Sequential methods for ocean data assimilation

From theory to practical implementations (II)

P. BRASSEUR

CNRS/LEGI, Grenoble, France

Pierre.Brasseur@hmg.inpg.fr

J. Ballabrera, L. Berline, F. Birol, J.M. Brankart, G. Broquet, V. Carmillet, F. Castruccio, F. Debost, D. Rozier, Y. Ourmières, T.Penduff, J. Verron

E. Blayo, S. Carme, Pham D.T., C. Robert

C.E. Testut, B. Tranchant, N. Ferry, E. Remy, M. Benkiran

F. Durand, L. Gourdeau, Ch. Maes, P. De Mey

A. Barth, C. Raick, M. Grégoire

L. Parent

2. Kalman Filter fundamentals
Assimilation cycle

$$\begin{cases} \boldsymbol{x}_i^a \\ \mathbf{P}_i^a \end{cases}$$

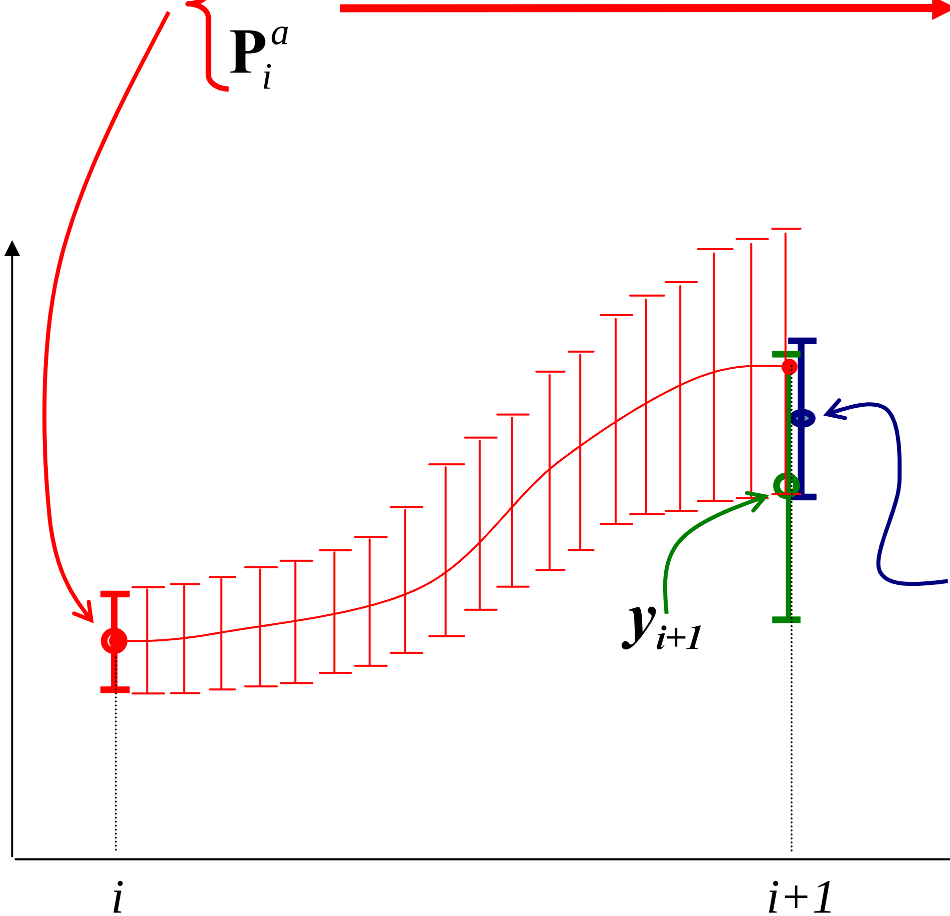
Forecast →

$$\begin{cases} \boldsymbol{x}_{i+1}^f = M\left(\boldsymbol{x}_i^a\right) \\ \mathbf{P}_{i+1}^f = \mathbf{M}\mathbf{P}_i^a\mathbf{M}^T + \mathbf{Q} \end{cases}$$

Analysis ↓

$$\mathbf{K}_{i+1} = \mathbf{P}_{i+1}^f\mathbf{H}^T(\mathbf{H}\mathbf{P}_{i+1}^f\mathbf{H}^T + \mathbf{R})^{-1}$$

$$\begin{cases} \boldsymbol{x}_{i+1}^a = \boldsymbol{x}_{i+1}^f + \mathbf{K}_{i+1}\left(\boldsymbol{y}_{i+1} - H\left(\boldsymbol{x}_{i+1}^f\right)\right) \\ \mathbf{P}_{i+1}^a = \left(\mathbf{I} - \mathbf{K}_{i+1}\mathbf{H}\right)\mathbf{P}_{i+1}^f \end{cases}$$

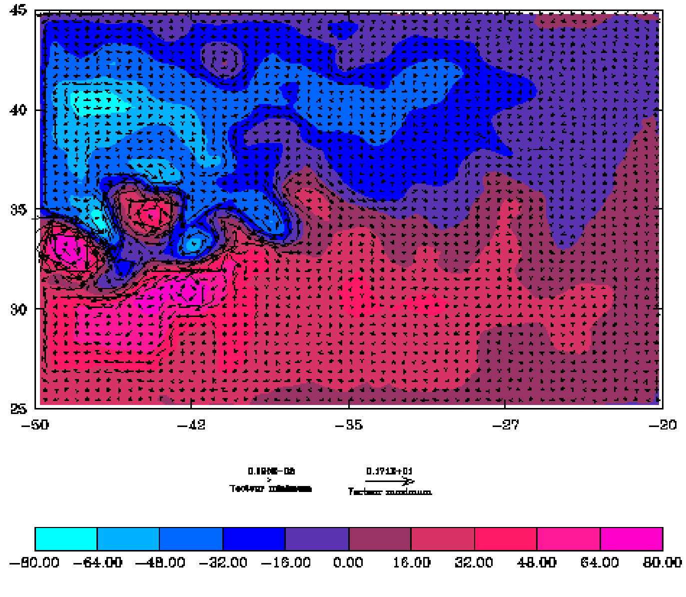


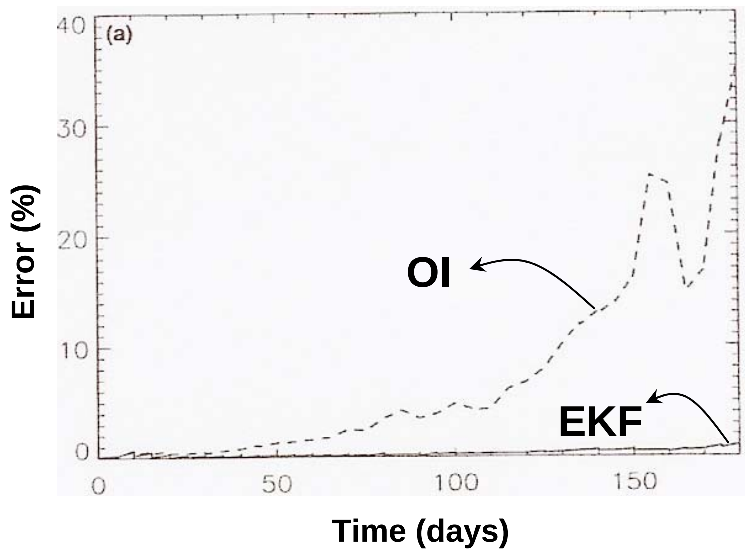
OI vs. EKF

- A full Kalman filter cannot be implemented into realistic ocean models (error forecast and analysis equations too expensive in CPU and memory requirements)

- « Optimal Interpolation » over-simplifies the propagation of errors by neglecting dynamical principles and statistical information

Idealized double-gyre model (Ballabrera *et al.*, 2001)





OUTLINE

- State-of-the-art
  1. Introduction
  2. Kalman filter: fundamentals
  3. *Applied* ocean data assimilation: specific issues
  4. Simplifications of the KF – Optimal Interpolation

- **Advanced issues**
  5. **Space reduction: state and error sub-spaces**
  6. **Low rank filters: SEEK and EnKF**
  7. **Consistency validation and adaptivity**
  8. **Improved temporal strategies : FGAT and IAU**

5. Space reduction
How ?

- **The concept of space reduction is introduced, with the objectives to :**
  - ✓ Substantially reduce the computational burden of a full Kalman filter, but
  - ✓ Preserve the essential properties of statistical estimation.
- **The reduction can be formulated in terms of <u>state space</u> of <u>error space</u>.**

- **State space reduction by :**
  - ✓ Selection of model state variables (T,S,psi)
  - ✓ Selection of grid points (coarse grid) or large-scale modes
  - ✓ KF for surface (observed) variables only + vertical extrapolation
  - ✓ …
- **Formally:** $w = \mathbf{T}\,x \quad \text{with} \quad dim\,\mathbf{T} = r \times n$

! $\mathbf{T}^{-I}$ *needed to project the KF estimate back to full space* !

5. Space reduction
Error covariance matrix decomposition

- **Properties:** covariance matrices are symetric , positive definite

$$\Rightarrow \mathbf{P} = \mathbf{L}\,\Lambda\,\mathbf{L}^T \quad \text{with} \quad \mathbf{L}: \text{ eigenvectors}$$

$$\Lambda = diag\{\lambda_i\}: \text{ eigenvalues}$$

- **Error sub-space** $\mathbf{S}$: defined as an approximation of $\mathbf{L}\sqrt{\Lambda}$ (limited to the dominant eigenmodes/eigenvalues which best represent the covariance $\mathbf{P}$)

<u>Low rank approximation:</u> $\mathbf{P}_O$ **specified as a low rank matrix**

$\mathbf{P}_O = \mathbf{S}_O \mathbf{S}_O^T$, with $\mathbf{S}_O$ of dim $n \times r$, $r \ll n = \dim(\boldsymbol{x})$

- Accurate specification of a full-rank $\mathbf{P}_O$ is impossible !
- Approximation done at initial assimilation time only
- Drastic simplification of analysis and forecast steps : $r \sim 10\text{-}100$

5. Space reduction
Error sub-space variants

$$\mathbf{P}_0 = \mathbf{S}_0 \, \mathbf{S}_0^T$$

r ~10-100

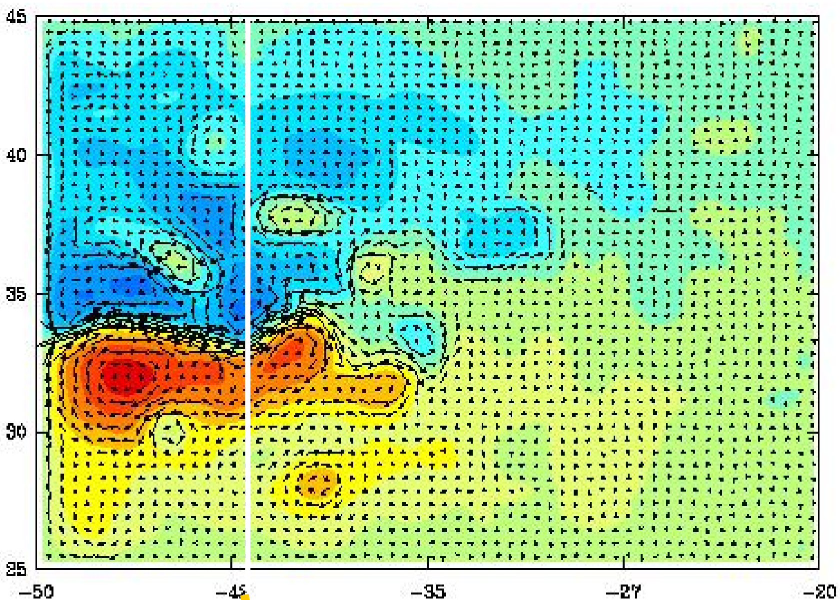
n ~10^6-10^8

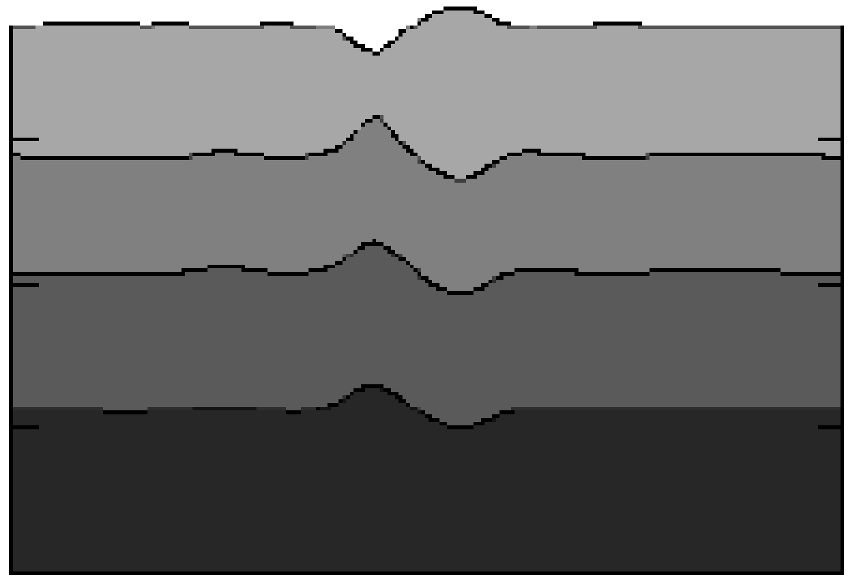
All covariance informations are contained in $\mathbf{S}_0$ *!*

Variants to build the sub-space :

- **EOFs from model variability**: global modes *well suited to basin-scale problems (e.g. Tropical Pacific) ; local modes well suited to eddy-resolving ocean DA problems (Atlantic basin HR models)*
- **Singular or breeding modes** : *for non-linear models*
- **Ensemble of Monte-Carlo perturbations of model states**
- **…**

- **Idealized double-gyre model (MICOM, 4 layers)**





Vertical section through dominant EOF

« 3D Cooper-Haines » mode

EOFs provide a robust description of the covariances between SLA and vertical displacements of isopycnals

- **Practical recipe :** to compute 3D, multivariate EOFs from a model run

- ✓ Sampling of historical sequence:

