Lucioles Sophia Antipolis
hicham.tber@sop-inria.fr

Data Assimilation in Oceanography

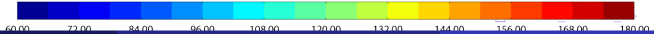
- Model $\mathcal{M}$ describing the distribution and evolution in space and time of the characteristics of the sea- State Variables Y-(velocity, temperature, pressure, ..)
- $\mathcal{M}$ depends, among others, on the initial conditions $Y_0 = X$.
- Observations : In-Situ, Spatial
- Data assimilation = estimating initial conditions (in our context)

Direct Model: OPA/NEMO

- Developed at LODYC-LOCEAN-Paris VI
- General ocean circulation model
- Configuration : ORCA 2^o : $i \times j \times k = 180 \times 149 \times 31$



Champ en couleur ( ): Min= 54.98, Max= 171.59, Int= 6.00



Optimal control theory summary

- Direct model
  The system state equation can be given by

$$\left\{ \begin{array}{c} y = \mathcal{M}(x) \\ y(0) = x \end{array} \right.$$

- Cost function:

$$2\mathcal{J}(x) = 2\mathcal{J}^0(x) + 2\mathcal{J}^b(x) = \|\mathcal{H}(y(x)) - y^0\|^2 + \|x - x^b\|^2$$

  $\mathcal{H}$: observation operator y^0 : observations x^b : background control
- minimization of $\mathcal{J}(x)$ with respect to the control vector x using gradient based algorithm

Incremental 4D-var

- Tangent linear approximation

$$\mathcal{M}(x+\delta x) = \mathcal{M}(x) + M\delta x,\ \delta x = x - x^b$$

$$\mathcal{H}(x+\delta x) = \mathcal{H}(x) + M\delta x,\ \delta x = x - x^b$$

- Cost function:

$$2\mathcal{J}(\delta x) = \|H.M\delta x - d\|^2 + \|\delta x\|^2$$

$d : y^o - \mathcal{H}(\mathcal{M}(x))$

- minimization of $\mathcal{J}(\delta x)$ with respect to the control vector δx using gradient based algorithm
- Quadratic cost function and shorter control vector. Nonlinearities taken by updating x^b.

Tangent and Adjoint Model Development for OPA

- Gradient based algorithm
- 4D-Var:

$$\mathcal{J}(x) \longrightarrow \mathcal{M}$$

$$\nabla\mathcal{J}(x) \longrightarrow M^T$$

- Incremental 4D-Var:

$$\mathcal{J}(\delta x) \longrightarrow M$$

$$\nabla\mathcal{J}(\delta x) \longrightarrow M \text{ and } M^T$$

$\mathcal{M}$: Model OPA
M : Tangent Linear Model of OPA
M^T : Adjoint Model of OPA

NEMO Tangent and Adjoint Model "NEMOTAM"
History

- 1992-94: 1st version developed by E. Greiner (LODYC) for OPA4.
  - Applied to 4D-Var with a tropical Atlantic configuration.
- 1995-96: Major rewrite for OPA7 by F. Van den Berghe (CETIIS) and A. Weaver (LODYC).
  - This version was never exploited scientifically.
- 1997-2001: Adapted to OPA8.0 - 8.1 by A. Weaver (LODYC-CERFACS).
  - Developed initially for 4D-Var with a tropical Pacific configuration (TDH).
  - Widely used, primarily for 4D-Var studies.
  - Applied to applications other than data assimilation (singular vectors / optimal perturbations).
- 2002-present: Developed for OPA8.2, free-surface version, by A. Weaver (CERFACS) and C. Deltel (LOCEAN).
  - 1st global ocean version (ORCA2¡).
  - Used for 4D-Var in the ENACT project.
  - Currently used by several groups for a variety of applications.

NEMOTAM Current Status

- Hand written tangent and adjoint model
- OPA 9.0 /NEMO: Major new version in Fortran 95.
- Development for OPA 9.0/NEMO using Automatic Differentiation.

Automatic Differentiation (AD) of Computer Programs

- Every programming language provides a limited number of elementary mathematical functions
- computer program, no matter how complicated, may be viewed as the composition of these so-called intrinsic functions

$$P = \{I_1; I_2; \ldots; I_{p-1}; I_p\} \text{ implement } F = f_p \circ f_{p-1} \circ \ldots f_1$$

- Derivatives for the intrinsic functions are combined using the chaine rule

$$F'(x_0 = x) = f'_p(x_{p-1}) \cdot f'_{p-1}(x_{p-2}) \cdots f'_1(x_0); \; x_i = f_i(x_{i-1})$$

Automatic Differentiation (AD) of Computer Programs

- But calculating and multiplying jacobians is too expensive

$$F'(x_0 = x) = f'_p(x_{p-1}) \cdot f'_{p-1}(x_{p-2}) \cdots f'_1(x_0)$$

- Tangent mode

$$y = F'(x) \cdot \dot{x} = f'_p(x_{p-1}) \cdot f'_{p-1}(x_{p-2}) \cdots f'_1(x_0)) \cdot \dot{x}$$

- Reverse mode

$$\overline{x} = F'^T(x) \cdot \overline{y} = f'^T_1(x_0) \cdots f'^T_{p-1}(x_{p-2}) \cdot f'^T_p(x_{p-1}) \cdot \overline{y}$$

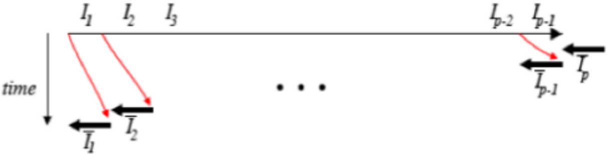
- Which mode ?

$$F : \mathbb{R}^m \longrightarrow \mathbb{R}^n$$

Tangent mode: $m \leq n$
Reverse mode: $m >> n$ (e.g. data assimilation)

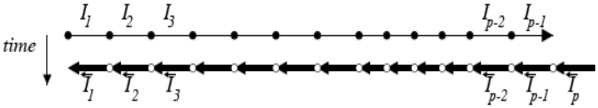
Automatic Differentiation (AD) of Computer Programs : Recompute vs Restore

- Form of adjoint code:

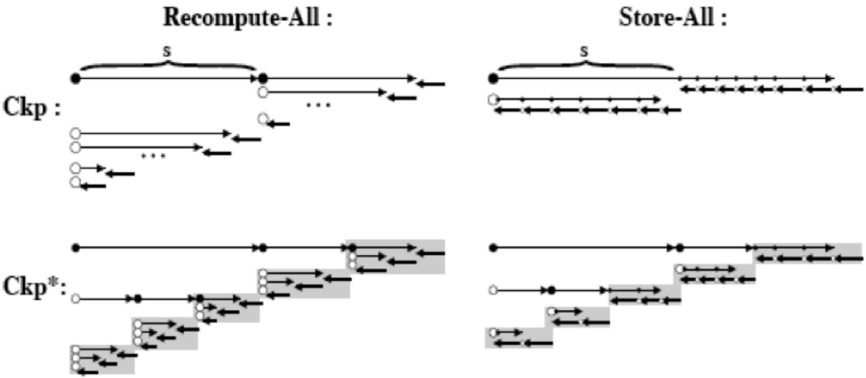


- Recompute all strategy:

- Restore all strategy:



Automatic Differentiation (AD) of Computer Programs : Checkpointing



Validation and Performances: Correctness test

- Compute a directional derivative $\nabla F(x)\dot{x}$ for a random direction $\dot{x}$ using finite difference and the tangent linear mode AD.
- Compute via the reverse mode a single adjoint $\bar{x} = \nabla F(x)^T \bar{y}$ for $\bar{y}$ equals $\dot{y}$ the output of the tangent linear mode.
- Check the following equality within the limits of the machine precision

$$\lim_{\varepsilon \longrightarrow 0} \left| \frac{F(x + \varepsilon \dot{x}) - F(x - \varepsilon \dot{x})}{2\varepsilon} \right|^2 = \dot{y}\dot{y} = \bar{x}\dot{x}$$

Validation and Performances : Correctness test

- Dot product test for 1000 iterations

Divided differences ($\varepsilon = 10^{-7}$)	4.405352760987440e+08
AD (tangent linear)	4.405346876439977e+08
AD (adjoint)	4.405346876439867e+08

Validation and Performances: Binomial Checkpointing

- Multilevel checkpointing e.g. MIT-gcm not optimal
- Optimal checkpointing 'treeverse/revolve'

