Comparison of sequential and variational reduced order data assimilation methods in the tropical Pacific ocean

Céline ROBERT[1], Eric BLAYO[1] and Jacques VERRON[2]

[1]LMC-IMAG, Grenoble, France
[2]LEGI, Grenoble, France

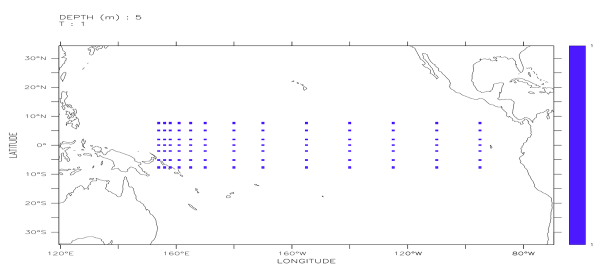
Objectives of the work

- **Synthesis of previous works performed on reduced-order data assimilation methods**
  - ▷ Reduced-order 4D-Var
  - ▷ Seek filter
- Implement and compare each method in the same realistic configuration
  - ▷ Tropical Pacific ocean
  - ▷ Assimilation of real temperature data profiles
- Perform an hybridation of the two methods in this same configuration.

1. Configuration of the experiments : Tropical Pacific ocean

1.1 Numerical model

Numerical model : OPA-TDH (Vialard *et al.*, 99)
Tangent linear model and adjoint model (Weaver and Vialard, 2003)
Atmospheric forcing : daily ERS-TAO winds and ECMWF fluxes
One year of experiment : 1993.



1.2 Data assimilation parameters

- Observed data : 3D temperature field from TAO/TRITON array
- Observation error : Diagonal error covariance matrix on observations **R**$= 1/\sigma_i^2 I_n$ with $\sigma_i = 0.5^o$ C

2. Incremental variational algorithm

2.1. Cost function

The cost function mesures the distance between the model solution and the observations. A regularisation term is added: the solution must be close to the background.

- $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_b$, $J_{tot}(\delta\mathbf{x}) = J_b(\delta\mathbf{x}) + J_o(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}\mathbf{B}^{-1}\delta\mathbf{x}^T + \|\mathbf{y} - H(\mathbf{x})\|^2_{obs}$
- Incremental formulation : $J_o = \frac{1}{2}\sum_{i=1}^{n}(\mathbf{d}_i - \mathbf{G}_i\delta\mathbf{x})^T\mathbf{R}_i^{-1}(\mathbf{d}_i - \mathbf{G}_i\delta\mathbf{x})$ with $\mathbf{d}_i = \mathbf{y}_i - H_i[\mathbf{x}^b(t_i)] = \mathbf{y}_i - G_i(\mathbf{x}^b)$

The dependance on δx is linear which allows simple descent solving method to find the minimum. This minimisation is performed using a gradient method, in which the gradient of the cost function is calculted using the **adjoint model**.

2.2. Reduced order approach

We look for an increment in a reduced dimension space spanned by a set of well-chosen vectors. The increment is calculated in this subspace by : $J_b(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{B}_w^{-1}\mathbf{w}$ with $\delta x^r = \sum_{i=1}^{r} w_i\mathbf{L}_i = \mathbf{L}\,\mathbf{w}$ where $\mathbf{w} = (w_1, ..., w_r)$

- Definition of **B** in the reduced space : $\mathbf{B}_w = E[(w - \overline{w})(w - \overline{w})^T]$
- Definition of **B** in the full space : $\mathbf{B}_r = E[(\delta x - \delta\overline{x})(\delta x - \delta\overline{x})^T]$

3. SEEK filter algorithm

Initialisation
The algorithm is initialised by a mean state of the model without assimilation.

$\mathbf{x}_{k=0} = \mathbf{x}_0$
$\mathbf{P}_{k=0} = \mathbf{S}^0\mathbf{S}^{0T}$

Forecast
Forecast is performed over 10 days.