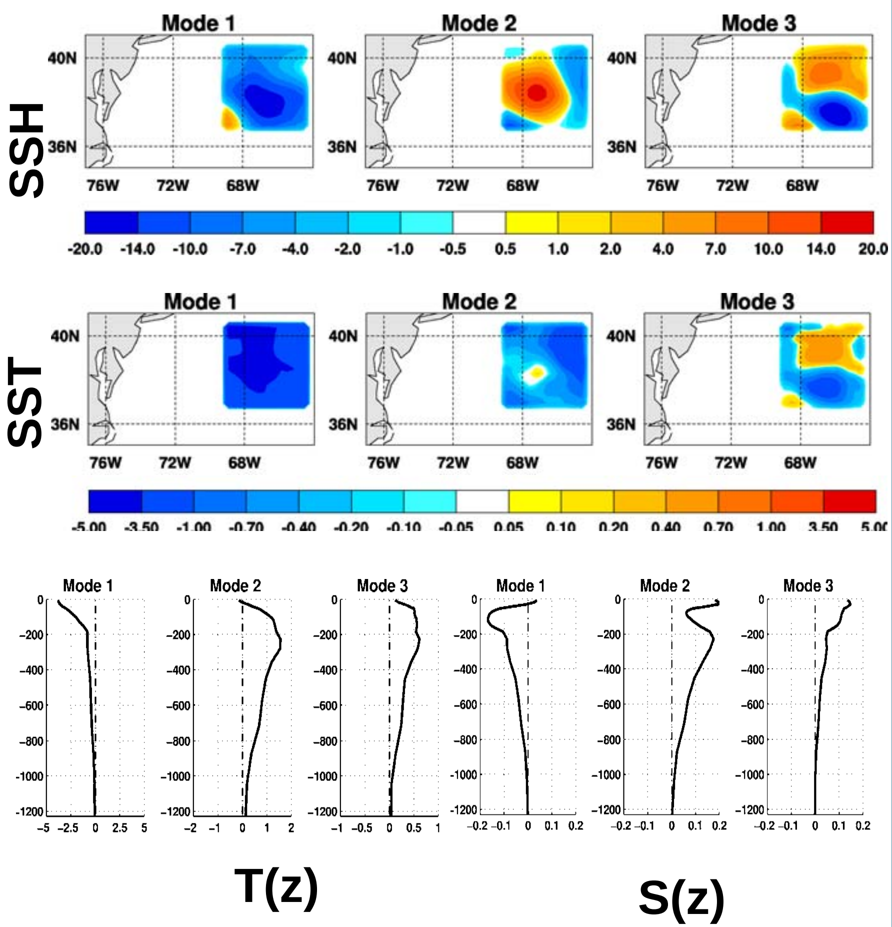
$$x^m(t_{i+1}) = M(t_i, t_{i+1})x^m(t_i) \quad , \quad i = 0,\dots,s-1$$
$$\Rightarrow \mathbf{X} = \left\{ x^m(t_i) - \overline{x^m(t_i)} \right\} \quad \dim n \times s$$

- ✓ Eigenmodes of « sample » matrix $\mathbf{XX}^T$ $(\dim n \times n)$ can be easily computed from the eigenmodes of $\mathbf{X}^T\mathbf{X}$ $(\dim s \times s)$ because

$$\mathbf{XX}^T\mathbf{L} = \mathbf{L}\,\Lambda \;\Leftrightarrow\; \mathbf{X}^T\mathbf{XX}^T\mathbf{L} = \mathbf{X}^T\mathbf{L}\,\Lambda = \mathbf{V}\,\Lambda$$
$$\Rightarrow \mathbf{L} = \mathbf{XV}\,\Lambda^{-1}$$

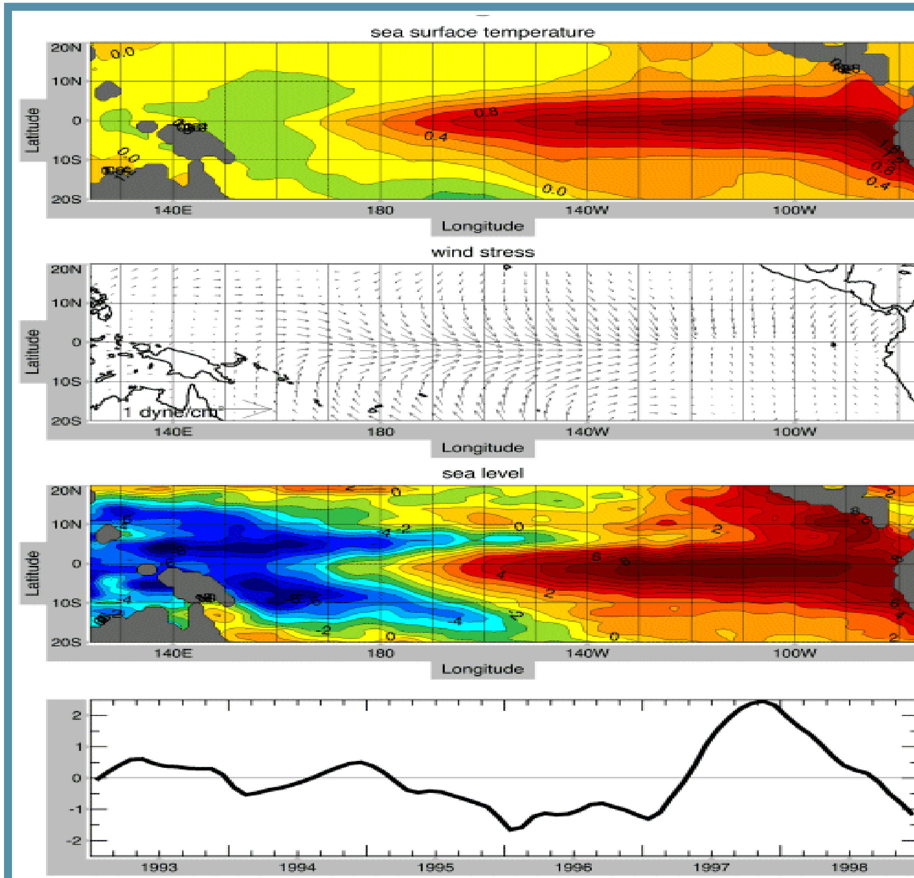
- ✓ Truncation to *r* dominant modes $\Rightarrow \mathbf{S}_0 \quad \text{and} \quad \mathbf{P}_0 \approx \mathbf{S}_0\mathbf{S}_0^T$

5. Space reduction
Error sub-spaces : examples



(from Testut et al., 2001)

Local EOFs



(from Picaut et al., 2001)

Global EOFs

6. Low rank Kalman filters
SEEK filter – forecast equation

Concept:

Use error space reduction $\mathbf{P}_i^a = \mathbf{S}_i^a \, \mathbf{S}_i^{a\,T}$ **to compute** $\mathbf{P}_{i+1}^f = \mathbf{M} \, \mathbf{P}_i^a \mathbf{M}^T + \mathbf{Q}$

$$\Rightarrow \quad \mathbf{P}_{i+1}^f = \left(\mathbf{M}\mathbf{S}_i^a\right)\left(\mathbf{M}\mathbf{S}_i^a\right)^T + \mathbf{Q}$$

- Time-evolving sub-space at moderate cost (max *r* model integrations)
- Model error parameterized in the evolving sub-space $\mathbf{Q} \div \mathbf{M} \, \mathbf{P}_i^a \mathbf{M}^T$ to preserve low rank

Variants to evolve the sub-space $\mathbf{S}_i^a \rightarrow \mathbf{S}_{i+1}^f$

- **« Extended » evolutive :** $\mathbf{S}_{i+1}^f = \mathbf{M}\mathbf{S}_i^a$ (*use tangent linear model : Pham et al., 1998)*
- **« Interpolated » evolutive :** $\mathbf{S}_{i+1}^f \div M\left(\boldsymbol{x}_i^a + \alpha \, \mathbf{S}_i^a\right) - M\left(\boldsymbol{x}_i^a\right)$ (*use non- linear model to update the error modes dynamically : Brasseur et al., 1999; Ballabrera et al., 2001)*
- **« Fixed basis » :** $\mathbf{S}_{i+1}^f = \mathbf{I} \, \mathbf{S}_i^a$ (*assume persistence or dominant model error to update the sub-space: Verron et al., 1999)*

Concept: **Use error space reduction** $\mathbf{P}_i^a = \mathbf{S}_i^a \mathbf{S}_i^{a^T}$ **to compute** $\mathbf{K}$

$$\mathbf{K}_{i+1} = \mathbf{P}_{i+1}^f \mathbf{H}^T (\mathbf{H}\mathbf{P}_{i+1}^f \mathbf{H}^T + \mathbf{R})^{-1}$$

$$\longrightarrow \quad = \ldots = \mathbf{S}_{i+1}^f \left[\mathbf{I} + \left(\mathbf{H}\mathbf{S}_{i+1}^f\right)^T \mathbf{R}^{-1} \left(\mathbf{H}\mathbf{S}_{i+1}^f\right)\right]^{-1} \left(\mathbf{H}\mathbf{S}_{i+1}^f\right)^T \mathbf{R}^{-1}$$

- More effective inversion: in **reduced space** instead of **observation space**, with r often much smaller than dim ($\boldsymbol{y}$)
- Increments are combinations of modes: $\boldsymbol{x}_{i+1}^a - \boldsymbol{x}_{i+1}^f = \mathbf{K}_i \left(\boldsymbol{y}_i - \mathbf{H}\,\boldsymbol{x}_{i+1}^f\right) = \boldsymbol{S}_{i+1}^f \boldsymbol{c}$

Variants to compute updates

- **« Global » analysis : the standard formulation ,** *requires regular data distributions in space to avoid spurious corrections at large distances*
- **« Local » analysis : define $\mathbf{H}$ as a « local » operator** to compensate for truncation errors and eliminate remote influence of data *(Brankart et al., 2003)*

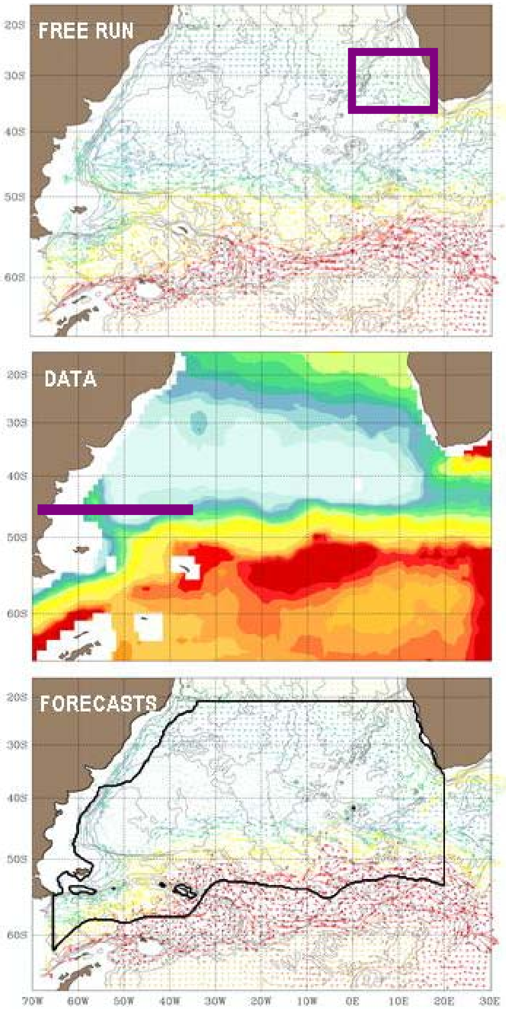
LEGI
LABORATOIRE DES ECOULEMENTS GEOPHYSIQUES ET INDUSTRIELS

6. Low rank Kalman filters
Local EOFs for mesoscale data assimilation

Free run

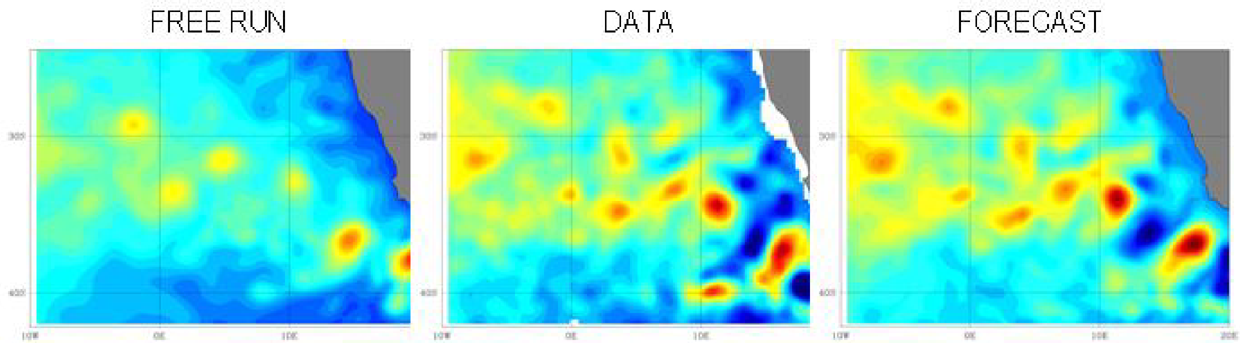
Sclim

Forecast

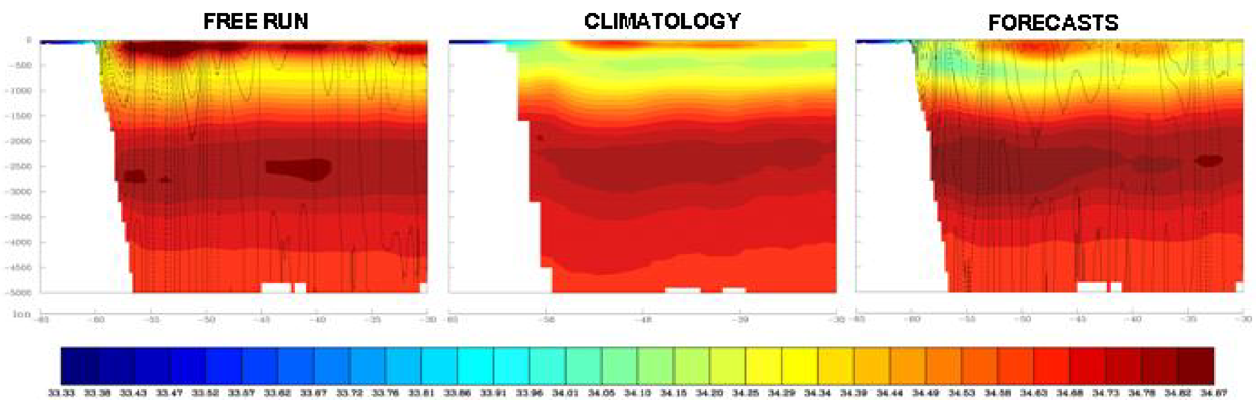


Circulation at 872 m

Agulhas Rings – June 25, 1995



Confluence – mean salinity at 45S



<u>South Atlantic hindcast</u> (Penduff *et al.*, 2003)

- ❑ **OPA model 1/3°, 1993-1996, 6h ECMWF**
- ❑ **Assimilation of SLA (T/P, ERS), SST (AVHRR)**

Concept:

Use an ensemble of r model states $x_i^{a,j}$ to specify the spread of possible initial conditions around the mean $\overline{x_i^{a,j}}$ and propagate each member individually (Evensen 1994).

