Variational assimilation of satellite and in situ observations in the regional Copernicus Services

N.Pinardi, University of Bologna

Contributors:

E.Clementi, J.Pistoia, A. Grandi, M. Drudi, R. Escudier, S. Masina

CMCC-Bologna

S.Ciliberti, B. Lemieux, L. Lima, G. Coppini

CMCC-Lecce

Summary of presentation

- The forecast uncertainty conundrum: statement of the problem
- The CMCC Ocean 3DVar system in the Mediterranean and Black Seas, regional systems of Copernicus
- Analysis and forecasts error structures
- The ensemble approach to understand uncertainty and advance data assimilation
- Outlook

Uncertainty in ocean forecasting: statement of the problem

- The aim is to have accurate forecasts as long as possible into the future. Limits to predictability are due to:
  - Inaccurate model representation of processes and model numerics;
  - Uncertainty in lateral fluxes, especially hydrology;
  - Inaccurate knowledge of initial conditions for predictive variables;
  - Atmospheric forcing uncertainties.
- Inaccurate model representation of processes and lateral fluxes, such as hydrology, are still the major source of uncertainty in ocean forecasting
- We concentrate here on the initial condition and the atmospheric forcing inaccuracies, hoping that models will come along and include better representation of ocean/shelf/coastal processes.

Uncertainty in ocean forecasting: initial condition uncertainty

- Data assimilation is the algorithm/methodology to reduce the initial condition uncertainties
- It started with objective analysis (70s-80s) and evolved thereafter to include Kalman Filters, 3Dvar/4Dvar and adjoint schemes. Practical applications of the theory force the assumption of Gaussian statistics and reduce the problem to a Least Square minimization problem.
- In Meteorology and Oceanography we have two independent starting estimates of the state of the system, the model and the observations
- Gauss, 1809 : … since our observations are nothing more than approximation to the truth,…. we need … a suitable combination of all observations and theory to approximate as much as possible the truth.
- Thus we can frame 'data assimilation' as a 'progressive refinement methodology' to obtain the best estimate of the present and past state of the ocean

The CMCC ocean 3DVAR scheme

(Dobricic and Pinardi, 2008, Storto et al., 2011)

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

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

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

misfit

In our system, the oceanic vector state is defined as:

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

The background error covariance matrix is defined as:

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

The CMCC ocean 3DVAR scheme

(Dobricic and Pinardi, 2008, Storto et al., 2015)

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

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

$$\mathbf{V} = \mathbf{V}_\eta \mathbf{V}_\mathbf{H} \mathbf{V}_\mathbf{V}^{\mathbf{t_S}}$$

$\mathbf{V}_\mathbf{V}^{\mathbf{t_S}}$ Vertical EOFs: bi-variate T-S for BS and tri-variate eta-T-S for Med.

"ts" is monthly

$\mathbf{V}_\eta$ - Dynamic Height operator (1000 m level of no motion)

$\mathbf{V}_H$ - Horizontal covariance (recursive filter)

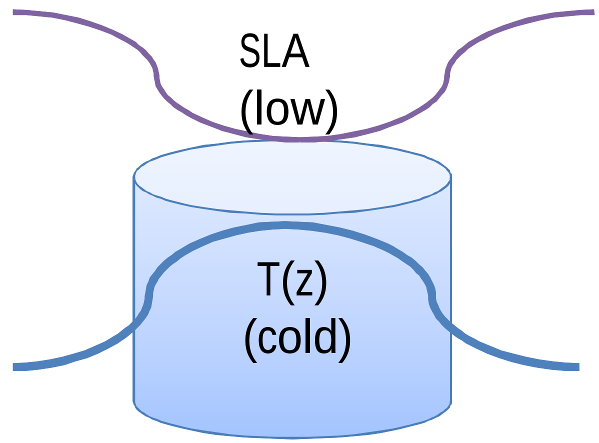
Key issues:

the vertical background error covariance matrix and the observational error covariance

The background vertical error covariance: the sea level anomaly problem

- We suppose that sea level represents modifications of the thermal and salinity structure of the water column

Cyclonic vortex: the SLA dnd T(z) covariance



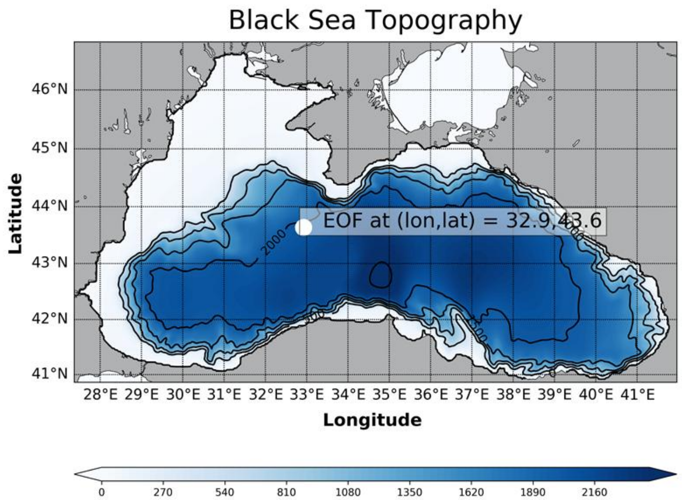
- Thus background error covariance should contain the “correct” vertical T,S correlation (water mass structure) correlated to surface sea level

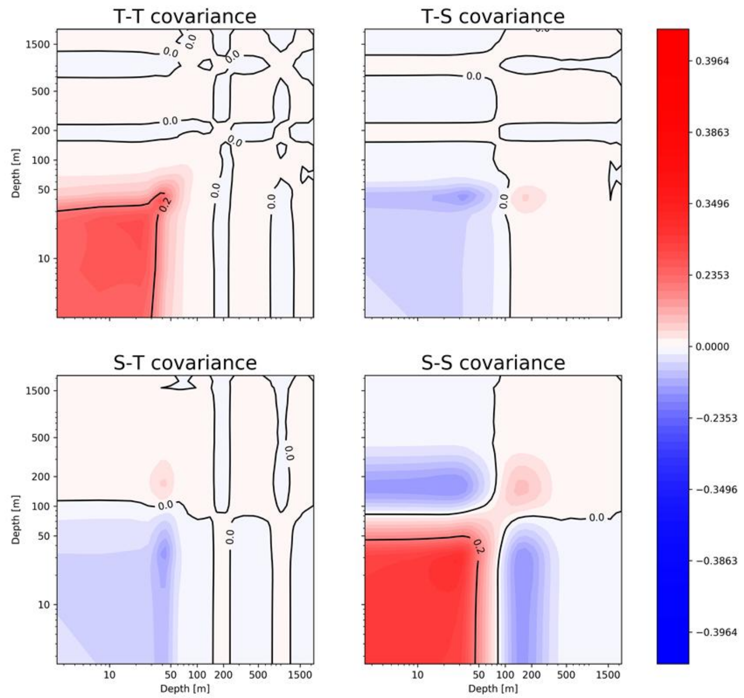
The background vertical error covariance: the example from the Black Sea

At each model grid point, a covariance matrix is built from Temperature, Salinity profile anomalies from a long model simulation

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

One C for each grid point:





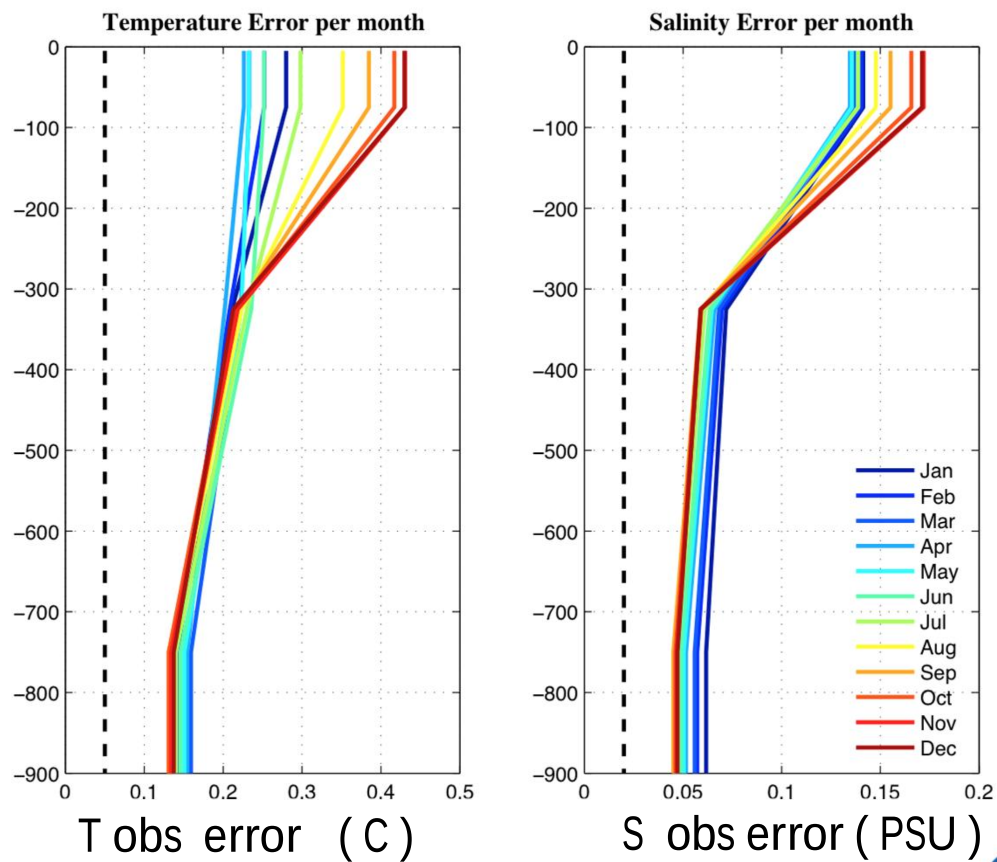
To filter the covariance, EOFs are computed at each grid point

The observational error covariance: example from the Mediterranean Sea

- Here the problem is to account for representativeness errors of the observations

