GMAO Ocean Data Assimilation

Michele Rienecker

Christian Keppenne, Robin Kovach, Shu-Chih Yang, Jossy Jacob

David Adamec

Watson Gregg, Lars Nerger

Global Modeling and Assimilation Office (GMAO)

NASA/Goddard Space Flight Center

JCSDA Workshop
Application of Remotely Sensed Observations in
Data Assimilation
August 1, 2007

Overall Goals:

- Improve ocean analyses for initialization of coupled forecasts.
- Ocean climate variability
- Ocean color Climate Data Records (CDRs)

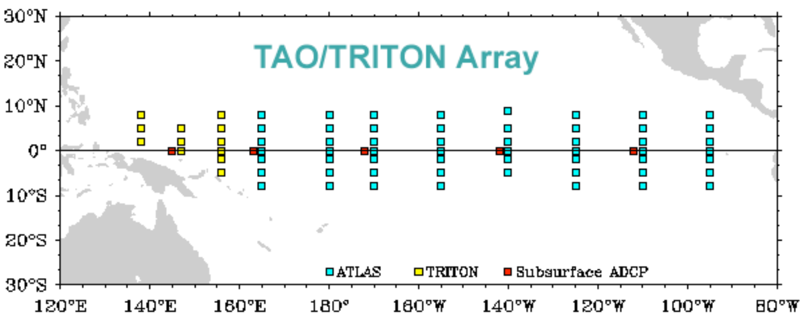
Methods:

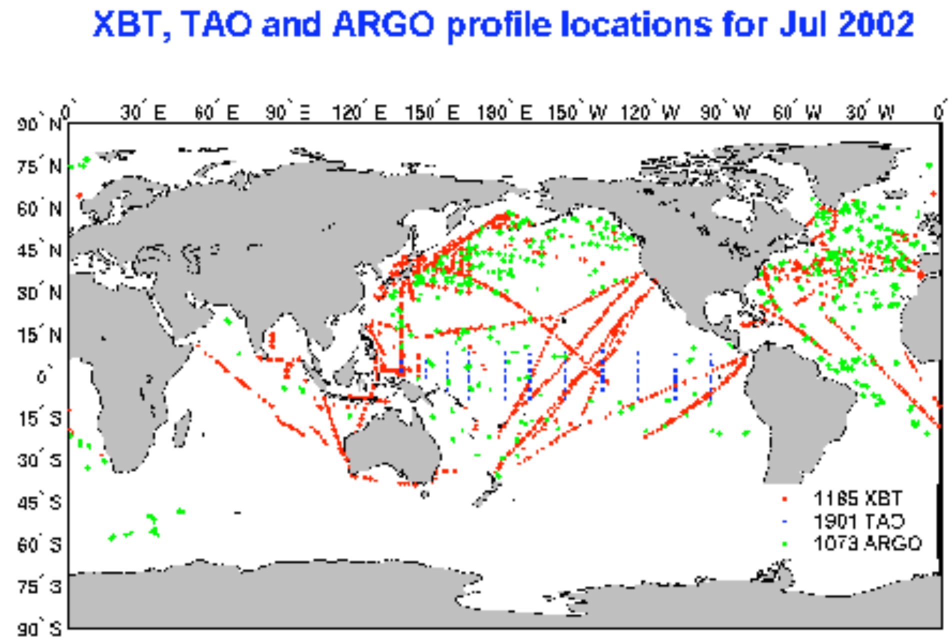
- Optimal Interpolation (univariate)
- Ensemble Kalman Filter (multi-variate, state-dependent; satellite altimetry)
- Bred-vectors to capture dominant growing modes of error
- SEIK filter (Ocean Color)

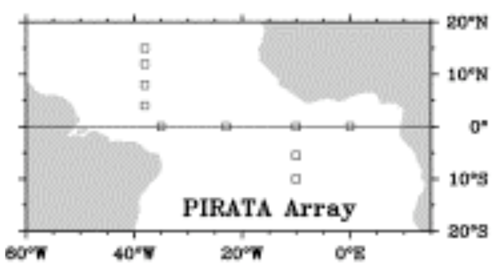
Model:

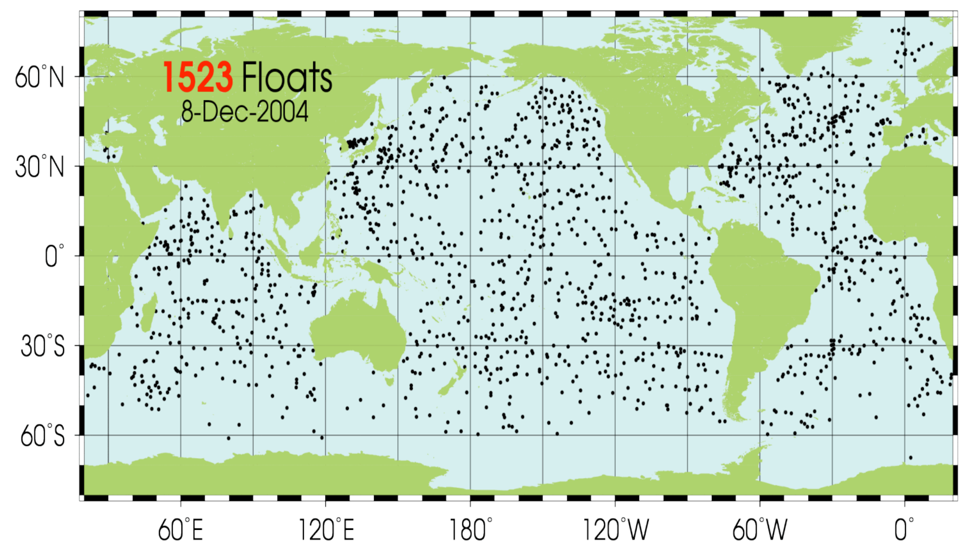
- Poseidon V4 and V5 - quasi-isopycnal model (Paul Schopf)
- MOM4 (GFDL)