Forecast equation:

$$x_{i+1}^{f,j} = M\left(x_i^{a,j}\right) + \eta^j \quad \text{with} \quad \overline{\eta^j \eta^{jT}} = \mathbf{Q} \quad , \quad j = 1,\ldots,r$$

This provides automatically: $\mathbf{P}_{i+1}^f = \dfrac{1}{r-1}\left(x_{i+1}^{f,j} - \overline{x_{i+1}^{f,j}}\right)\left(x_{i+1}^{f,j} - \overline{x_{i+1}^{f,j}}\right)^T$

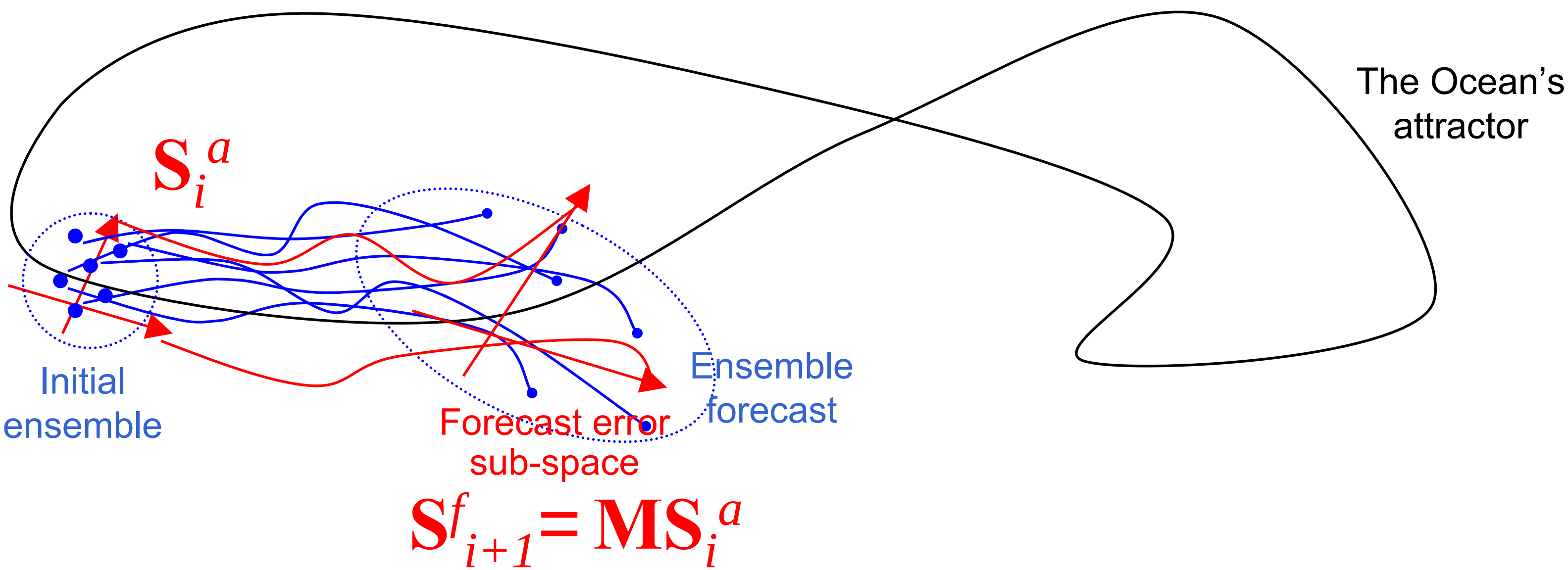
Analysis equation:

$$x_{i+1}^{a,j} = x_{i+1}^{f,j} + \mathbf{K}_{i+1}\left(\tilde{y}_{i+1} - H\left(x_{i+1}^{f,j}\right)\right) \quad , \quad j = 1,\ldots,r$$

This provides automatically: $\mathbf{P}_{i+1}^a = \dfrac{1}{r-1}\left(x_{i+1}^{a,j} - \overline{x_{i+1}^{a,j}}\right)\left(x_{i+1}^{a,j} - \overline{x_{i+1}^{a,j}}\right)^T$

<u>Same philosophy</u> :

Sequential corrections along privileged directions of error growth



Differences between SEEK, EnKF, EnKS : Brusdal et al., JMS, 2003
SEEK, SEIK and EnKF intercomparison : Nerger et al., 2004

❑ **A major difficulty with DA schemes is the specification of background and observation error statistics, which are critical to the analysis step.**

❑ **Filter error statistics :** $\overline{\boldsymbol{\varepsilon}^o} = 0 \quad \overline{\boldsymbol{\varepsilon}^f} = 0 \qquad \mathbf{R} = \overline{\boldsymbol{\varepsilon}^o \boldsymbol{\varepsilon}^{oT}} \qquad \mathbf{P}^f = \overline{\boldsymbol{\varepsilon}^f \boldsymbol{\varepsilon}^{fT}}$

❑ **Innovation « seen » by the filter:**

$$\mathbf{d}_i = \mathbf{y}_i - \mathbf{H}\mathbf{x}_i^f = (\mathbf{H}\mathbf{x}_i^t + \boldsymbol{\varepsilon}_i^o) - \mathbf{H}\mathbf{x}_i^f = \boldsymbol{\varepsilon}_i^o - \mathbf{H}\boldsymbol{\varepsilon}_i^f$$

❑ **Residuals « seen » by the filter:**

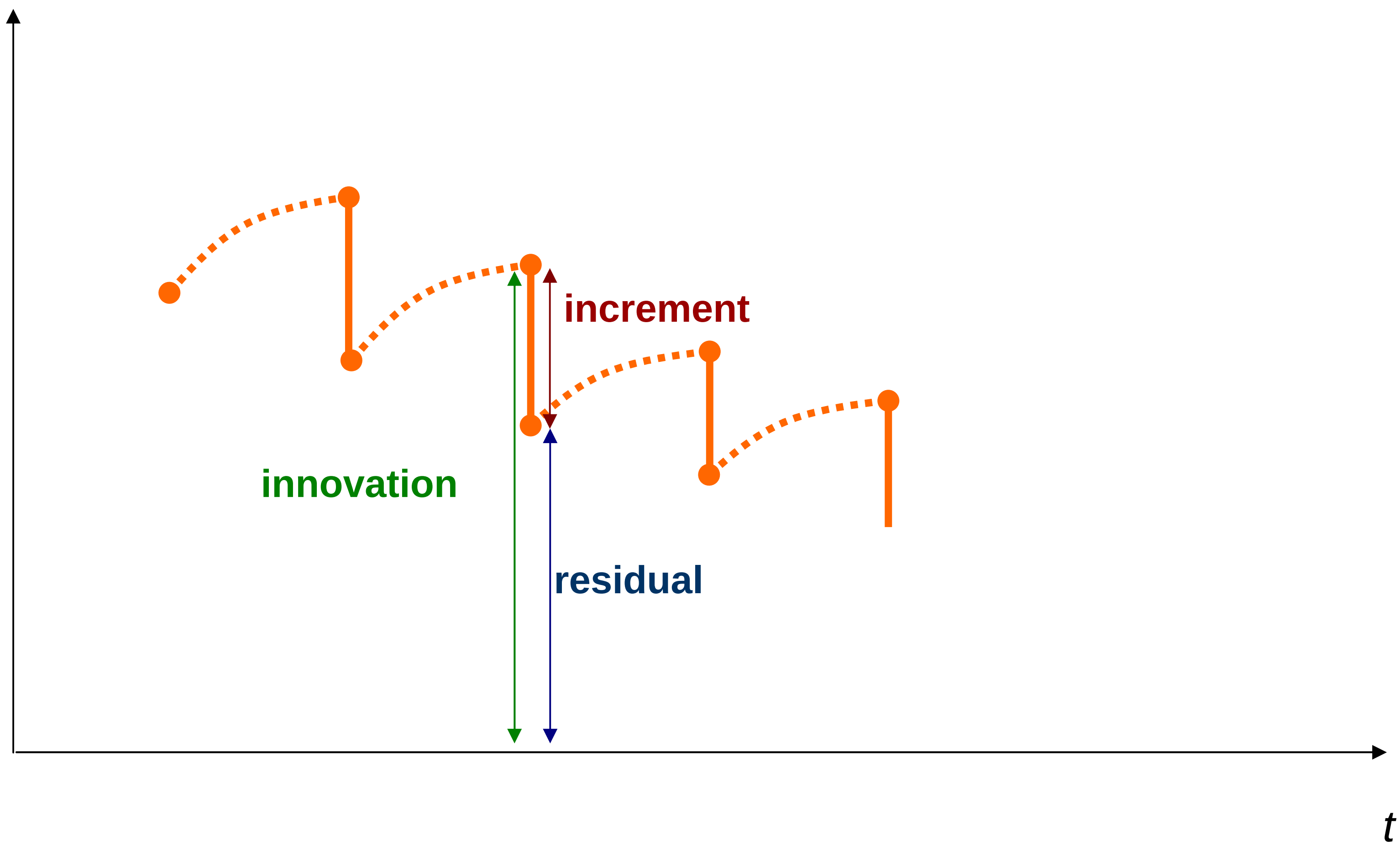
$$\mathbf{r}_i = \mathbf{y}_i - \mathbf{H}\mathbf{x}_i^a = (\mathbf{H}\mathbf{x}_i^t + \boldsymbol{\varepsilon}_i^o) - \mathbf{H}\mathbf{x}_i^a = \boldsymbol{\varepsilon}_i^o - \mathbf{H}\boldsymbol{\varepsilon}_i^a$$

❑ **Increments computed by the filter:**

$$\boldsymbol{x}_i^a - \boldsymbol{x}_i^f = \mathbf{K}_i \left(\boldsymbol{y}_i - \mathbf{H}\, \boldsymbol{x}_i^f \right)$$

7. Validation of DA systems - Adaptivity
Innovation & residual statistics

observation misfit



- **During the assimilation process, « anomalies » can be detected between the innovation sequence and the prior statistical assumptions (in a KF context).**

❑ **Unbiased innovation sequence :** $\overline{\mathbf{d}_i} = 0$

❑ **Unbiased residuals (** $\overline{\mathbf{r}_i} = 0$ **) or increments** $\mathbf{H}\overline{(x_i^a - x_i^f)} = 0$

❑ **Consistent error covariances :**

$$\overline{\mathbf{d}_i \mathbf{d}_i^T} = \mathbf{R} + \mathbf{H}\mathbf{P}_i^f \mathbf{H}^T$$

$$\overline{\mathbf{r}_i \ \mathbf{r}_i^T} = \mathbf{R} - \mathbf{H}\mathbf{P}_i^a \mathbf{H}^T$$

❑ χ_p^2 **distribution of** (Bennett, 1992) **:**

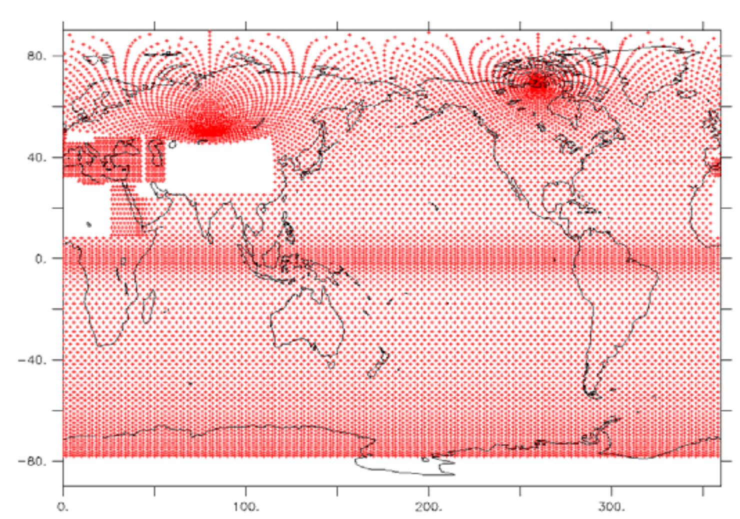
$$J_i = \mathbf{d}_i^T (\mathbf{H}\mathbf{P}_i^f \mathbf{H}^T + \mathbf{R})^{-1} \mathbf{d}_i$$

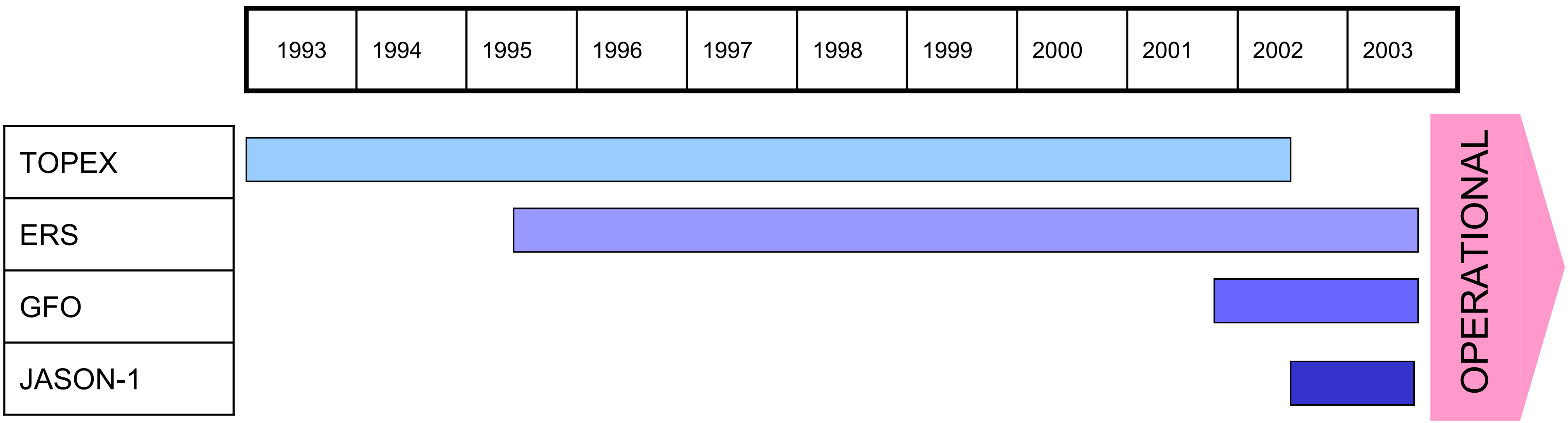
or, for low-rank $J_i = \mathbf{d}_i^T \left\{ (\mathbf{H}\mathbf{S}_i^f)(\mathbf{H}\mathbf{S}_i^f)^T + \mathbf{R} \right\}^{-1} \mathbf{d}_i$

