Variational Ocean Data Assimilation for the Mediterranean Forecasting System

N.Pinardi[1], J. Pistoia[2], A. Grandi[2]
S.Dobricic[3], C.Wikle[4], R. Milliff[5], M.Berliner[6]

[1]*Univ. of Bologna, Department of Physics and Astronomy, Bologna*
[2]*Istituto Nazionale di Geofisica e Vulcanologia, Bologna*
[3]*Joint Research Center, Ispra, Varese*
[4]*Dept. of Statistics, Univ. of Missouri, USA*
[5]*CIRES, University of Colorado, USA*
[6]*Dept. of Statistics, Ohio State Univ., USA*

August 2015

Summary

- Main aims of data assimilation for ocean predictions
- The Mediterranean Forecasting System
- The Data Assimilation scheme
- Analysis and forecast error structure
- New developments
  - The forecast uncertainty conundrum: ensemble forecasting with wind distributions
  - High frequency error covariance matrix estimates with BHM

Data assimilation for predictions

- Bjerknes (1914) described the two conditions that should be fulfilled in order to solve the prediction problem in atmosphere and oceans
  - I- Know the present state of the system as accurately as possible
  - II- Know the laws of physics that regulate the time evolution of the basic field state variables, i.e. have predictive models
- In order to solve the prediction problem the scientific approach should consider 3 partial problems
  - Comp.1: The observational network: QC and real time data management
  - Comp.2: The diagnostic/data assimilation: selection of algorithms that produce best initial conditions
  - Comp.3: The prognostic component

Mediterranean Forecasting System

MFS OGCM and atmospheric forcing

RT data, quality control and pre-processing

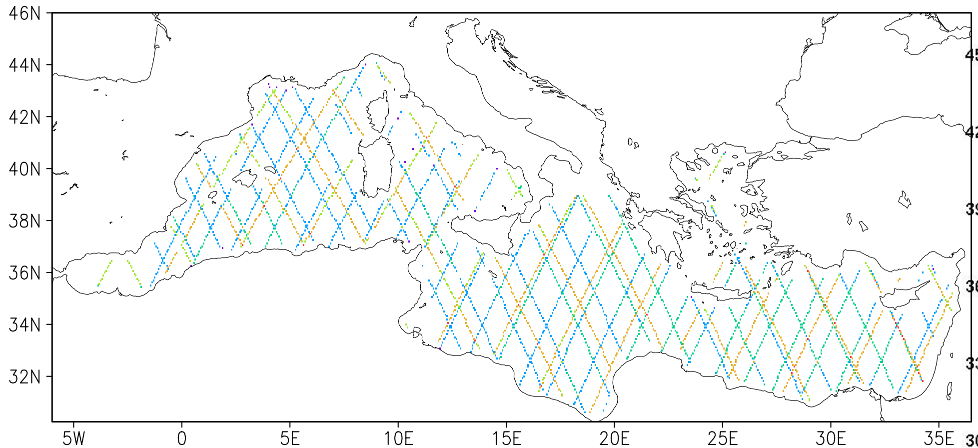
Sequential daily assimilation cycle with best atmospheric forcing

Initial condition for the forecast (analysis)

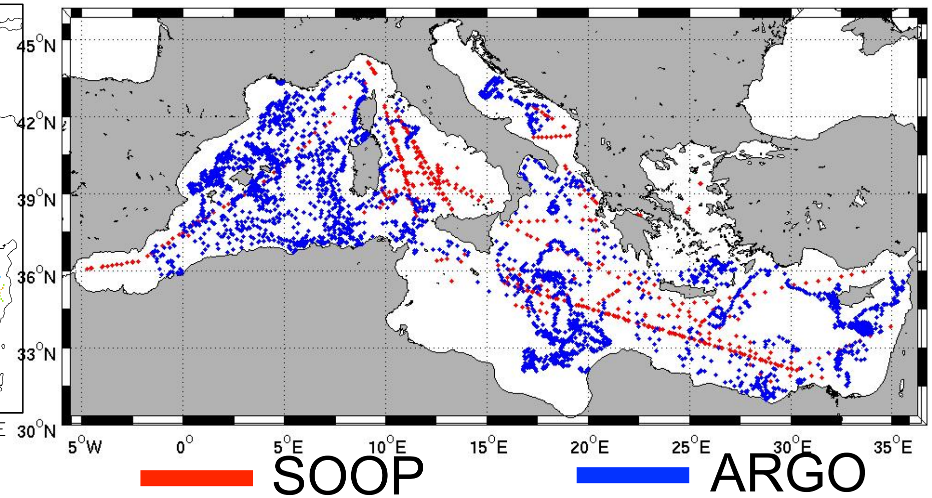
Daily 10 days forecasts

Real time observational component

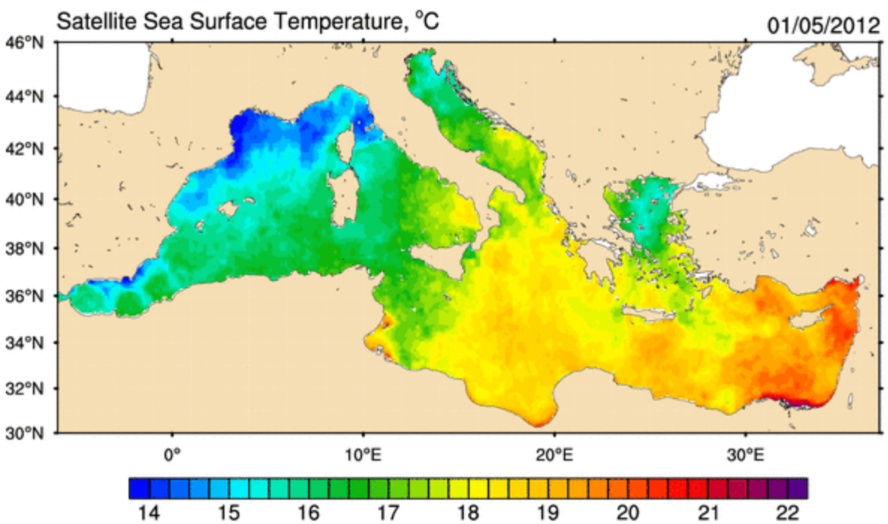
Multisatellite along track sea level

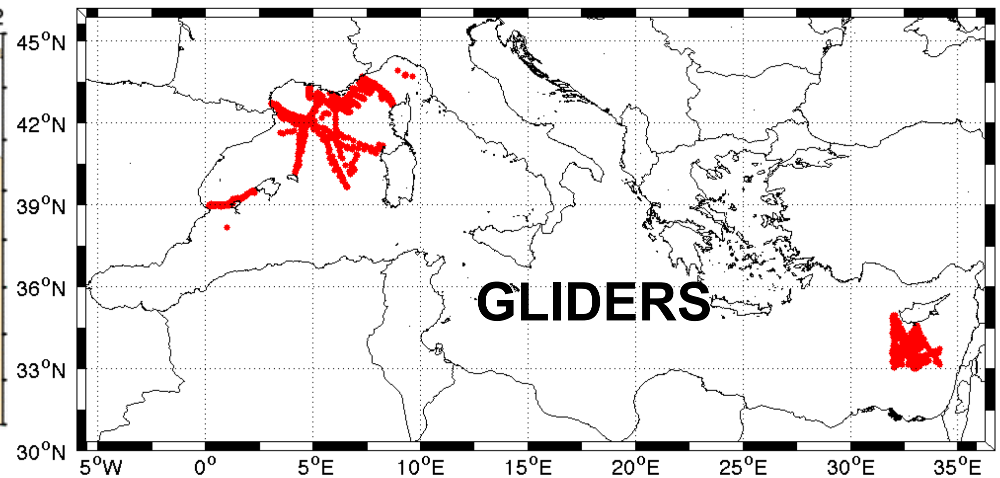


2008-2011 coverage

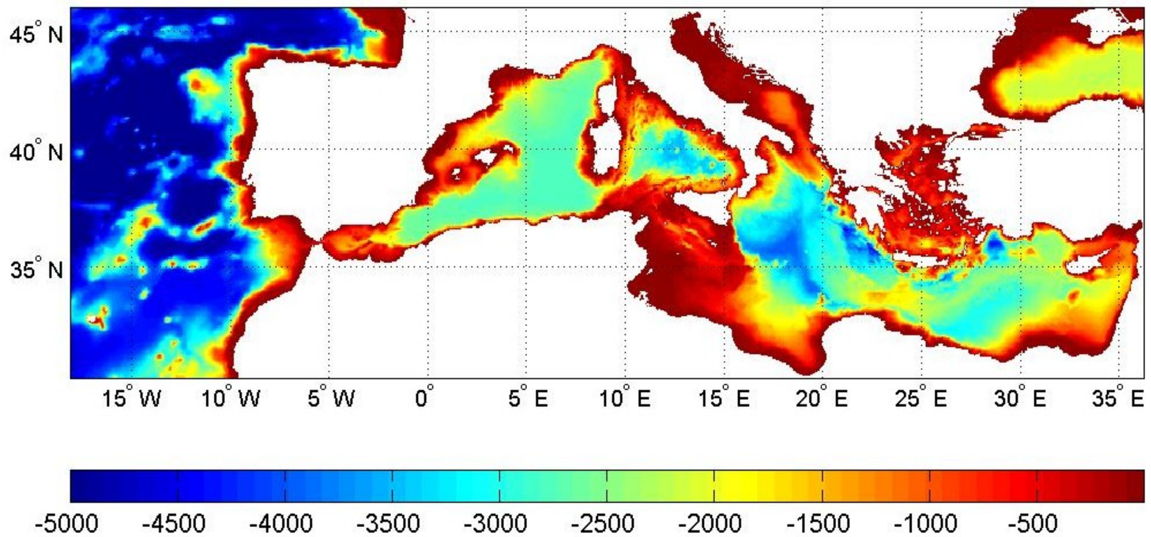


Multi-sensor daily OI SST





Modelling component



MFS-16L72 NEMO v3.4

Explicit free surface, bilaplacian viscosity and diffusion

1/16º x 1/16º horizontal resolution (6.5 km),

72 unevenly partial steps vertical levels(1.5m-300m)

River runoff (seas. clim.) and surface atmo. pressure included

Air-sea heat, water and momentum fluxes all interactive

Coupling with WWIII model (now only for surface momentum drag)

The ocean 3DVAR scheme (Dobricic and Pinardi, OM, 2008)

A cost function, linearized around the background state, is minimized:

$$J = \frac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}[\mathbf{H}(\delta\mathbf{x}) - \mathbf{d})]^T\mathbf{R}^{-1}[\mathbf{H}(\delta\mathbf{x}) - \mathbf{d})]$$

$$\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_b \qquad \mathbf{d} = [H(\mathbf{x_b}) - \mathbf{y_o}] \text{ misfit}$$

The oceanic vector state is defined:

$$\mathbf{x} = [\mathbf{u},\mathbf{v},\eta,\mathbf{T},\mathbf{S},]^T$$

The background error covariance matrix is defined as:

$$\mathbf{B} = \mathbf{V}\mathbf{V}^T$$

The ocean 3DVAR scheme (Dobricic and Pinardi, OM, 2008)

$$\mathbf{B} = \mathbf{V}\mathbf{V}^T$$

The key choice: $\mathbf{V}$ is modeled as a sequence of linear operators:

$$\mathbf{V} = \mathbf{V_D}\mathbf{V_{uv}}\mathbf{V_h}\mathbf{V_H}\mathbf{V_V^{ts}}$$

$\mathbf{V_V^{ts}}$ - Vertical EOFs for T, S cov.