Ocean in situ observations
TAO/TRITON/PIRATA moorings
+ XBTs
+ ARGO

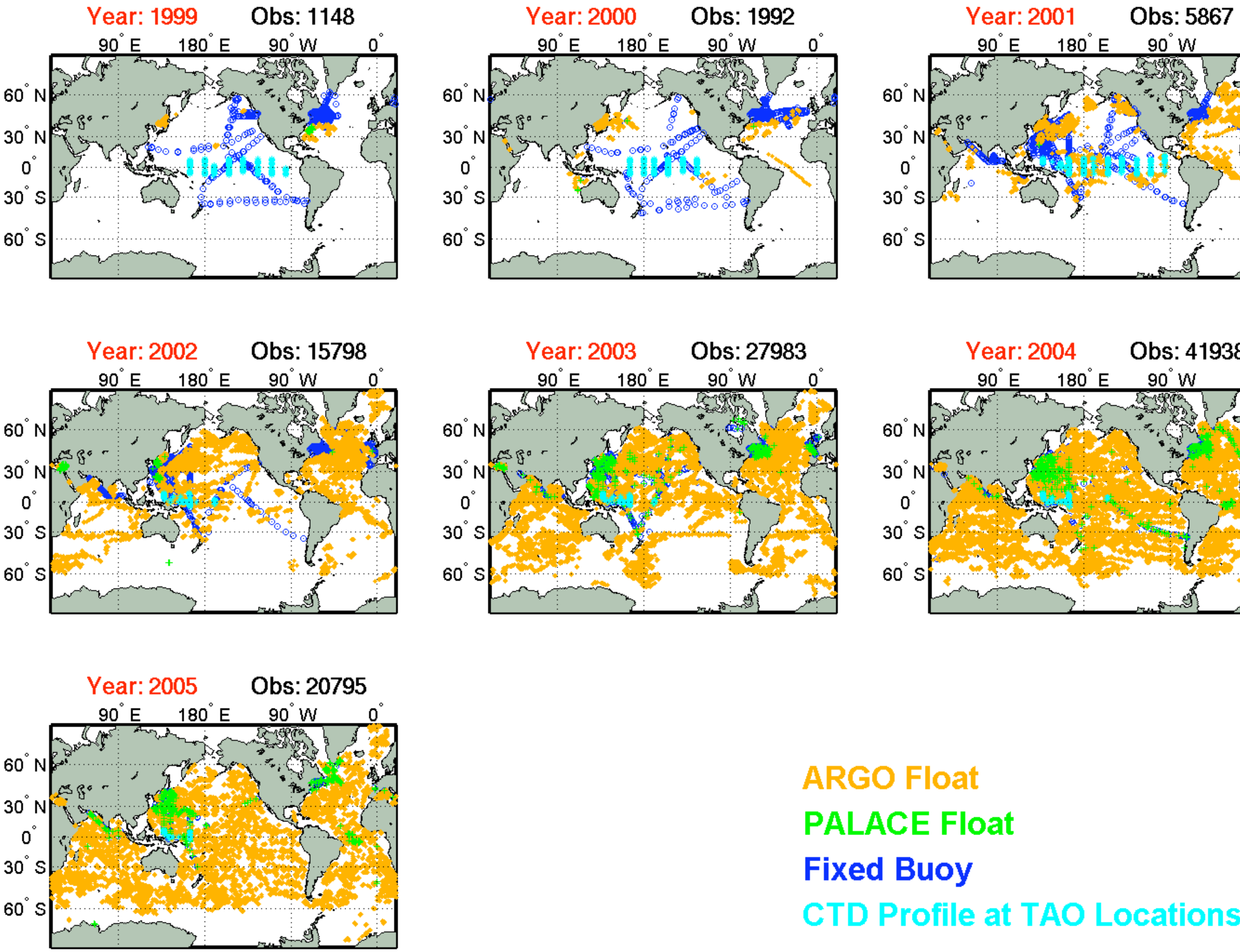


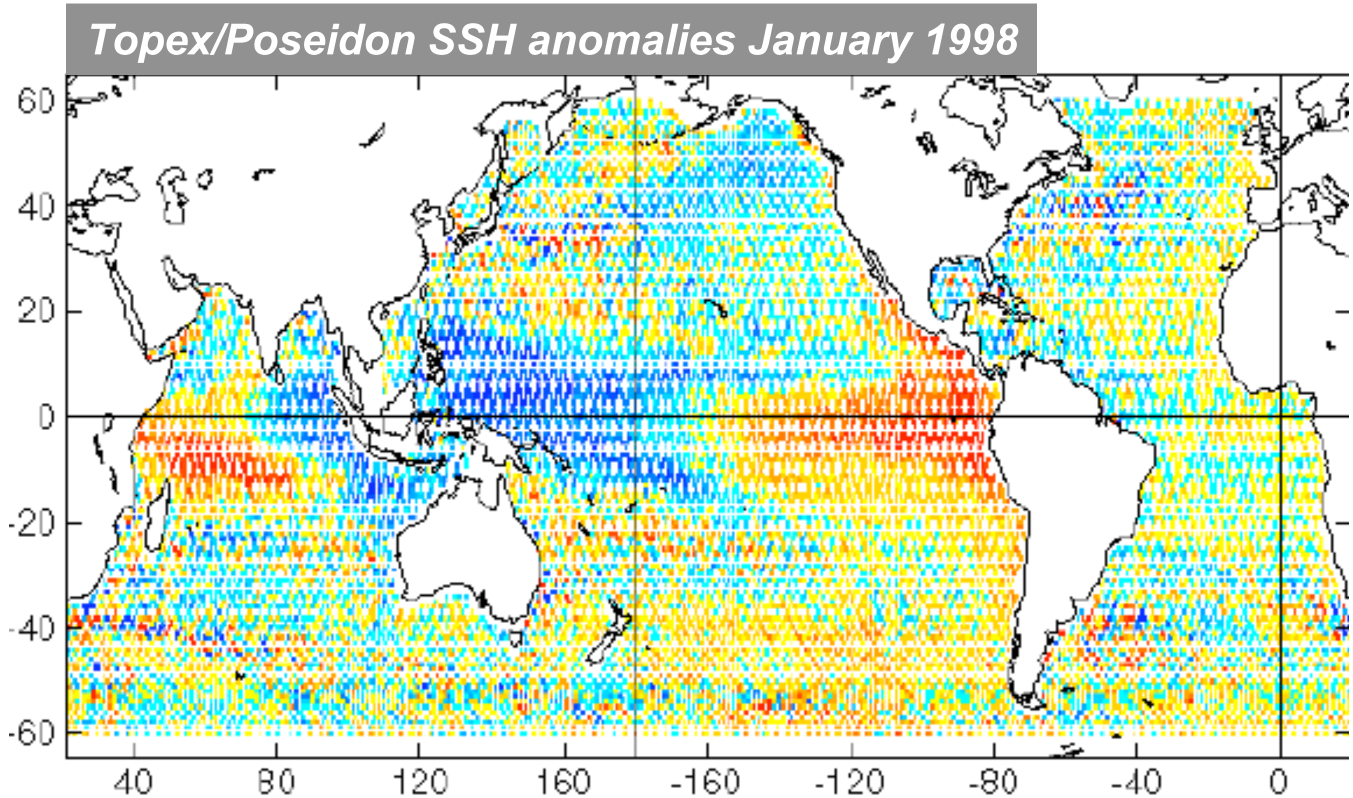






Salinity Profiles per annum





Surface Height
TOPEX: August 1992-2005
JASON: December 2001---
JASON-2: June 2008

Surface Winds
SSM/I: July 1987 ---
NSCAT: Aug 1996 -- June1997
QuikSCAT: June 1999 ---

Sea Surface Temperature
AVHRR: 1982 ---
MODIS: 2000 ---
TMI: 1997 ---
Aqua/AMSR-E: 2002 ---

Surface Salinity
Aquarius: 2010

Ocean Color
CZCS: Oct 1978 -- June 1986
SeaWiFS: August 1997 --
MODIS: 2000---

Optimal Interpolation (univariate)

- Fixed Gaussian covariances: x_s=20º, y_s=5º, z_s=100m; more isotropic with increasing latitude