7. Validation of DA systems - Adaptivity

MERCATOR Global : 1993-2002 analysis

- 1 year spin-up (1992)
- 11 years of **weekly** assimilation of SLA





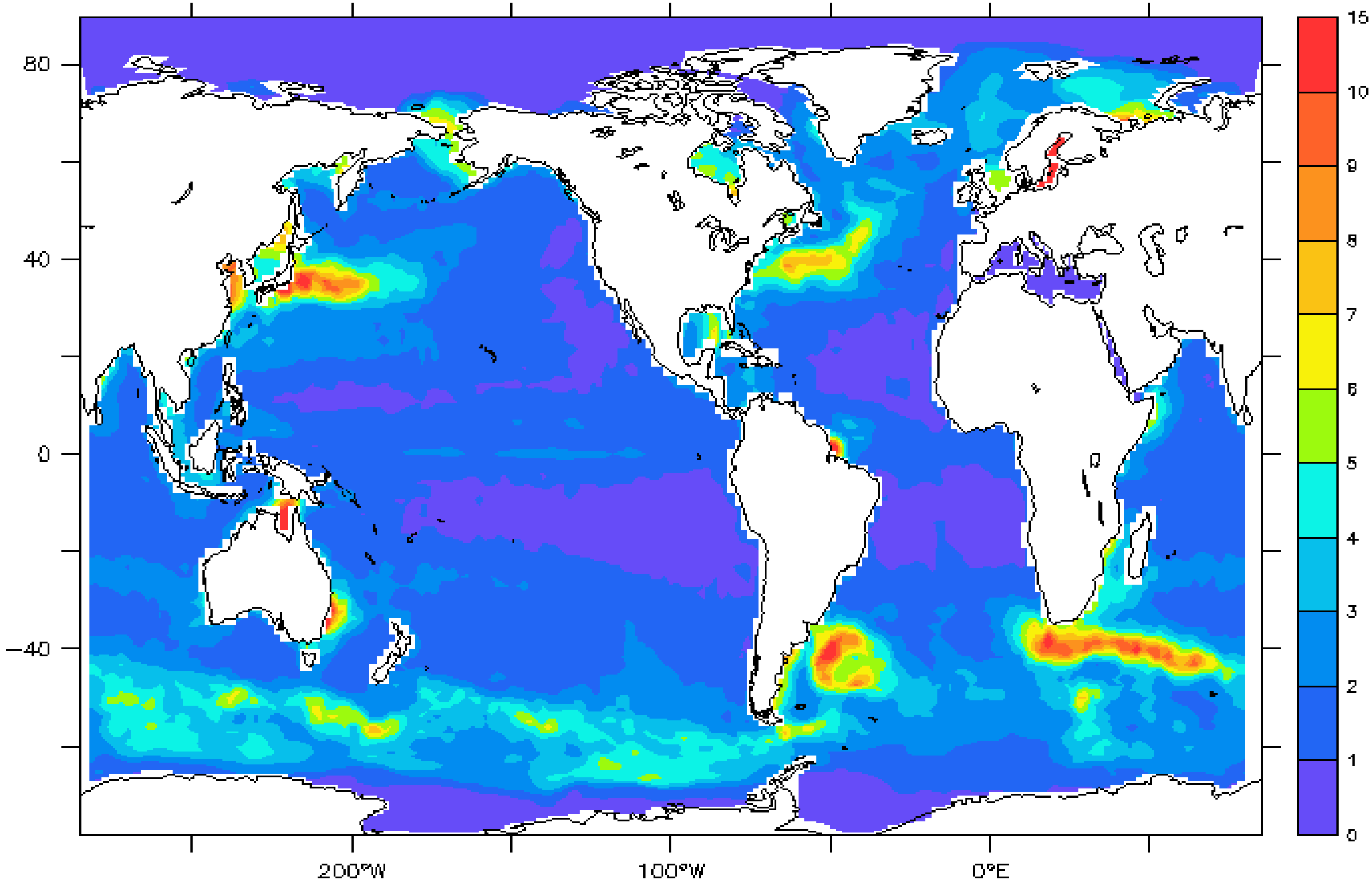
3D state estimation : SAM-1 and SAM-2

- **Optimal interpolation : SAM-1**
  1. SOFA + Cooper/Haines mode (open ocean attractor):
     2D statistical estimation + vertical adjustement : **SAM-1v1**
  2. SOFA + multivariate 1D vertical EOFs (from model or data variability):
     2D +1D statistical estimation : **SAM-1v2**

- **Reduced-order Kalman filter : SAM-2**
  1. SOFA + EOFs 3D (multivariate model variability): inversion in observation space : **SAM-2v0**
  2. SEEK + EOFs 3D (multivariate model variability): inversion in error sub-space: **SAM-2v1**

7. Validation of DA systems - Adaptivity
MERCATOR Global : 1993-2002 analysis

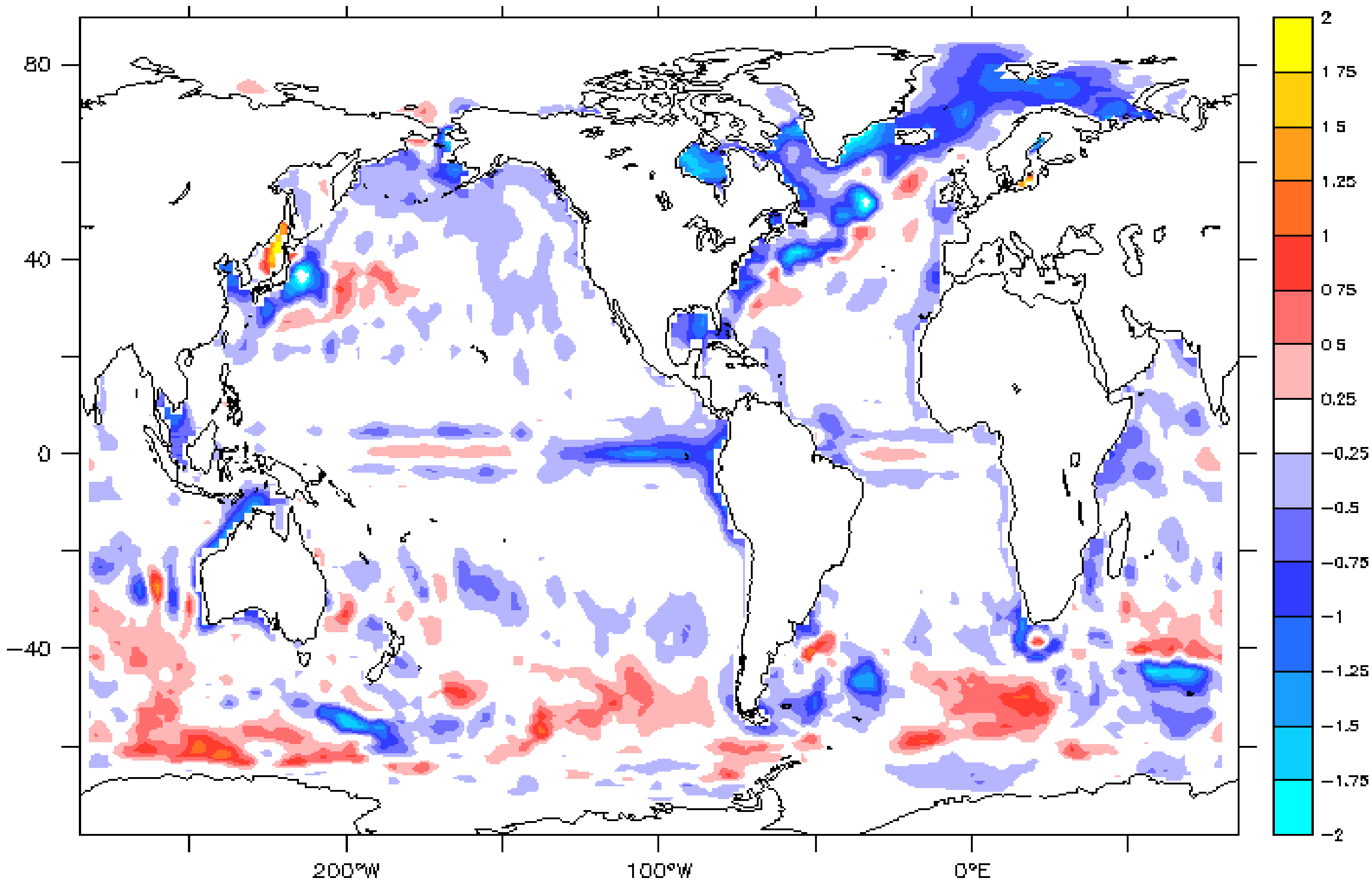
Standard deviation of SLA increment for 1993-2002, (cm)



Mercator global ocean prototype (Ferry *et al.*, 2004)

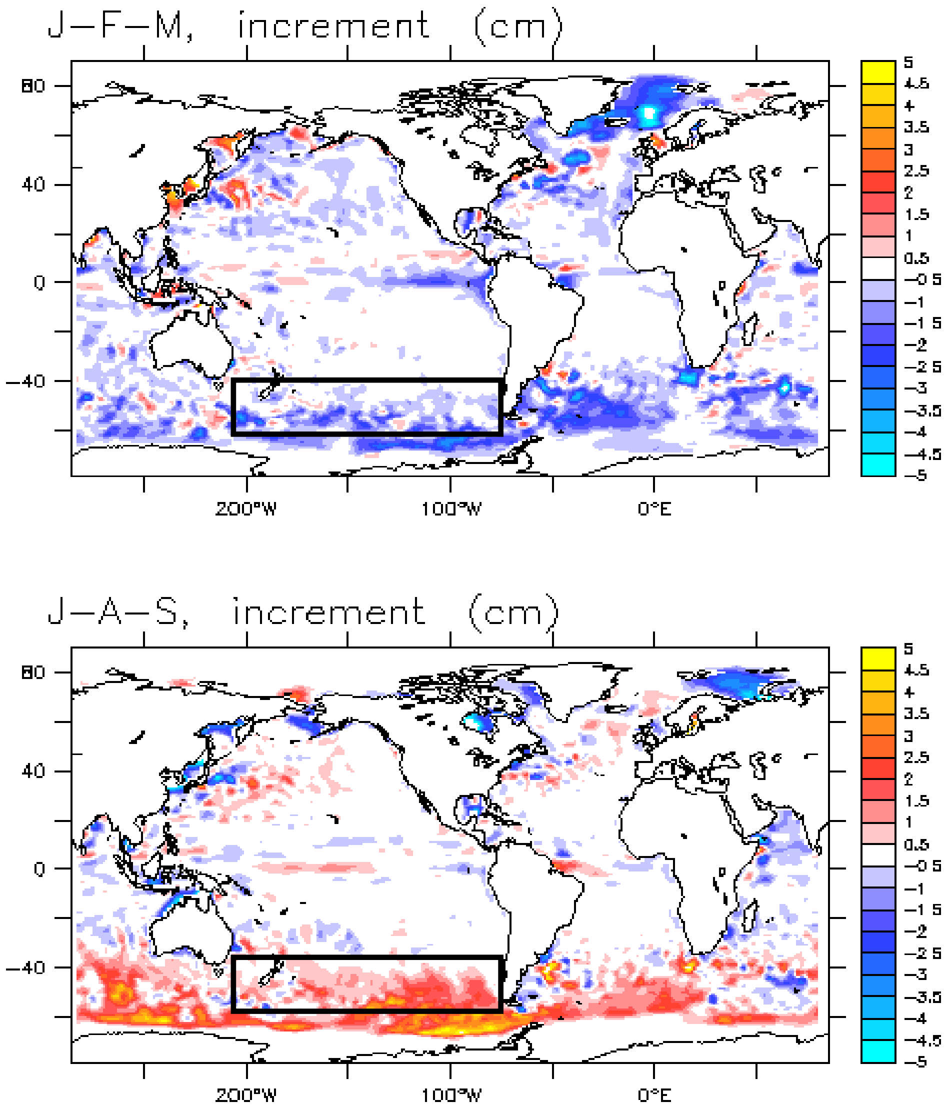
7. Validation of DA systems - Adaptivity
Detection of system biases

Mean SLA increment, 1993-2002, (cm)

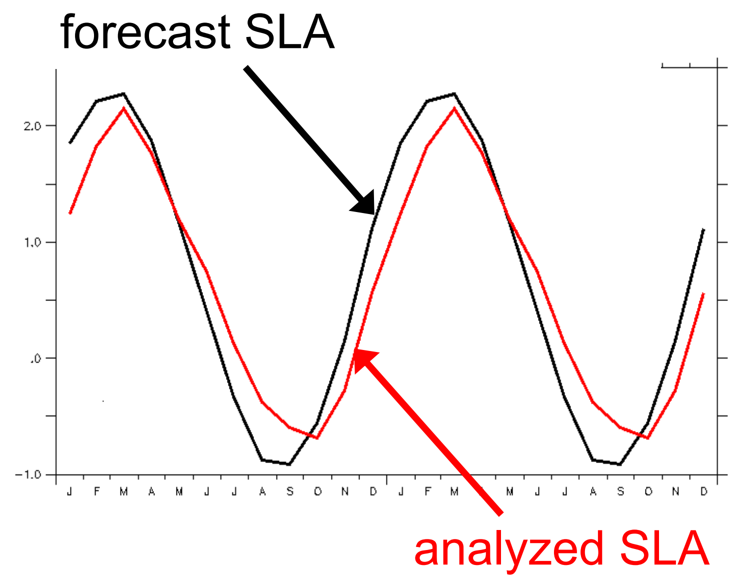


Mercator global ocean prototype (Ferry *et al*., 2004)