$\mathbf{x}^f_{k+1} = M_{k,k+1}[\mathbf{x}^a_k]$
$\mathbf{P}^f_{k+1} = \frac{1}{\rho}\mathbf{S}^f_{k+1}(\mathbf{S}^f)^T_{k+1}$

Analysis
Analyses are performed every 10-days, with observations from the past 10-days

$\mathbf{P}^f_k = \mathbf{S}^f_k\mathbf{S}^{fT}_k$
$\mathbf{K}_k = \mathbf{S}^f_k[I + (\mathbf{H}_k\mathbf{S}^f_k)^T\mathbf{R}_k^{-1}(\mathbf{H}_k\mathbf{S}^f_k)]^{-1}(\mathbf{H}_k\mathbf{S}^f_k)^T\mathbf{R}_k^{-1}$
$\mathbf{x}^a_k = \mathbf{x}^f_k + \mathbf{K}_k[\mathbf{y}_k - \mathbf{H}_k\mathbf{x}^f_k]$
$\mathbf{P}^a_k = \mathbf{S}^f_k[I + (\mathbf{H}_k\mathbf{S}^f_k)^T\mathbf{R}_k^{-1}(\mathbf{H}_k\mathbf{S}^f_k)]^{-1}\mathbf{S}^{fT}_k$

Basis evolution
In case of basis evolution, each vector is propagated over the cycle by the fully non-linear model.

$\mathbf{P}^f_{k+1} = \mathbf{P}^a_k = \mathbf{S}^k_a\mathbf{S}^{kT}_a$
$[\mathbf{S}^f_{k+1}]_r = M[\mathbf{x}^a_k + [\mathbf{S}^a_k]_r] - M[\mathbf{x}^a_k]$

Link Seek-Reduced 4D-Var :

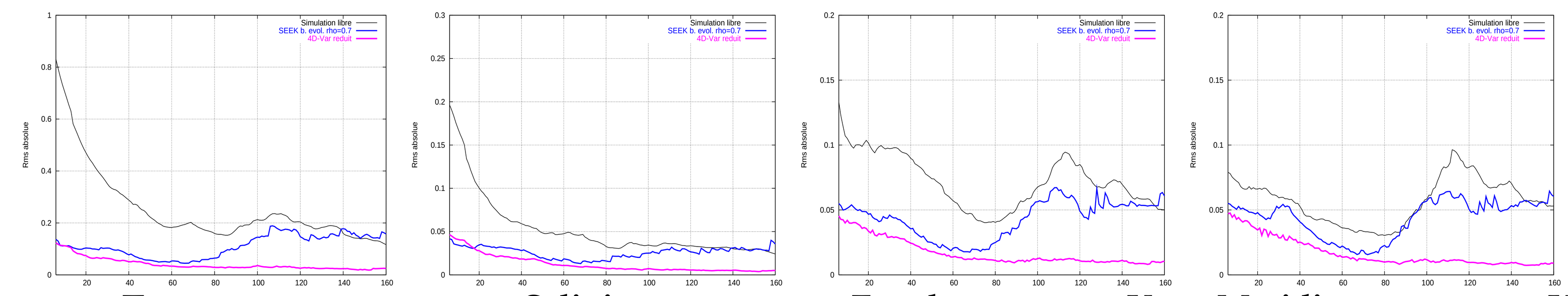
The reduced-dimension error space is the same for the Reduced 4D-Var and the Seek filter.

$\mathbf{P}_{k=0} = \mathbf{S}^0\mathbf{S}^{0T}$

$\Rightarrow \mathbf{P}_{k=0} = \mathbf{L}\Lambda\mathbf{L}^T \quad \equiv \mathbf{B}_r$

4. Twin experiments

The observations are generated by the non-linear model at TAO points. The experiments are performed over the year 1993.
The initial basis is composed of the thirty first modes of an Eofs analysis of a free run over the year 1993. The Seek filter runs with a forgetting factor equal to 0.7. The basis is fixed. Rms error values are calculated on the 2nd level of the model : 15 m. depth.

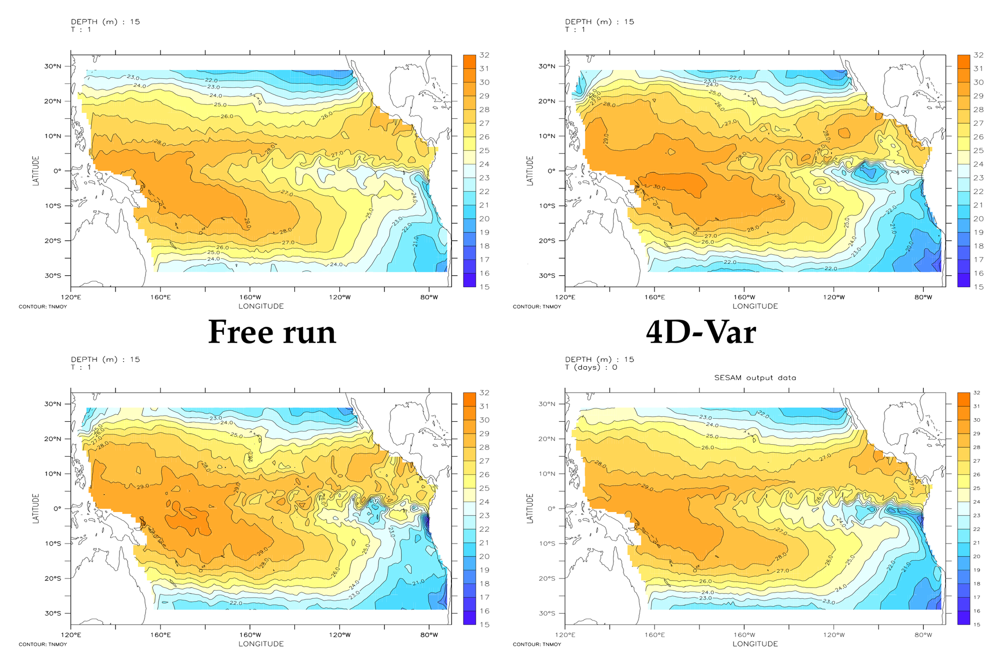


Temperature **Salinity** **Zonal component U** **Meridian component V**

For each variable, the error level is very low. We can see that Reduced-4D-Var obtain the best error level at 15 m. However, the cost of the Seek filter is ten times cheaper than the cost of the Reduced 4D-Var.

5. True experiments

The observations are real profiles from TAO array and XBT. The experiments are performed over the year 1993.
The Seek filter runs with a forgetting factor equal to 1. The basis is fixed.
Analysis : these figures shows the annual mean of the analysed states.