$\mathbf{V}_H$ - Horizontal covariances (rec. filter).

$\mathbf{V}_\eta$ - Barotropic model for eta covar.

$\mathbf{V}_{uv}$ - Diagnostic u and v covariances

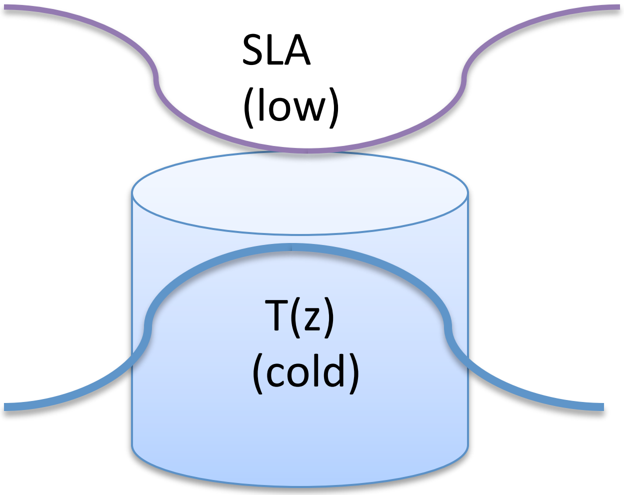
$\mathbf{V}_D$ - Divergence damping filter (coastal constraint).

Key issue for initialization purposes: the vertical structure of the error covariance matrix

A key assimilation problem in the ocean

- Large data sets are at the surface of a fluid heated/cooled above: need to preserve stability to prevent vertical mixing after correction
- Need to “project” satellite altimetry into correct T,S correction profiles

Cyclonic vortex: the SLA dnd T(z) covariance

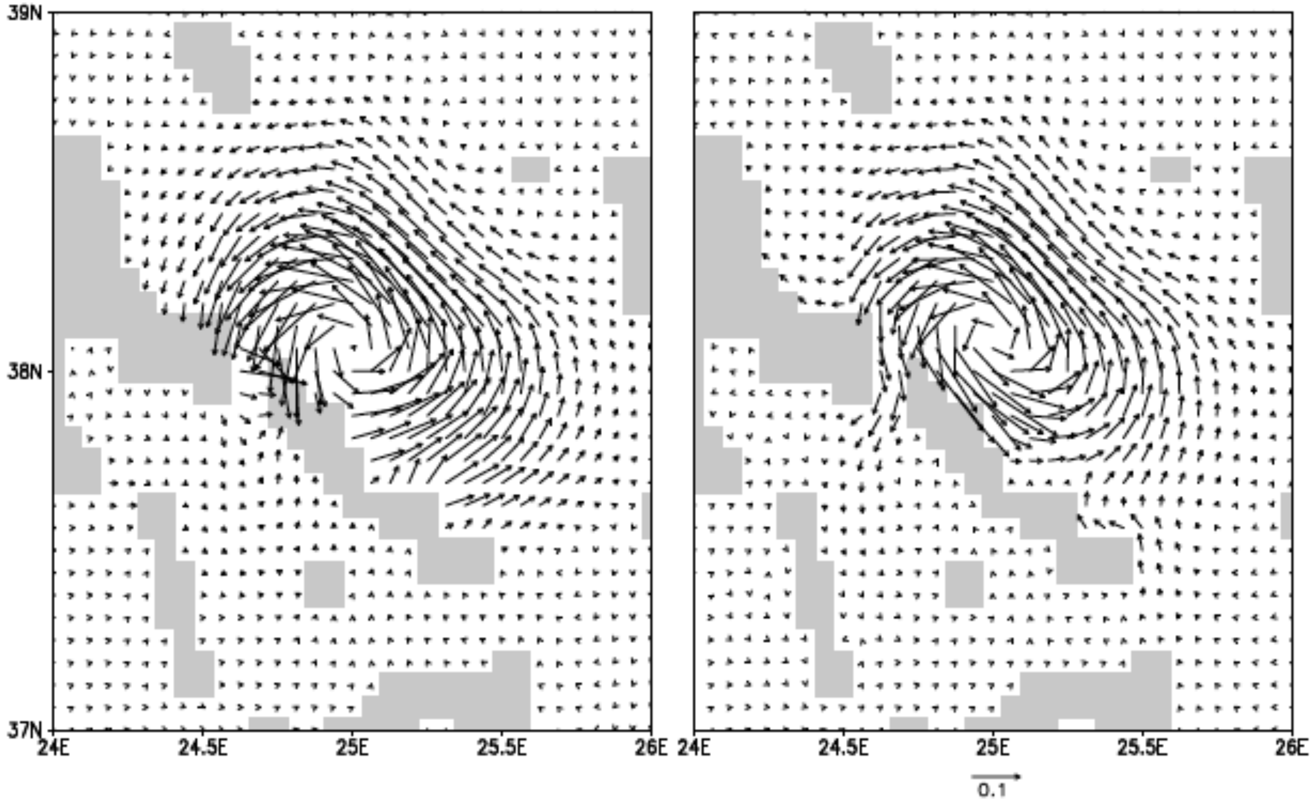


- Thus background error covariance part due to T,S crucial

A key assimilation problem in the ocean

- The coastal constraint: divergence dumping to have flow parallel to coasts and not across

After assimilation without filter   After assimilation with filter



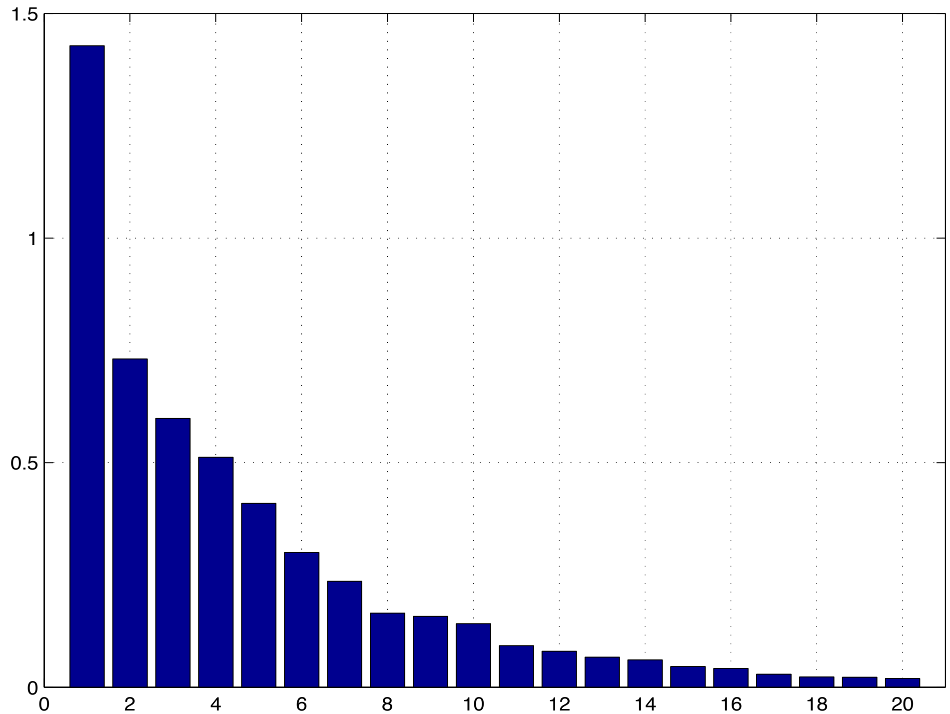
The vertical error covariance: vertical eigenvectors and eigenvalues

A data matrix is composed with Temperature and Salinity anomalies from a long simulation (10-20 years) or re-analysis for each season and several regions

$$\mathbf{A} = \left[\frac{h_i\mathbf{dT}_i}{H\sigma_T}\cdots\cdots\frac{h_i\mathbf{dS}_\mathbf{i}}{H\sigma_S}\right]\begin{bmatrix}\dfrac{h_i\mathbf{dT}_i}{H\sigma_T}\\ \cdot\\ \cdot\\ \dfrac{h_i\mathbf{dS}_\mathbf{i}}{H\sigma_S}\end{bmatrix}$$

The matrix is then decomposed in singular values

$$\mathbf{A} = \mathbf{V}_{\mathbf{V}}^{\mathbf{t}\,\mathbf{s}}\mathbf{L}\mathbf{S}^{\mathbf{T}}$$



First 20 eigenvectors over about 60 possible

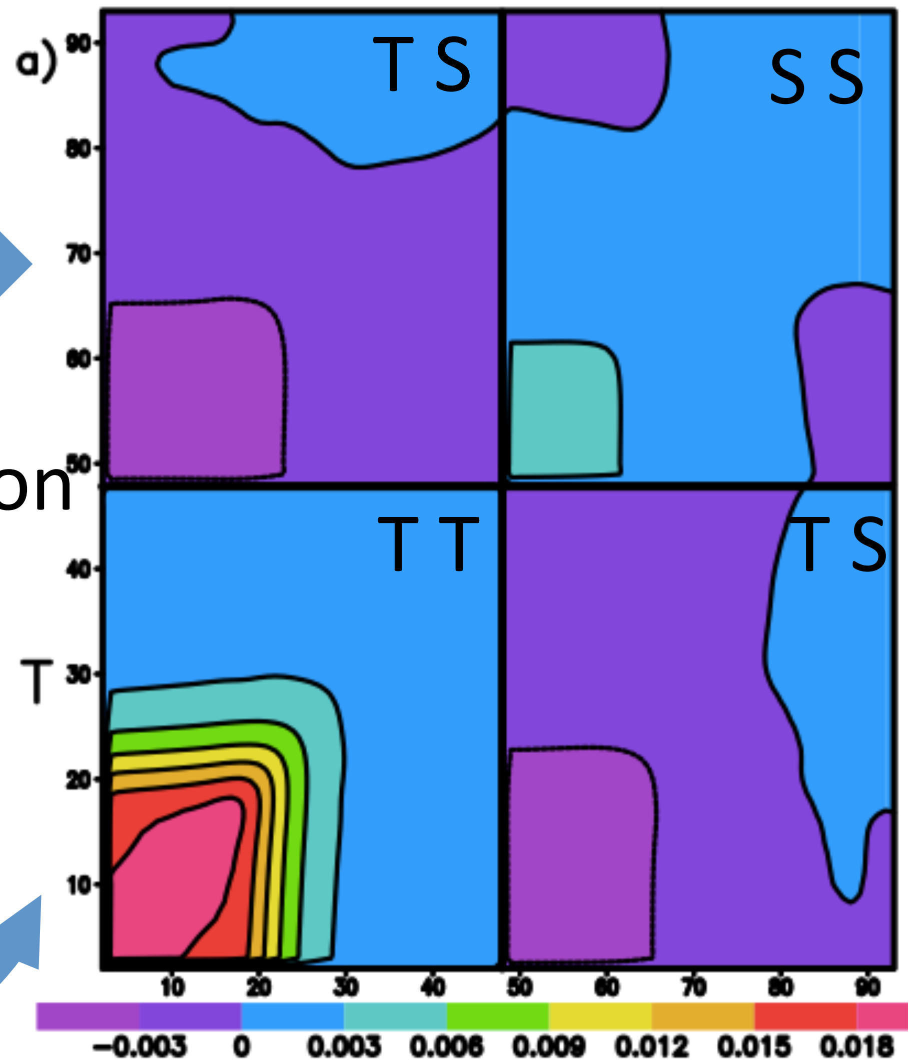
Operational Vertical error covariance matrix structure