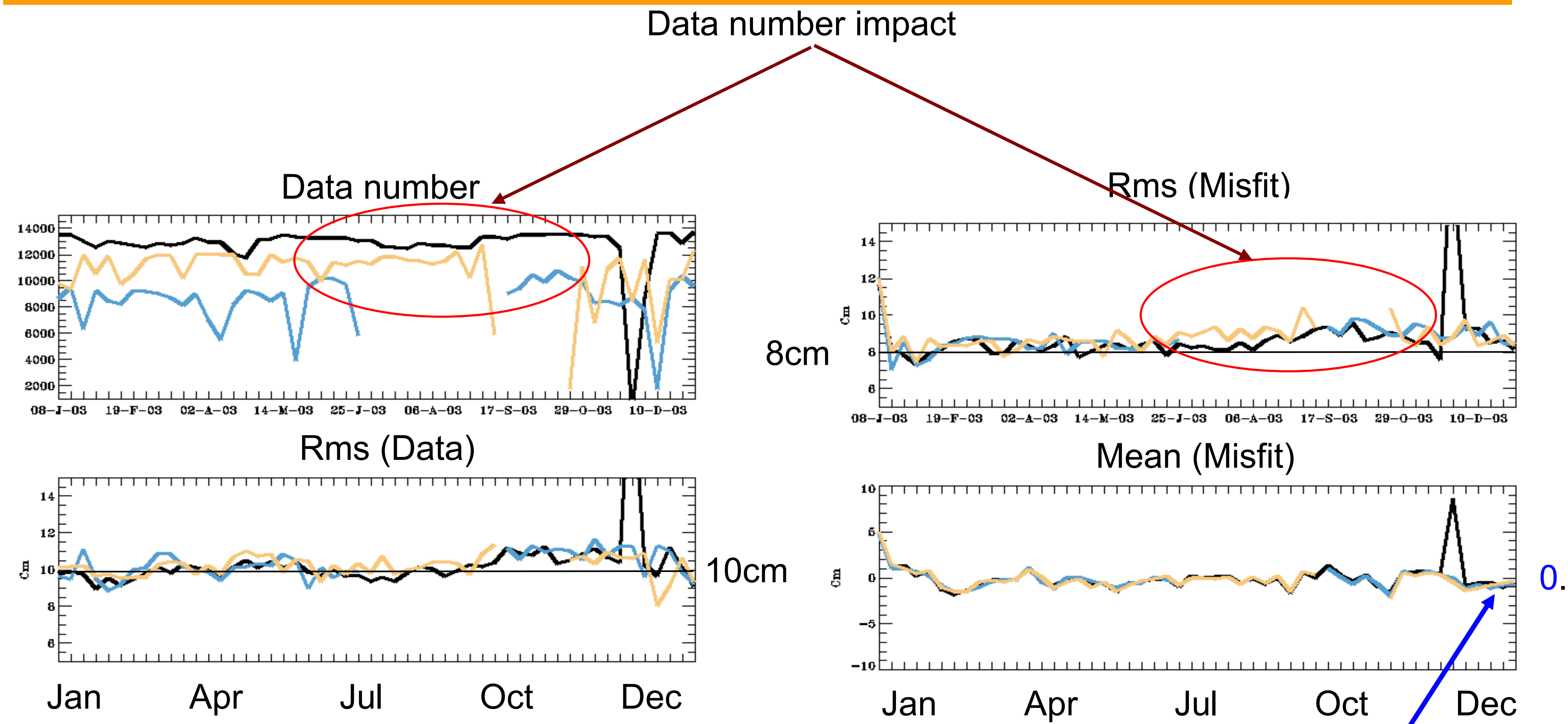
7. Validation of DA systems - Adaptivity
Analysis of system biases



❑ Mean seasonal SLA increment



7. Validation of DA systems - Adaptivity
Assimilation diagnostics



Black : Jason1
Blue : Ers2 (Envisat)
Orange : Gfo

No bias

7. Validation of DA systems - Adaptivity *Variants*

Concept: **« on line » modification of prior statistics ($\mathbf{R}$, $\mathbf{Q}$, …) in order to better match the statistics of the innovation sequence**

- Simple adaptive schemes can be implemented *easily*, and *at low cost*, into operational systems

Adaptivity variants

- **« Adaptive sub-space » :** use residual innovation to generate new error modes and refresh the sub-space intermittently (Brasseur et al., 1999; Durand et al., 2003)
- **« Adaptive variance » :** tune model error parameterization to improve the fit between innovation and filter statistics *(Brankart et al., 2003)*

Parameterization example : $\mathbf{P}_{i+1}^{f} \approx \alpha \, \mathbf{P}_{i}^{a}$

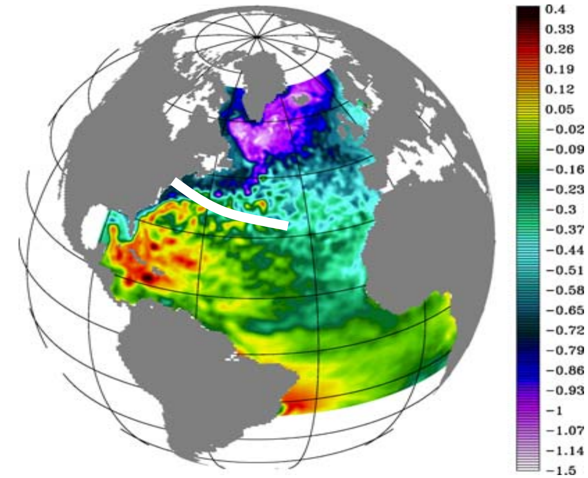
α → determined using innovation statistics history

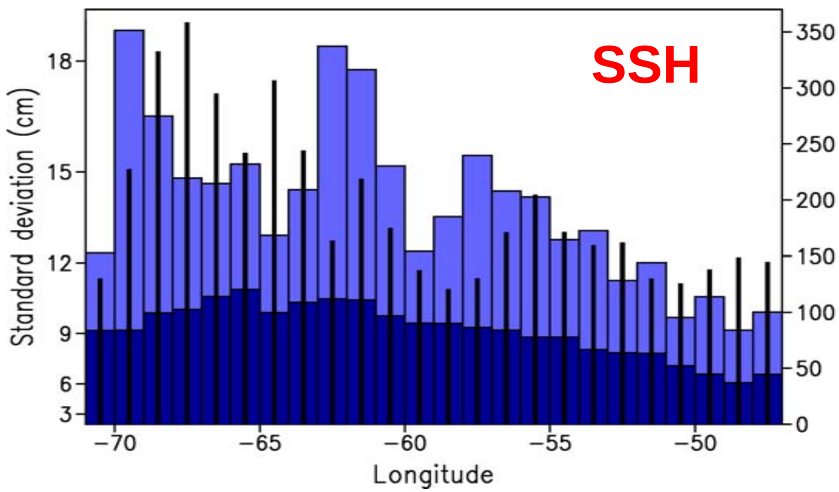
7. Validation of DA systems - Adaptivity
Example (I)

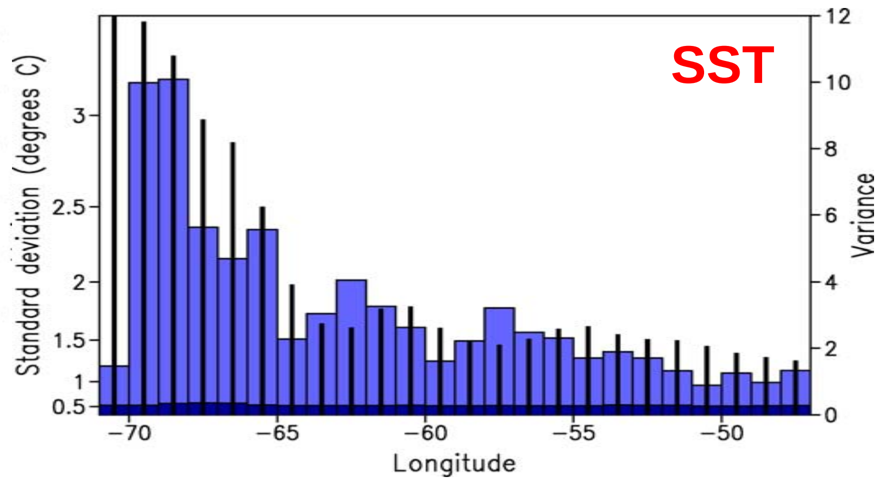
Comparison between 2 estimations of the forecast error variance in a zonal section crossing the Gulf Stream (HYCOM model, Brankart et al., 2003):

(i) from the filter (blue histograms) and
(ii) from innovation sequence (black bars).

$$tr(\overline{\mathbf{d}_i^f \mathbf{d}_i^{f^T}}) \approx tr(\mathbf{HP}^f\mathbf{H}^T) + tr(\mathbf{R})$$

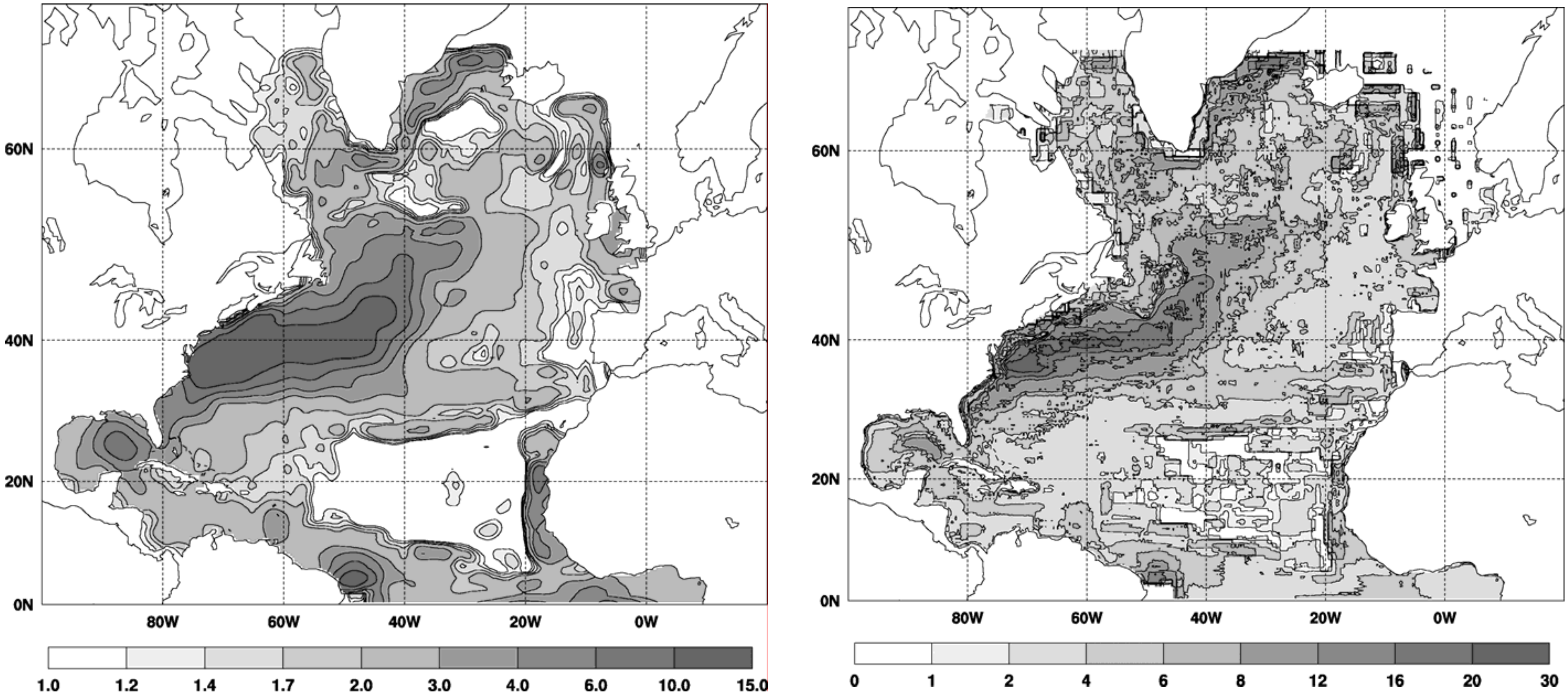






$diag\ (\overline{\mathbf{d}_i^f \mathbf{d}_i^{f^T}})$ $\quad$ diag $\mathbf{R}$ $\quad$ diag $\mathbf{P}^f$

7. Validation of DA systems - Adaptivity
Example (ii)



Model error amplification α

Forecast error estimates

8. Improved temporal strategies
Distributed observations

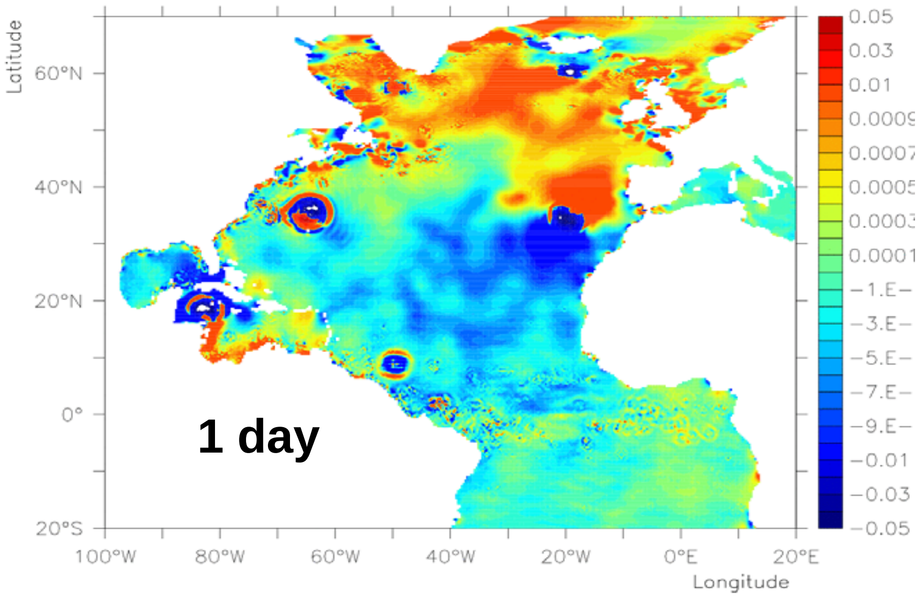
- Ocean observations are continuously distributed in time during the assimilation period; however, it is impossible to rigorously incorporate the data at their exact acquisition time. Therefore, **intermittent data assimilation schemes are *approximate*.**

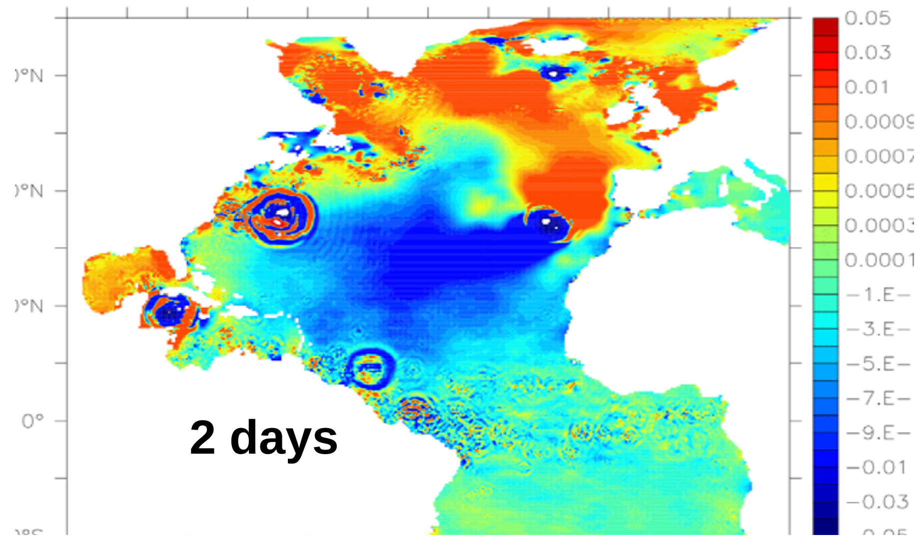
- Typical length of assimilation periods:
  - 3-7 days for mesoscale ocean current predictions;
  - 30 days for initialisation of seasonal climate predictions.

