Adaptive parameterisation of error statistics in ensemble or reduced order square root filters

E. Kpemlie, N. Freychet, E. Cosme, J-M. Brankart, P. Brasseur

LEGI,CNRS/ University of Grenoble BP 53X, 38402 Grenoble Cedex 9, France

emmanuel.kpemlie@hmg.inpg.fr

INTRODUCTION

A good representation of state error statistics in Kalman filters is often difficult to obtain because the nature of errors is complex, poorly known and multi-source. Also, these statistics must be represented with a limited number of error modes or ensemble members, what can easily lead to the filter divergence.

A recipe consists in adjusting the statistics online during the assimilation process. This is called “adaptive filtering”.

A variety of strategies have been developed [*4,5*] since the seminal approach introduced for atmospheric data assimilation by ***Dee (1995)***.

In a recent paper by ***Brankart et al. (2010),*** the formulation of Dee (1995) is:

- advantageously formulated for a reduced order square root filter,
- extended to the adaptive adjustment of observation error statistics,
- designed to estimate optimal adaptive parameters at a reasonable computational cost.

The objective of the present study is to apply and evaluate the formulation developed in [*2*] in a more realistic oceanographic context. The benefits of adaptive filtering is illustrated with results from twin experiments with a high resolution Ocean model.

METHOD

1. Kalman filter analysis equations

$d_k = y_k - H_k x_k^f$ (innovation)

$C_k = H_k P_k^f H_k^T + R_k$ (innovation cov matrix)

$K_k = P_k^f H_k^T C_k^{-1}$ (Kalman gain)

$x_k^a = x_k^f + K_k d_k$ (filter analysis)

$P_k^a = (I + K_k H_k) P_k^f$ (analysis cov matrix)

2. The adaptive scheme of Dee *[1]*

- P_k^f is written $P_k^f(\alpha_k)$ and α_k is determined to get the best fit between $<d_k d_k^T>$ and C_k.
- This is done by minimizing the likelihood function for α_k which models the optimization problem: $J_k(\alpha_k) = d_k^T C_k(\alpha_k)^{-1} d_k + \ln |C_k(\alpha_k)|$

3. An extension of the scheme of Dee (*Brankart et al., 2010*)

- Set $P_k^f = P_k^f(\alpha_k)$ and $R_k = R_k(\beta_k)$ α_k and β_k: random vectors of parameters
- Reformulate the cost function by adding prior pdf p_0 (α_k and β_k non-constant in time) and diffusion parameter f to assume the effect of time in the estimation of innovation statistics:

$$J_k(\alpha_k,\beta_k) = -\ln[p_0(\alpha_k,\beta_k)] + \frac{1}{2}\sum_{k'=1}^{k} f^{k-k'}\left[d_{k'}^T C_{k'}^{-1}(\alpha_k,\beta_k) d_{k'} + \ln|C_{k'}(\alpha_k,\beta_k)|\right]$$

  - With f = 0 and without the first term: solution of *Dee [1]*

- Take advantage of the reduced order approach of SEEK or ETKF [*3*]: by setting

i) $P_k^f = S_k^f S_k^f$ *ii)* $\delta_k = (H_k S_k^f)^T R_k^{-1} d_k$ *iii)* $U_k \Lambda_k^{-1} U_k^T = (H_k S_k^f)^T R_k^{-1} (H_k S_k^f)$

  - where δ_k is the projection of the vector d_k, U_k and Λ_k^{-1} result from SVD of the right hand side term of (*iii*)

- The cost function is written then: $\boxed{J_k(\alpha_k,\beta_k) = -\ln[p_0(\alpha_k,\beta_k)] + \frac{1}{2}\sum_{k'=1}^{k} f^{k-k'}\left[J_{k,k'}^A + J_{k,k'}^B\right]}$

With $J_{k,k'}^A = d_{k'}^T R_{k'}^{-1} d_{k'} - \delta_{k'}^T U_{k'}\left[I + \Lambda_{k'}^{-1}\right]^{-1} U_{k'}^T \delta_{k'}$ and $J_{k,k'}^B = \ln|R_{k'}| + tr\{\ln[I + \Lambda_{k'}^{-1}]\}$

- From innovation space (*n x n*), J_k is defined in reduced space (*r x r*), with *r* < *n*
- These formulas avoid the computation of the inverse and the determinant of matrix C_k^{-1} (Λ^{-1} is diagonal)

✓ **The adaptive scheme developed is implemented with the SEEK filter** *[3]*.