- Temperature (T) and Salinity (S) assimilated separately

Ensemble Kalman Filter (multi-variate, state-dependent; satellite altimetry)

- Surface height is a diagnostic - calculate $\langle\delta SSH, \delta T(z)\rangle$ and $\langle\delta SSH, \delta S(z)\rangle$ to "project" corrections to surface height anomalies through the water column
- Temperature data used to update salinity and currents
- Salinity data used to update temperature and currents

Observations

- Instrument error and Representation error
- Synthetic salinity used to constrain water masses

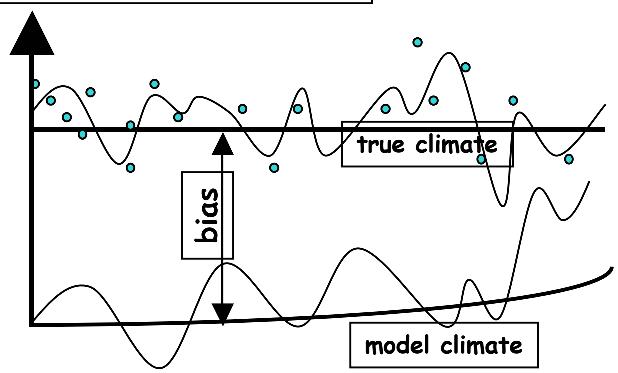
Surface Forcing

- One of the major source of errors!
- Heat, fresh water and momentum fluxes

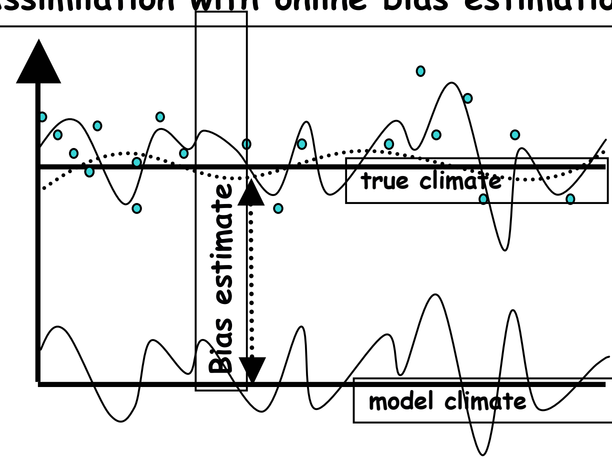
We assimilate:

In-situ temperature profiles
In-situ salinity profiles from Argo floats
Synthetic salinity profiles from observed T(z) and climatological T-S relations
T/P and Jason-1 SSH anomalies ⇒ Bias must be accounted for when assimilating SSH

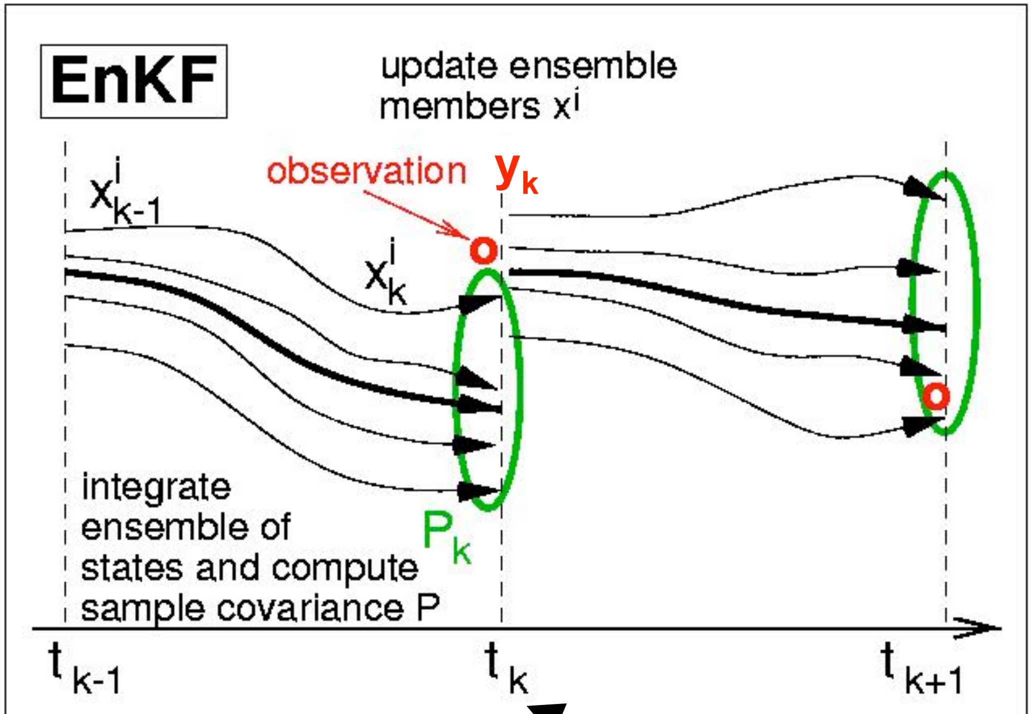
a) Standard assimilation



b) Assimilation with online bias estimation (OBE)



Side by side estimation of:
- Unbiased error
- Climatological error (bias)



x_k^i state vector (T, S, u,v, SSH)

P_k state error covariance

R_k observation error covariance

Propagation $\mathbf{t_{k-1}}$ to $\mathbf{t_k}$:

$$\mathbf{x_k^{i-} = f(x_{k-1}^{i+}) + w_k^i}$$

w = model error

Update at $\mathbf{t_k}$:

$$\mathbf{x_k^{i+} = x_k^{i-} + K_k(y_k^i - x_k^{i-})}$$

for each ensemble member i=1…N

$$\mathbf{K_k = P_k (P_k + R_k)^{-1}}$$

with $\mathbf{P_k}$ computed from ensemble spread

Compactly supported EnKF (bias estimation omitted)

$$\boldsymbol{x}_{\mathrm{i,k}}^{\mathrm{f}} = \boldsymbol{M}(\boldsymbol{x}_{\mathrm{i,k-1}}^{\mathrm{a}}, \boldsymbol{f}_{\mathrm{k-1}}) + \boldsymbol{N}_{\mathrm{i,k-1}}, \qquad \mathrm{E}(\boldsymbol{N}_{\mathrm{i,k-1}}\boldsymbol{N}_{\mathrm{i,k-1}}^{\mathrm{T}}) \approx \boldsymbol{Q}_{\mathrm{k-1}}, \qquad \mathrm{i} = 1,\ldots,\mathrm{n}, \quad \text{(1a)}$$