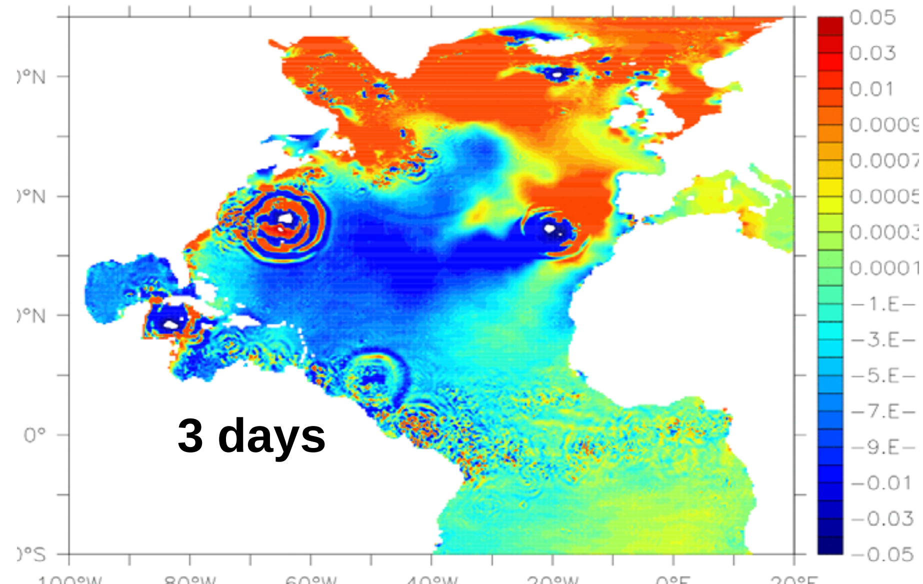
- Two related problems arise with intermittent corrections: shocks to the model, and data rejection.

Strategies to alleviate these problems ?

8. Improved temporal strategies
« Shocks » to model forecasts

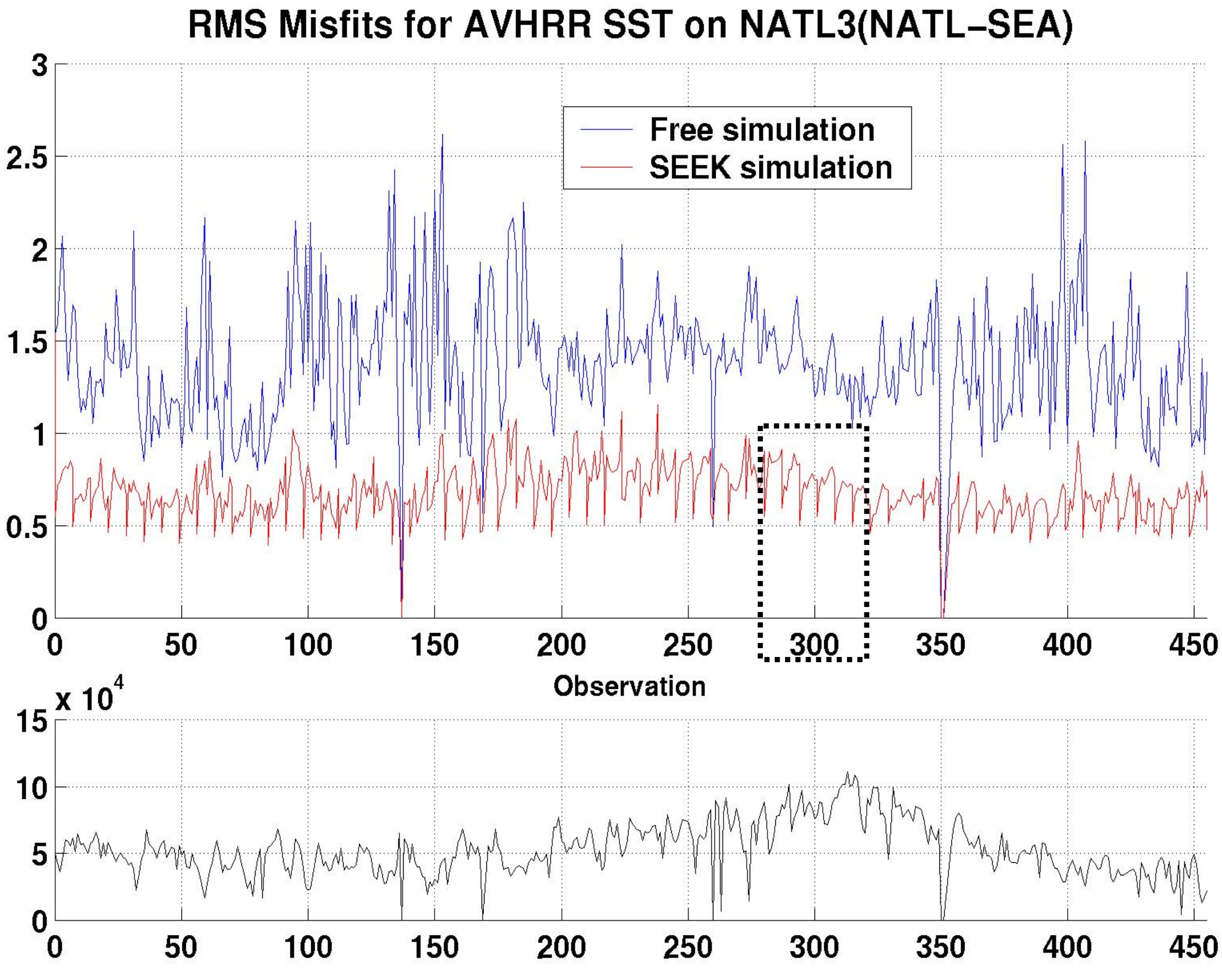


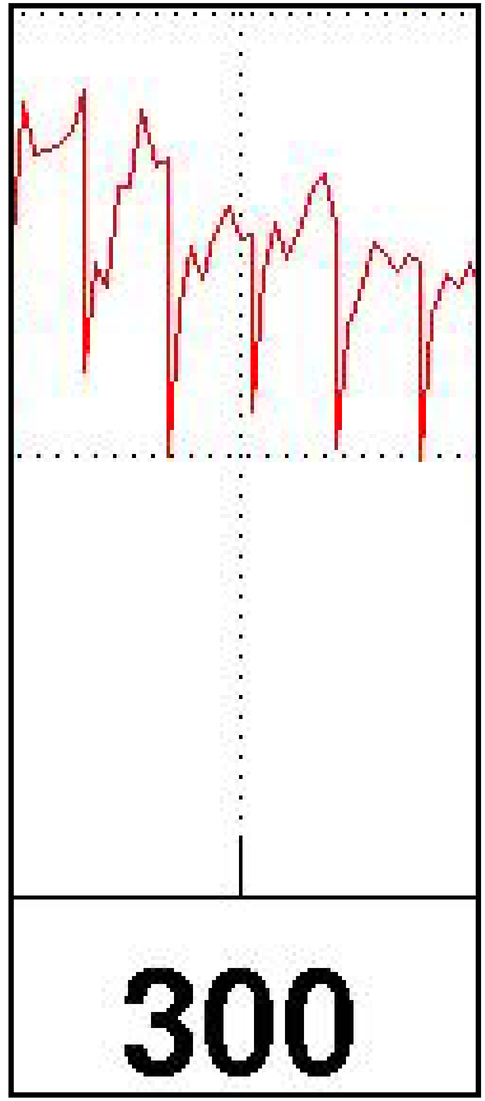




Assimilation of isolated T/S profiles: SSH increment after 1, 2, 3 days of model forecast

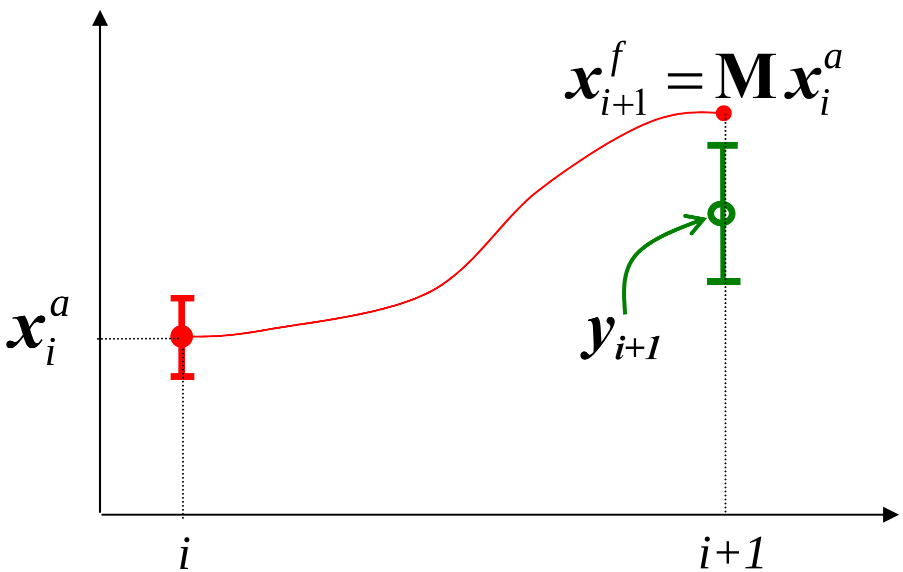
8. Improved temporal strategies
« Rejection » of SST data assimilation



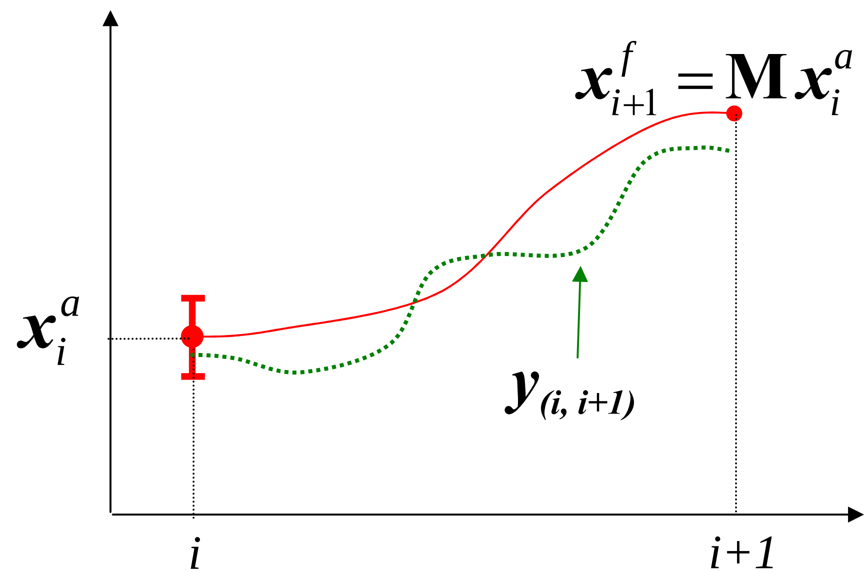


8. Improved temporal strategies
Distributed observations

- **Discrete DA problem**

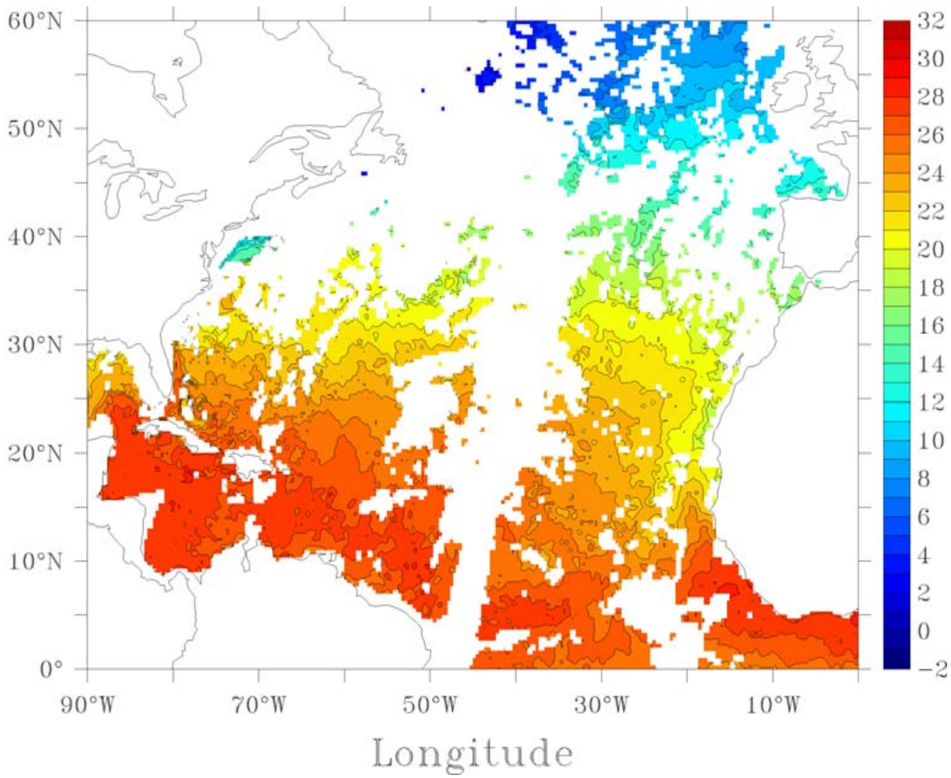


- **Continuous DA problem**

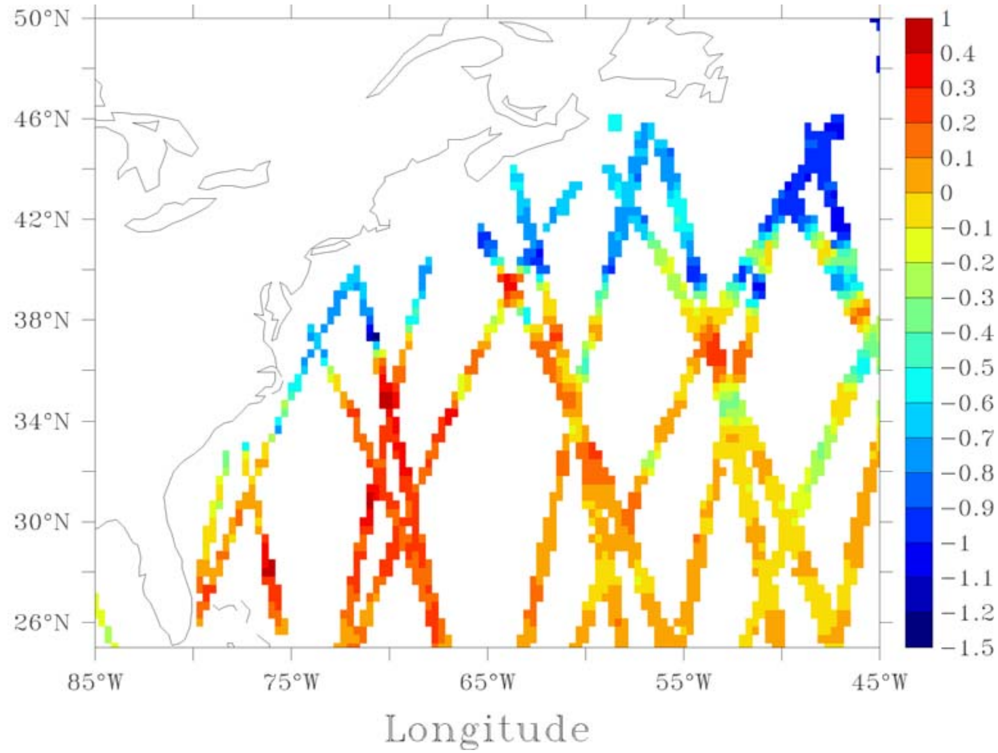


8. Improved temporal strategies
Example: composite data sets

3-day composite AVHRR SST
December 20-21-22,1992



3-day composite SLA
December 20-21-22,1992



- The observation vector y_{i+1} contains informations related to different instants.