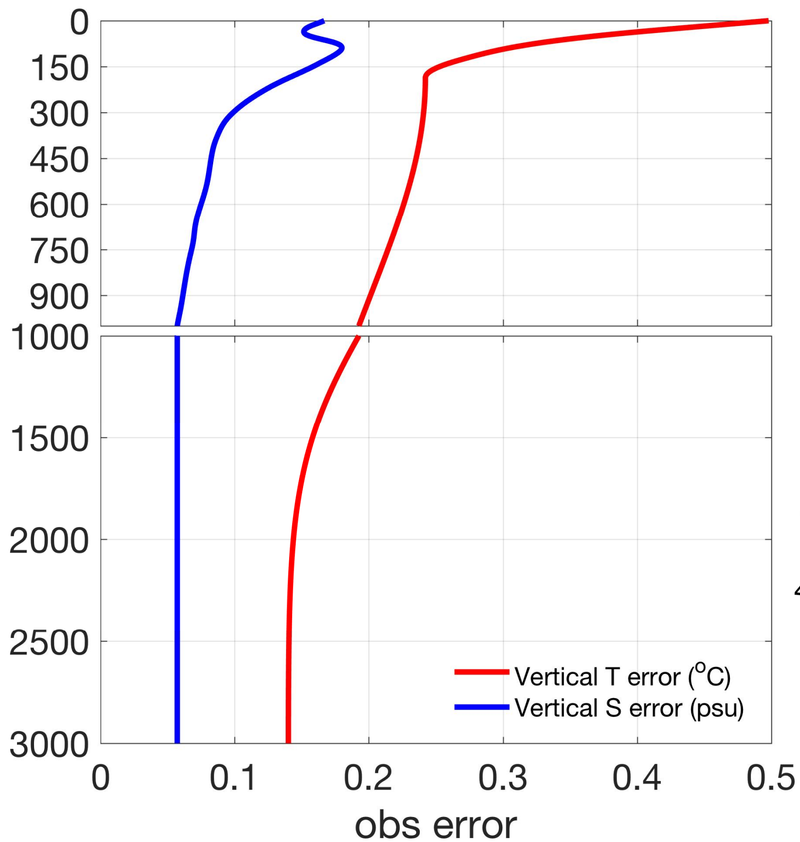
Desroziers (2005) designed a method to find the best representativeness error after first analysis is done for several years

$$E\left(d_b^o d_b^{oT}\right) = R + \alpha HBH^T$$

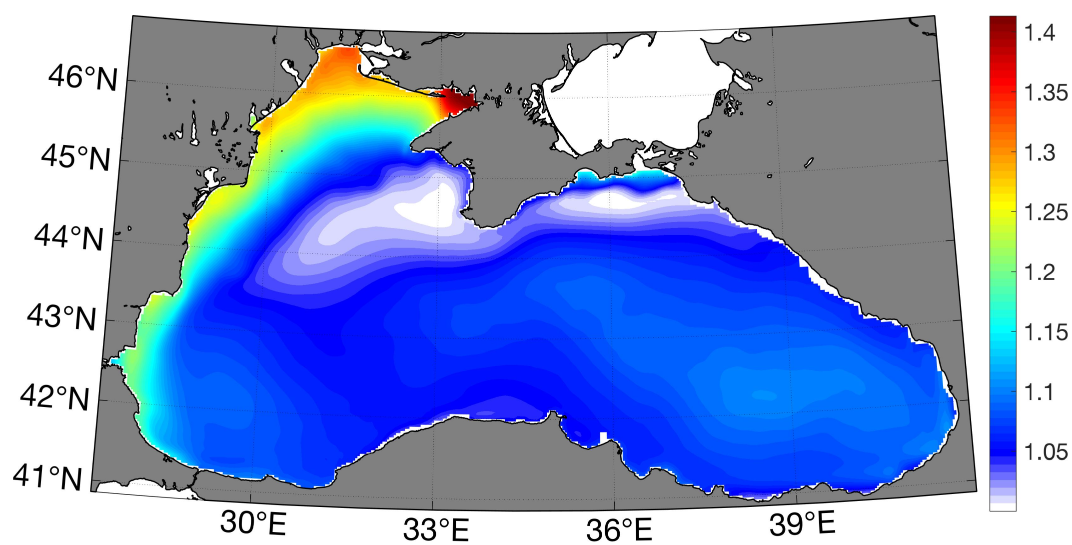


The observational error covariance: example from the Black Sea

Vertical error distribution in the BS-PHY system

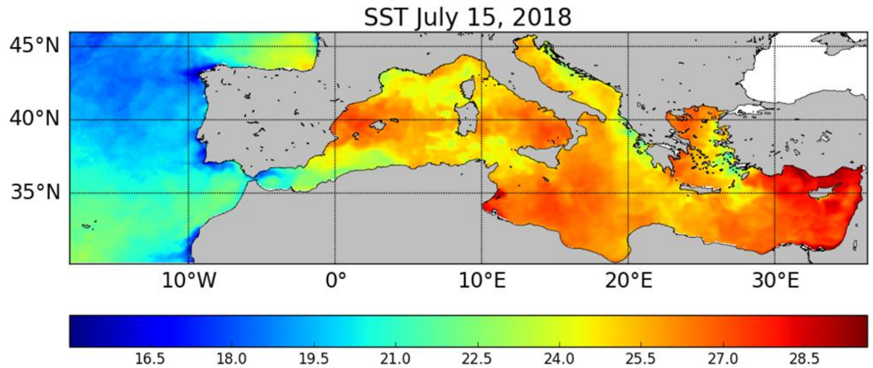


Horizontal distribution of representativeness error

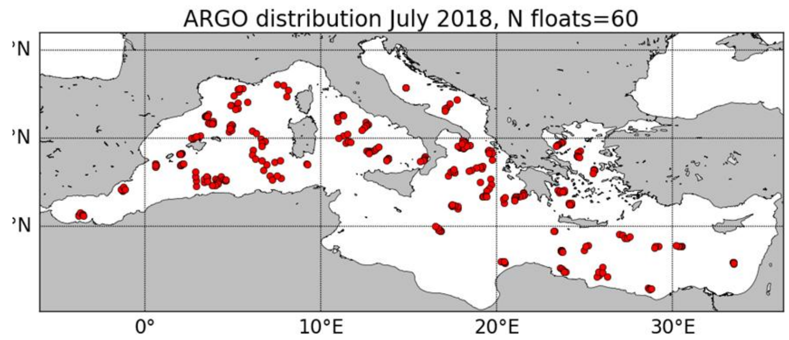


Real time observations for assimilation in the Mediterranean Sea

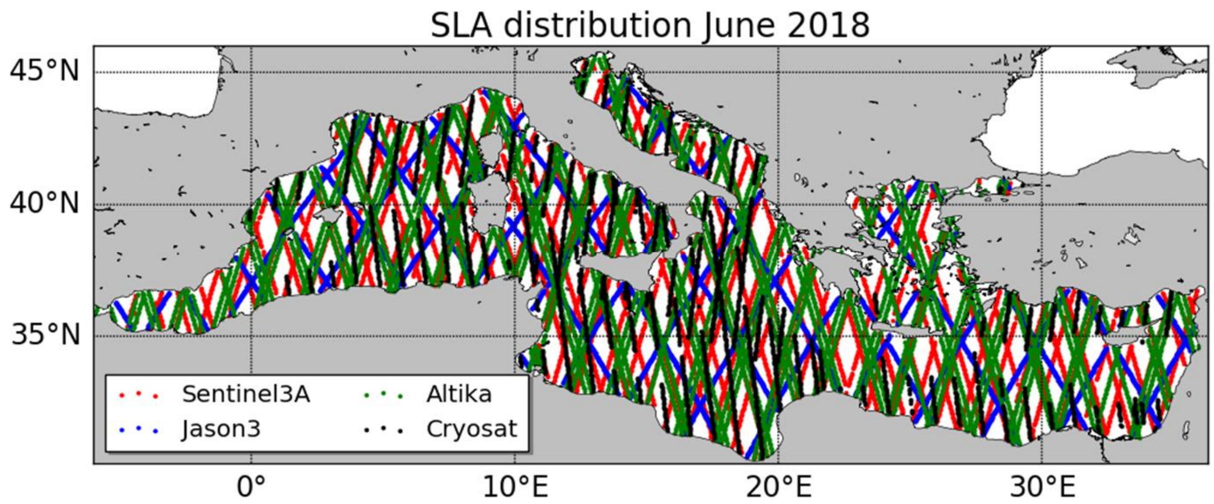
Multi-sensor daily L4 SST



ARGO (July 2018)