$$\boldsymbol{S} = \{\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_{\mathrm{n}}\} = \{\boldsymbol{H}(\Phi(\boldsymbol{x}_1^{\mathrm{f}} - \bar{\boldsymbol{x}}^{\mathrm{f}})), \boldsymbol{H}(\Phi(\boldsymbol{x}_2^{\mathrm{f}} - \bar{\boldsymbol{x}}^{\mathrm{f}})), \ldots, \boldsymbol{H}(\Phi(\boldsymbol{x}_{\mathrm{n}}^{\mathrm{f}} - \bar{\boldsymbol{x}}^{\mathrm{f}}))\} \quad \text{(1b)}$$

$$\boldsymbol{H}\boldsymbol{P}^{\mathrm{f}}\boldsymbol{H}^{\mathrm{T}} = \frac{1}{\mathrm{n}-1}\boldsymbol{S}\boldsymbol{S}^{\mathrm{T}}, \quad \text{(1c)}$$

$$\boldsymbol{a}_{\mathrm{i}} = [\boldsymbol{C} \bullet (\boldsymbol{H}\boldsymbol{P}^{\mathrm{f}}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R})]^{-1}(\boldsymbol{y} + \boldsymbol{e}_{\mathrm{i}} - \boldsymbol{H}(\boldsymbol{x}_{\mathrm{i}}^{\mathrm{f}})), \qquad \mathrm{E}(\boldsymbol{e}_{\mathrm{i}}\boldsymbol{e}_{\mathrm{i}}^{\mathrm{T}}) \approx \boldsymbol{R}, \qquad \mathrm{i} = 1,\ldots,\mathrm{n}, \quad \text{(1d)}$$

$$\boldsymbol{x}_{\mathrm{i,l}}^{\mathrm{a}} = \boldsymbol{x}_{\mathrm{i,l}}^{\mathrm{f}} + \frac{1}{\mathrm{n}-1}\sum_{\mathrm{j=1}}^{\mathrm{n}}(\Phi(\boldsymbol{x}_{\mathrm{j,l}}^{\mathrm{f}} - \bar{\boldsymbol{x}}_{\mathrm{l}}^{\mathrm{f}}))\boldsymbol{s}_{\mathrm{j}}^{\mathrm{T}}(\boldsymbol{c}_{\mathrm{l}} \bullet \boldsymbol{a}_{\mathrm{i}}), \qquad \mathrm{i} = 1,\ldots,\mathrm{n}. \quad \text{(1e)}$$

Compensating for the effects of small ensemble size:
Φ: smoothing operator for small scales
C: Compact support operator (Schur product) from Gaspari and Cohn (1985)
Variance inflation to avoid filter collapse

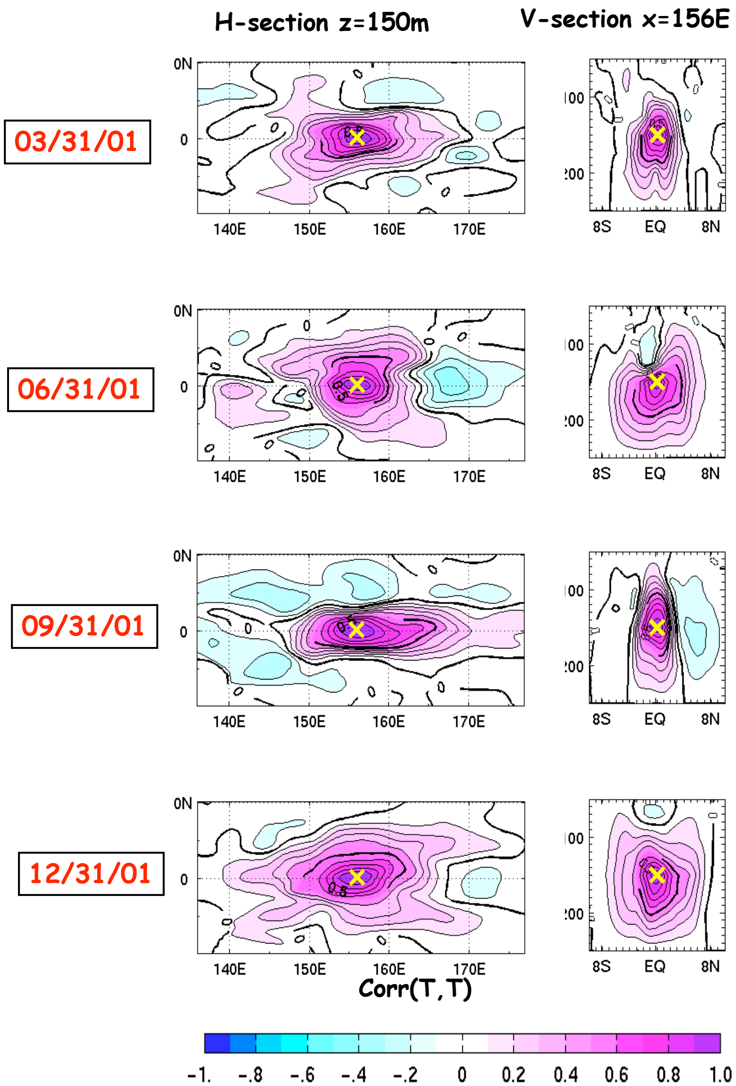
Ocean state-dependent covariances with the EnKF

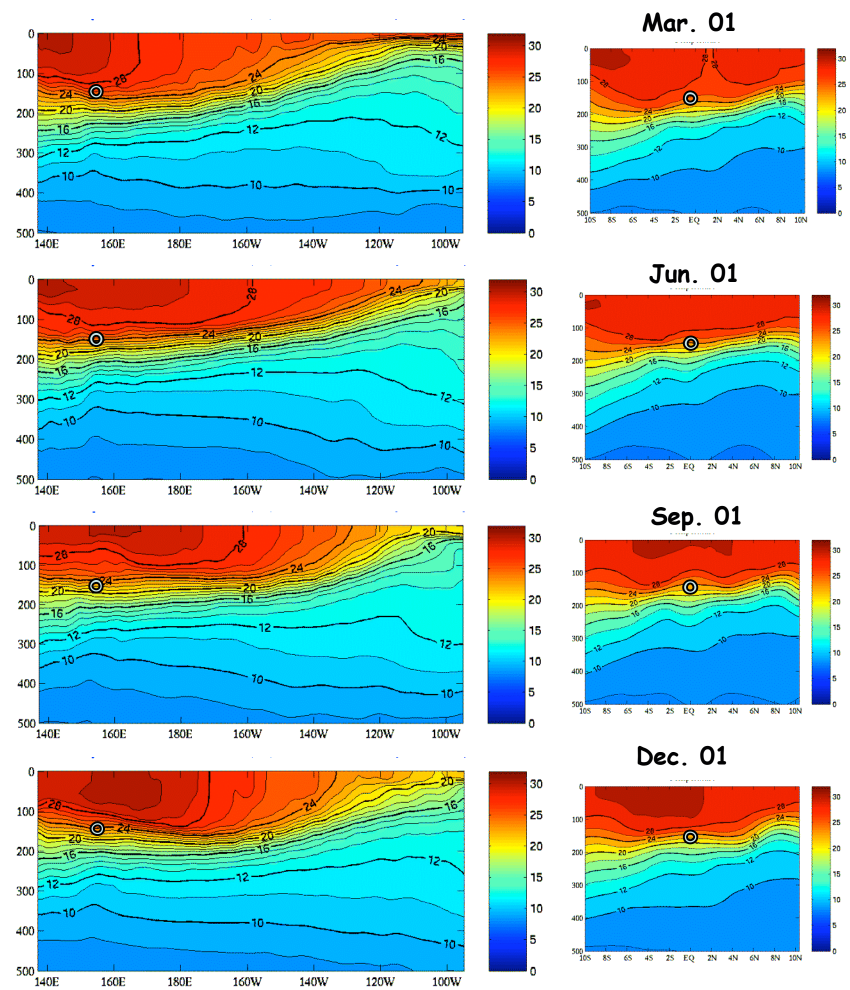
Temporal evolution of Kalman gain for T obs.