8. Improved temporal strategies
Towards a time-continuous DA scheme

2 possible modifications of KF :

- **FGAT (First Guess at Appropriate Time)**
- **IAU (Incremental Analysis Update,** Bloom *et al.*, 1996**)**

$$\left\{\begin{array}{l} x_i^a \\ P_i^a \end{array}\right. \xrightarrow{\text{Forecast 1}} \left\{\begin{array}{l} x_{i+1}^f = M\left(x_i^a\right) \\ P_{i+1}^f = \mathbf{M}P_i^a\mathbf{M}^T + \mathbf{Q} \end{array}\right.$$

Analysis

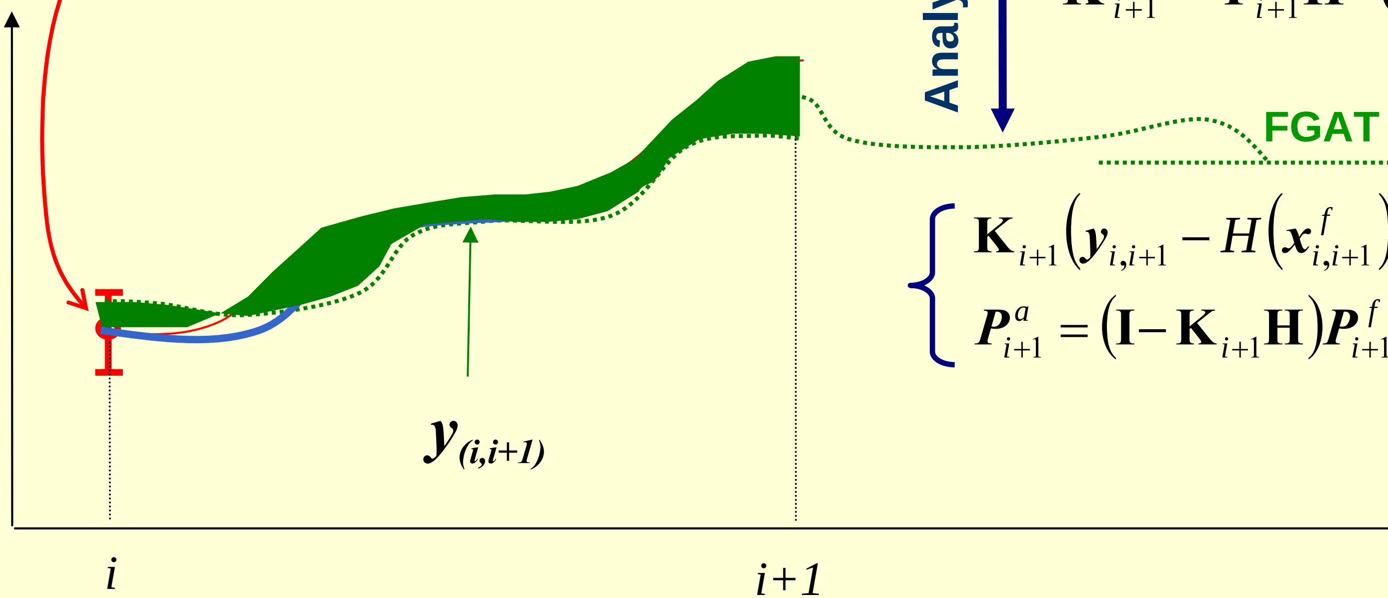
$$\mathbf{K}_{i+1} = \mathbf{P}_{i+1}^f\mathbf{H}^T(\mathbf{H}\mathbf{P}_{i+1}^f\mathbf{H}^T + \mathbf{R})^{-1}$$

FGAT

$$\left\{\begin{array}{l} \mathbf{K}_{i+1}\left(y_{i,i+1} - H\left(x_{i,i+1}^f\right)\right) = c_{i,i+1} \\ P_{i+1}^a = \left(\mathbf{I} - \mathbf{K}_{i+1}\mathbf{H}\right)P_{i+1}^f \end{array}\right.$$

Forecast 2
IAU

$$\tilde{x}_{i+1}^f = \tilde{M}\left(x_i^a\right)$$



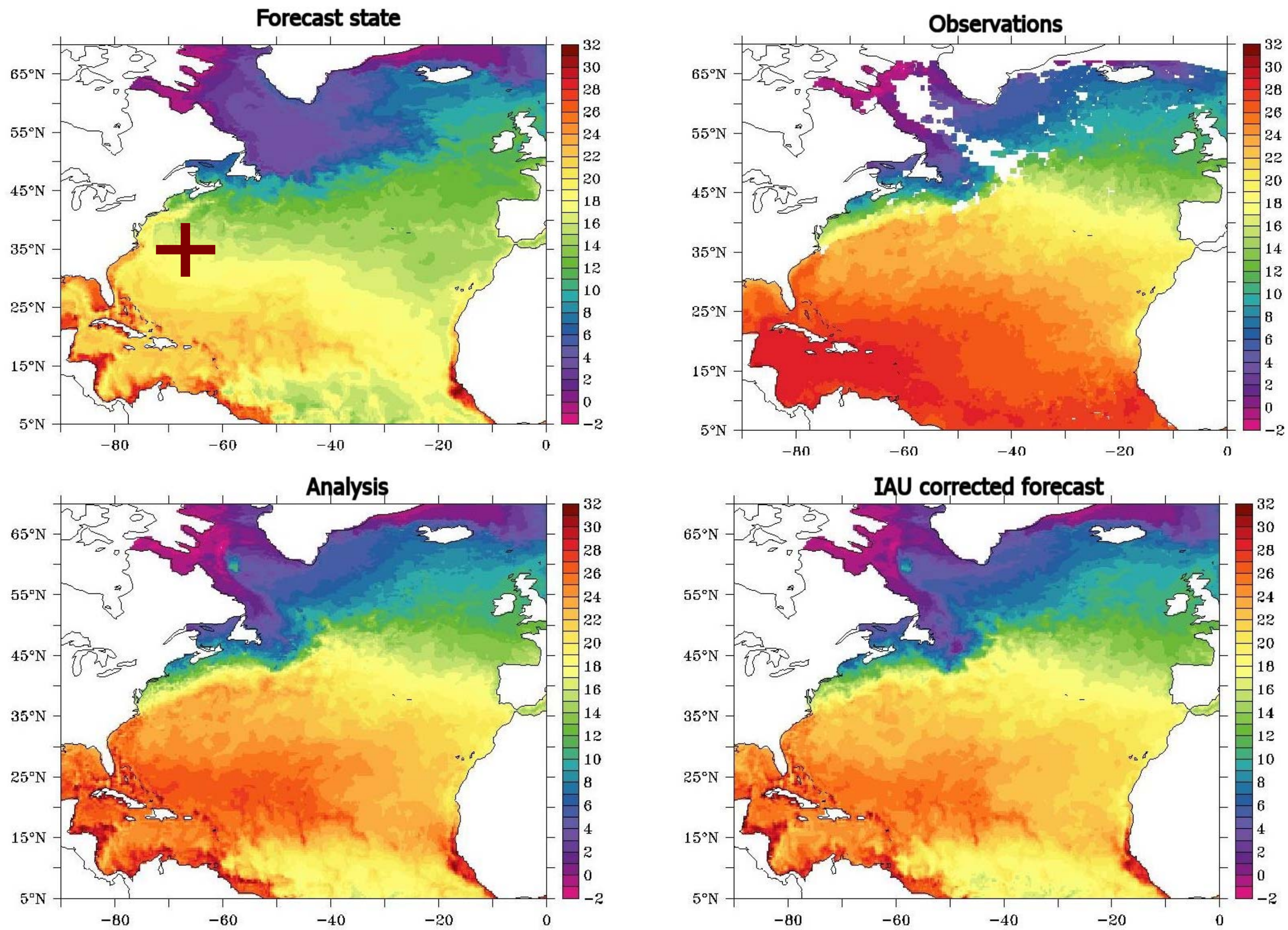
8. Improved temporal strategies
IAU implementation

- **Implementation of Incremental Analysis Update in OPA primitive equation model** :
  - Compute innovation using SST/SLA data and FGAT scheme ;
  - Compute Kalman gain and analysis increment at the end of assimilation window using SEEK algorithm;
  - Divide temperature and salinity increments by the number of model time steps in assimilation window $\Rightarrow \left(\frac{\delta T}{l}, \frac{\delta S}{l} \right)$
  - Integrate the OPA model on $(t_i\ ,\ t_{i+1})$ once again, with modified equations for temperature and salinity, i.e :

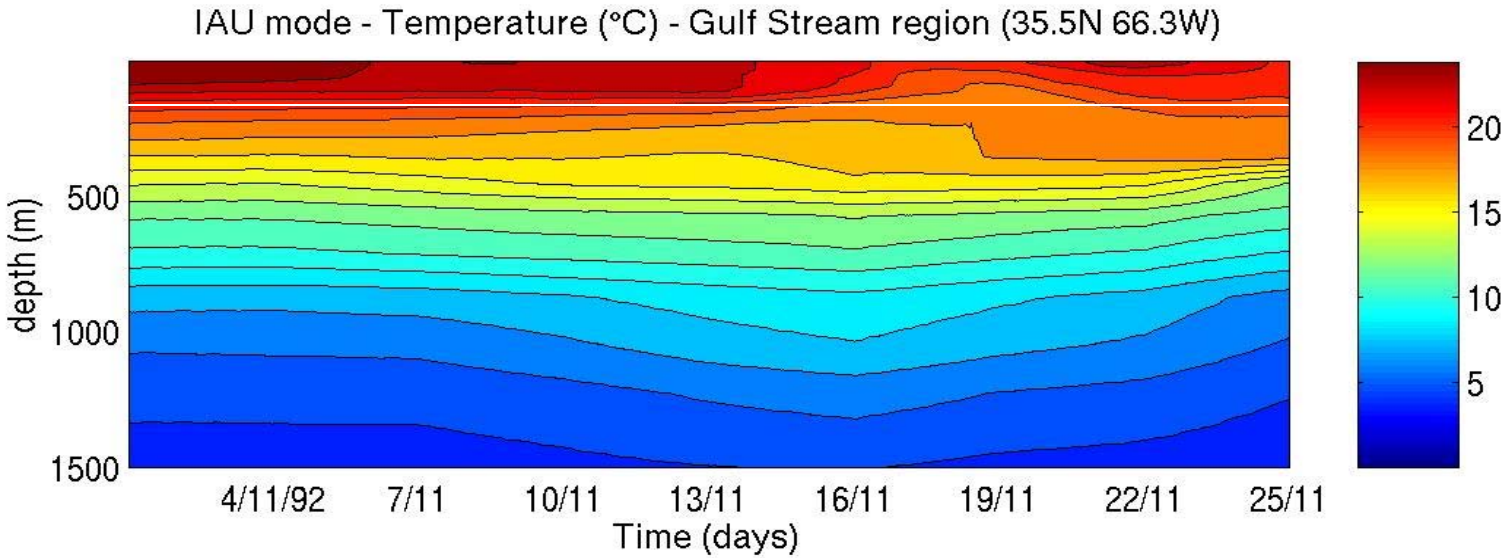
$$\frac{\partial T}{\partial t} + \boldsymbol{u}.\nabla^h T + w\frac{\partial T}{\partial z} = D^h(T) + \frac{\partial}{\partial z}\left(\tilde{\lambda}\frac{\partial T}{\partial z} \right) + \frac{\delta T}{l}$$

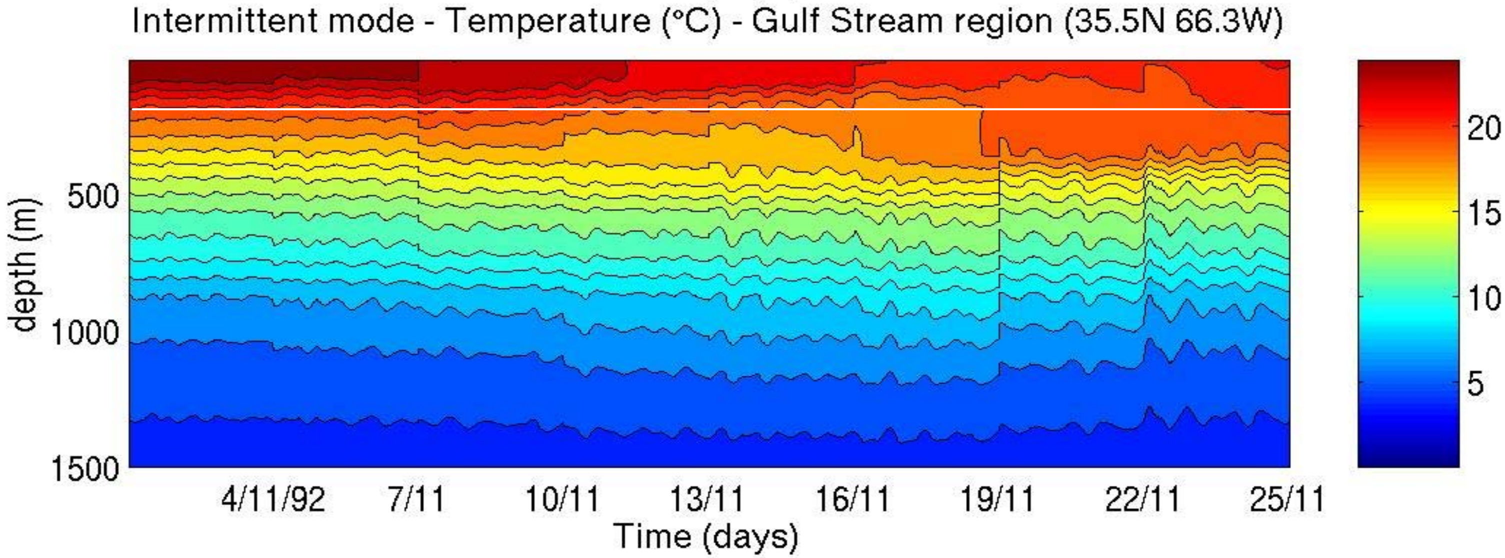
8. Improved temporal strategies
IAU – example (i)