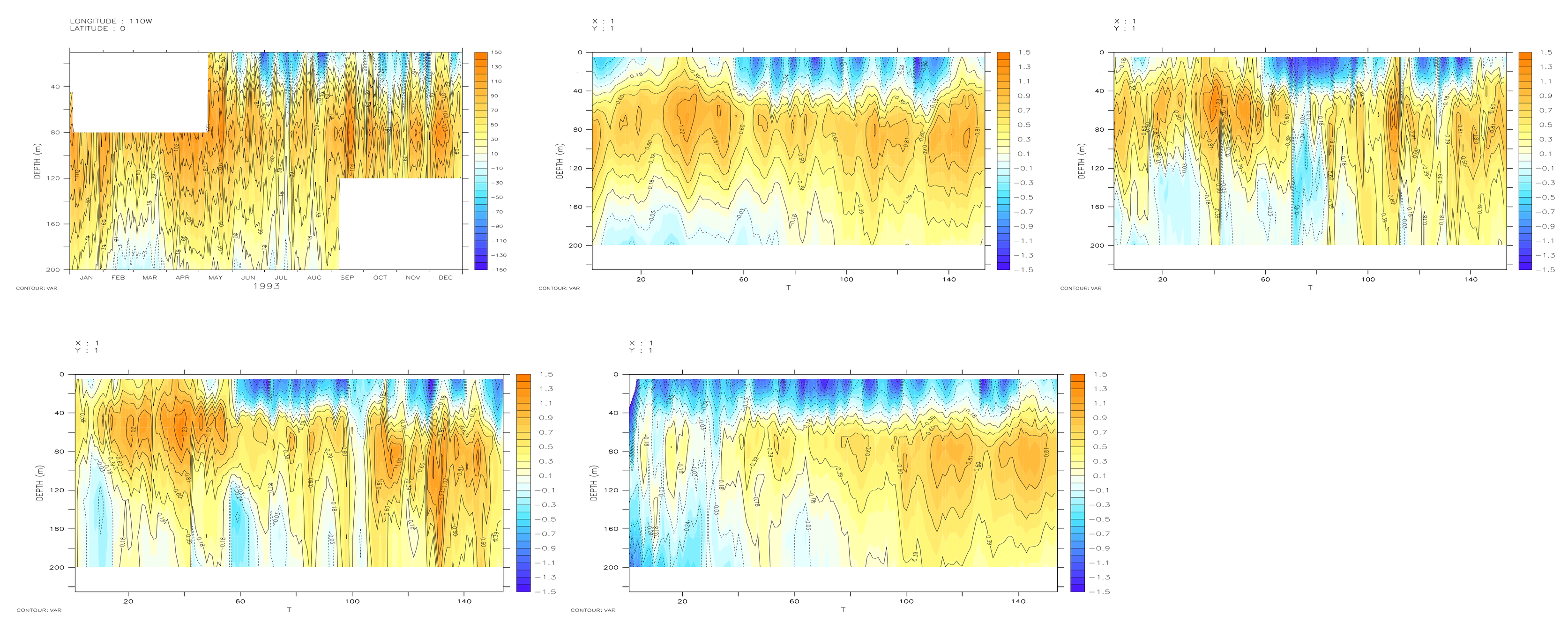


Free run **4D-Var** **R-4D-Var/4D-Var** **Seek filter**

The Reduced 4D-Var alone doesn't work well in this real data experiment. The Seek filter tends to produce low-gradient states. The best solution is obtained by a mixed algorithm Reduced 4D-Var/4D-Var, combining 10 iterations of Reduced 4D-Var followed by 10 iterations of full 4D-Var.

6. Validation of the results

We compare model forecasts with (U,V,T,S) data profiles at TAO points. Globally the model without assimilation provides a solution which is too smooth. Data assimilation with the Seek filter tends to improve these results. The best solution is provided by the mixed 4D-Var (but for a 10 times greater computation cost).
We can see on these figures a vertical profile of the zonal component of the velocity at (110°W 0°N), as a function of time.



6. Conclusions

Reduced 4D-Var is a very good alternative to classical 4D-Var in twin experiments : it leads to a better identification of the true state for a cheaper computation cost.
When assimilating real data, the Reduced 4D-Var alone doesn't lead to realistic solution while the Seek filter does. Improved results seem to be obtained with a Mixed Reduced 4D-Var/4D-Var approach. These conclusions need to be confirmed in an El Niño situation in which the Seek filter is known to be robust.

References

[1] S. Durbiano, Vecteurs caractéristiques de modèles océaniques pour la réduction d'ordre en assimilation de données, Phd Thesis, LMC-UJF, (2001)

[2] L.Parent, Assimilation de données dans l'océan PAcifiwue Tropical sur la période 1994-1998, Phd Thesis,LEGI-UJF,, (2000)

[3] D.T. Pham, J. Verron, M.C. Roubaud, A singular evolutive extended Kalman filter for data assimilation in oceanograhy, Inverse Problems, Vol. 14, 1998.

[4] A.T. Weaver, J. Vialard,D.L.T. Anderson, Three- and four-dimensional variational assimilation with a general circulation model of the tropical Pacific Ocean, Part 1 : formulation, internal diagnostics and consistency checks, Monthly Weather Review, 2002.

Acknowledgements

This work, perfomred in LMC-IMAG and LEGI laboratories is supported by CNES.