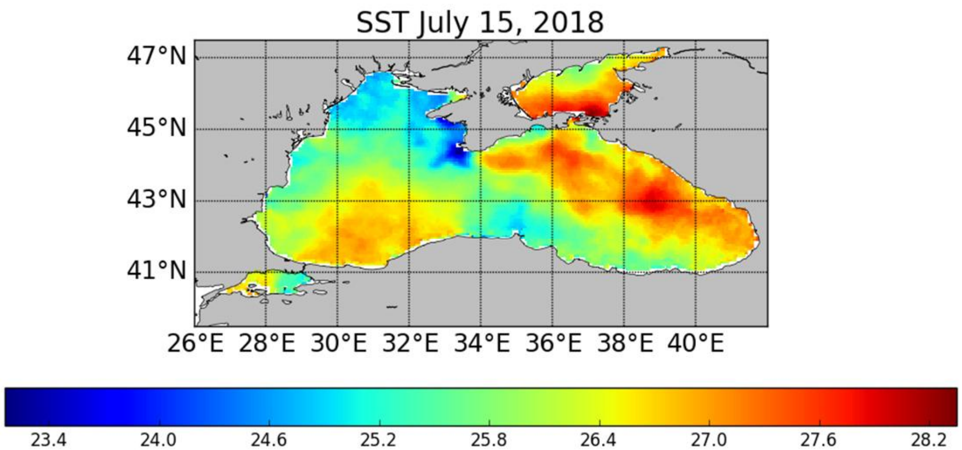
Multisatellite along track sea level



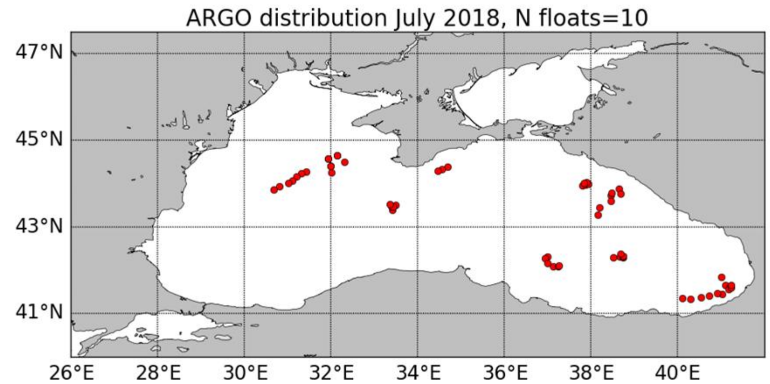
Time period covered : Jul 2018

Real time observations for assimilation in the Black Sea

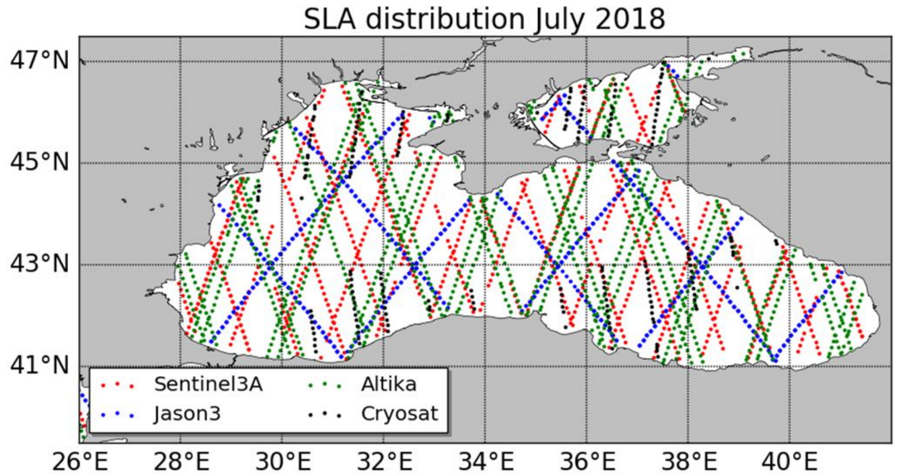
Multi-sensor daily L4 SST



ARGO

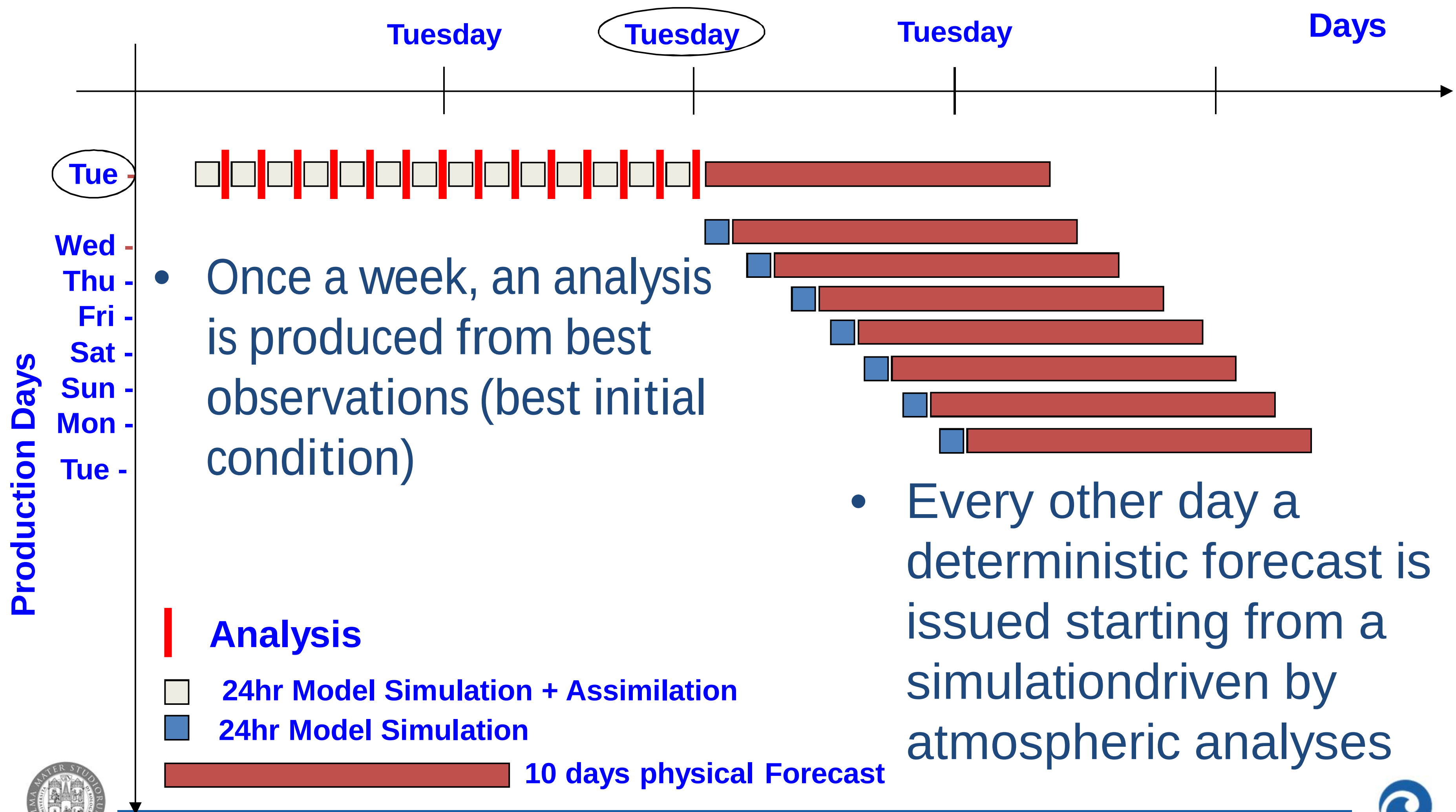


Multisatellite along track sea level



Time period covered : Jul 2018

The Mediterranean analysis/forecast production system



- Once a week, an analysis is produced from best observations (best initial condition)

- Every other day a deterministic forecast is issued starting from a simulationdriven by atmospheric analyses

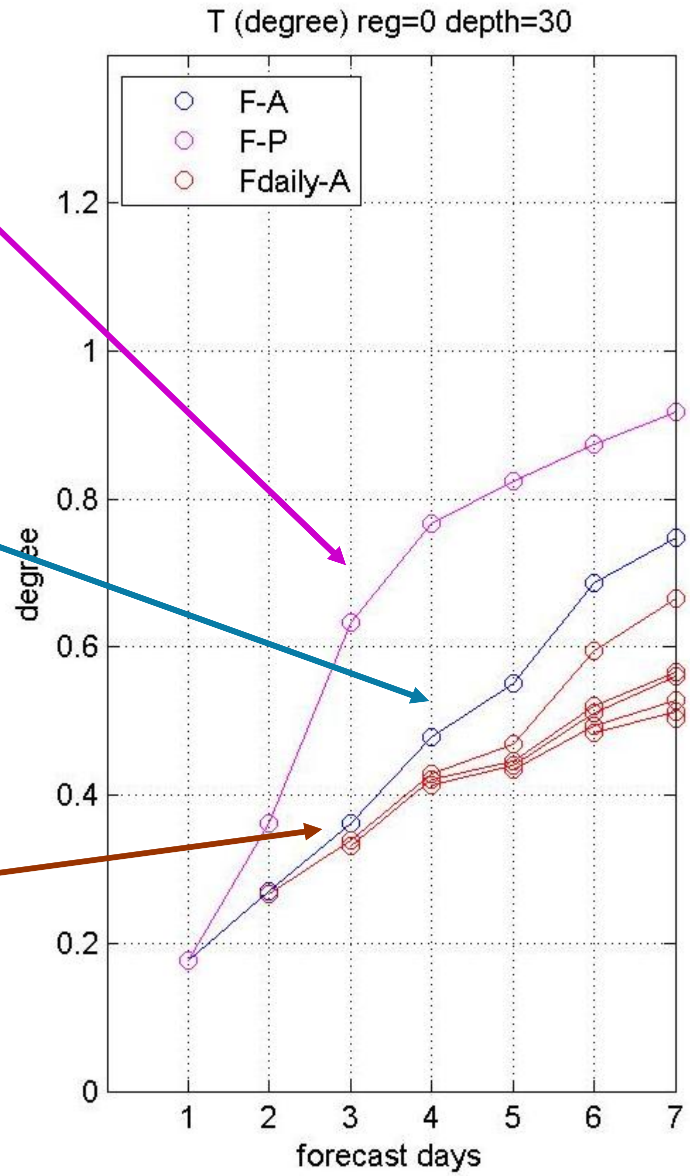
Ocean forecast error at 30 m: the effect of atmospheric forcing errors

Control:

Forecast-persistence

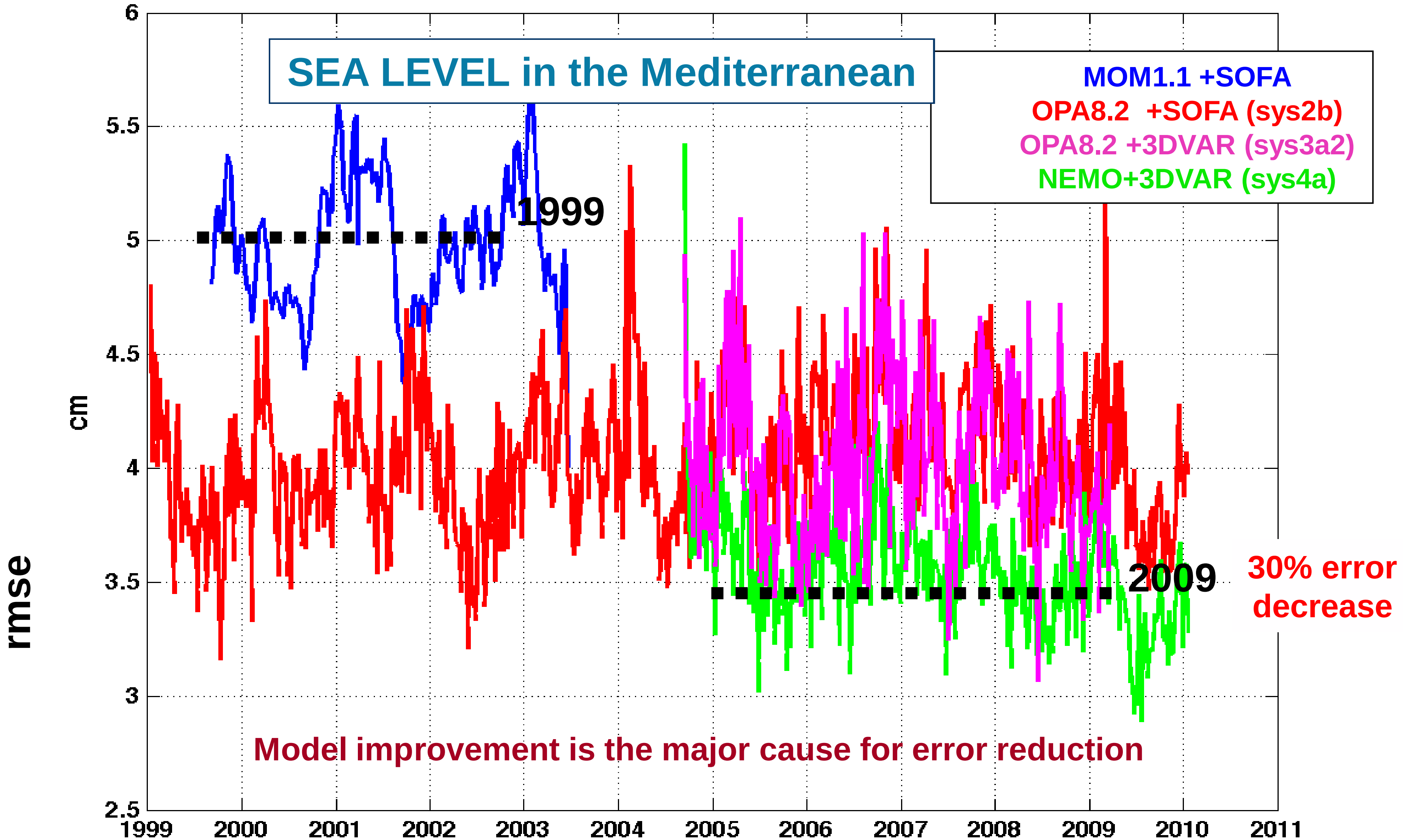
Weekly forecast

Forecasts started from simulations driven by atmospheric analyses



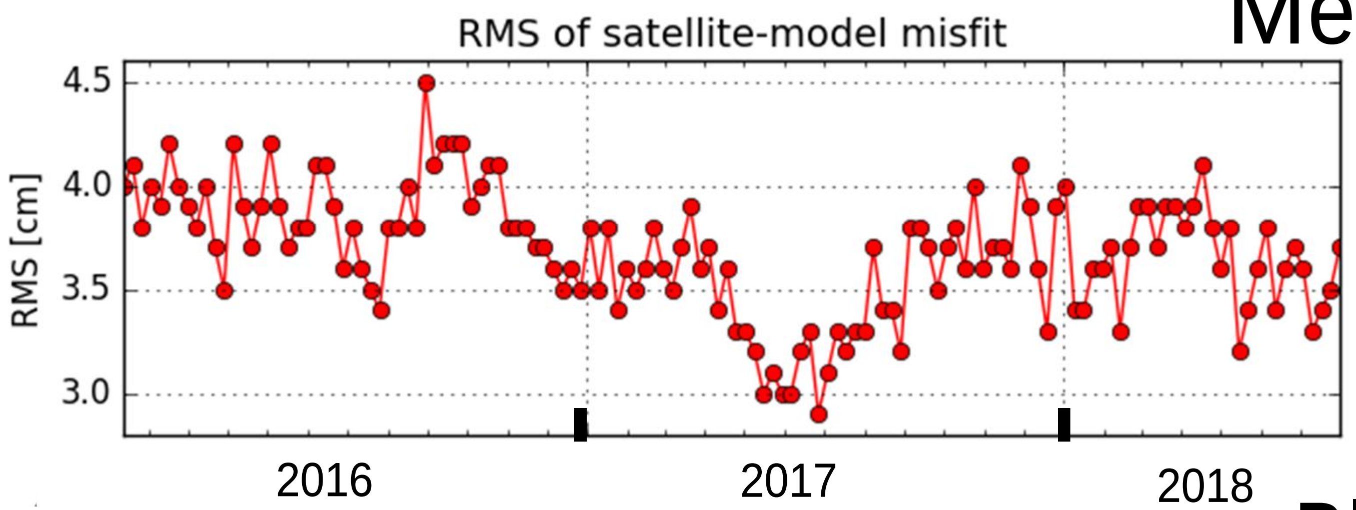
Forecast error at day 7 decreased by 30% only reducing atmospheric forcing error

How did the error decrease in 10 years?