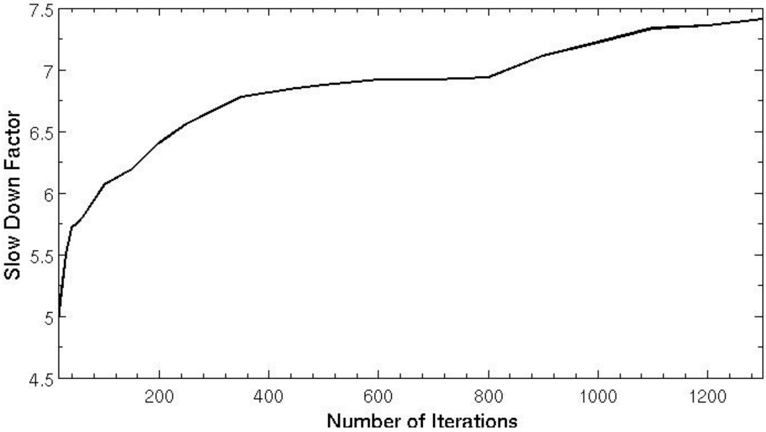
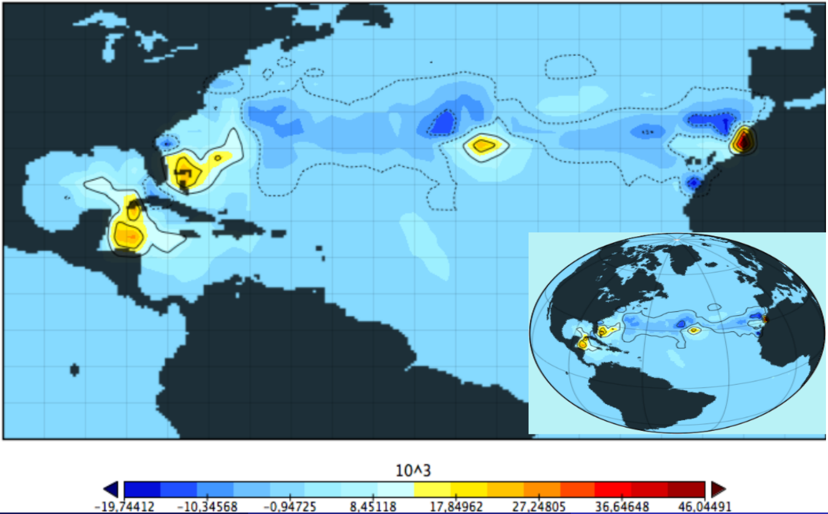


Figure: Slow dow factor vesus Max number of iterations

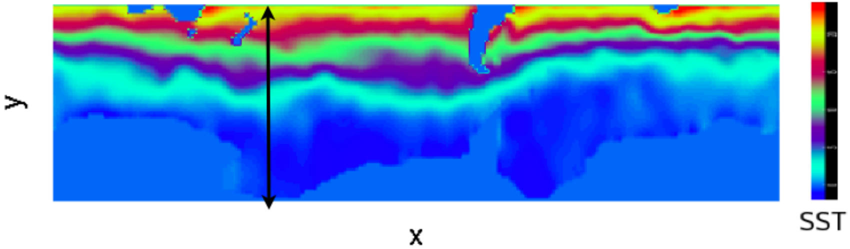
Validation and Performances: Sensitivity Analysis

- The sensitivity of the north atlantic heat transport at 29^0N, to changes in temperature at the ocean surface.

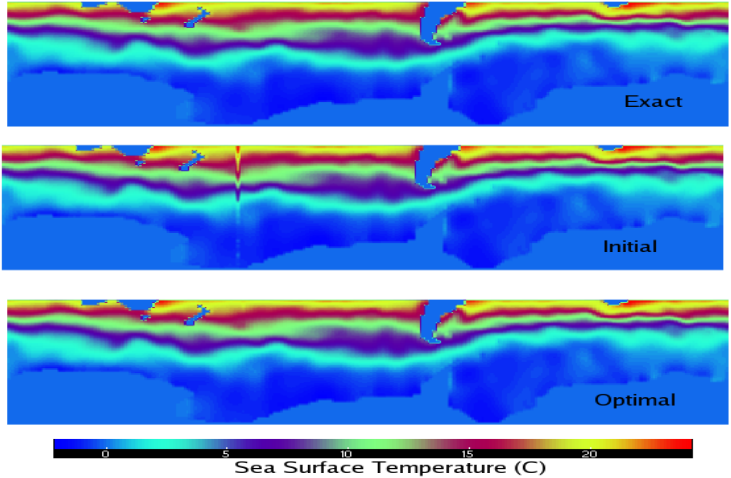


Twin Experiment

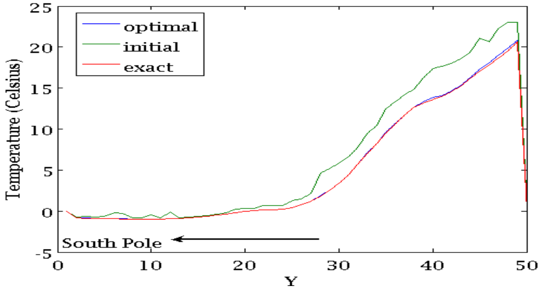
- Zoom on Antarctic
- Fully nonlinear approach
- Distributed observations
- Estimating sea surface temperature at x=60 (longitude)

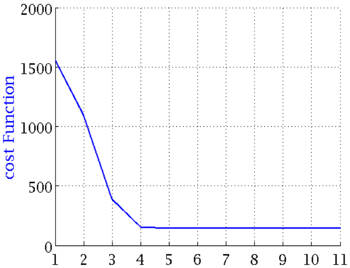


Twin Experiment: Results



Twin Experiments: Results





In Progress

- PCG Solver
- Wind stress assimilation for a stationary solution