- **SST estimates – 05/11/92**



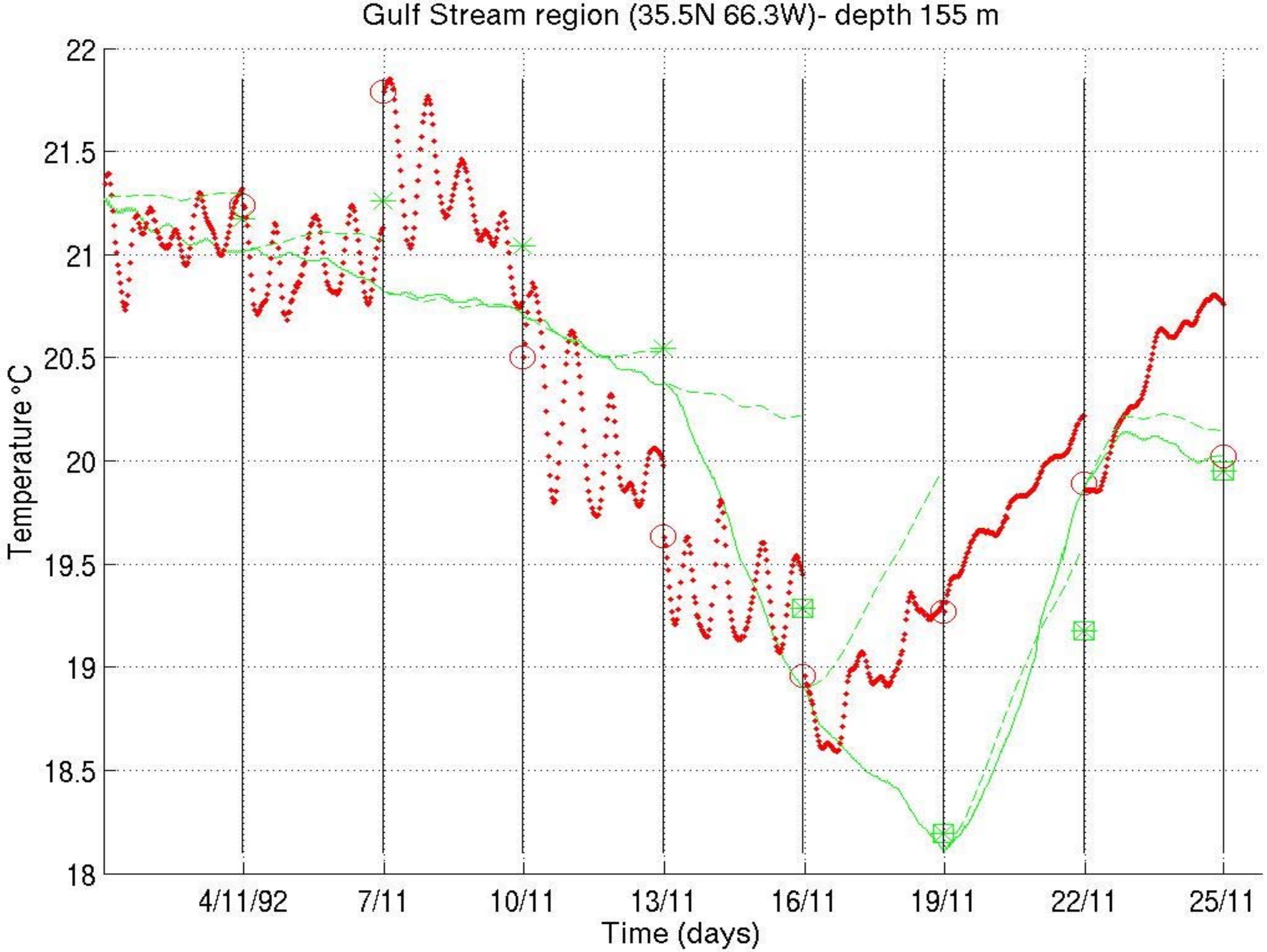
8. Improved temporal strategies
IAU – example (ii)





8. Improved temporal strategies
IAU – example (ii)

- **Temperature “mooring” – 11/92**



9. Statistical filter with inequality constraints

Motivations

- Inequality constraints are inherent to ocean models

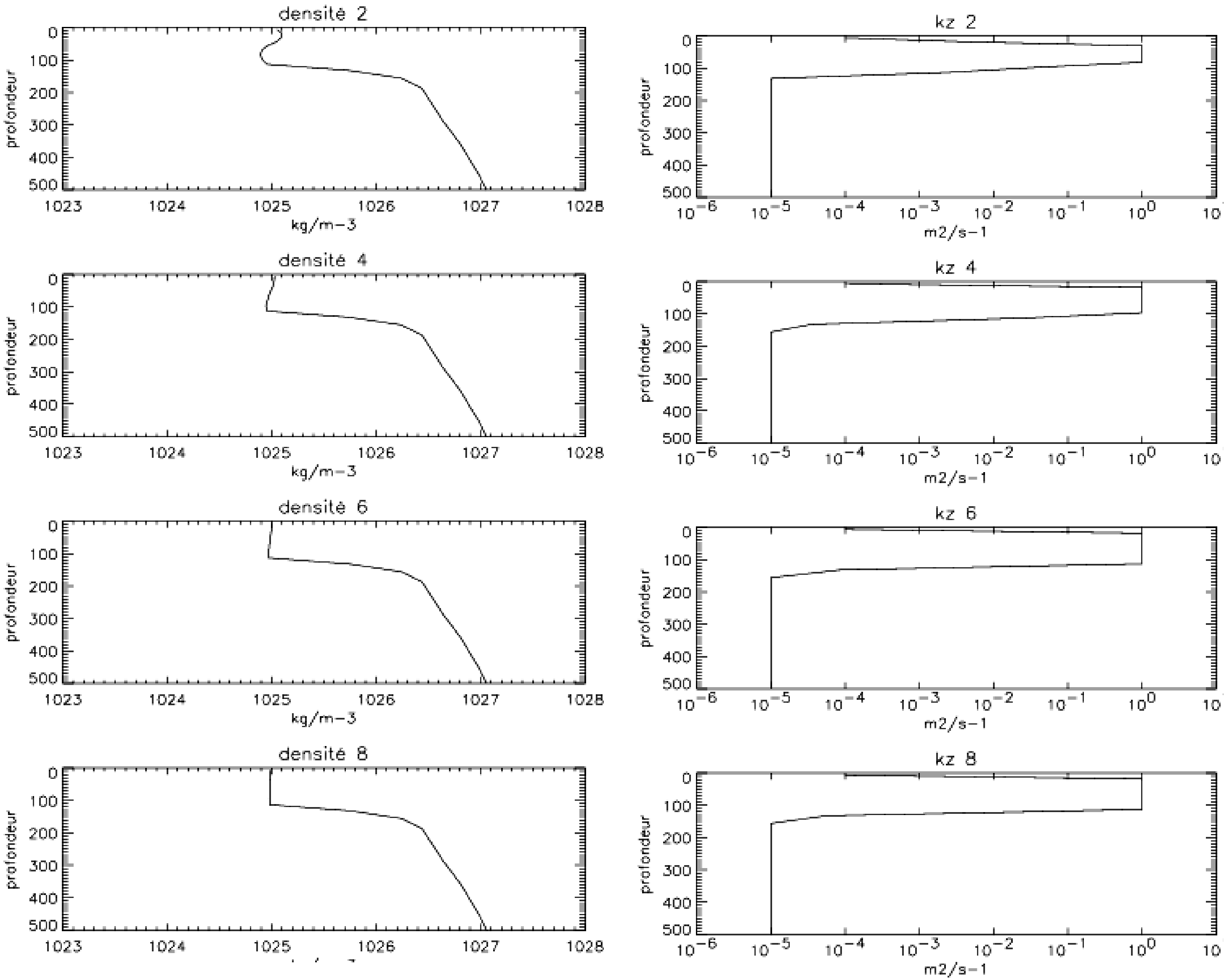
- Examples
  - Concentrations in biogeochemical models must be positive
  - T must be larger than freezing temperature
  - Static stability (a non-linear combination of T/S vertical gradients) must be verified at every assimilation step
  - …

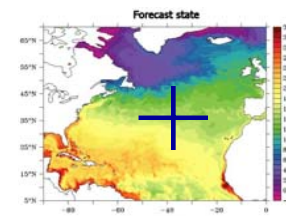
- The traditional Kalman filter framework with gaussian statistics doesn’t guarantee equality/inequality constraints

- Empirical correction schemes can be implemented after the statistical analysis step to restore the constraints

9. Kalman filter with inequality constraints

DA-induced static instability





References
Ensemble methods

- Brusdal K., Brankart J.M., Halberstadt G., Evensen G., Brasseur P., van Leeuwen P.J., Dombrowsky E. and Verron J., 2003: A demonstration of ensemble-based assimilation methods with a layerd OCGM from the perspective of operational ocean forecasting systems, *J. Mar. Syst.*, 40-41, 253-289.

- Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998 : Analysis scheme in the ensemble Kalman filter*, Mon. Weather Rev.*, **126**, 1719-1724.

- Evensen G., 1994:Sequential data assimilation with a non-linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res*., 99(C5), 10143-10162.

- Heemink A., M. Verlaan and A. Segers, 2001: Variance Reduced Ensemble Kalman Filtering, *Mon. Weather Rev.*, **129**, 1718-1728.

- van Leeuwen P.J., 2001 : An Ensemble Smoother with Error Estimates, *Mon. Weather Rev*., **129**, 709-728.

References
SEEK filter

- Ballabrera-Poy J., Brasseur P. and Verron J., 2001: Dynamical evolution of the error statistics with the SEEK filter to assimilate altimetric data in eddy-resolving ocean models, *Q. J. R. Met. Soc*., 127, 233-253.
- Brankart J.-M., Testut C.-E., Brasseur P. and Verron J., 2003: Implementation of a multivariate data assimilation scheme for isopycnic coordinate ocean models: Application to a 1993-96 hindcast of the North Atlantic Ocean circulation, *J. Geophys. Res*., 108, 3074, doi :10.1029/2001JC001198.
- Brasseur P., J. Ballabrera and J. Verron, Assimilation of altimetric observations in a primitive equation model of the Gulf Stream using a Singular Extended Kalman filter. *J. Mar. Syst.*, 22, 269-294, 1999.
- Pham D.T., J. Verron and M.C. Roubaud, Singular evolutive extended Kalman filter with EOF initialization for data assimilation in oceanography. *J. Mar. Syst.* , 16, 323-340, 1998.
- Testut C.-E., Brasseur P., Brankart J.-M., and Verron J., 2003 : Assimilation of sea-surface temperature and altimetric observations during 1992-1993 into an eddy-permitting primitive equation model of the North Atlantic Ocean, *J. Mar. Syst.*, 40-41, 291-316.
- Verron J., Gourdeau L., Pham D.T., Murtugudde R., and Busalacchi A., 1999: An Extended Kalman Filter to Assimilate Satellite Altimeter Data into a Non-linear Numerical Model of the Tropical Pacific: Method and validation, *J. Geophys. Res*., 104(C3), 5441-5458.

References
Error sub-space methods

- Fukumori I. and Malanotte-Rizzoli P., 1995: An approximate Kalman filter for ocean data assimilation: an example with an idealized Gulf Stream model, *J. Geophys. Res*., **100**(C4), 6777-6793.

- Lermusiaux, P. F. J. , 2001 : Evolving the subspace of the three-dimensional ocean variability: Massachusetts Bay*, J. Marine. Sys*., **29**,385-422.

- Malanotte-Rizzoli P., Fukumori I. and Young R., 1996: A Methodology for the Construction of a Hierarchy of Kalman Filters for Nonlinear Primitive Equation Models. In: P. Malanotte-Rizzoli (Ed.): *Modern Approaches to Data Assimilation in Ocean Modelling*, Elsevier Science B.V.

- Verlaan M. and A. W. Heemink, 1997: Tidal flow forecasting using reduced rank square root filters, *Stoch. Hydrology and Hydraulics*, 11, 349-368.

References
Other issues

- Bloom S. C., Takacs L. L., Da Silva A. M. et Ledvina D., 1996: Data assimilation using incremental analysis updates. *Mon. Weather Rev.,* 124(6) : 1256-1271.
- Cohn S.E. and Parrish D.F., 1991: The behaviour of forecast error covariances for a Kalman filter in two dimensions, *Mon. Wea. Rev*., **119**, 1757-1785.
- Dee, D.P., and A.M. Da Silva, 1998: Data assimilation in the presence of forecast bias. *Q. J. Roy. Met. Soc.,* **117,** 269-295.
- Fukumori, I., 2002: A partitioned Kalman filter and smoother. *Mon. Weather Rev.,* **130,** 1370-1383.
- Ide K., A.F. Bennett, P. Courtier, M. Ghil and A.C. Lorenc, 1997: Unified notation for Data Assimilation: Operational, Sequential and Variational, *J. Met. Soc. Japan,* **75**, 181-189.
- Miller R.N., 1986: Toward the application of the Kalman filter to regional open ocean modelling, *J. Phys. Oceanogr*., **16**, 72-86, 1986.
- Miller R.N., 2002 : Ensemble Generation for Models of Multimodal Systems, *Mon. Weather Rev*., **130**, 2313-2333.
- Weaver A.T., Vialard J. and D.L.T. Anderson, 2003: Three- and four-dimensional variational assimilation with an ocean general circulation model of the tropical Pacific Ocean. Part 1: formulation, internal diagnostics and consistency checks. *Mon. Weather Rev*., **131**, 1360-1378.

Conclusions

- **A fact** : A limited number of schemes have been successfully developed from theoretical basis to operational implementations (Mercator, HYCOM)

- **Learnings** : Specification of adequate error statistics (sub-space, statistical models etc.) is a central issue. Simplified KF (e.g. SEEK with fixed basis) have been very effective to test different statistical models.

- **A word of caution** : There is no generic method that can be considered as a « plug-and-play » solution. Each particular DA problem requires a good degree of understanding and *ad hoc* customization.

- **The future ?** : the next challenge to DA could be to combine local and global inversions (i.e. hybrid 4D-VAR / KF methods).

5. Global prototype
SAM-1, univariate

Schematically: