Ocean Data Assimilation into the GEOS-5 Coupled Model with the GMAO Coupled Ensemble Analysis System

Christian Keppenne[1,2,*], Guillaume Vernieres[1,2], Michele Rienecker[1], Robin Kovach[1,2], Jossy Jacob[1,2], Atanas Trayanov[1,2], & Jelena Marshak[1]

[1]NASA GMAO
[2]SAIC

Also see poster PO35M-04

*contact: christian.keppenne@nasa.gov

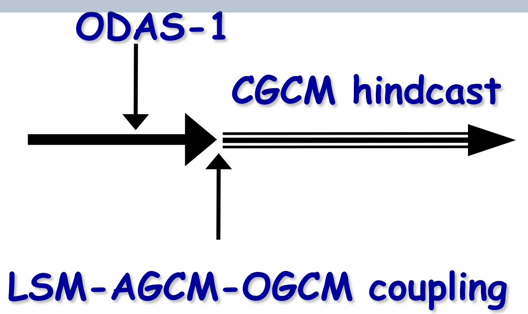
Outline:

- ODAS-1 and uncoupled ensemble assimilation
- Coupled ensemble assimilation with ODAS-2
  - In situ ARGO data
  - Sea-level anomalies
- Outlook

Ocean Sciences Meeting
Portland, OR, February 24, 2010

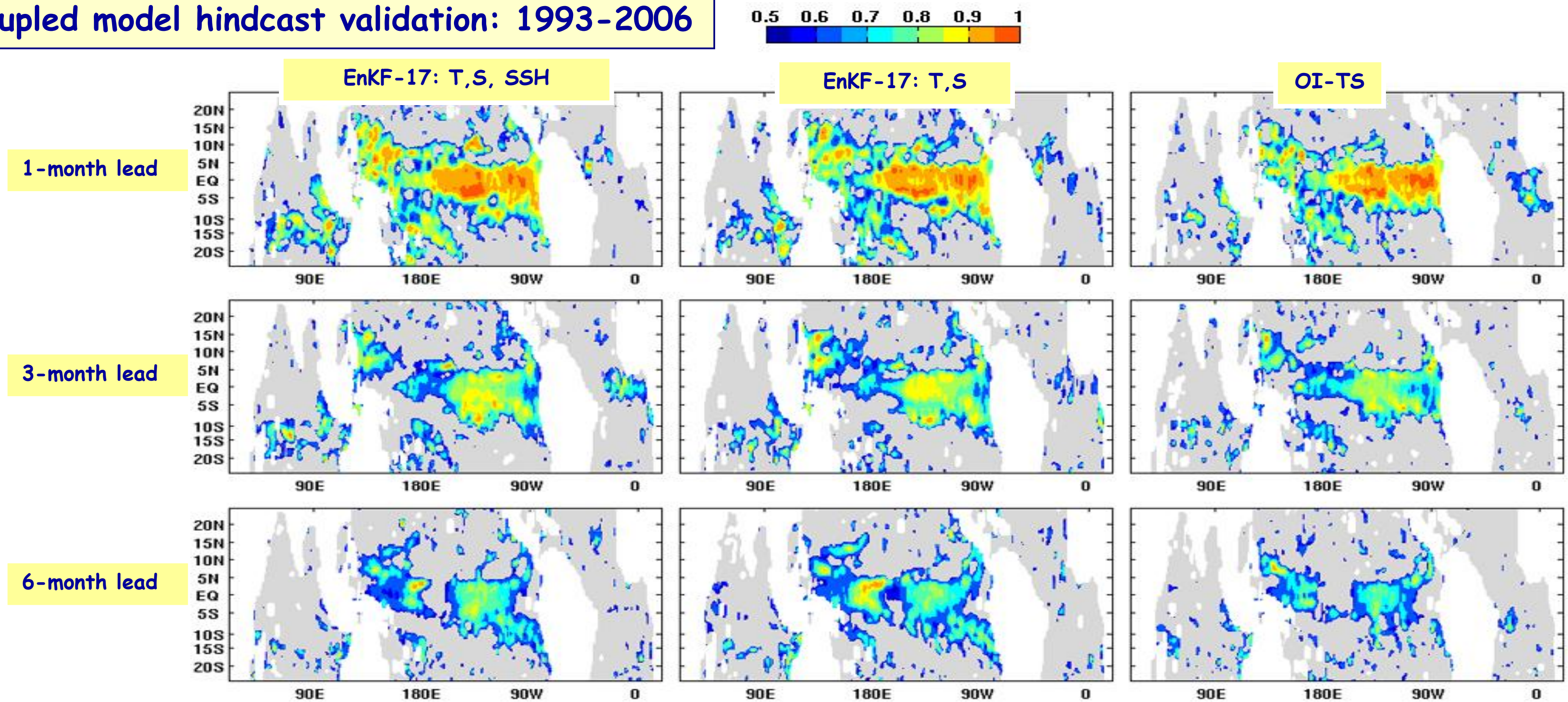
ODAS-1 and uncoupled data assimilation

GMAO CGCM coupled forecasting system
- GMAO AGCM: 90 x 144 x L34 (2° x 2.5° x L34)
- Poseidon v4 OGCM: 538 x 572 x L27 ((1/3° x 5/8° x L27)



ODAS-1 EnKF: Topex SSH, ARGO T&S, TAO T, XBT T, Pirata T
ODAS-1 OI-TS: ARGO T&S, TAO T, XBT T, Pirata T + synthetic S

Coupled model hindcast validation: 1993-2006



Possibly due to coupling/initialization shock, the coupled forecast improvement due to the SSH assimilation is lost after 6 months

SSH anomaly correlation: July-start hindcasts

Coupled data assimilation with EnKF and GMAO ODAS-2

Portability

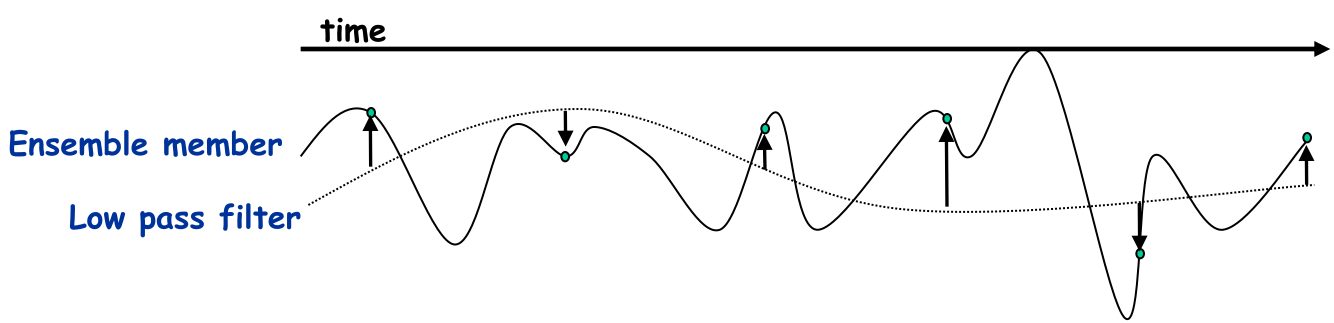
- Implemented as Earth System Modeling Framework (ESMF) gridded component
- Completely model independent

  current model configuration:
  - OGCM: MOM 4.1 - 720x410x40 resolution
  - AGCM: GEOS-5 – 288x144x72 resolution

3 Sources of background-error covariance information

$$P^{f} = P^{dyn} + P^{stat} + P^{func}$$

1) P^{dyn} : Dynamic, state-dependent error covariances from ensemble integration (EnKF-nxm)
   - Current state of n ensemble members minus low pass filter
   - m-1 recent past states of each ensemble member minus a low pass filter

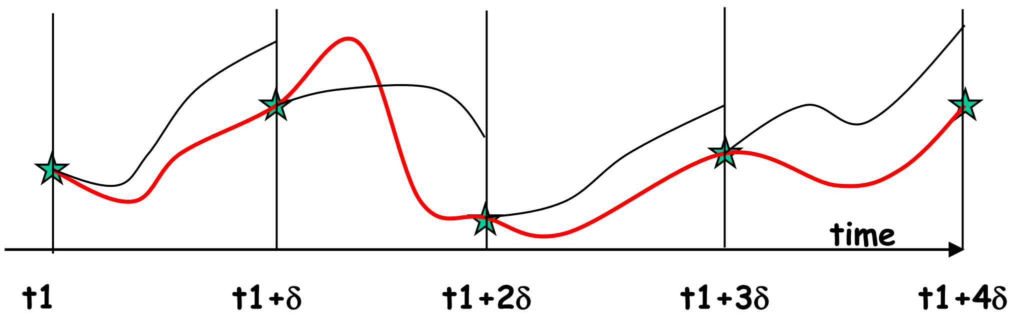


Coupled data assimilation with EnKF and GMAO ODAS-2

3 Sources of background-error covariance information

2) P^{stat} : Static,state-independent "error EOFs" and/or bred vector perturbations

3) P^{func} : Functional, idealized pseudo-Gaussian univariate covariance



"Error EOFs" are calculated from a time series of differences between a coupled model run constrained by replaying the GMAO atmospheric analysis and a sequence of unconstrained short-term forecasts

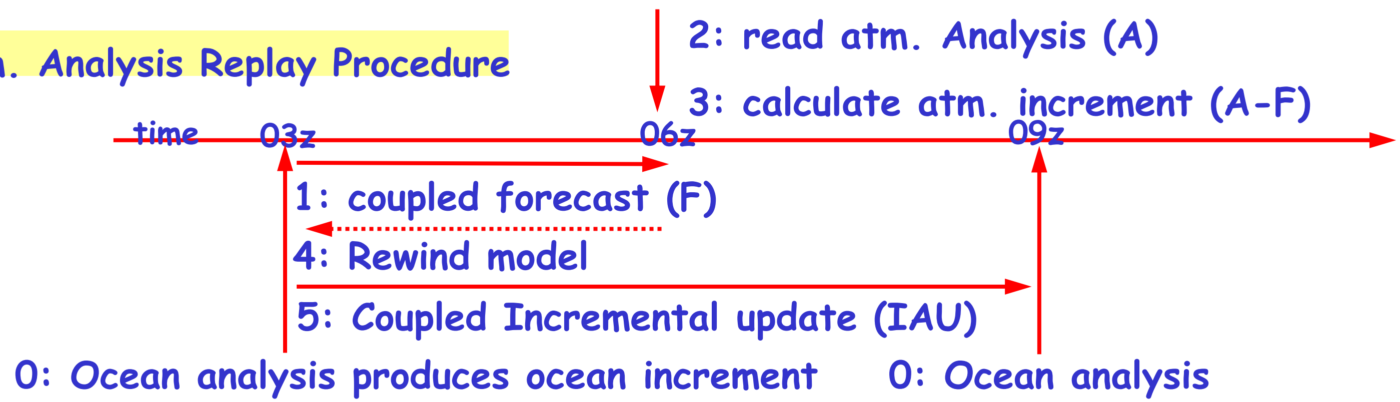
(see poster PO35M-04 for details)

Coupled data assimilation with EnKF and GMAO ODAS-2

Atmosphere constrained with "replay" of GMAO atmospheric analysis

- The offline-computed GMAO atmospheric analysis is read during the ensemble integration and used to calculate an ensemble of atmospheric increments.
- Incremental analysis updating (IAU) applies the atmospheric increments incrementally
- Incremental updating of ocean model ensemble may or may not occur at the same time

Atm. Analysis Replay Procedure



Flow-adaptive error covariance localization: in (x, y, z, t, ρ = neutral density) space

Adaptive error-covariance inflation/deflation

iteratively solves for the background error-covariance inflation/deflation needed to explain a pre-specified fraction of the global innovation variance

Use of hybrid particle pre-filter (HPF)

Step 1: HPF Pre-filter reorganizes the ensemble of models prior to each analysis

Step 2: Augmented EnKF assimilates the data using localized EnKF algorithm

EnKF assimilation experiment: in situ ARGO data

- CGCM configuration
  - GEOS-5 AGCM 288x181x72
  - MOM-4 OGCM 720x410x40
- Data
  - Daily assimilation of ARGO T profiles 04/01/06 – 05/31/06 (active data set)
  - ARGO S profiles used for validation (passive data set)

- Initial condition
- 03/01/06 coupled model restart from single coupled model run with atm. Anal. Replay

- Ensemble initialization (03/01/06 – 04/01/06)
  - initial perturbation from linear combinations of model signal EOFs
  - daily perturbations with 1% of initial perturbation amplitude

- Assimilation (04/01/06-05/31/06)
  - CE-16: 16-member control ensemble – no assimilation
  - EnKF-16x11: 16-member EnKF augmented with 10 latest past instances (lag = 1 day)
  - HPF-16: 16-member hybrid reordering particle filter – sees ARGO T profiles
  - HPEnKF-16x11: HPF-16 used as pre-filter prior to each EnKF-16x11 analysis

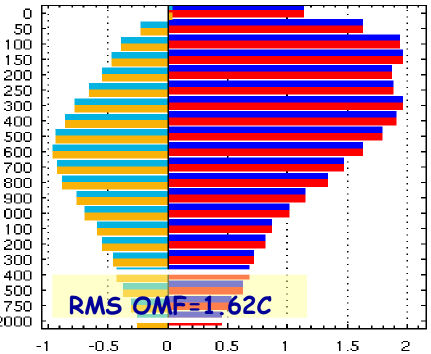
- Computing resources for 16 integration streams
  - 120-node (960 cores, i.e. 12x5 = 60 cores per ensemble member) on NASA Center for Computational Sciences (NCCS) Discover Linux cluster (11-Gflops/ second/core, 4GB RAM/core)

EnKF Assimilation of ARGO T

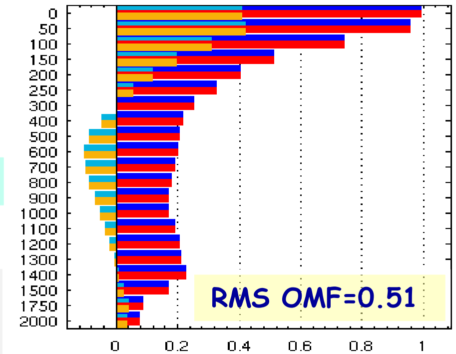
OMF and OMA statistics for May 2006 for (left) active T data (K) and (right) passive S data (PSU)

mean OMF
RMS OMF
mean OMA
RMS OMA

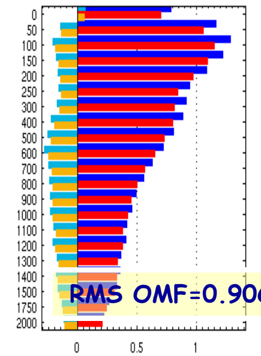
CE-16 T



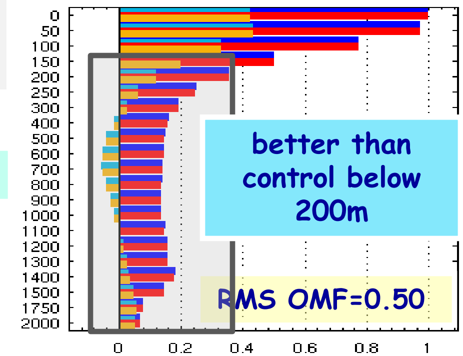
CE-16 S



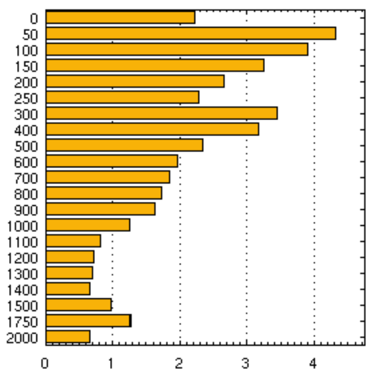
EnKF-16x11 T



EnKF-16x11 S

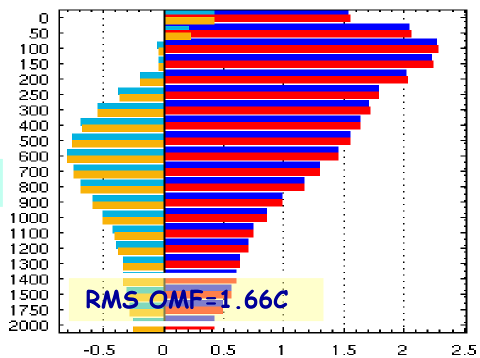


better than control below 200m

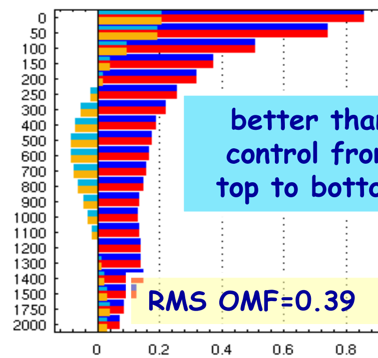


S data count (T data have same distribution)

HPF-16 T

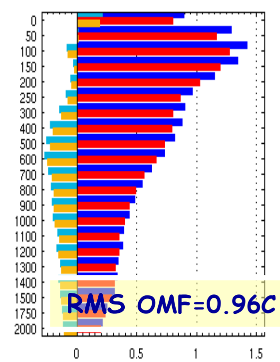


HPF-16 S

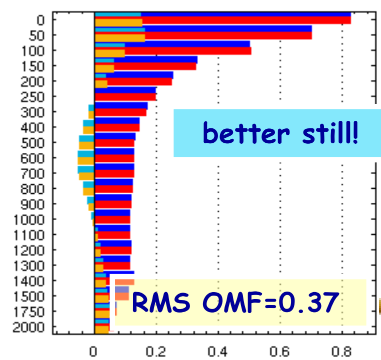


better than control from top to bottom

HPEnKF-16x11 T



HPEnKF-16x11 S

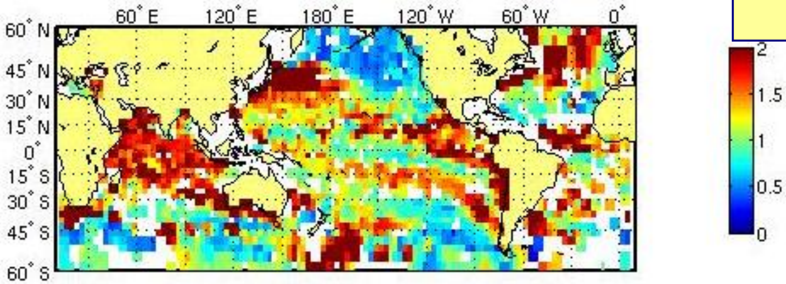


better still!

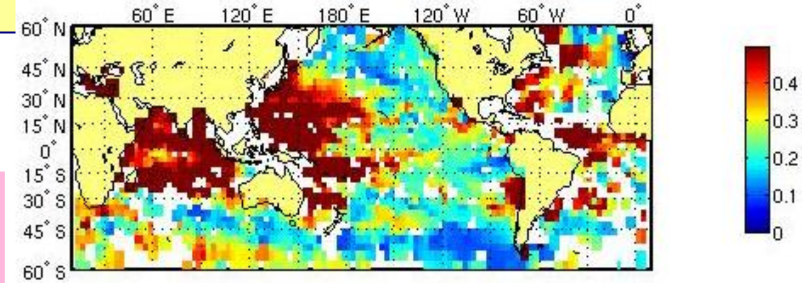
EnKF assimilation of ARGO-T

RMS T OMF - May 2006 (ARGO T active)

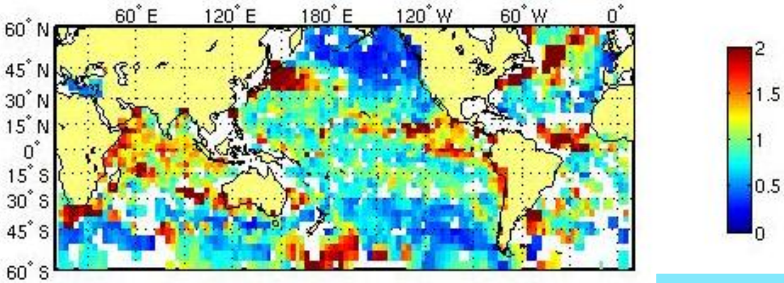
RMS S OMF - May 2006 (ARGO S passive)



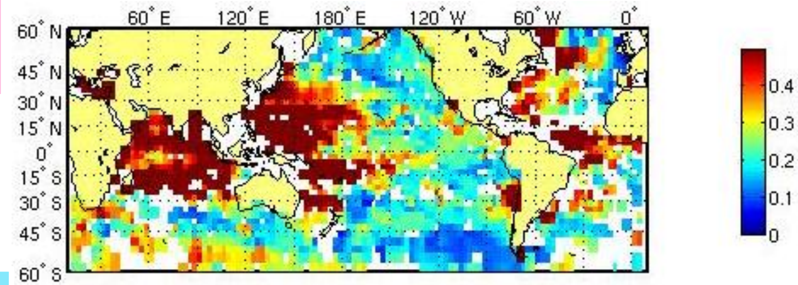
CE-16



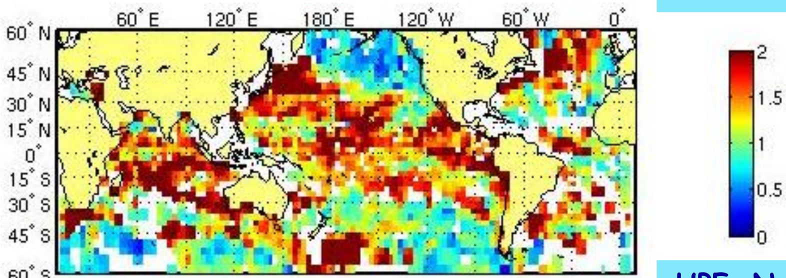
RMS OMF statistics binned to 3x3 deg. boxes in 0-3000m range for (left) active T data (K) and (right) passive S data (PSU)



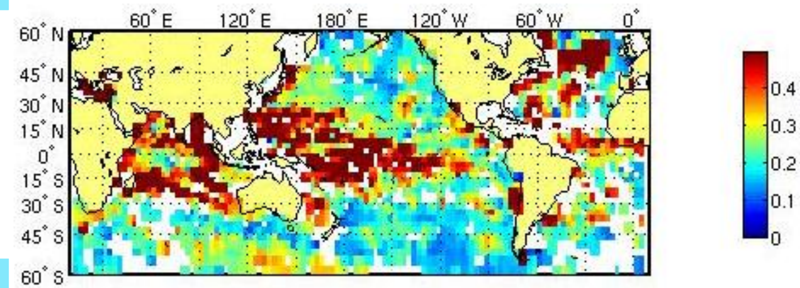
EnKF-16x11



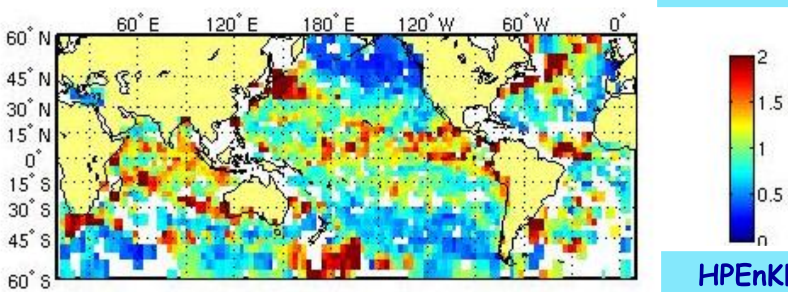
EnKF: Major T improvement but insignificant change for S



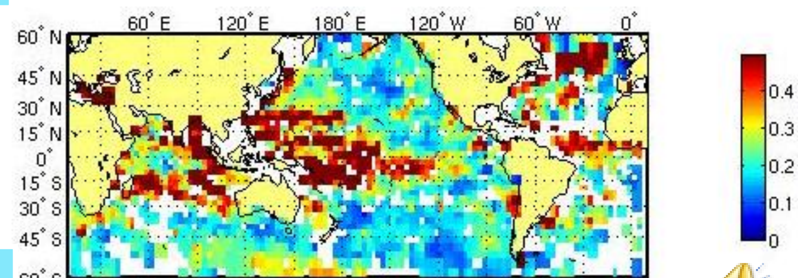
HPF-16



HPF: No notable T improvement but major improvement for S



HPEnKF-16x11

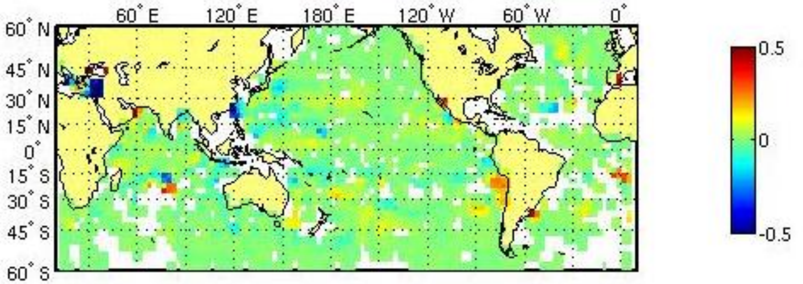


HPEnKF: Major improvements for both T and S

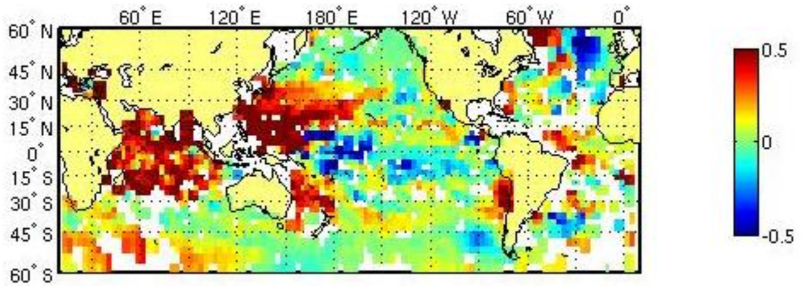
EnKF assimilation of ARGO-T

Salinity changes over control ensemble for May 2006. Warm colors denote areas where the analysis is closer to the passive S ARGO data than the control ensemble (improvement). Cold colors denote areas where the control ensemble resembles the ARGO DATA more closely (analysis is worse).

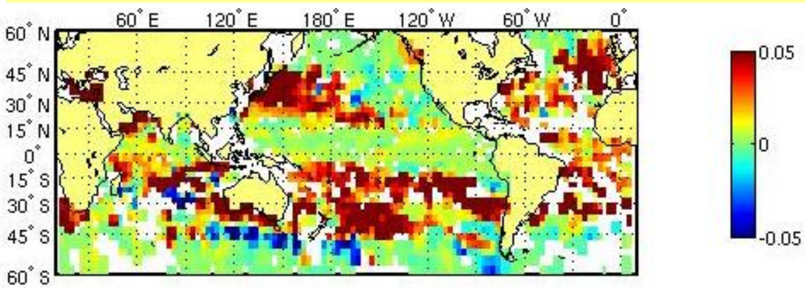
CE-16 RMS OMF – ENKF-16x11 RMS OMF: z<200m



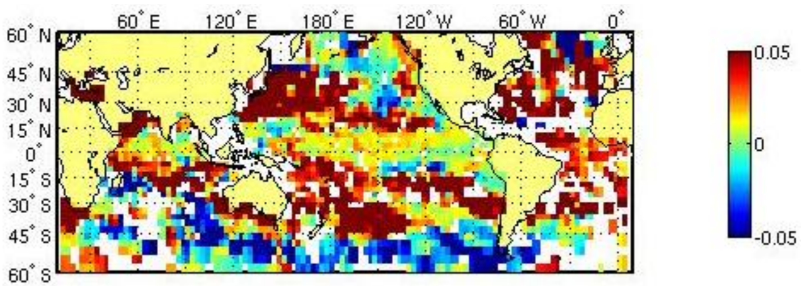
CE-16 RMS OMF – HPENKF-16x11 RMS OMF: z<200m



CE-16 RMS OMF – ENKF-16x11 RMS OMF: z>200m



CE-16 RMS OMF – HPENKF-16x11 RMS OMF: z>200m



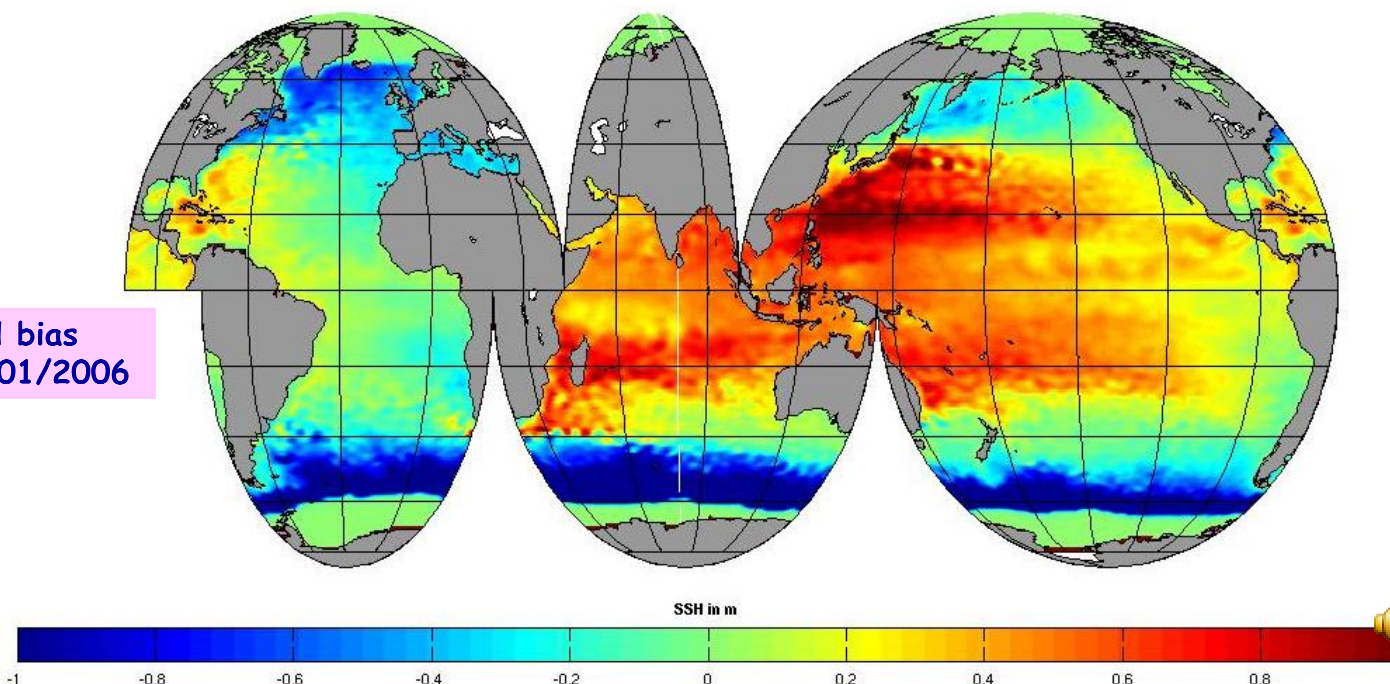
SSH assimilation with AEnKF: challenges and initial results

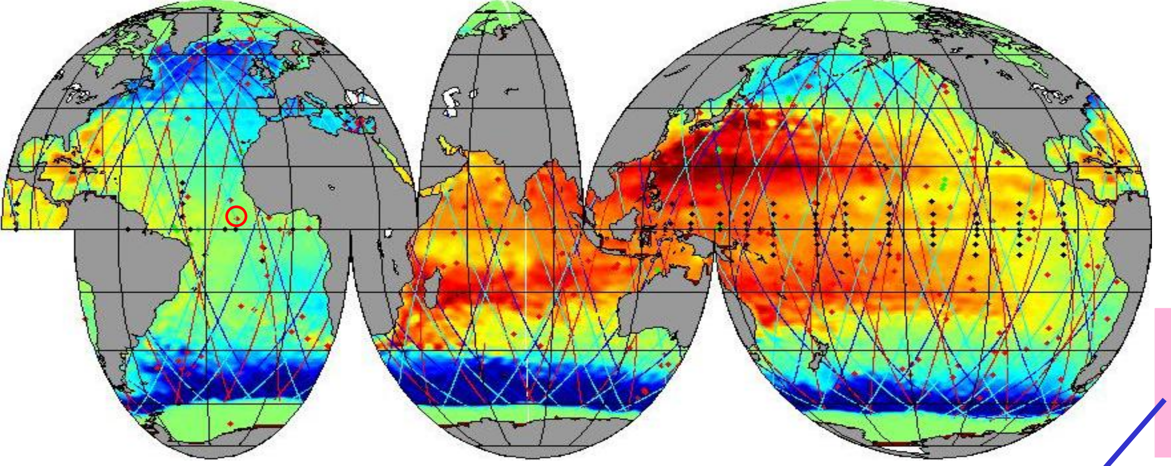
- Challenge: model climatology changes as the data are assimilated
- Solution: online bias estimation is a must when assimilating SSH anomalies!

$$\eta = \int_z f(\rho(z))dz$$

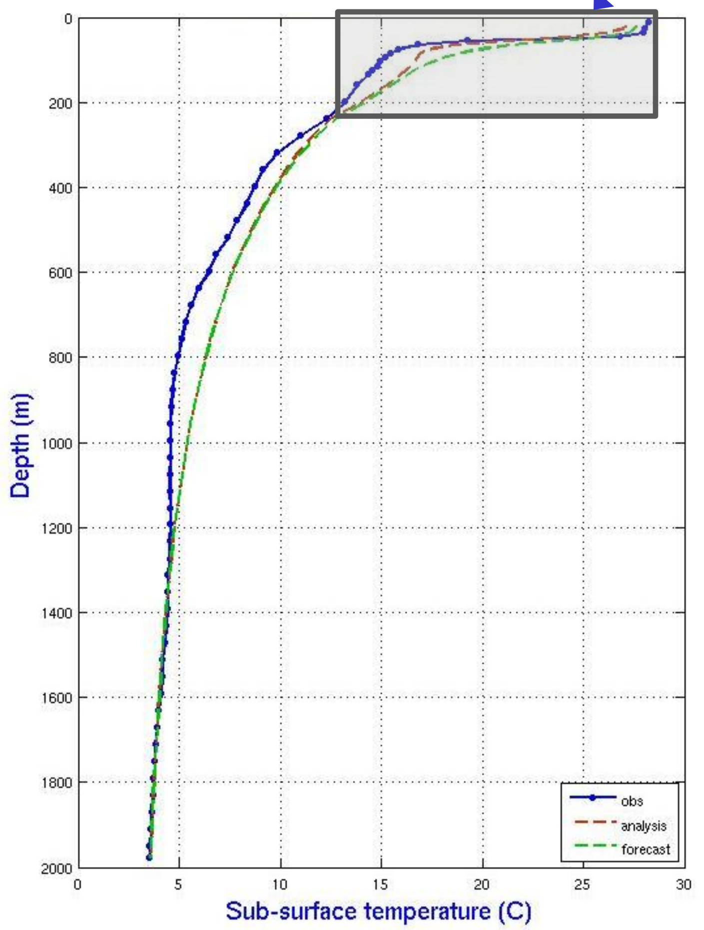
- Challenge 1: must estimate $\rho(z)$ from a scalar η measurement
- Challenge 2: given the estimated $\rho(z)$, must derive the proper $T(z)$ and $S(z)$ distributions
- Tentative solution (two-step correction):
  - Step 1: use <SSH, ρ> covariances to construct density increment
  - Step 2: get T and S increments from <T, ρ> and <S, ρ> covariances

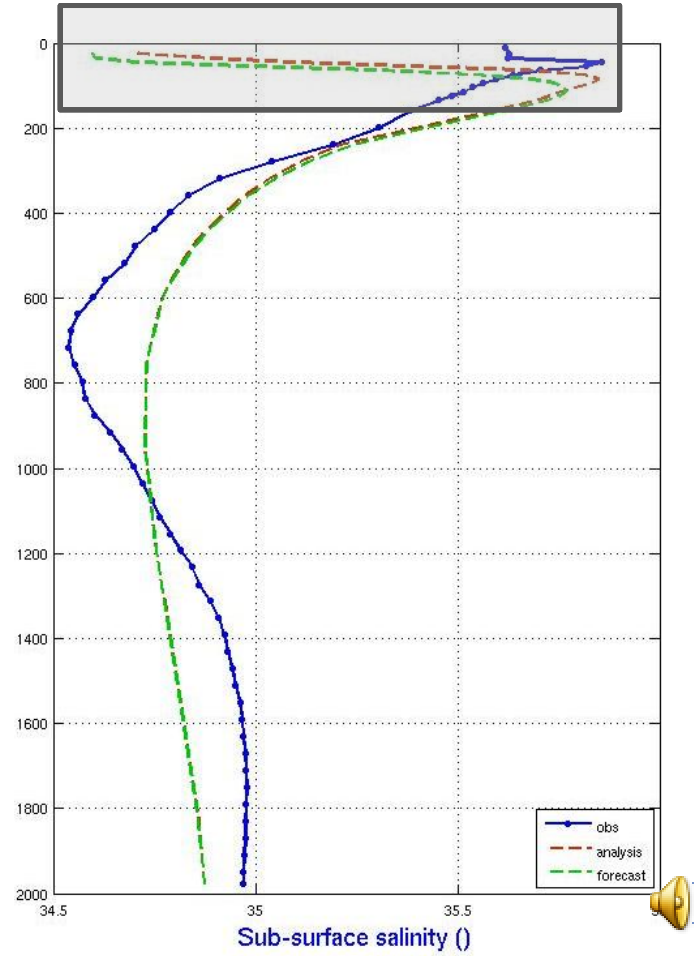
MOM-4 SSH bias estimate on 04/01/2006

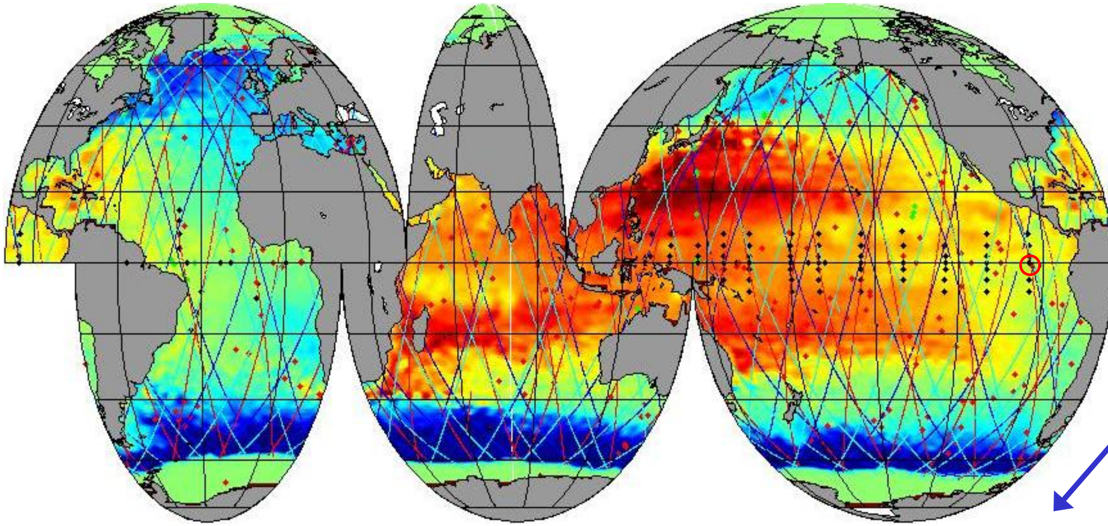




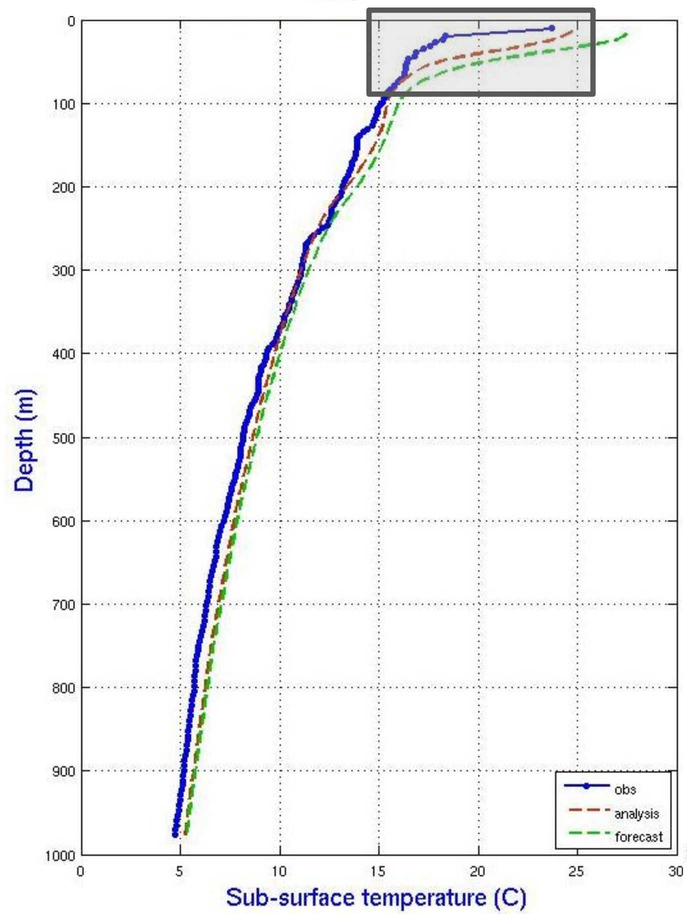
Splitting a ρ increase between its cooling component and its salting component

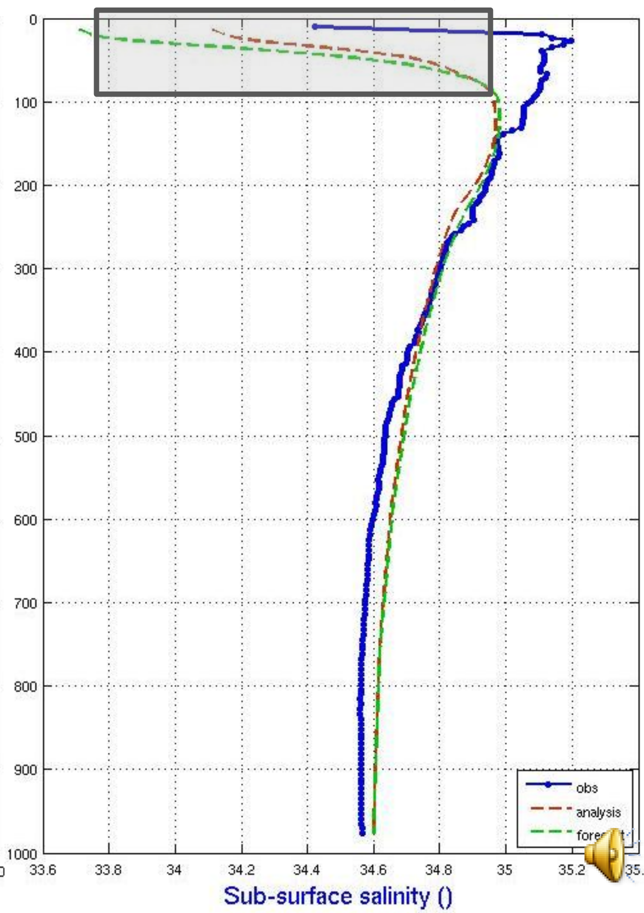






Splitting a ρ increase between its cooling component and its salting component





Outlook

• Especially when the analysis step is preceded by the application of the hybrid particle pre-filter (HPF), the ODAS-2 augmented EnKF provides an effective means to assimilate ocean observations into ensembles of operational-resolution coupled models on cost-effective scalable Linux clusters.

• With the "replay" of the atmospheric analysis into the ensemble of coupled models, the HPEnKF system can readily be used to initialize ensembles of coupled model forecasts, thereby avoiding undesirable coupling shocks such as those present in the first generation GMAO coupled forecasting/GMAO ODAS-1 system.

• The implementation of the system as an ESMF gridded component renders it model-independent. As such, it is readily portable to ensembles of various OGCM/AGCM combinations at a cost no greater than that of writing ESMF wrappers for the models of choice (only necessary if models are not ESMF-ready).