$$\mathbf{B} = \mathbf{V}\mathbf{V}^T$$

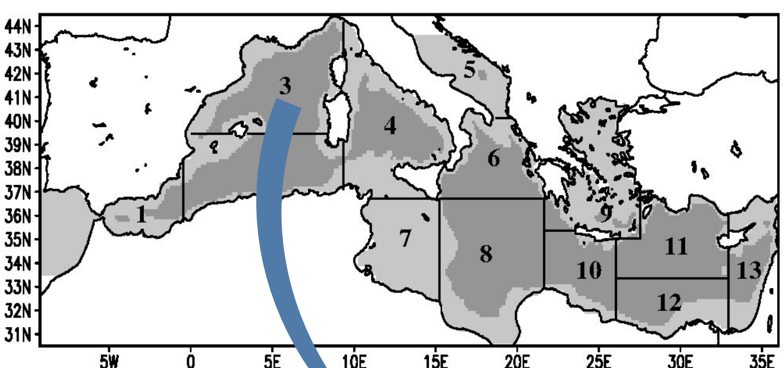
$$\mathbf{V} = \mathbf{V}_\mathbf{D}\mathbf{V}_\mathbf{uv}\mathbf{V}_\mathbf{h}\mathbf{V}_\mathbf{H}\mathbf{V}_\mathbf{V}^{\mathbf{t}_\mathbf{S}}$$