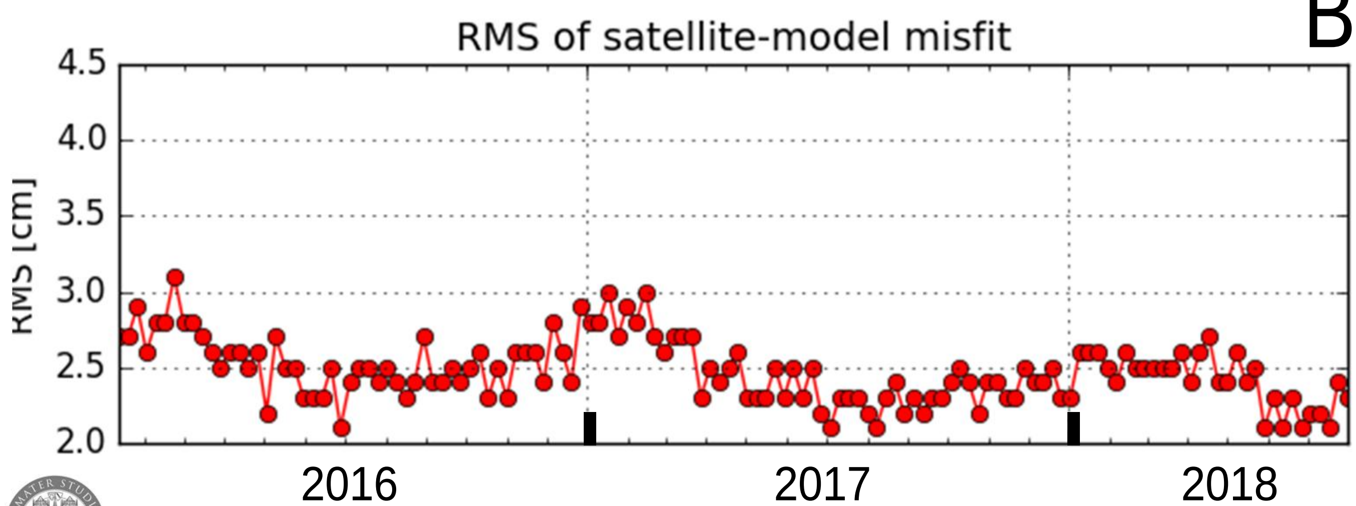


RMSE time history: Sea Level Anomaly

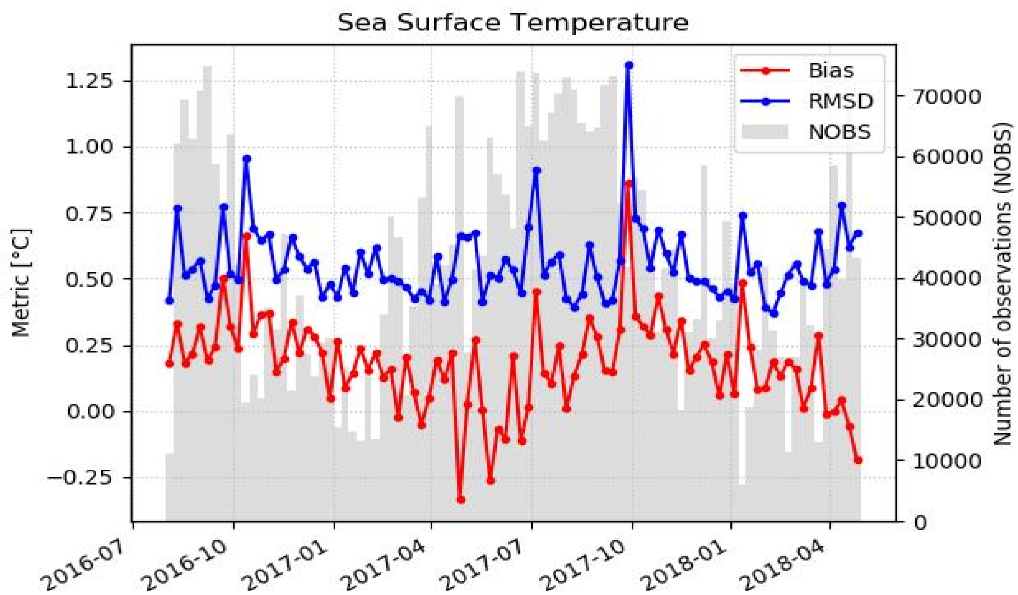
Mediterranean



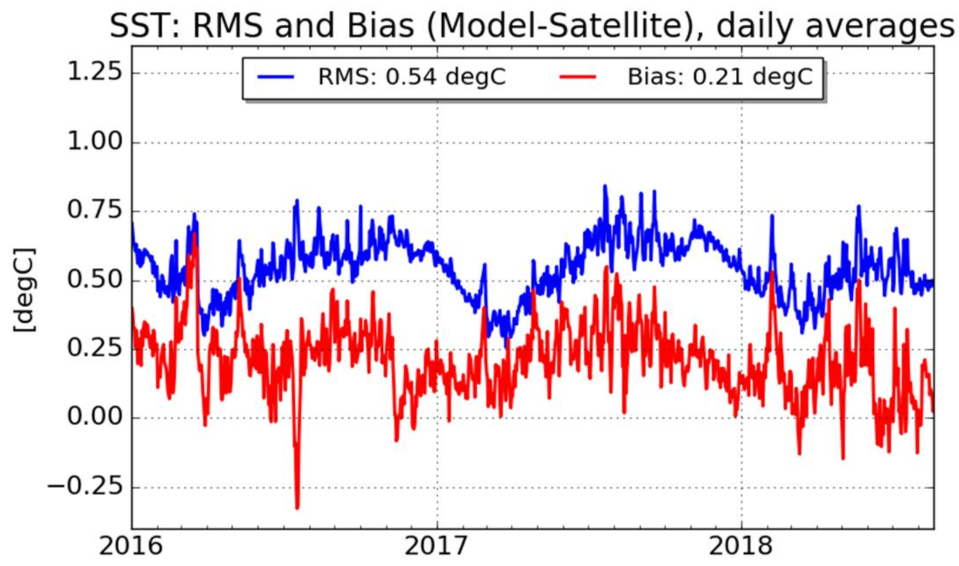
Black Sea



RMSE time history: SST



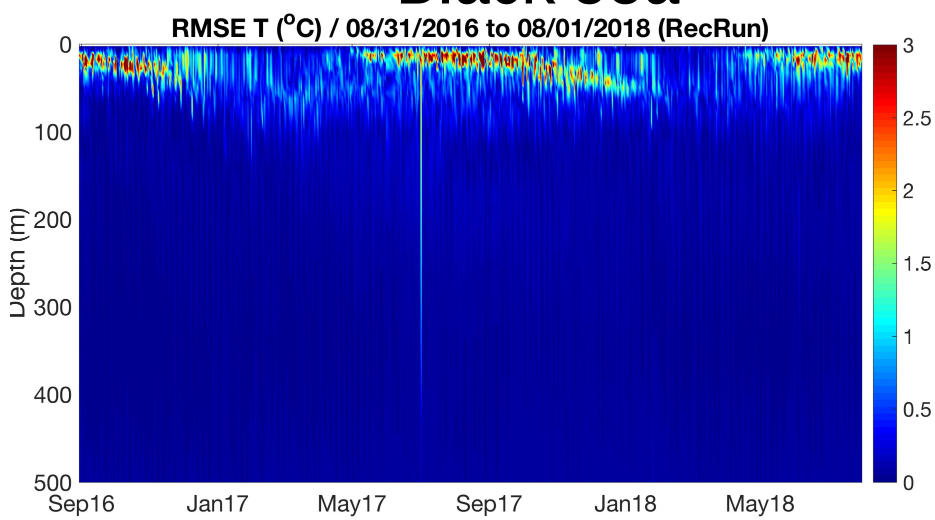
Black Sea



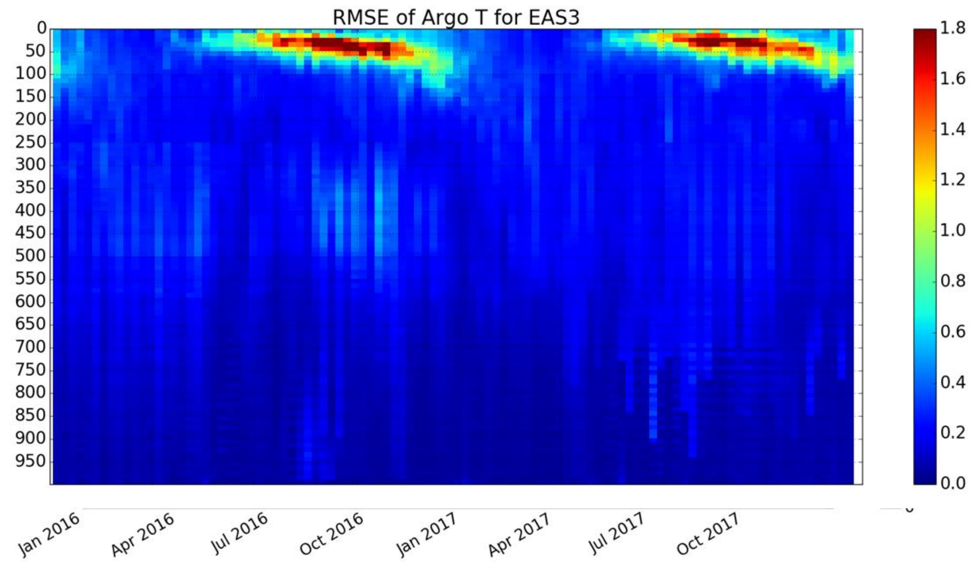
Mediterranean Sea

The vertical structure of the error for the Med and Black Sea

Black Sea



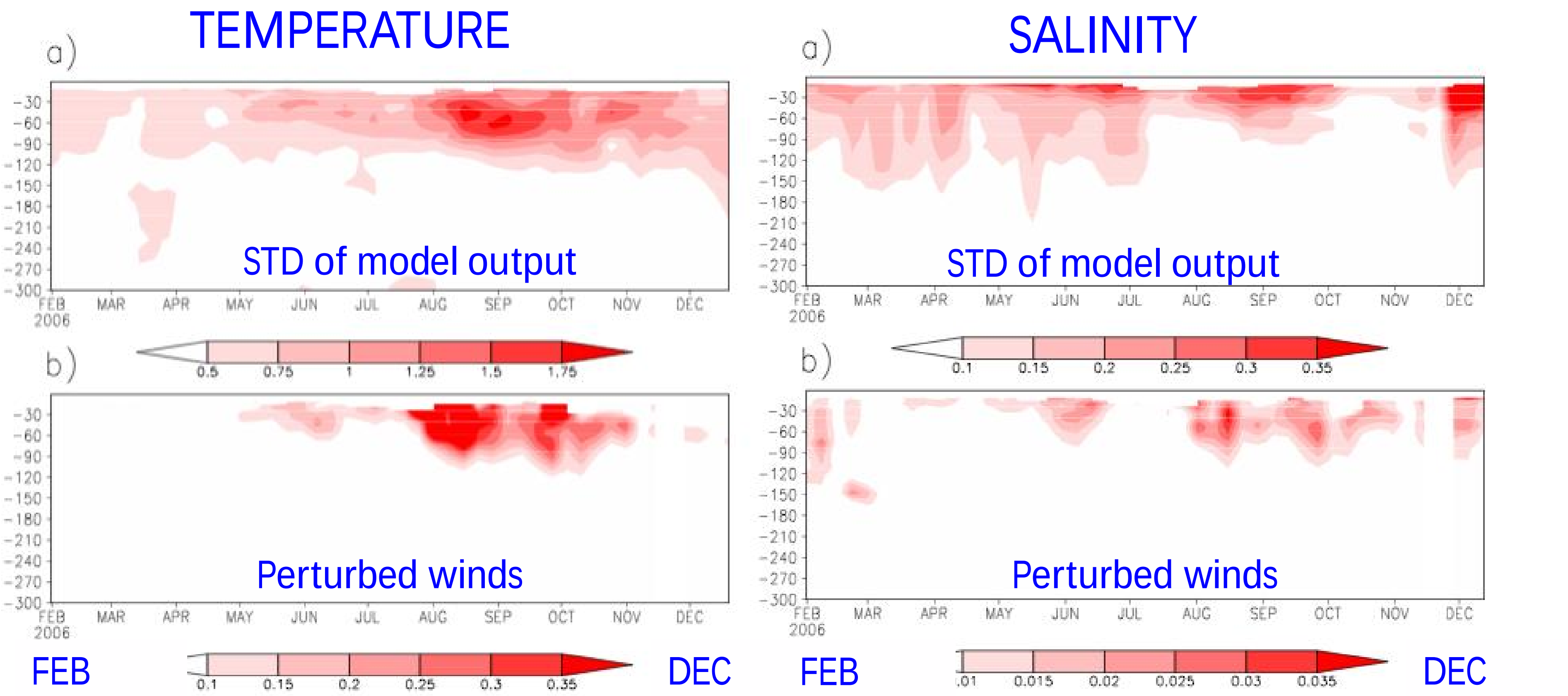
Med Sea



Three major uncertainties are concurring:

1) Surface atmospheric forcing
2) Mixing parameterization (including waves)
3) Nonlocal advection

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



Answer: uncertainty in atmospheric forcing projects on the vertical structure of the temperature & salinity errors

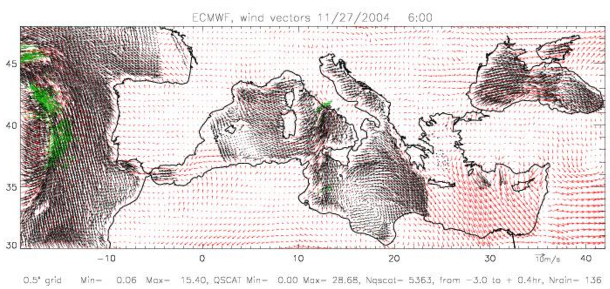
Ensemble methods for ocean forecasting

- Ensemble methods can provide information about background fields uncertainty and thus on how to build the background error covariance matrix
- However, if it is done randomly, large (almost impossible) number of members are required
- Strategy for the ocean: since vertical structure of the error is connected to atmospheric forcing uncertainty, perturb the atmospheric winds (and by consequence momentum, heat and water fluxes)

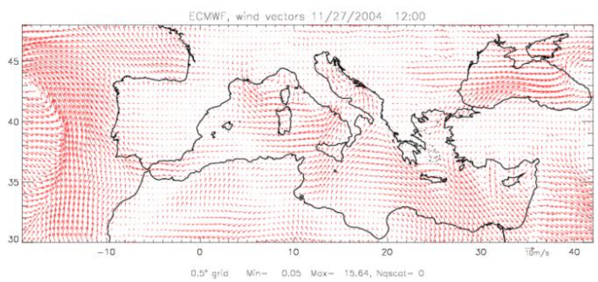
How do we perturb the atmospheric forcing for ocean uncertainties?