✓ In the scalar case without prior pdf p_0, the optimal α_k is $\alpha_k^* = \frac{d_k d_k^T - R_k}{H_k P_k^f H_k^T}$ which can be negative when the system is not accurately initialized. To prescribe a positive prior is essential. Here a Gamma distribution with mode 1 and parameter $\gamma(\mu,\theta)$ is taken as the prior: $p_0(\alpha_k) \approx \alpha_k^{\mu-1} \exp(-\alpha_k/\theta)$

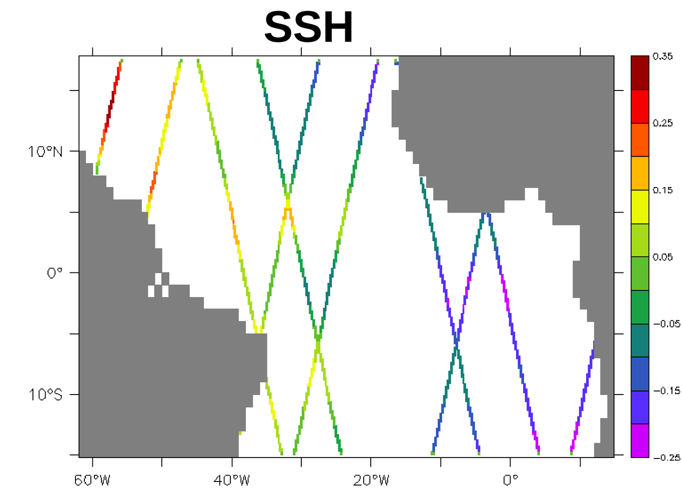
TWIN EXPERIMENTS

Physical state variables: **SSH, T**emperature**, S**alinity**, U, V**

Observation variable: **SSH** (along-track satellite, type Jason), **T** and **S** (spare vertical profile: in situ data type, ARGO float)

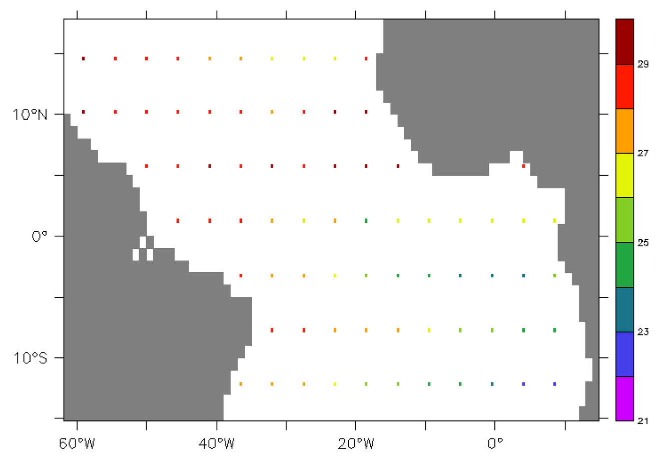
Experiment field: The adaptive filter is tested with a regional, **Tropical Atlantic** configuration of NEMO model at a resolution of **¼ degree**. See ***Freychet's poster in OS4.1***

SSH



Observed variable for 1 day

T



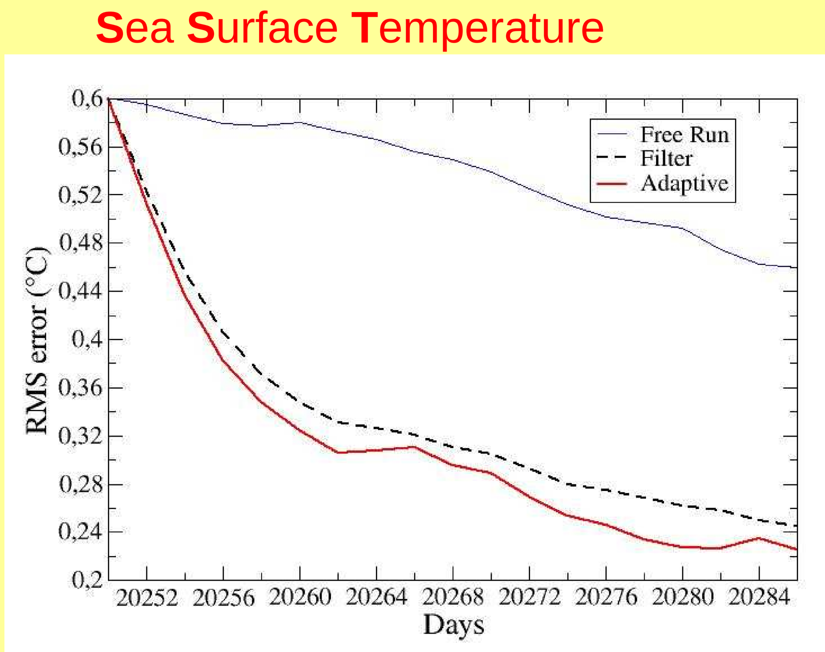
Experiment protocol: Local analysis with evolutive error statistics

- True state (**ref. state**): year 2005 - Initial state (**free run**): year 2003
- Initial state error: **39 modes** from an EOF analysis of a 6-year free run
- $\boldsymbol{P} = \alpha \boldsymbol{P}^f$ and only α is found
- Observations errors R: * **SSH** (0.3 m) * **T** (3 °C) * **S** (3)

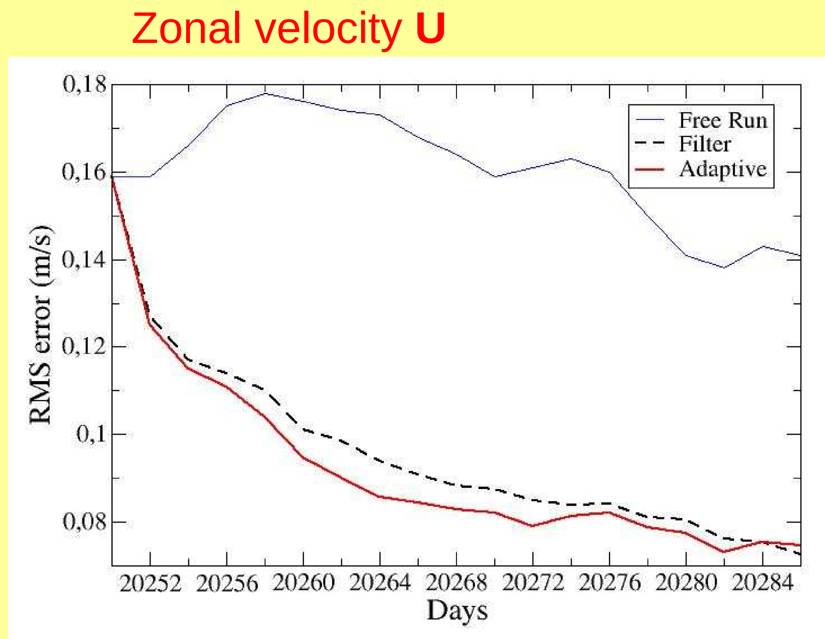
RESULTS

- 1 month run and assimilation cycle every 2 days: from 23/06/05 to 29/07/05
- Filter analysis step compared with adaptive analysis step: *RMS* error computed

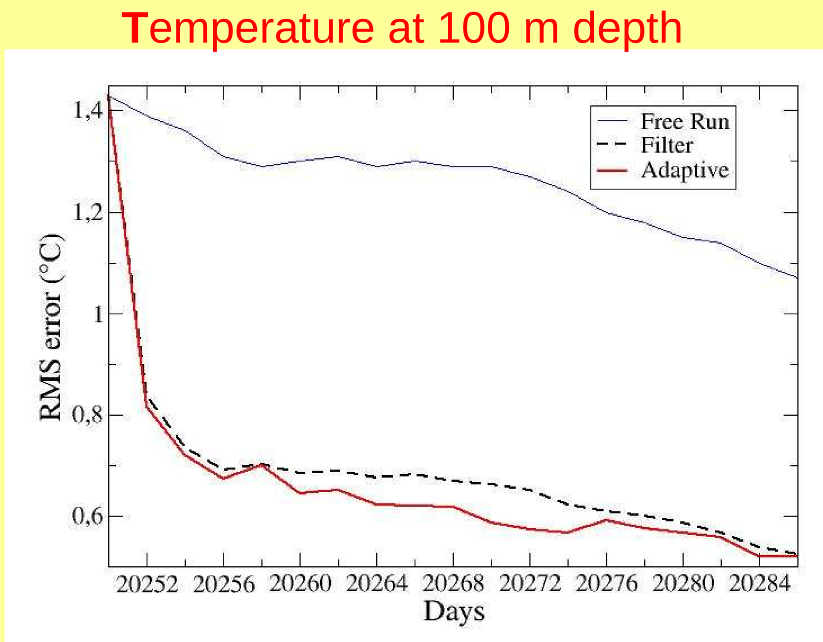
Sea **S**urface **T**emperature



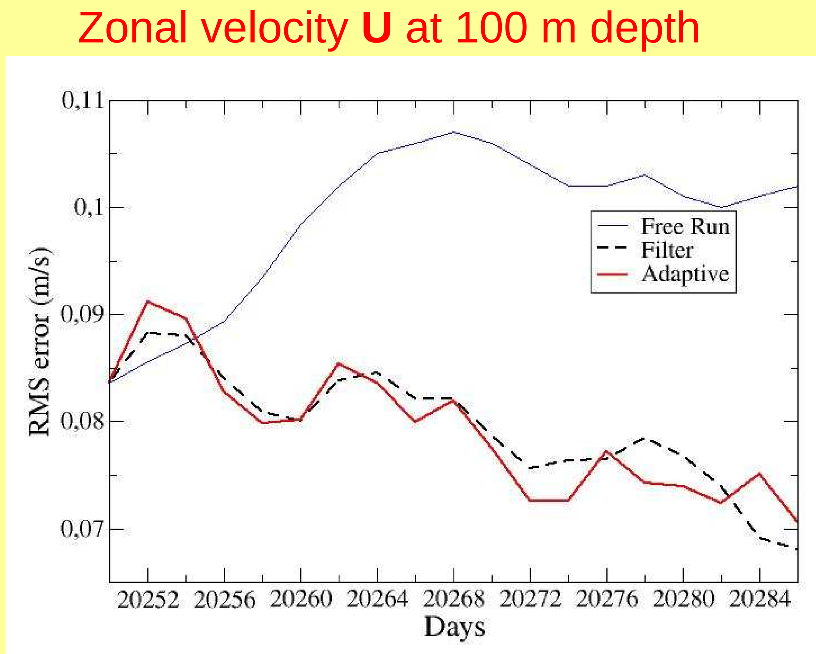
Zonal velocity **U**



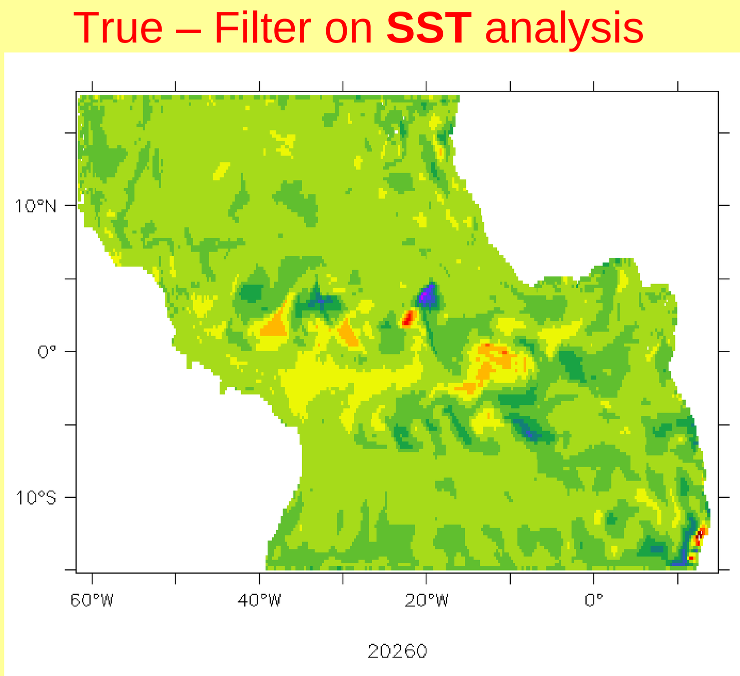
Temperature at 100 m depth



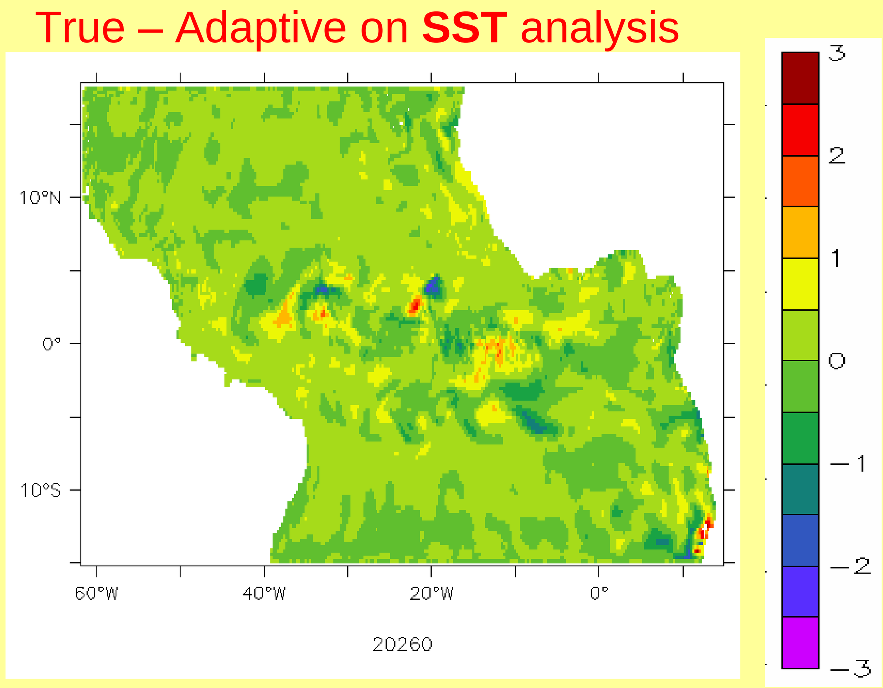
Zonal velocity **U** at 100 m depth



True – Filter on **SST** analysis



True – Adaptive on **SST** analysis



Similar results are found on **SSH, S**alinity and Meridional velocity **V**

CONCLUSIONS

- A general adaptive scheme is implemented with a square root filter
- The numerical cost is low due to the square root formulation
- Results with a realistic configuration are encouraging and valuable

PERSPECTIVES

- Compute α and β simultaneously
- Experiment more complex models $P^f(\alpha)$ and $R(\beta)$ to further improve the filter analyses.

REFERENCES

[1] Dee, Monthly Weather Rev., 123, 1128-1145 (1995)

[2] Brankart et al., Monthly Weather Rev., 138(3), 932-950 (2010)

[3] Pham et al., Journal Marine Sys., 16, 323-340 (1998)

[4] Testut et al., Journal Marine Sys., 40-41, 291-316 (2003)

[5] Hoang et al., Tellus, 57A, 153-170 (2005)