$$\mathbf{C} = \mathbf{V}_V^{t_S}\mathbf{V}_V^{t_S\,\mathbf{T}}$$

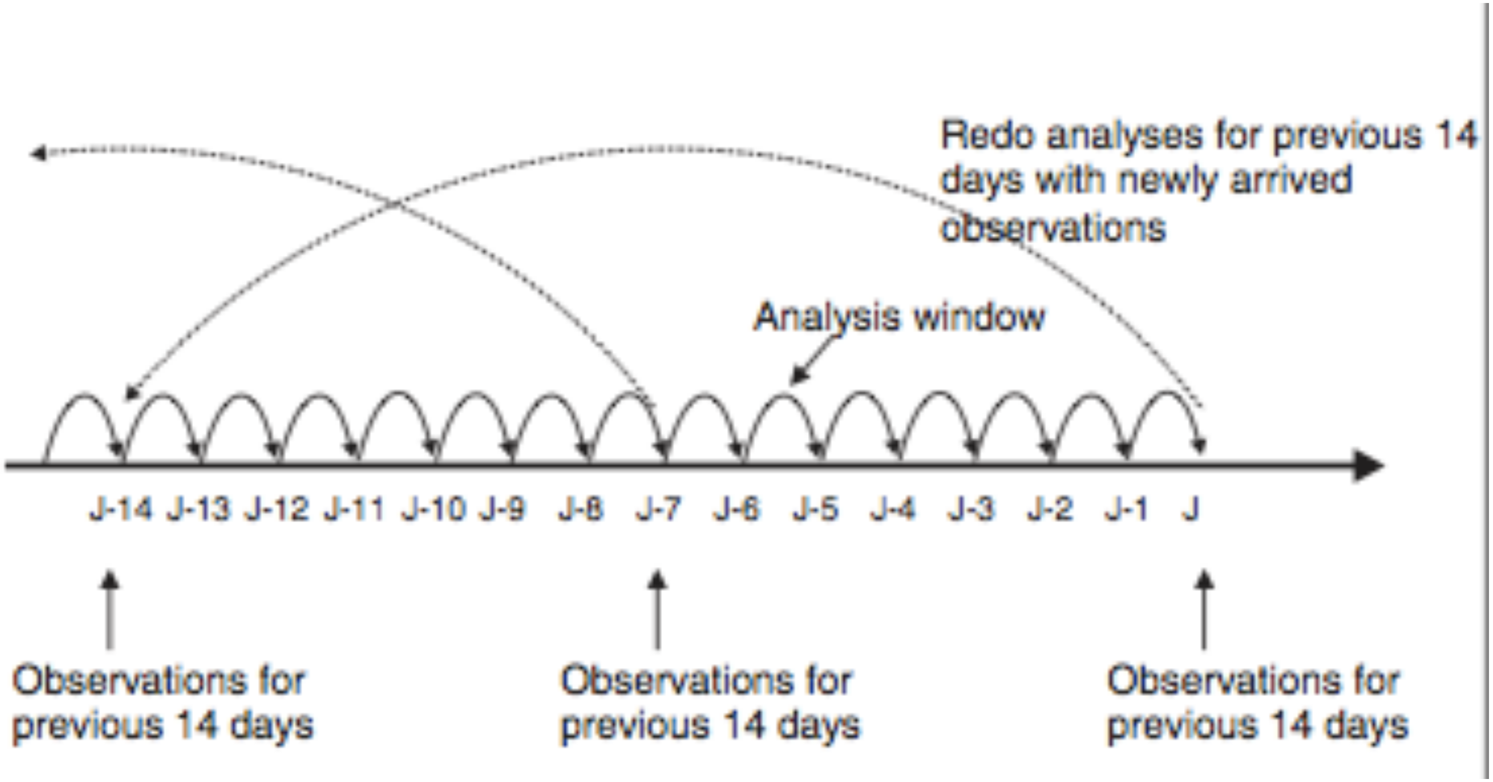
C



One C for region & season

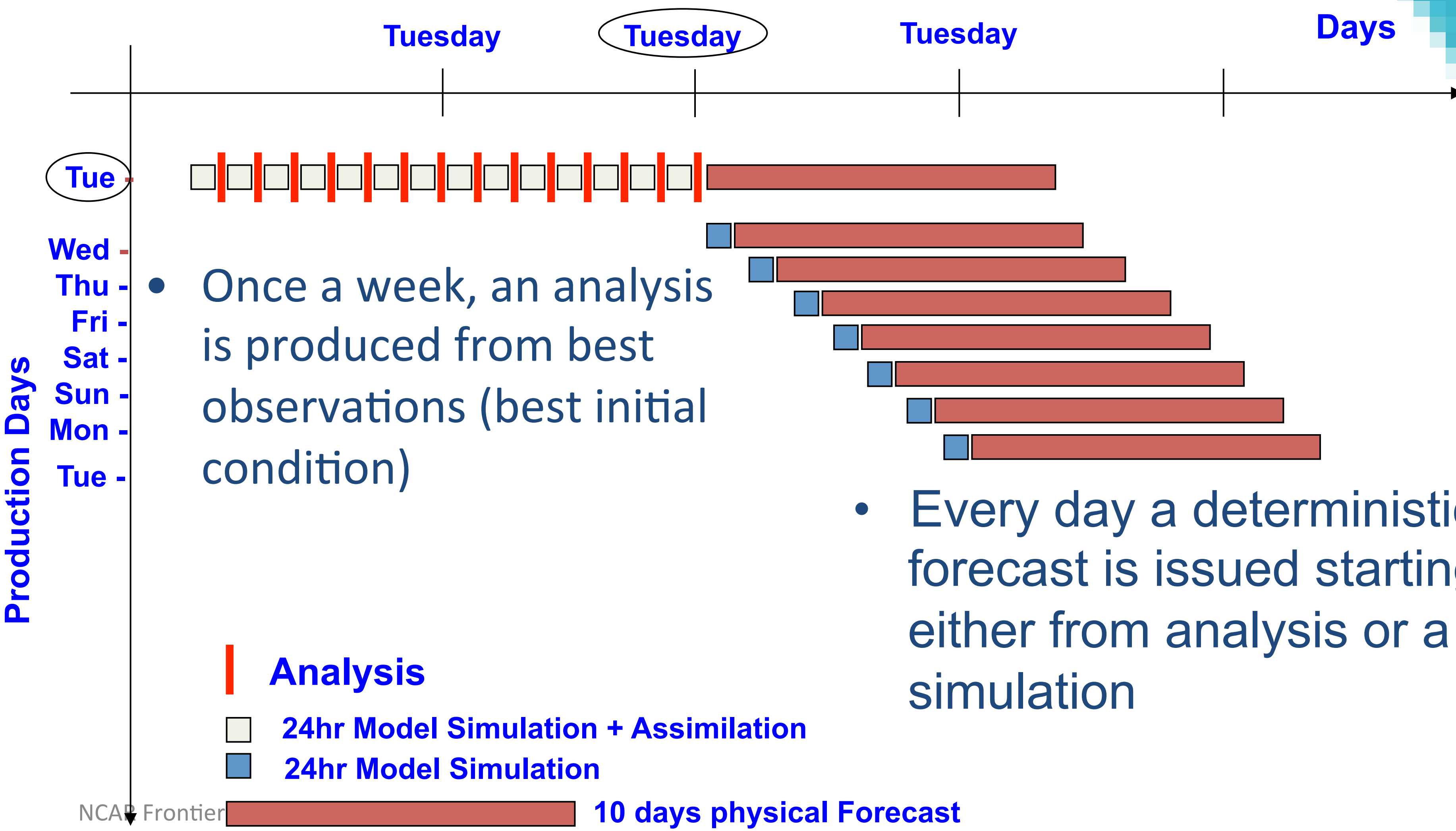


The data are assimilated weekly with a daily window



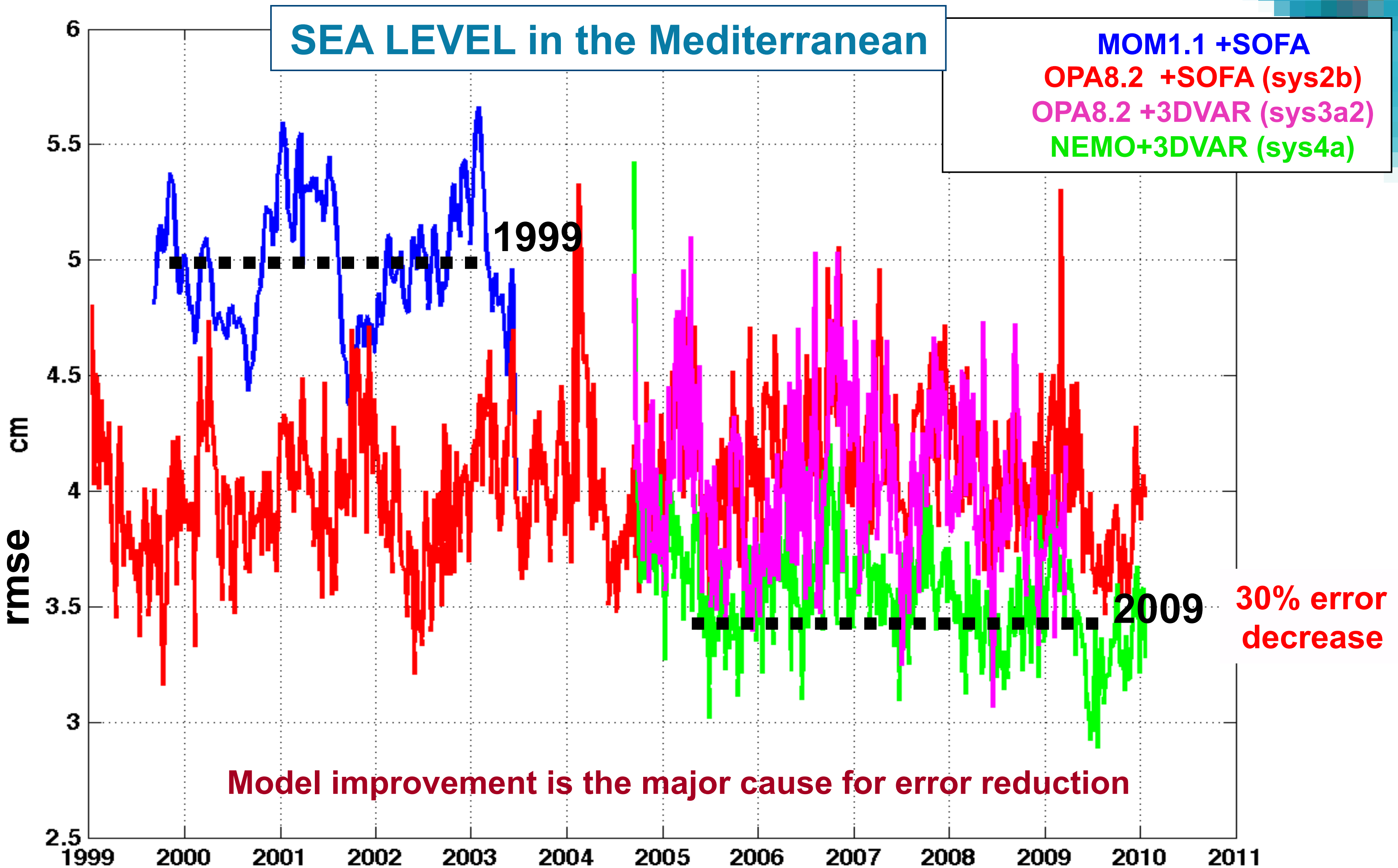
Weekly assimilation cycle because data of higher quality is available

The MFS deterministic forecast production system

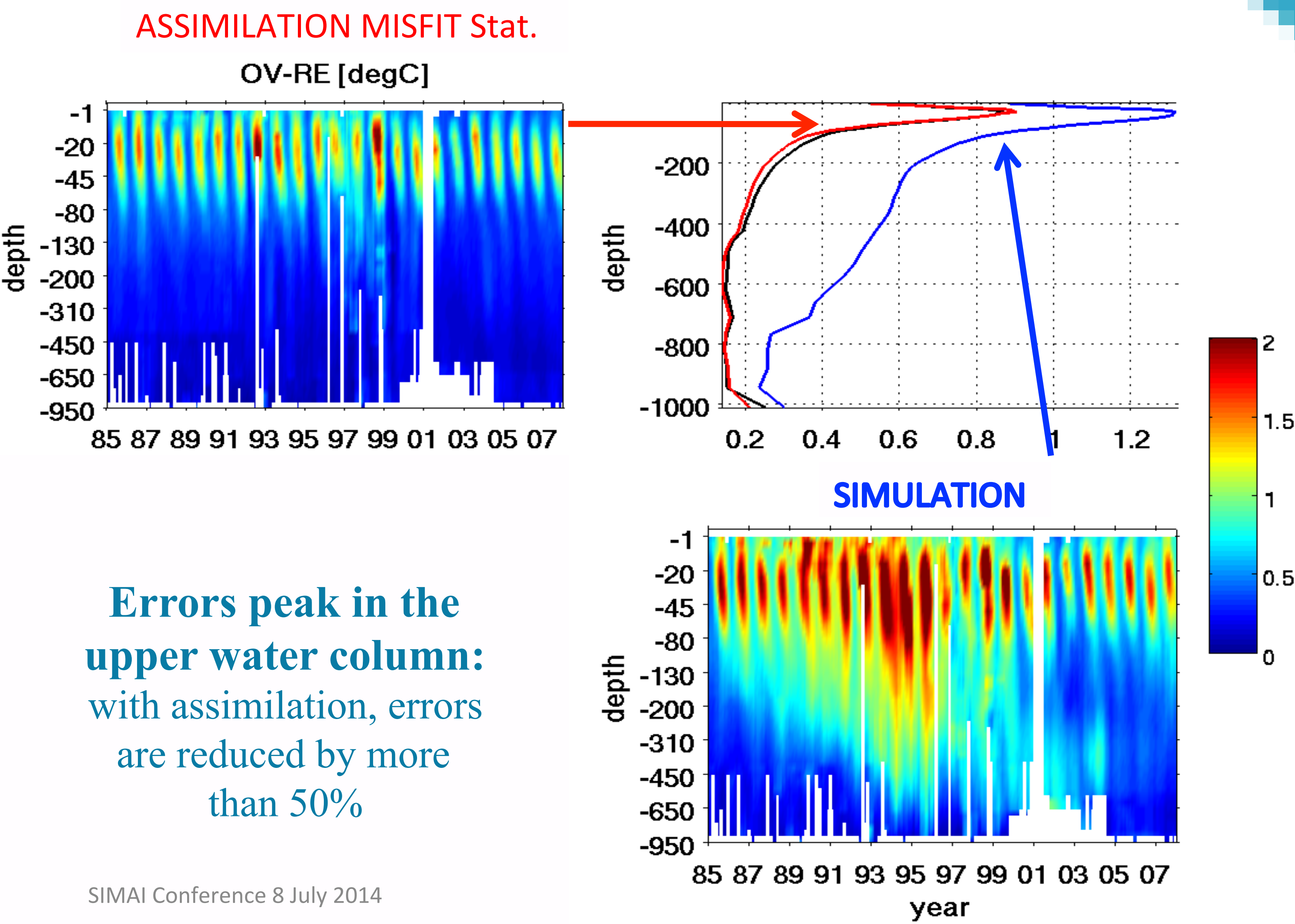


- Once a week, an analysis is produced from best observations (best initial condition)
- Every day a deterministic forecast is issued starting either from analysis or a simulation

How did the error decrease in 10 years?



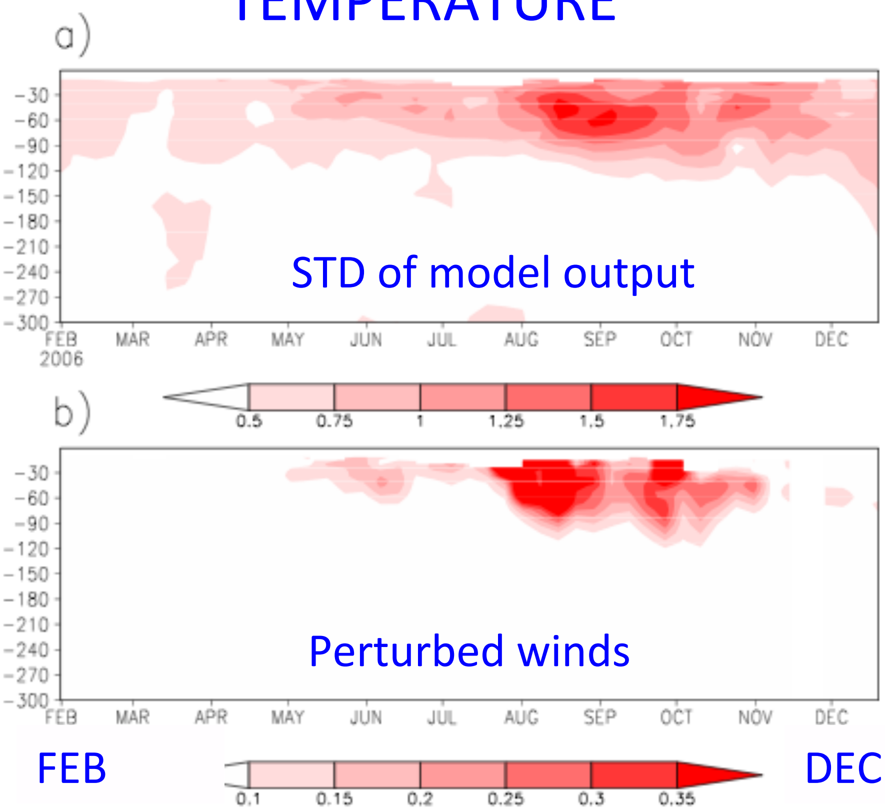
The error structure for temperature



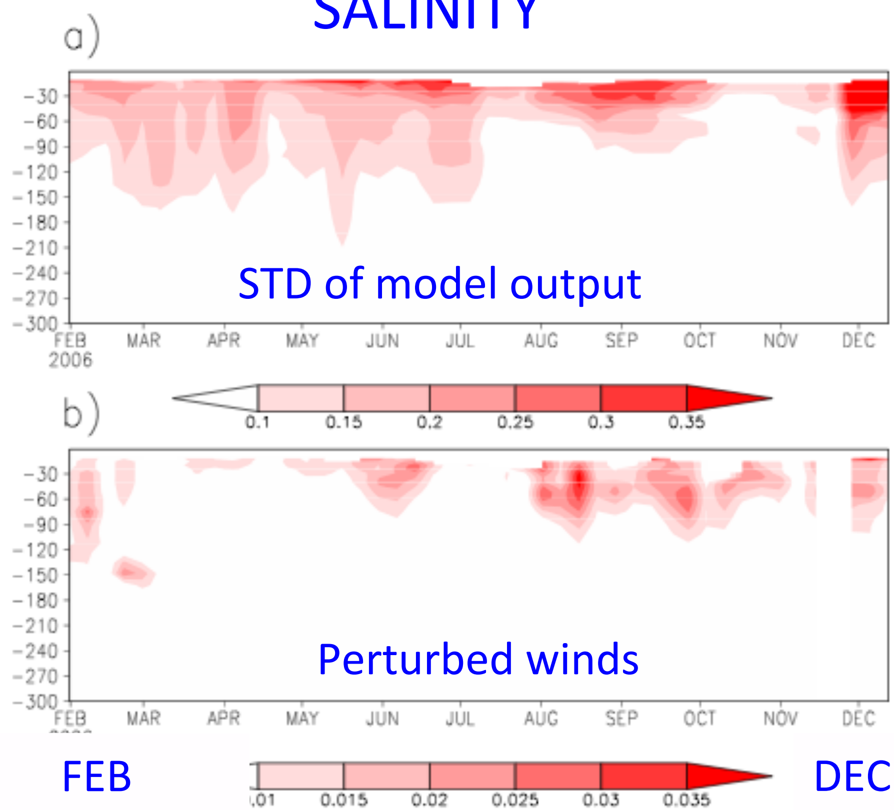
Errors peak in the upper water column:
with assimilation, errors are reduced by more than 50%

What is this vertical error variance in T and S due to?