- Perturb winds. Method explored: Bayesian Hierarchical Model (BHM)
- Conceptual and implementation blocks:

Data Stage: 2 types of data
Scatterometer winds and ECMWF analyses/forecasts



QSCAT
ASCAT



ECMWF

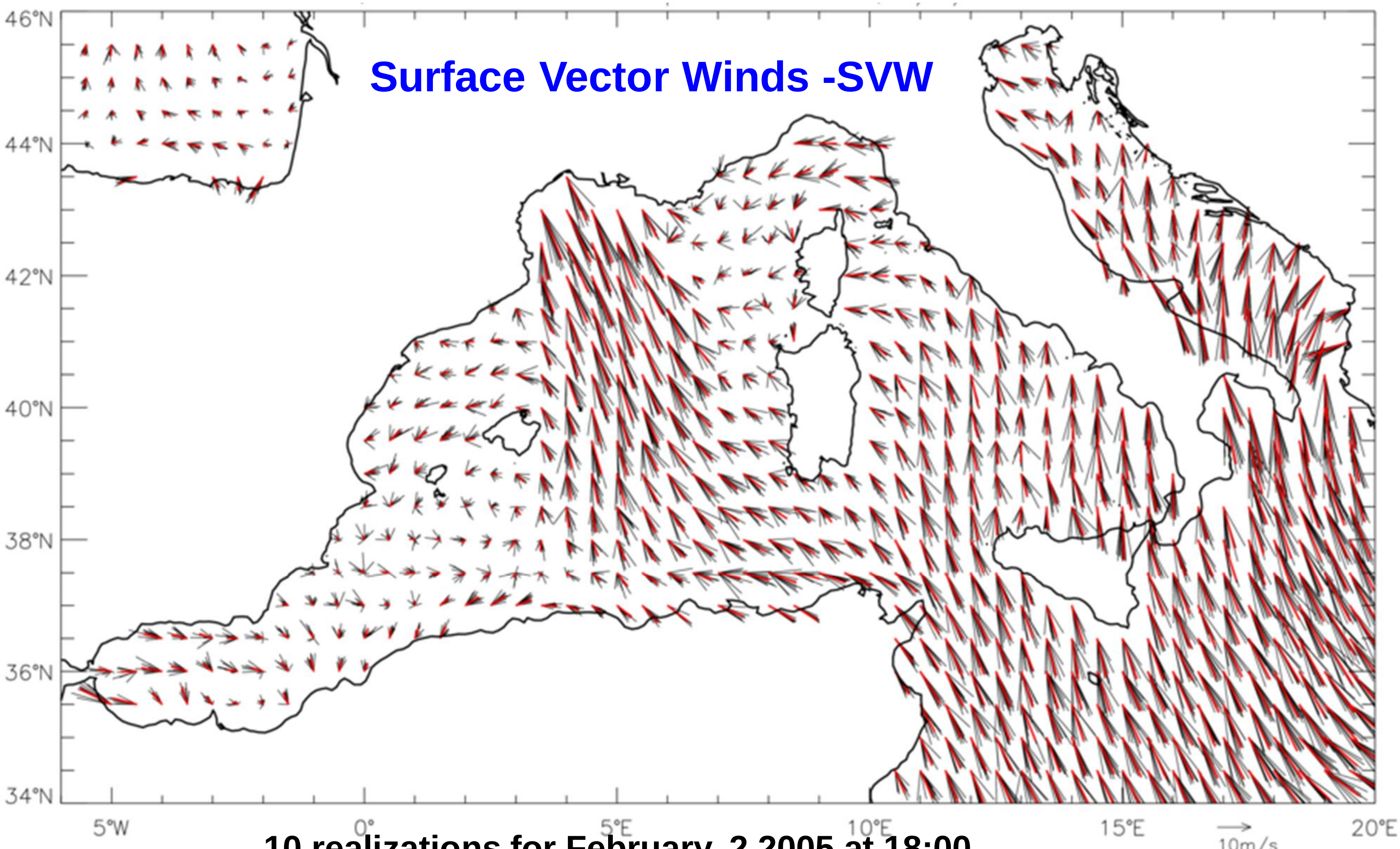
Process model stage:
Raylegh friction surface model translated into a stochastic finite difference equation

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

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

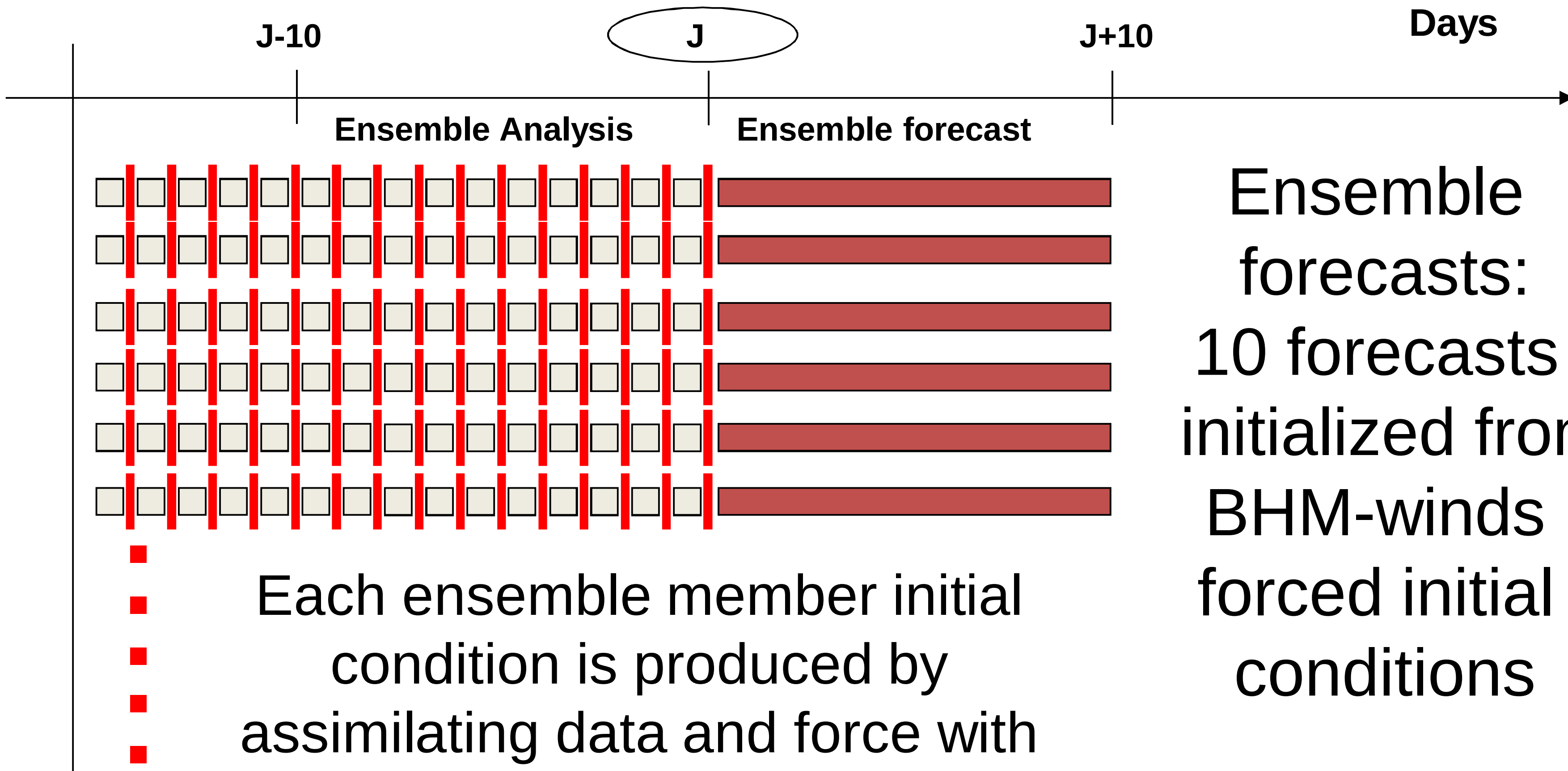
$$\begin{aligned} U_t &= \theta_{uy} D_y P_t + \theta_{ux} D_x P_t + \epsilon_u \\ V_t &= \theta_{vx} D_x P_t + \theta_{vy} D_y P_t + \epsilon_v \end{aligned}$$

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)

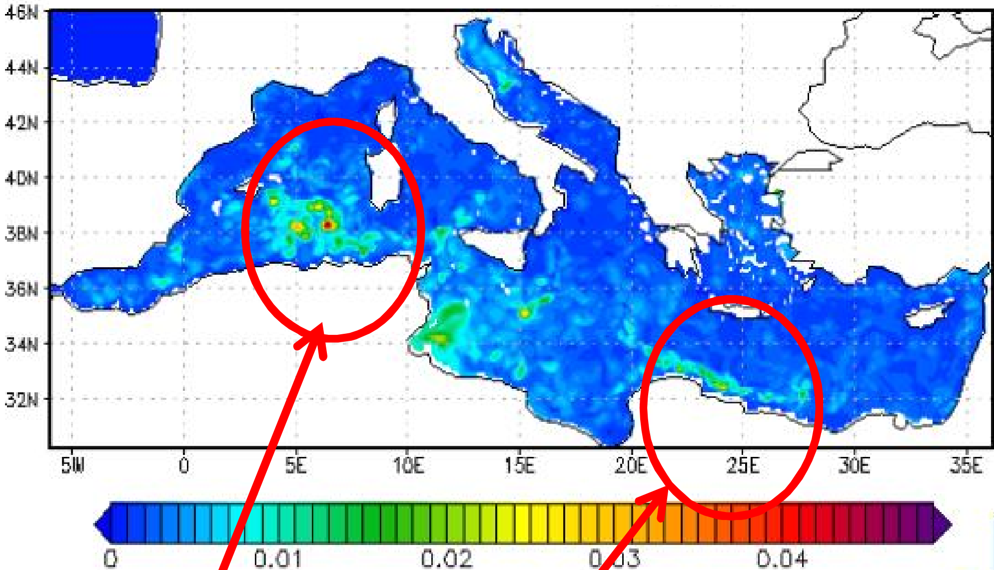


Ensemble forecasts: 10 forecasts initialized from BHM-winds forced initial conditions

