CADNA for simulation and data assimilation: a user perspective

Julien Brajard, Pacôme Eberhart, Baptiste Landreau, Pierre Fortin, Fabienne Jézéquel

June 27, 2019

IPSL, Sorbonne Université, NERSC, LIP6

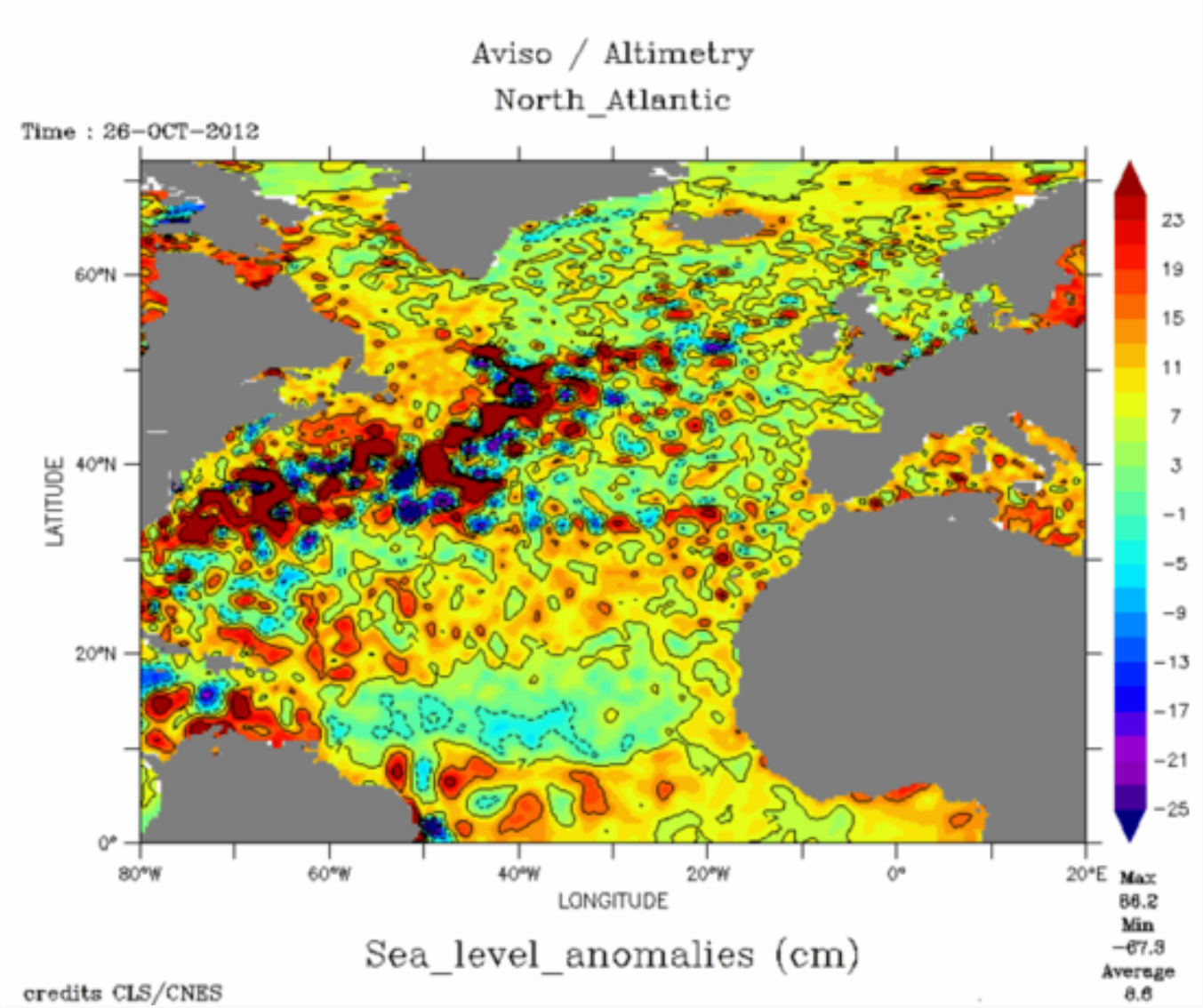
Earth System Modelling

Expensive simulation of the Earth System
(Ocean/Land/Atmosphere/Ice)