TEMPERATURE



SALINITY



Errors in atmospheric forcing are projecting on the vertical structure of the temperature & salinity errors

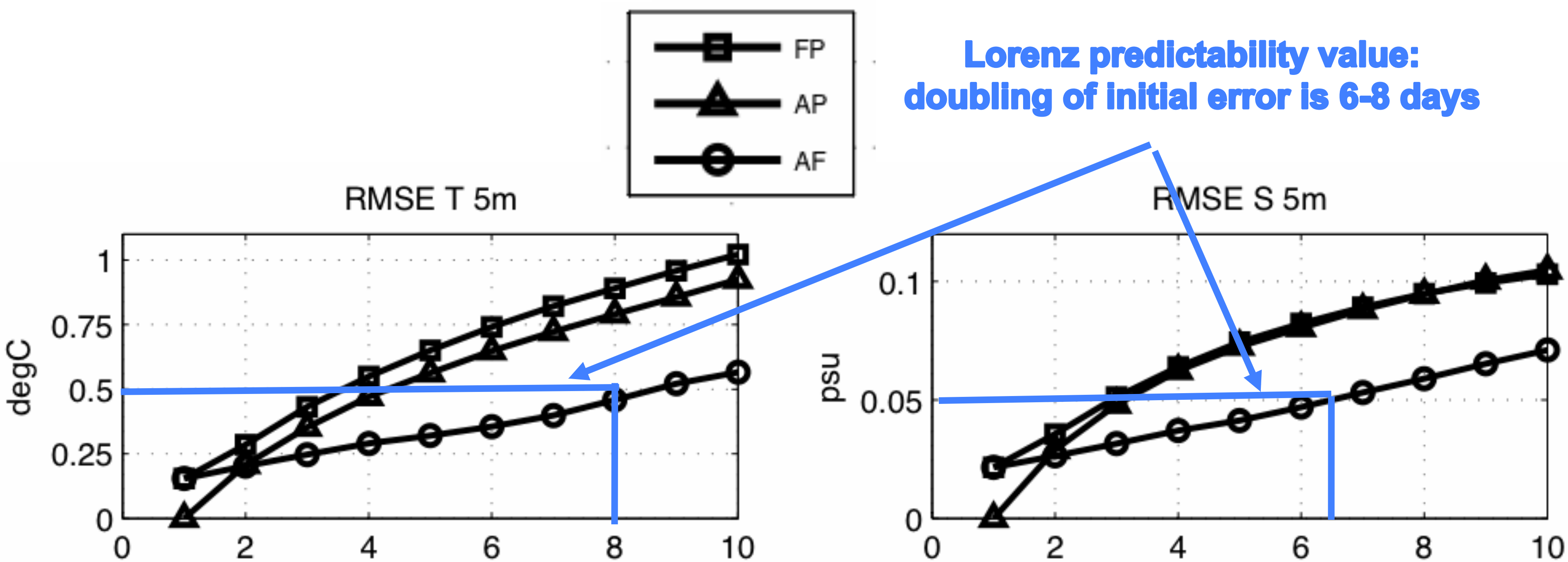
Predictability time for T and S at the surface

Analysis-forecast

$$AF_c(t) = \left\langle \sqrt{\frac{\sum_{1}^{N}\left(X_{FC}(t) - X_{AN}(t)\right)^2}{N}} \right\rangle$$

Analysis-Persistence

$$AP(t) = \left\langle \sqrt{\frac{\sum_{1}^{N}\left(X_{AN}(t) - X_{AN}(t = \mathrm{d}1)\right)^2}{N}} \right\rangle$$



(2006) Forecast days

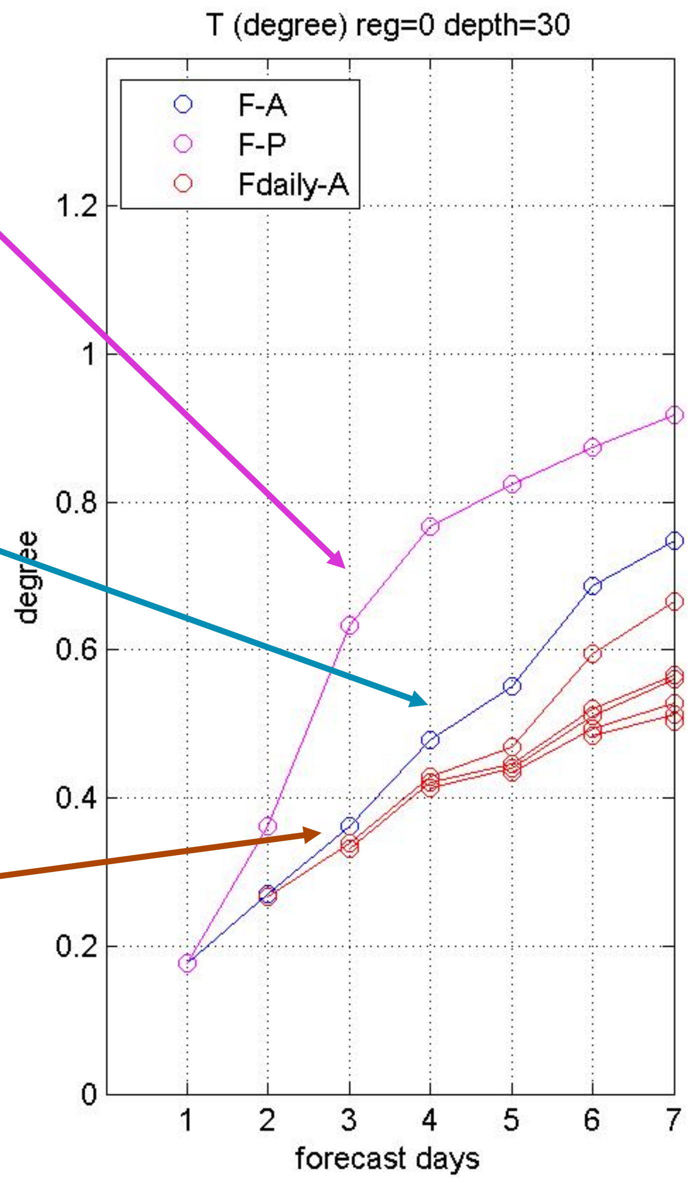
Ocean forecast error at 30 m: the effect of atmospheric forcing errors

Control:

Forecast-persistence

Weekly forecast

Forecasts started from simulations driven by atmospheric analyses



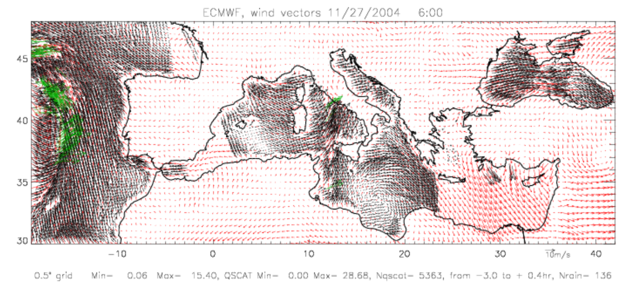
Forecast error at day 7 decreased by 30% only reducing atmospheric forcing error

The forecast uncertainty conundrum

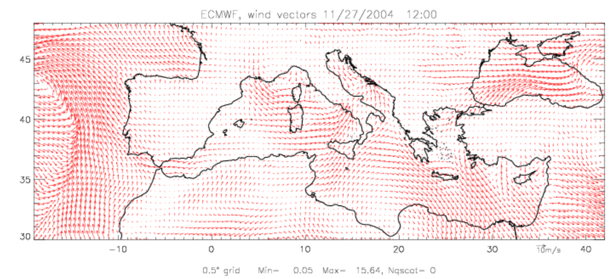
- Shukla (2005): 'The largest obstacles in realizing the potential predictability of weather and climate are inaccurate models and insufficient observations, rather than an intrinsic limit of predictability'
- Uncertainty of ocean forecasts depends on:
  - Ocean Initial condition errors
  - Atmospheric forcing errors
  - Model errors (Physics, numerics)
- Hypothesis:
  - We use ensemble forecasting as a means to test ocean predictability issues
  - We concentrate on atmospheric wind forcing errors and how they affect the initial condition and forecast errors

Building the wind distributions using Bayesian Hierarchical Modelling (BHM-SVW)

Conceptual and implementation blocks:



QSCAT
ASCAT



ECMWF

Data Stage: 2 types of data

Scatterometer winds and ECMWF analyses/forecasts

Process model stage:

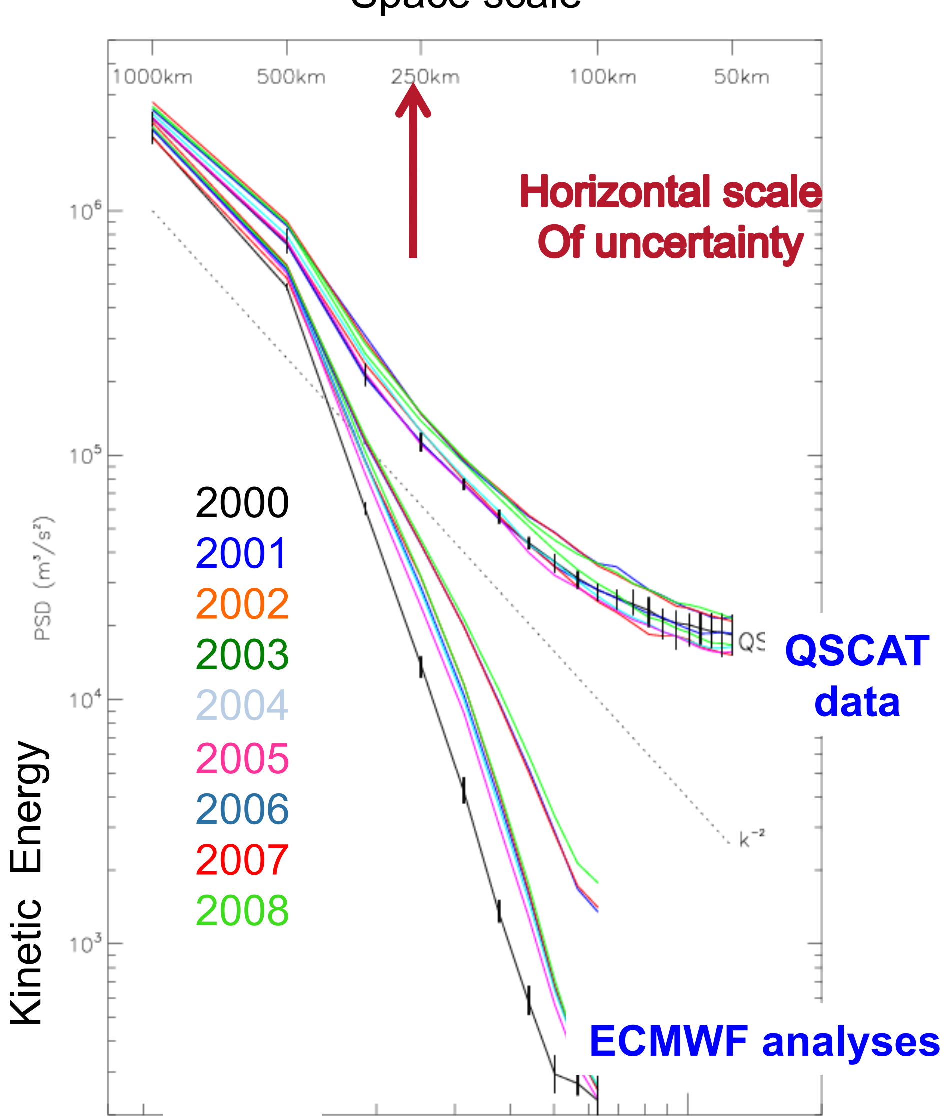
Raylegh friction surface model translated into a stochastic finite difference equation