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

The Ocean Ensemble Forecast with BHM winds: the spread

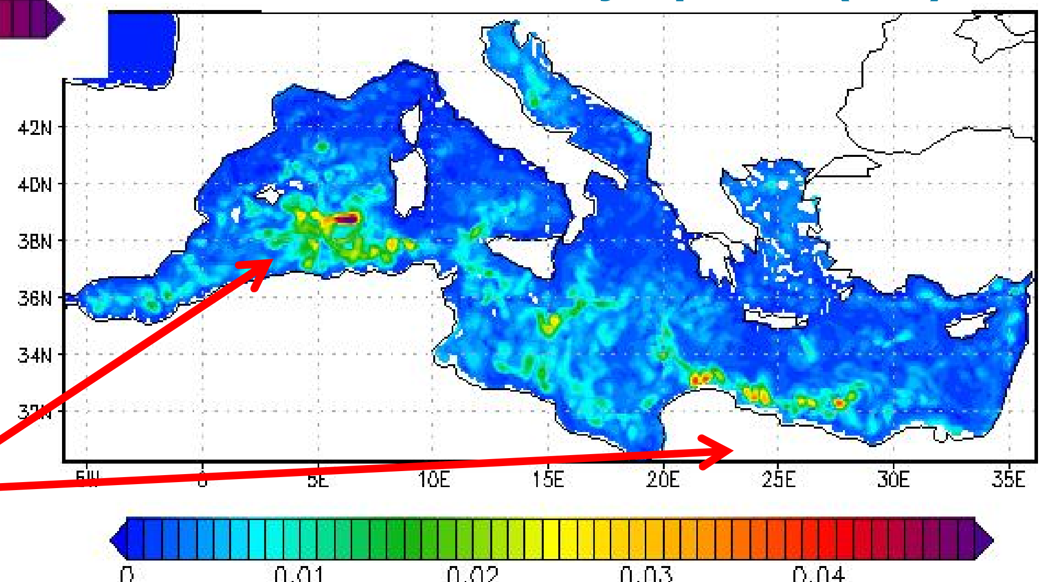
Initial condition spread (std)



Sea Surface Height

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

10-th fcst day spread (std)



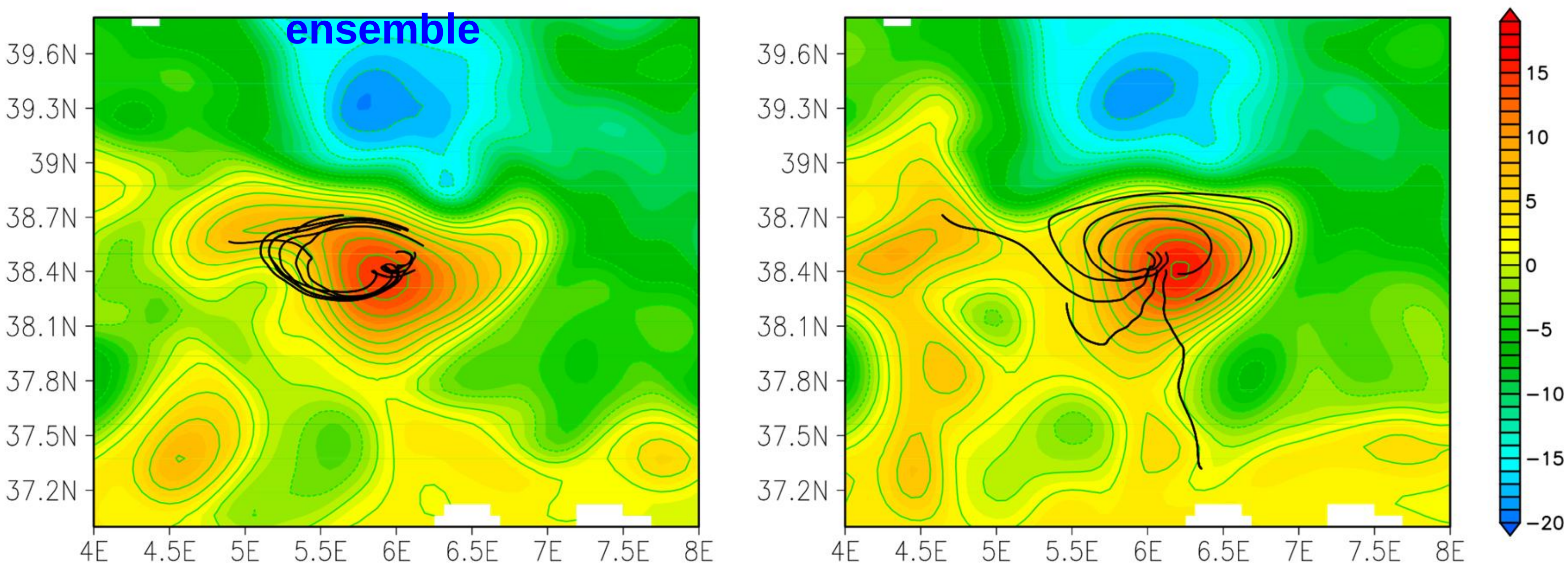
Uncertainty is amplified during the 10 days of forecast

The Ocean Ensemble Forecast with BHM winds: how does it compare with ECMWF-EPS ensemble members?

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 ocean mesoscales

How to advance?

- Assimilate more observational data (gliders, HF radar obs, wave obs for example)
- Better quantify the background error covariance matrix
  - Use ensemble perturbation methods that depends on atmospheric forcing uncertainties
  - use hybrid error covariance matrix estimates to capture the multi-scale character of the error covariance matrix

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

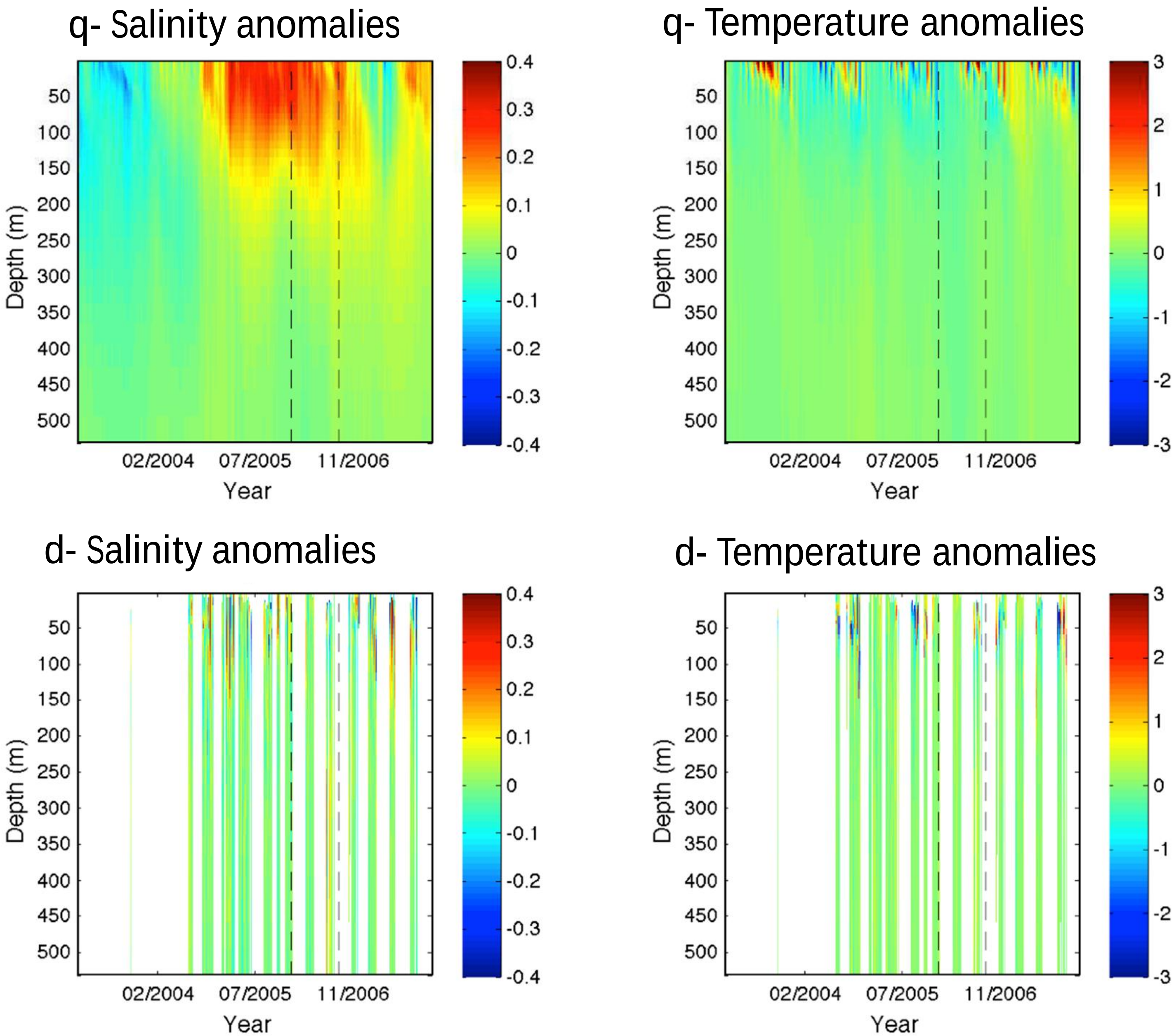
- Data stage: $$q_t | e_t \sim N\left(H_{qt} e_t, \Sigma_{qt}\right)$$ $$d_t | e_t \sim N\left(H_{dt} e_t, \Sigma_{dt}\right)$$
- Process model: the vertical structure is given by the seasonal vertical EOFs but we estimate with an AR model 5-days amplitudes (Beta)

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

- 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

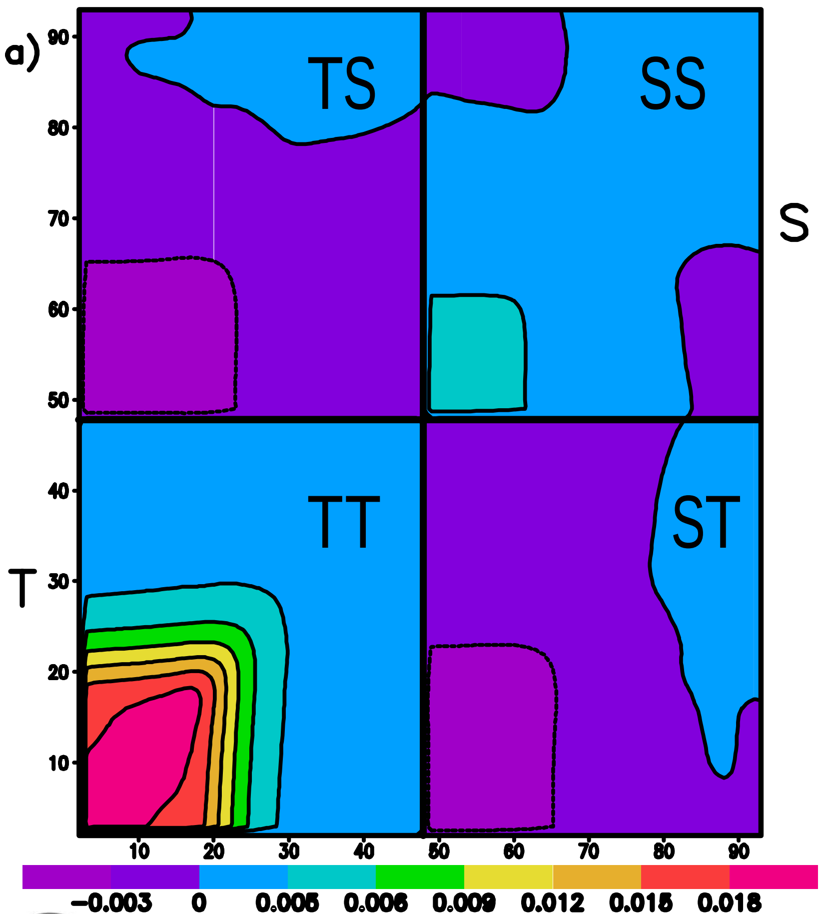


Model anomalies vertical structure

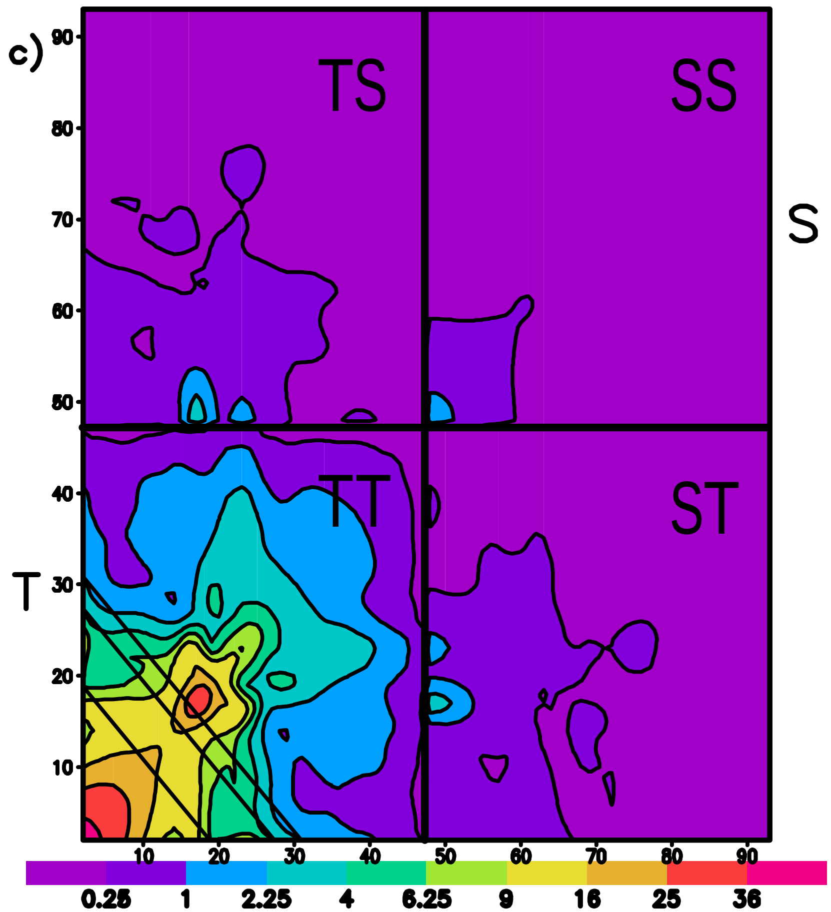
Misfit vertical structure

The high frequency error covariance matrix

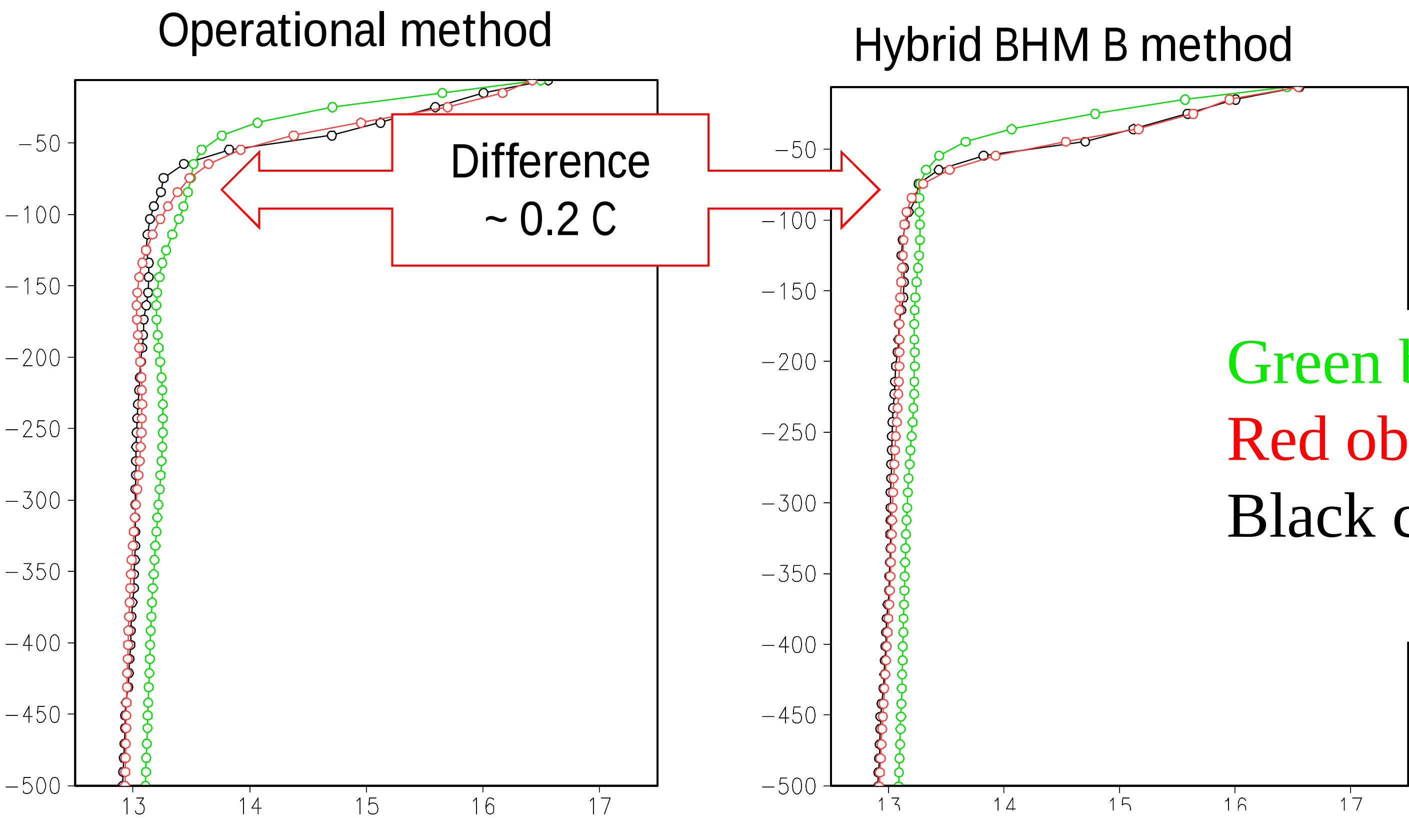
SEASONAL winter



10 Feb, 2006 C from BHM



Improvements on the assimilation due to high frequency error covariance



Green background
Red observations
Black corrected profile

Outlook

- An operational ocean 3DVAR assimilation system associated with the regional Med and Black Sea CMEMS forecasting centers has been developed by analyzing the specific ocean observation requirements
- Model improvements still provide the major source of improvements for analyses
- Errors in Temp and salinity peak between the 20-100 meter layer mainly due to atmospheric forcing uncertainties
- The future:
  - BHM winds perturbations offer a way to quantify the short term forecasting uncertainties and they will be used to help in the background error covariance definition
  - High frequency background error covariance matrix (hybrid DA) will be constructed from model variance information and misfits