Ocean vorticity

source: E3SM

Earth System Observation



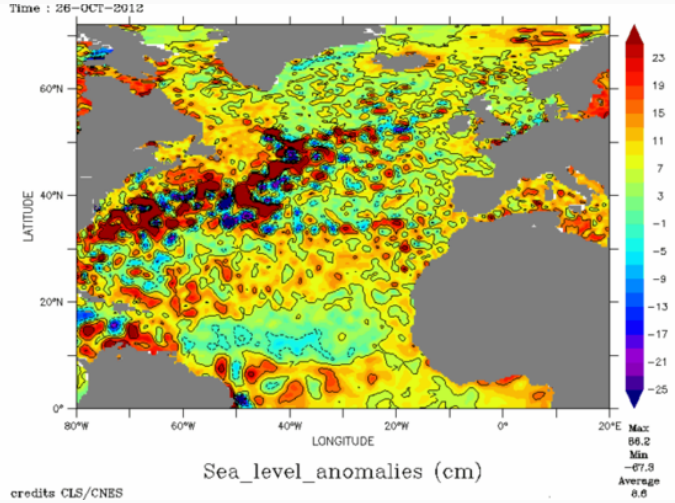
source: Aviso

Data assimilation

Model

+

Observation

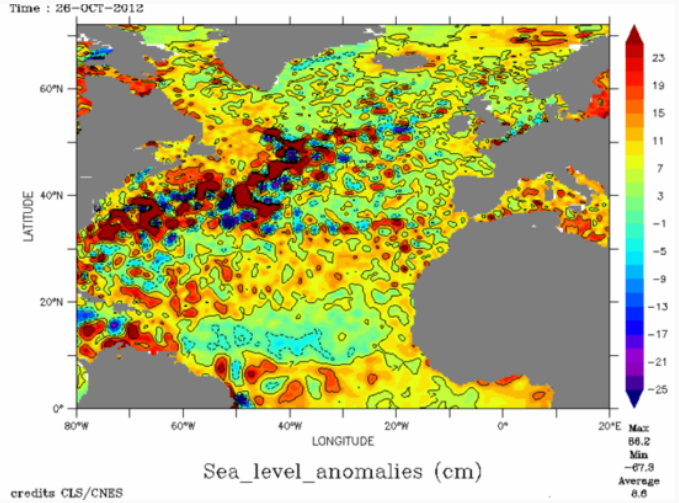


Data assimilation

Model

+

Observation



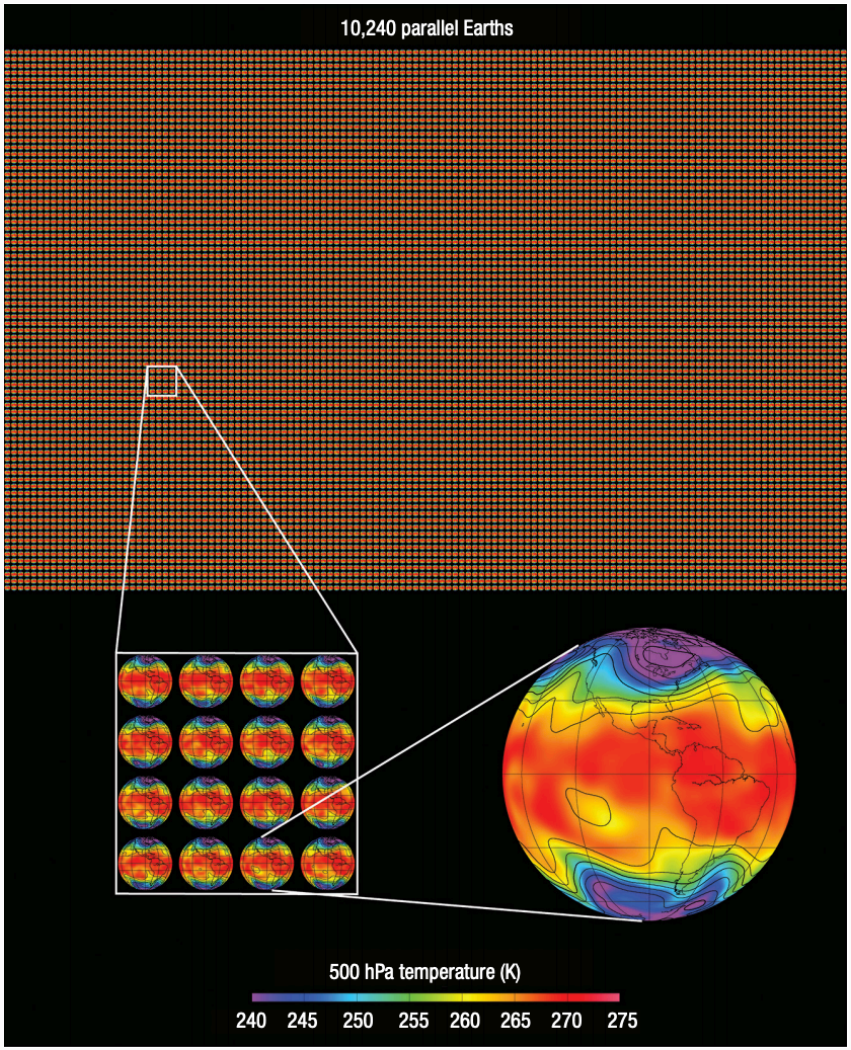
Data Assimilation

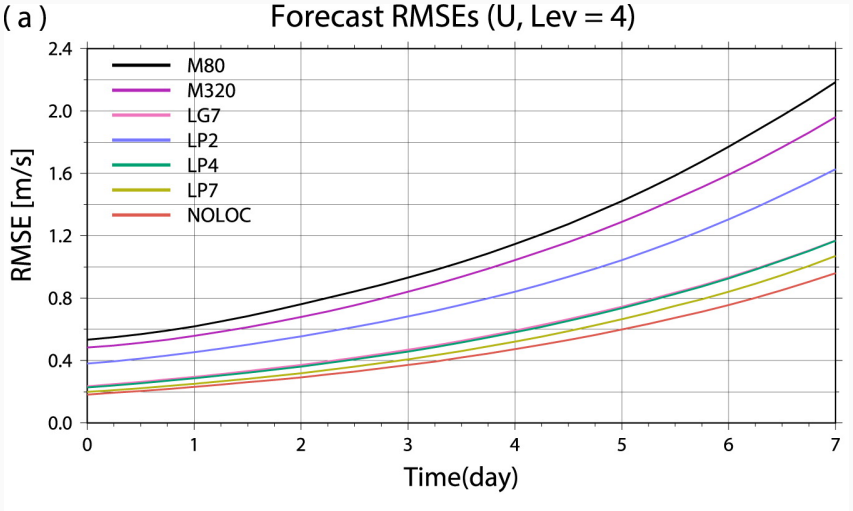
High Performance Computing

Simulation and data assimilation codes are HPC codes:

- Highly parallel computation on several cores
- High-resolution model
- Ensemble run (i.e. run several perturbed simulations).
- Toward hybrid architecture CPU/GPU

Illustration of ensemble size





from 80 to 10,000 members, forecast accuracy improves from 1 day to 5 days.

Kondo et al., 2016

Reduced precision

AGU100 ADVANCING EARTH AND SPACE SCIENCE

Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2018MS001341

Key Points:
- Lowering precision could accelerate an ensemble Kalman filter
- The level of precision used should fit the level of model error
- We perform tests with a spectral dynamical core

Choosing the Optimal Numerical Precision for Data Assimilation in the Presence of Model Error

Sam Hatfield[1], Peter Düben[2], Matthew Chantry[1], Keiichi Kondo[3], Takemasa Miyoshi[4], and Tim Palmer[1]

[1]Atmospheric, Oceanic and Planetary Physics, University of Oxford, Oxford, UK, [2]European Centre for Medium Range Weather Forecasts, Reading, UK, [3]Japan Meteorological Agency, Meteorological Research Institute, Tsukuba, Japan, [4]RIKEN Center for Computational Science, Kobe, Japan

Objective

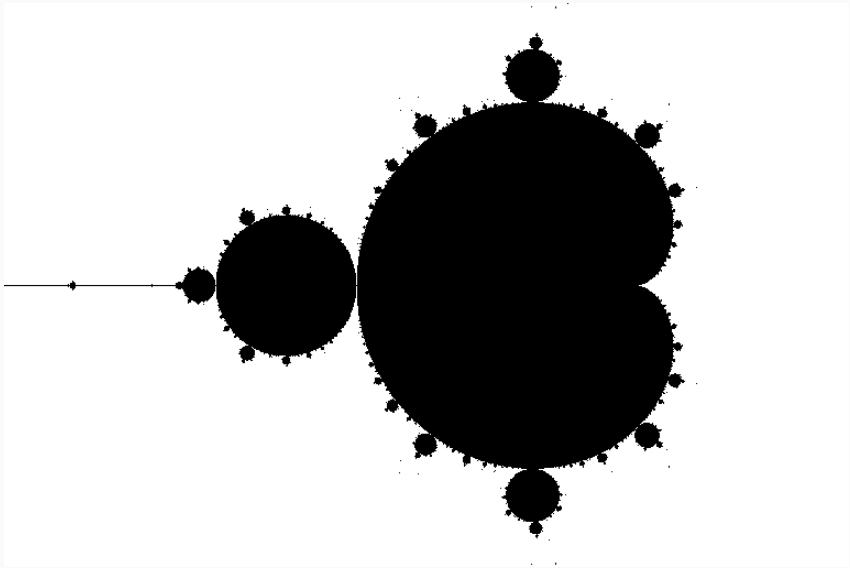
Estimating round-off errors using CADNA for simulation and Data assimilation code in HPC codes.

Available CADNA versions

- Sequential CADNA (no parallelism, CPU)
- CADNA-OpenMP
- CADNA-GPU

One example of diagnostics: Unstable test

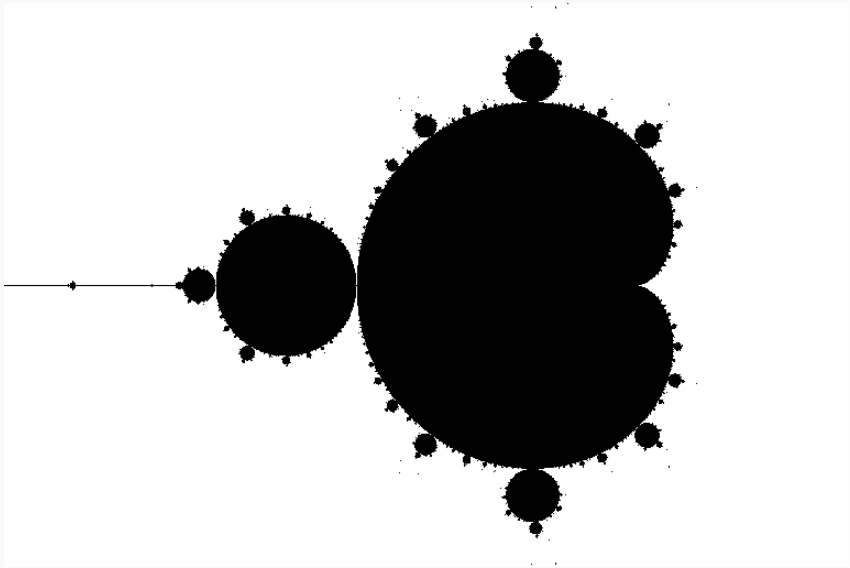
Detect unstable tests (`if z > 0`)



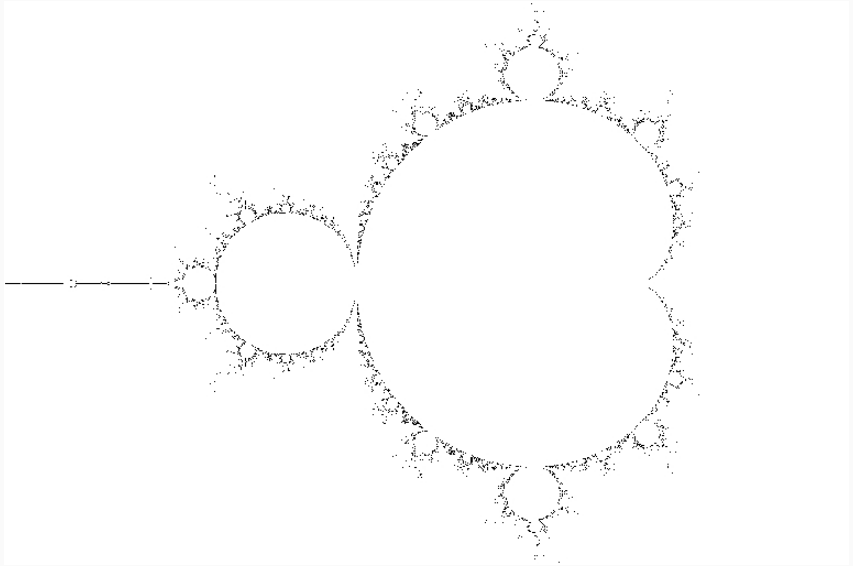
Mandelbrot set (computed on GPU)

One example of diagnostics: Unstable test

Detect unstable tests (`if z > 0`)



Mandelbrot set (computed on GPU)

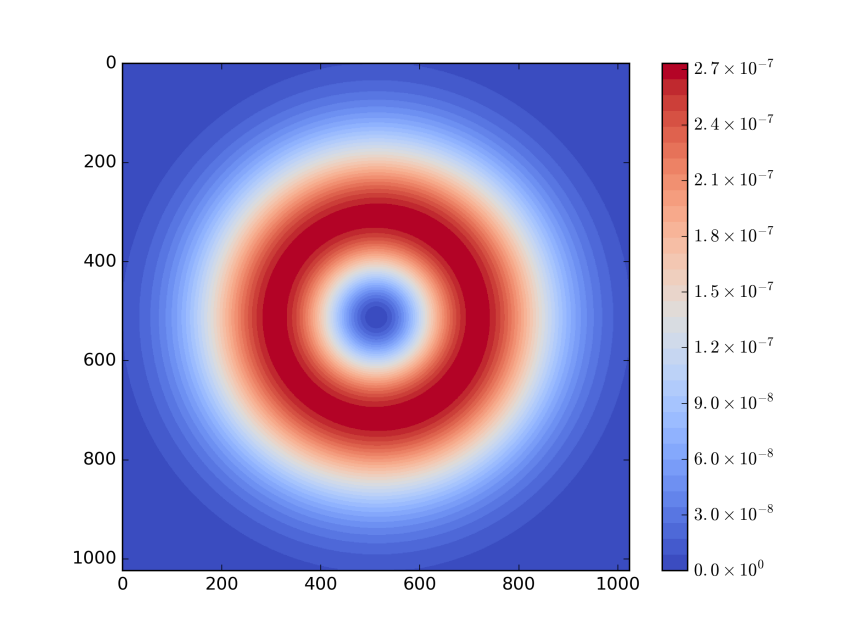


Unstable tests

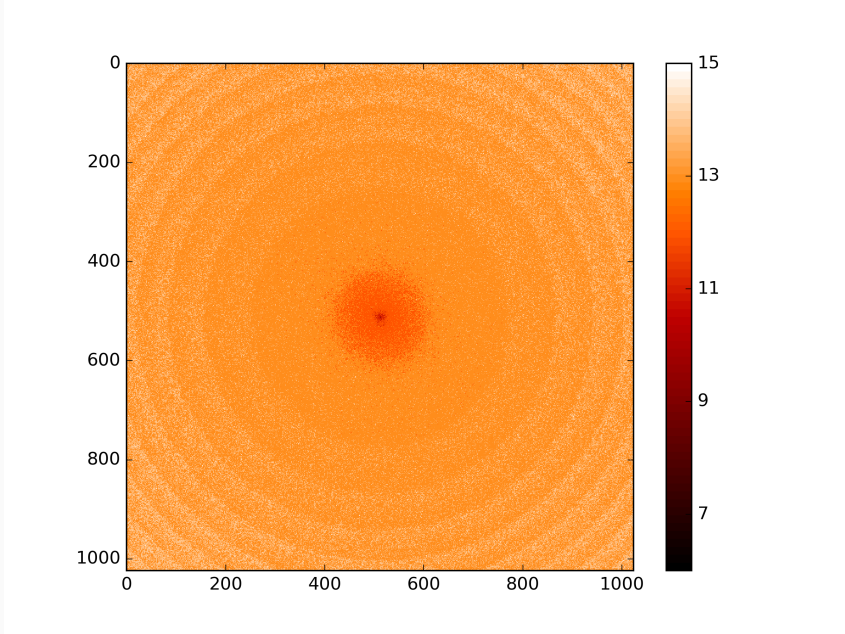
Eberhart et al., 2018

Number of exact significant digit

Application on shallow-water simulation (motion of a flow)



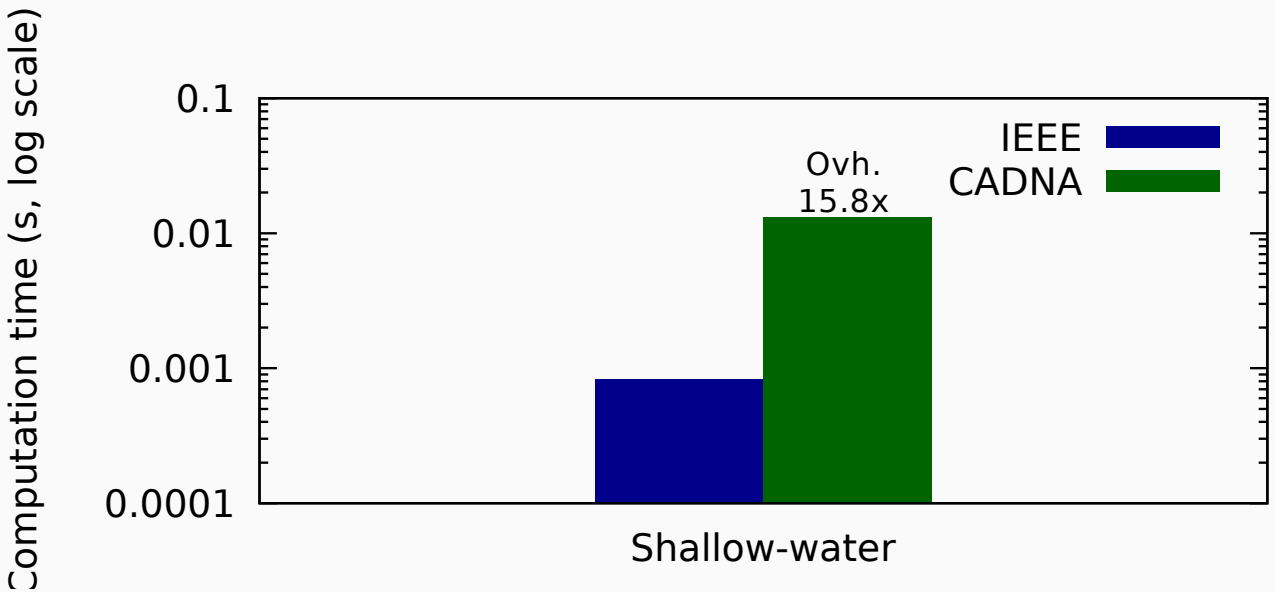
Velocity of the fluid



Number of exact significant digit

Eberhart et al., 2018

CADNA's overhead on GPU



Adjoint operator in Data assimilation

Data assimilation minimize the misfit between a model $\mathcal{M}$ and observations $\mathbf{y}_k$ at time t_k:

$$J(\mathbf{x}_0) = \frac{1}{2} \sum_{k=0}^{K} \|(\mathbf{y}_k - \mathbf{H} \circ \mathcal{M}_{k:0}(\mathbf{x_0})\| + R(\mathbf{x}_0)$$

where:

- $\mathbf{x}_0$ is the state of the model (e.g. velocity field)
- $\mathbf{H}$ is the observation operator
- R is a regularization term

Minimization using a gradient descent technique. The gradient is defined by

$$\nabla_{\mathbf{x}_0} J(\mathbf{x_0}) = -\sum_{k=0}^{K} \mathbf{M}_{k:0}^{\mathrm{T}} \mathbf{H}^{\mathrm{T}} \left[\mathbf{y}_k - \mathbf{H} \circ \mathcal{M}_{k:0}(\mathbf{x_0})\right] + \nabla_{\mathbf{x}_0} R(\mathbf{x}_0)$$

where $\mathbf{M}_{k:0}^{\mathrm{T}}$ is the adjoint operator

Assessing the validity of an adjoint code in 4D-VAR

In practice, two codes are available:

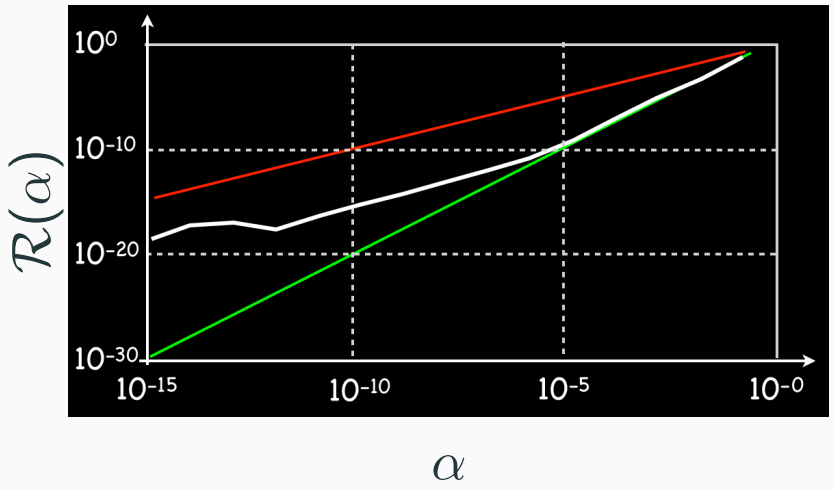
- A code for integrating the model between t_0 and t_k: $\mathcal{M}_{k:0}$
- A code for applying the adjoint operator: $\mathbf{M}_{k:0}^{\mathrm{T}}$

A sanity check is performed using the residual quantity defined by:

$$\begin{aligned}
\mathcal{R}(\alpha) &= J(\mathbf{x}_0 + \alpha\delta\mathbf{x}) - J(\mathbf{x}_0) - \alpha\nabla_{\mathbf{x}_0} J(\mathbf{x_0}).\delta\mathbf{x} \\
&= \mathcal{O}(\alpha^2) \text{ if adjoint code correct} \\
&= \mathcal{O}(\alpha) \text{ if adjoint code uncorrect}
\end{aligned}$$

Result on a land surface model

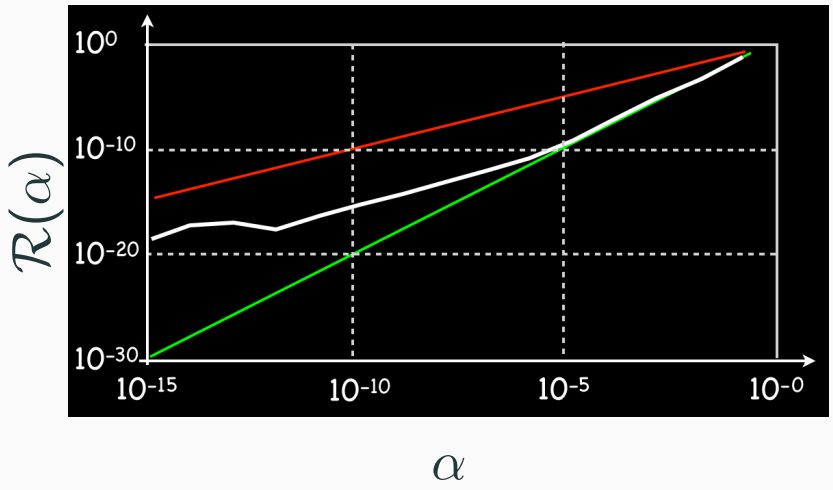
Assess the numerical validity of the residual $\mathcal{R}(\alpha)$ with CADNA



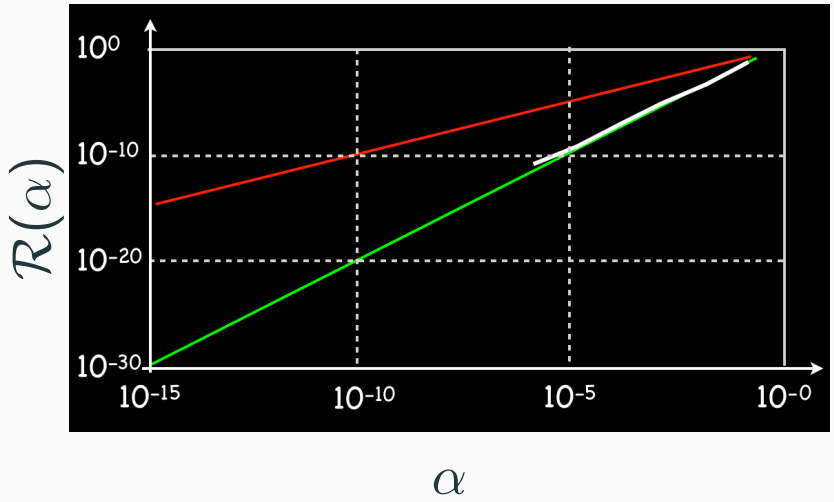
without CADNA

Result on a land surface model

Assess the numerical validity of the residual $\mathcal{R}(\alpha)$ with CADNA



without CADNA

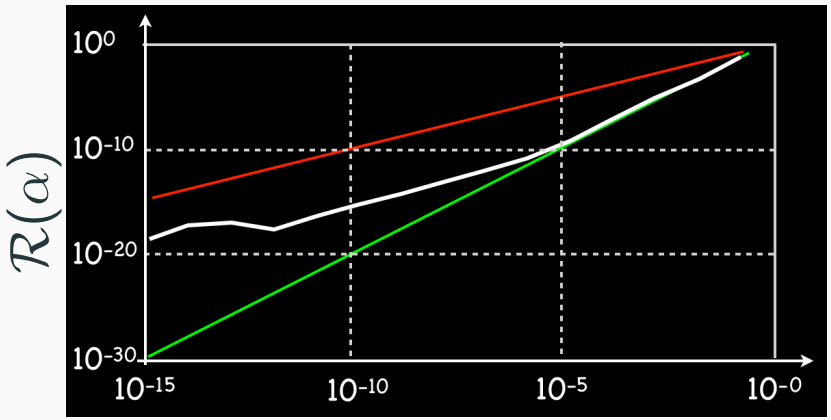


with CADNA sequential

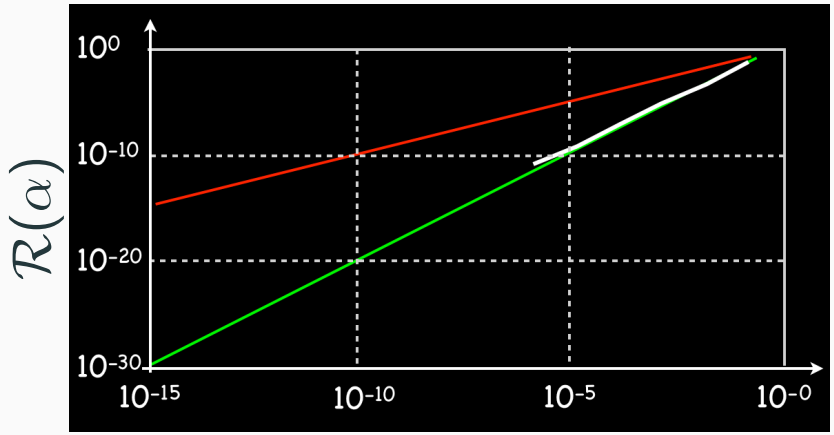
Brajard et al., 2013

Result on a land surface model

Assess the numerical validity of the residual $\mathcal{R}(\alpha)$ with CADNA



without CADNA



with CADNA sequential

Brajard et al., 2013

Numerical result using shallow-water with CADNA-OpenMP

	IEEE	CADNA
Serial	3.446611873236805E-06	3.4461E-06
OpenMP - 16 treads	3.446619149194419E-06	3.446E-06

Eberhart et al., 2016

Conclusions

CADNA:

- enables numerical validity of the code (number of exact significant digits, unstable tests, ...)
- can be used in an HPC context (with an overhead)
- can be applied for simulation and data assimilation codes.

`http://cadna.lip6.fr/`


Julien Brajard, Pei Li, Fabienne Jézéquel, Hector-Simon Benavidès, and Sylvie Thiria.
Numerical validation of data assimilation codes generated by the yao software.
In *SIAM Annual Meeting*, 2013.


Pacôme Eberhart, Julien Brajard, Pierre Fortin, and Fabienne Jézéquel.
High performance numerical validation using stochastic arithmetic.
Reliable Computing, 21:35–52, 2015.


Pacôme Eberhart, Julien Brajard, Pierre Fortin, and Fabienne Jézéquel.
Estimation of round-off errors in openmp codes.
In *International Workshop on OpenMP*, pages 3–16. Springer, 2016.


Pacôme Eberhart, Baptiste Landreau, Julien Brajard, Pierre Fortin, and Fabienne Jézéquel.
Improving cadna performance on gpus.
In *2018 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW, PDSEC)*, pages 1016–1025. IEEE, 2018.

julien.brajard@sorbonne-universite.fr, julien.brajard@nersc.no

Julien Brajard, Pei Li, Fabienne Jézéquel, Hector-Simon Benavidès, and Sylvie Thiria.
Numerical validation of data assimilation codes generated by the yao software.
In *SIAM Annual Meeting*, 2013.

Pacôme Eberhart, Julien Brajard, Pierre Fortin, and Fabienne Jézéquel.
High performance numerical validation using stochastic arithmetic.
Reliable Computing, 21:35–52, 2015.

Pacôme Eberhart, Julien Brajard, Pierre Fortin, and Fabienne Jézéquel.
Estimation of round-off errors in openmp codes.
In *International Workshop on OpenMP*, pages 3–16. Springer, 2016.

Pacôme Eberhart, Baptiste Landreau, Julien Brajard, Pierre Fortin, and Fabienne Jézéquel.
Improving cadna performance on gpus.
In *2018 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW, PDSEC)*, pages 1016–1025. IEEE, 2018.

julien.brajard@sorbonne-universite.fr, julien.brajard@nersc.no

---

Call for papers

2-4 October 2019

https://sites.google.com/view/climateinformatics2019