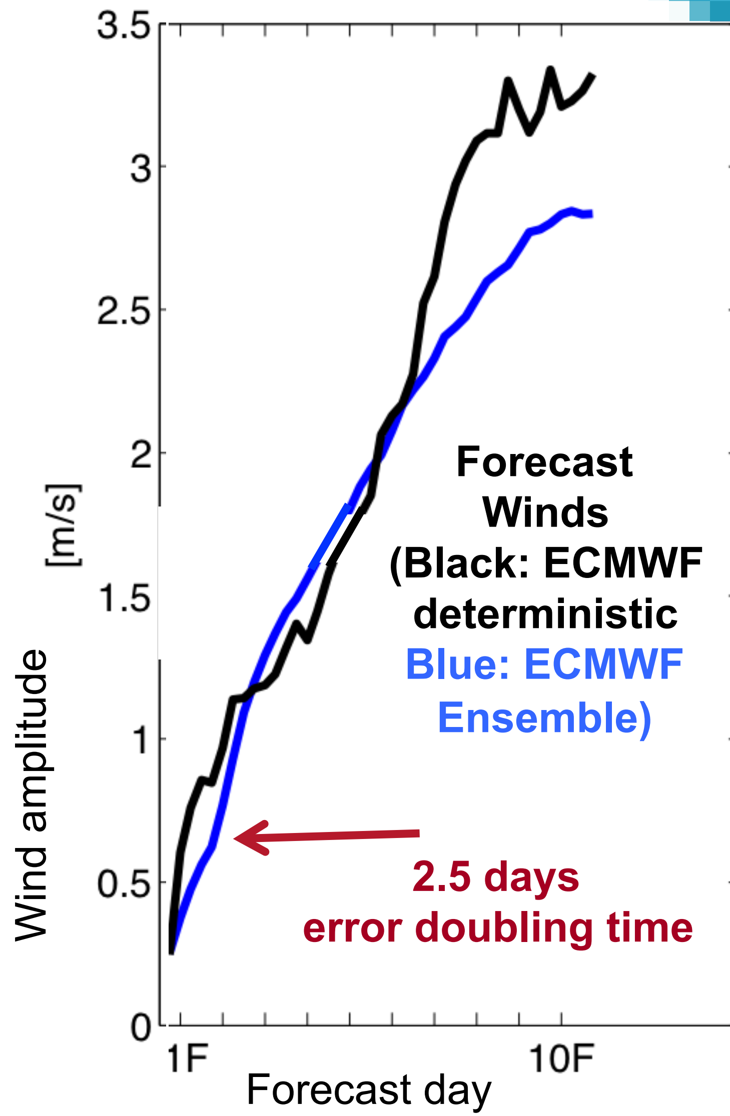
$$u = -\frac{f}{\rho_0\left(f^2+\gamma^2\right)}\frac{\partial p}{\partial y} - \frac{\gamma}{\rho_0\left(f^2+\gamma^2\right)}\frac{\partial p}{\partial x}$$

$$v = \frac{f}{\rho_0\left(f^2+\gamma^2\right)}\frac{\partial p}{\partial x} - \frac{\gamma}{\rho_0\left(f^2+\gamma^2\right)}\frac{\partial p}{\partial y}$$

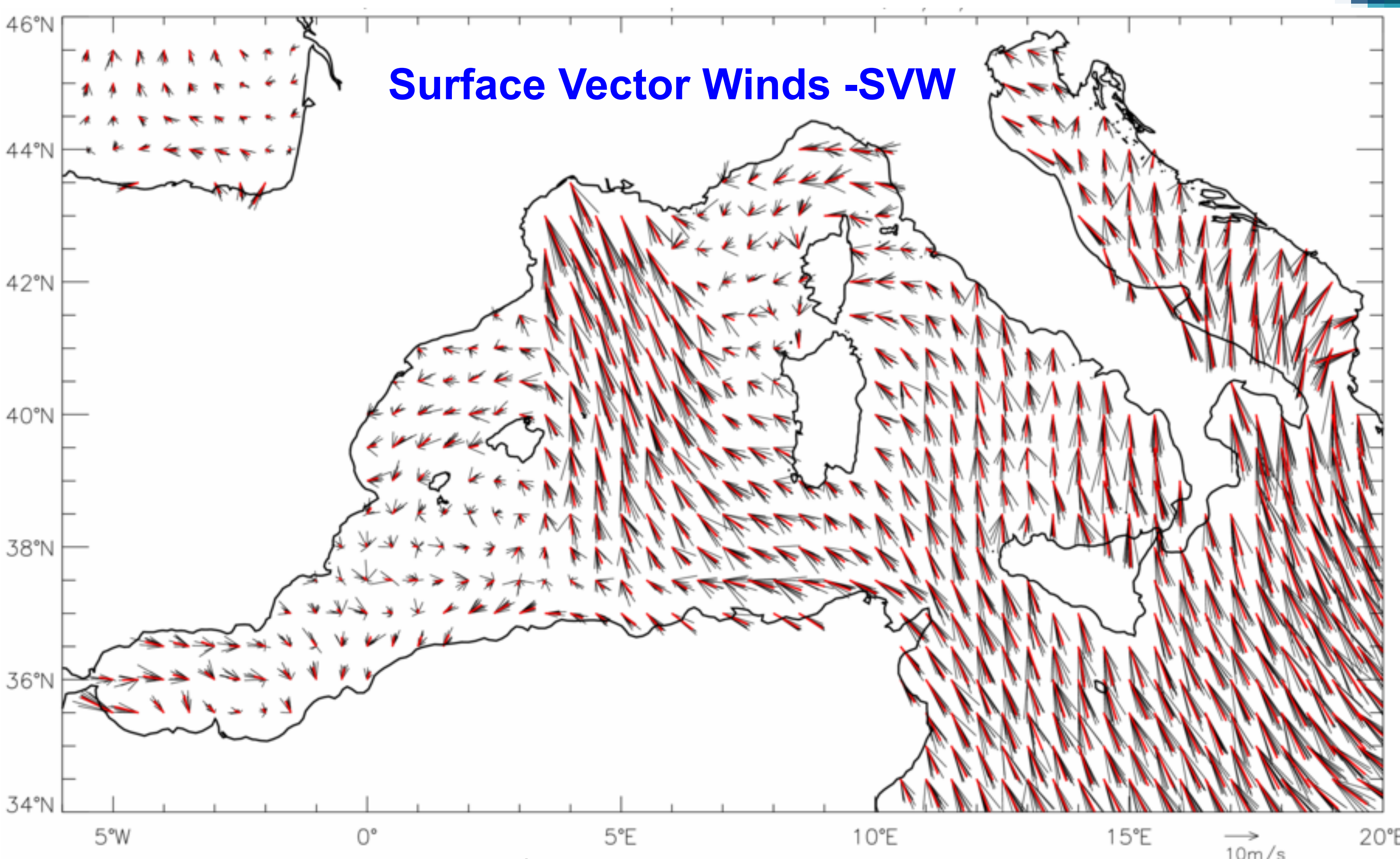
$$\begin{aligned} U_t &= \theta_{uy} D_y P_t + \theta_{ux} D_x P_t + \epsilon_u \\ V_t &= \theta_{vx} D_x P_t + \theta_{vy} D_y P_t + \epsilon_v \end{aligned}$$

What is the uncertainty in the winds ? (Milliff et al., 2011)



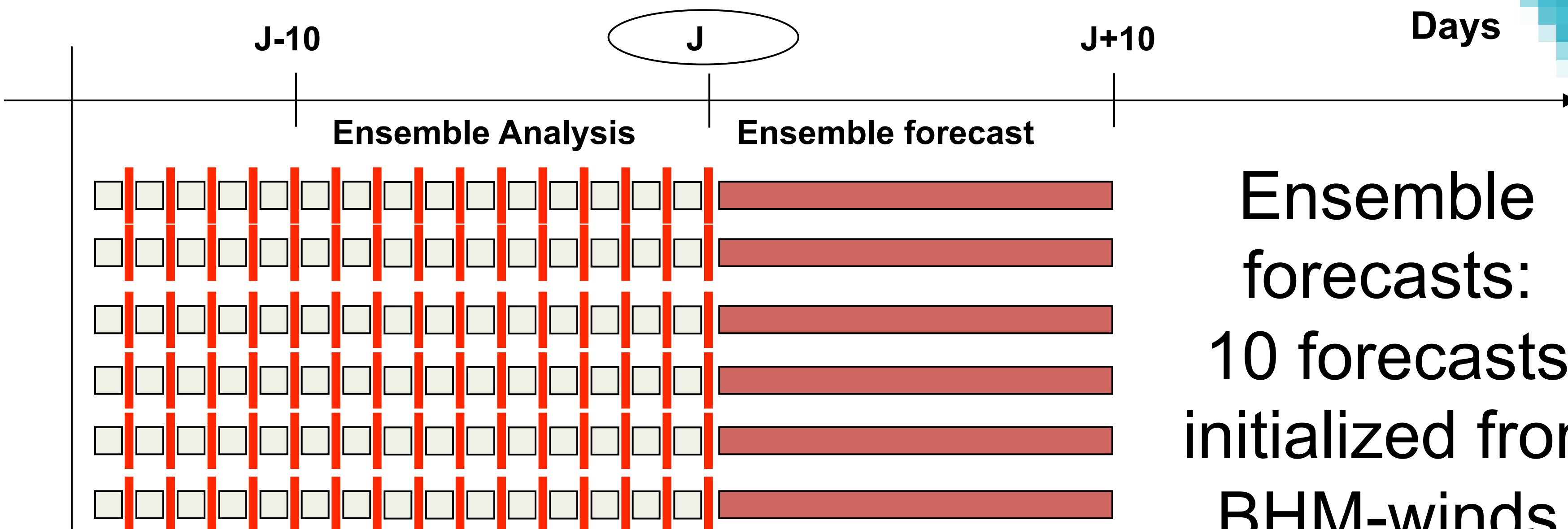


Posterior distributions of winds from a Bayesian Hierarchical Model (Milliff et al., 2011)



10 realizations for February, 2 2005 at 18:00

The Ocean Ensemble Forecast with BHM winds (Pinardi et al., 2011)

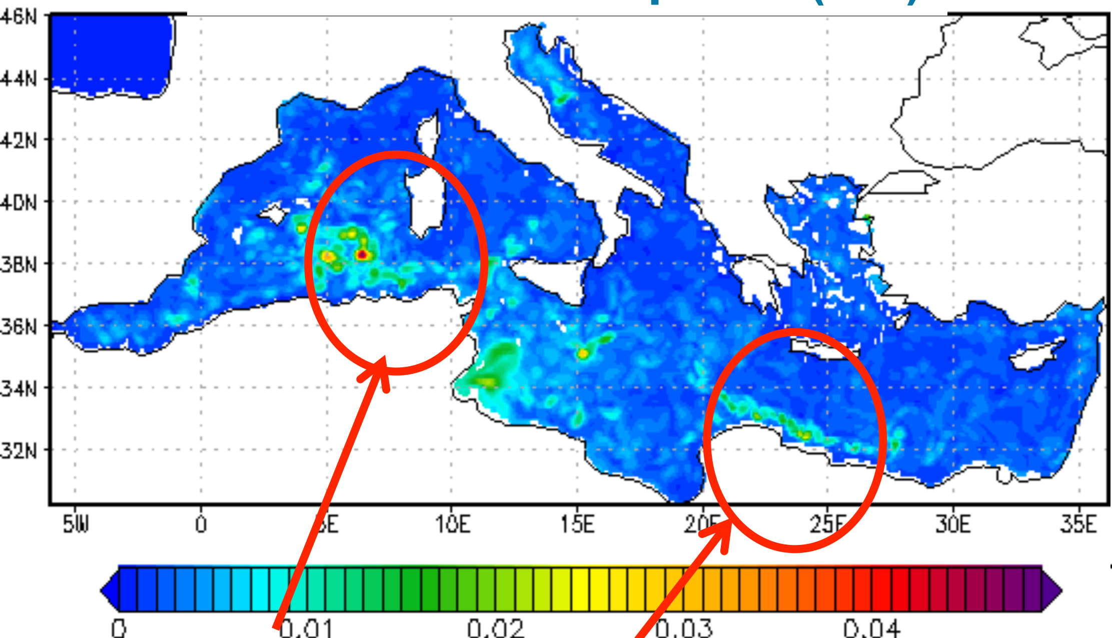


Ensemble forecasts: 10 forecasts initialized from BHM-winds forced initial conditions

Each ensemble member initial condition is produced by assimilating data and force with BHM-SVW realizations

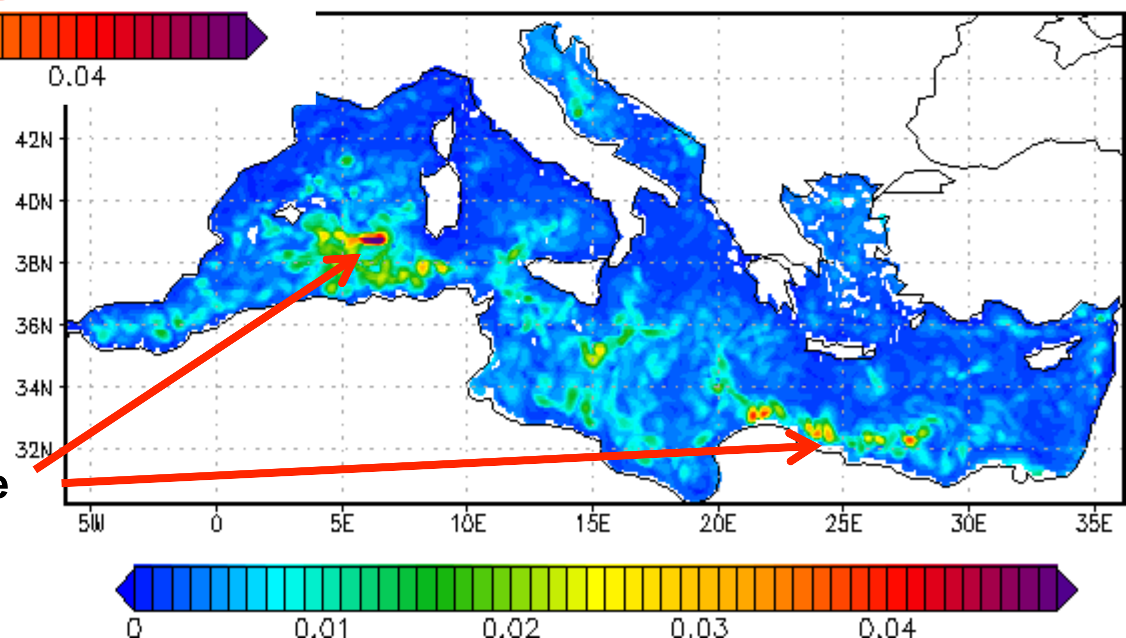
Ensemble forecast: initial condition and last forecast day spread

Initial condition spread (std)



Sea Surface Height

10-th fcst day spread (std)

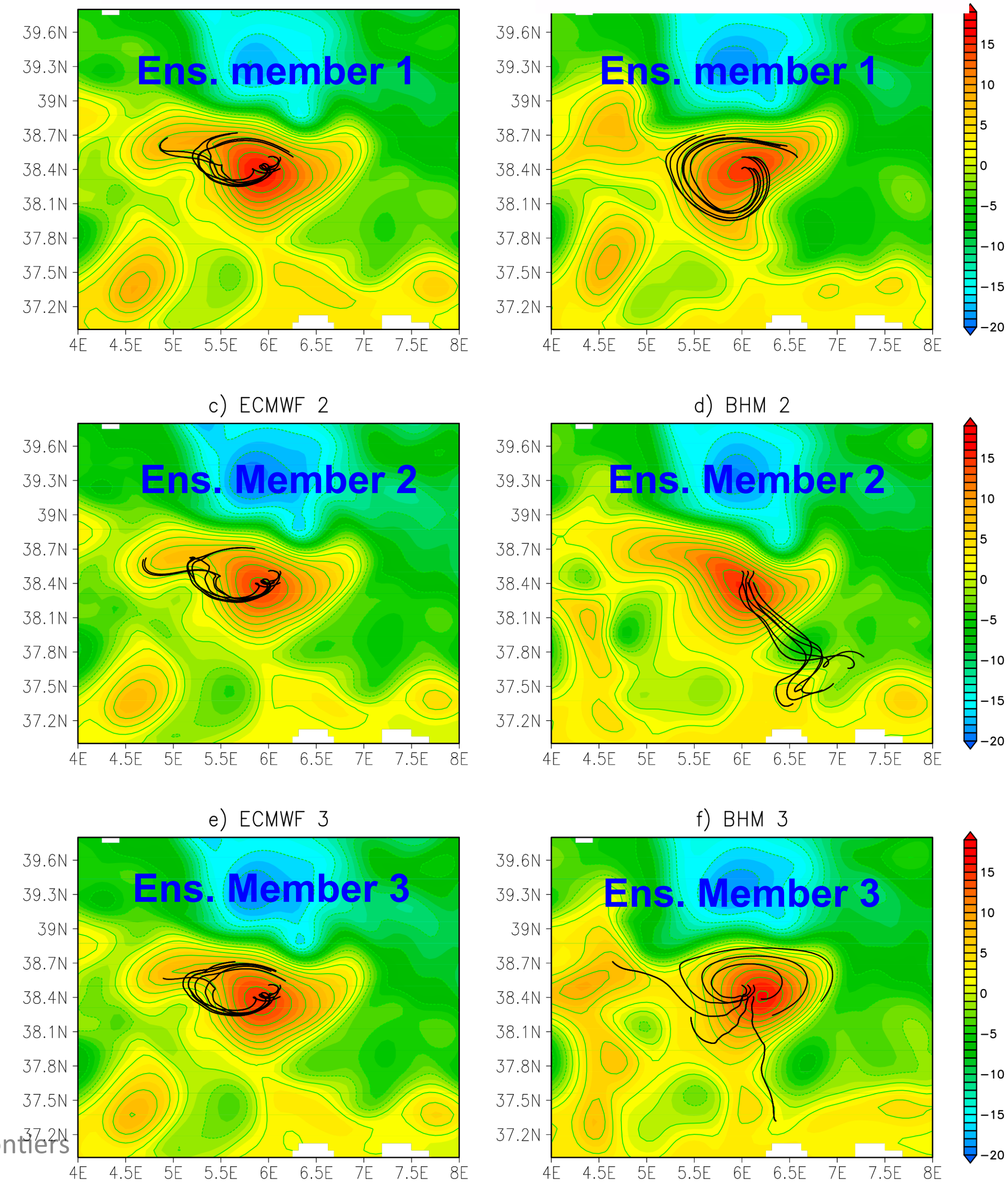


Uncertainty is concentrated at the mesoscales. Sea level spread is comparable to observed sea level error

Uncertainty is amplified during the 10 days of forecast

The forecast spread at 10F

ECMWF-EPS forced ensemble | **BHM-SVW ensemble**



ECMWF Ensemble Prediction System (EPS) forcing is not effective to produce flow field changes at the mesoscales

Intermediate conclusions

- Temperature and salinity vertical error structure is largely connected to uncertainties in wind forcing
- Sea surface height error largely connected to mesoscale eddies position and strength
- Background error covariance should be obtained by perturbing winds in addition to simple random-like perturbation as in traditional EKF literature (Evensen, 2003)

High frequency error covariance matrix estimates with BHM (Dobricic et al., QJRMS, 2015)

- Estimate with a Bayesian Hierarchical Model (BHM) the time varying vertical error covariance matrix C by using misfits (d) and model stand. dev. (q) for T,S
- To estimate the error covariance we use a Bayesian Hierarchical Model (BHM) approach:
  - Data stage model
  - Process model
  - Parameter models

High frequency error covariance matrix estimates with BHM

The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

- Data stage: $$q_t \mid e_t \sim N\left(H_{qt} e_t, \Sigma_{qt}\right)$$ $$d_t \mid e_t \sim N\left(H_{dt} e_t, \Sigma_{dt}\right)$$
- Process model: the vertical structure is given by the seasonal vertical EOFs but we estimate with an AR model 5-days amplitudes (Beta)

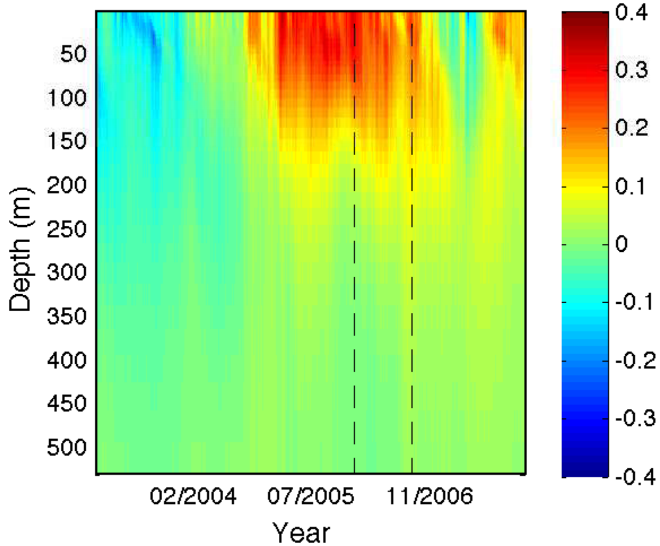
$$e_t = V_{t_s} \beta_t + \eta_t \ldots; \quad \eta_t \sim N\left(0, \tau_t I\right) \quad \beta_t \approx N\left(0, \Lambda_t \Gamma_t \Gamma_t^T \Lambda_t\right)$$

- Finally we can write B_{V_t} as:

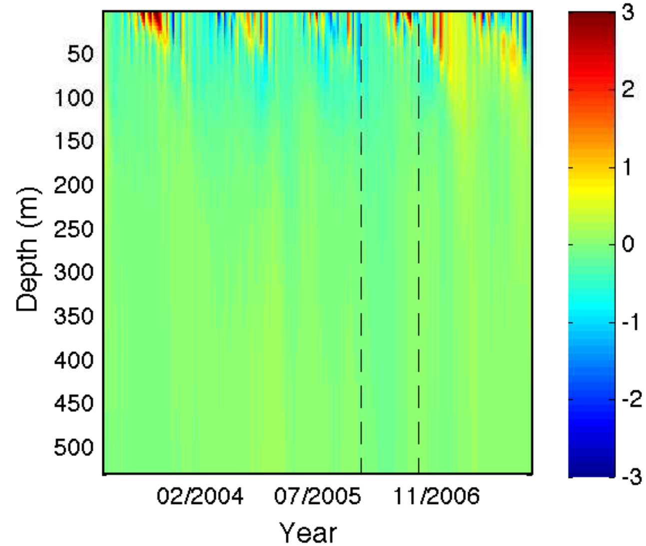
$$B_{Vt} = V_{ts} \Lambda_t \Gamma_t \Gamma_t^T \Lambda_t V_{ts} + \tau_t I$$

The data stage sets

q- Salinity anomalies

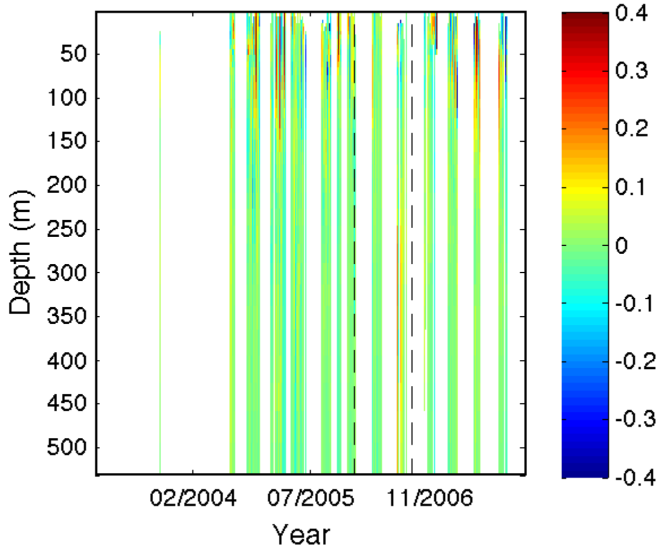


q- Temperature anomalies

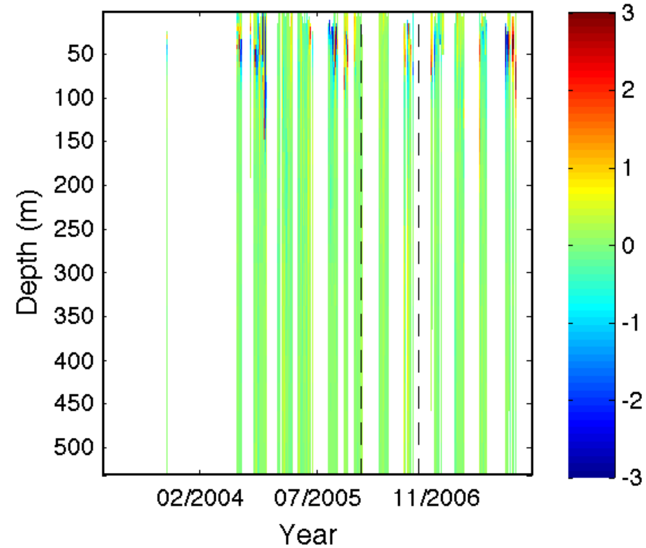


Model anomalies vertical structure

d- Salinity anomalies

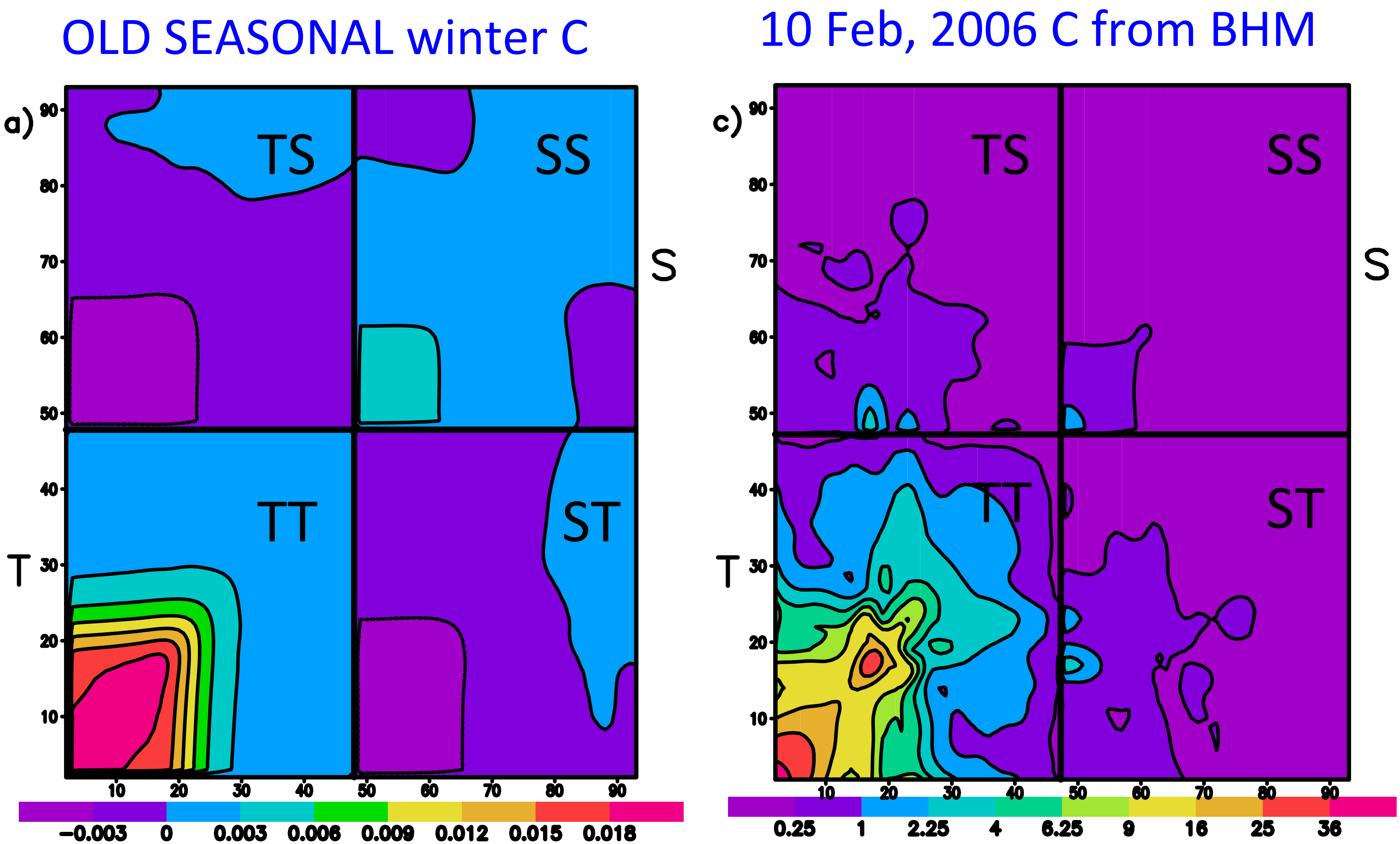


d- Temperature anomalies

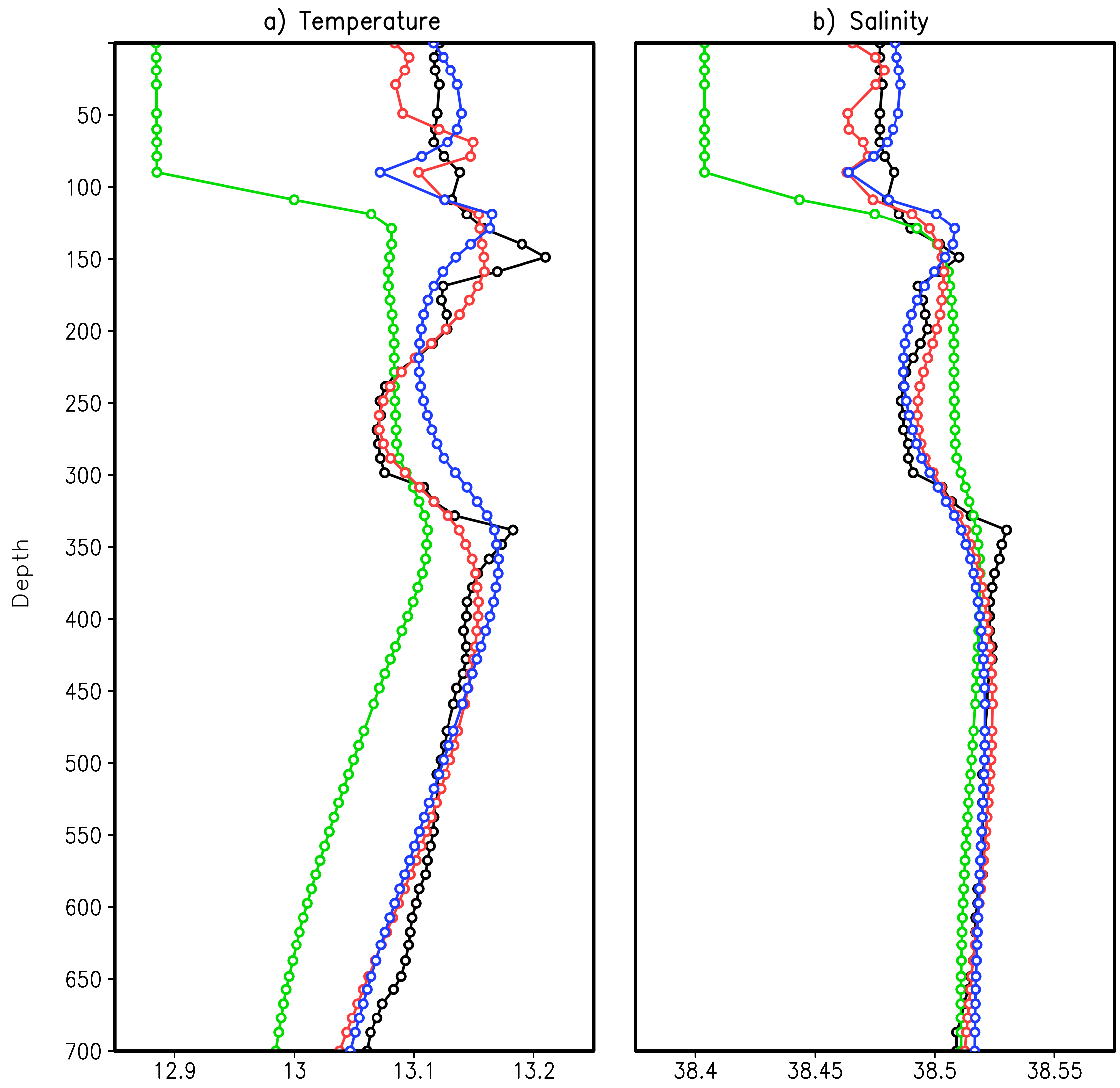


Misfit vertical structure

The high frequency error covariance matrix



Improvements on the assimilation due to high frequency error covariance



Green background
Black observation
Blue old method
Red BHM method

Conclusions

- An operational ocean 3DVAR assimilation system has been used to study analysis errors, forecast errors and different choices of background error covariance matrices
- Model improvements still provide the major source of improvements for analyses
- Errors in Temp and salinity peak between the 20-100 meter layer and vary seasonally mainly due to atmospheric forcing errors
- Ocean Ensemble forecasting with BHM winds offer a way to quantify the short term forecasting uncertainties
- High frequency background vertical covariance matrix can be constructed from model variance information and misfits and it improves the model analyses