Christian Keppenne

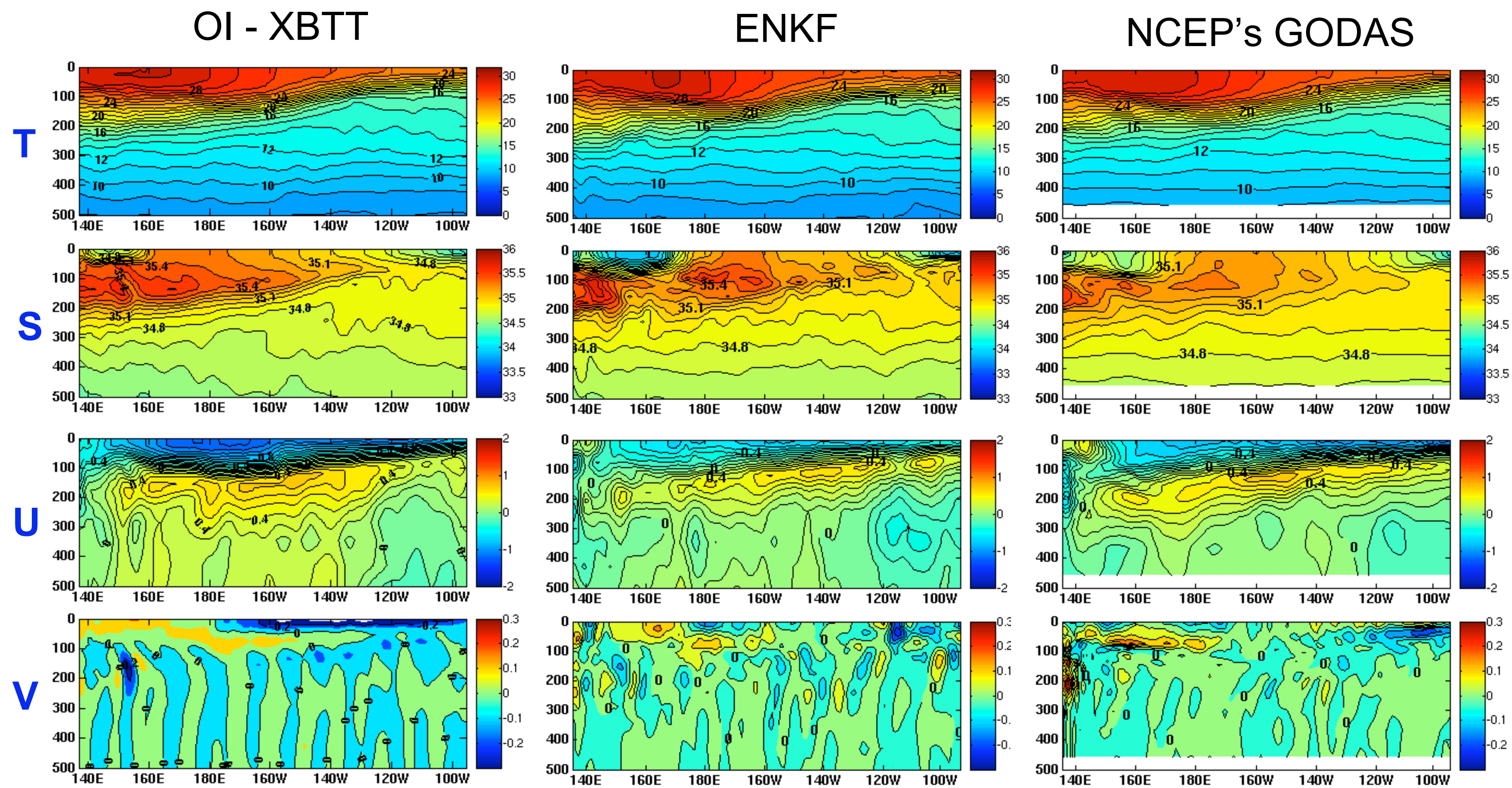
EnKF-33: filter

Schur(C,P) @(0N,156E,150m)

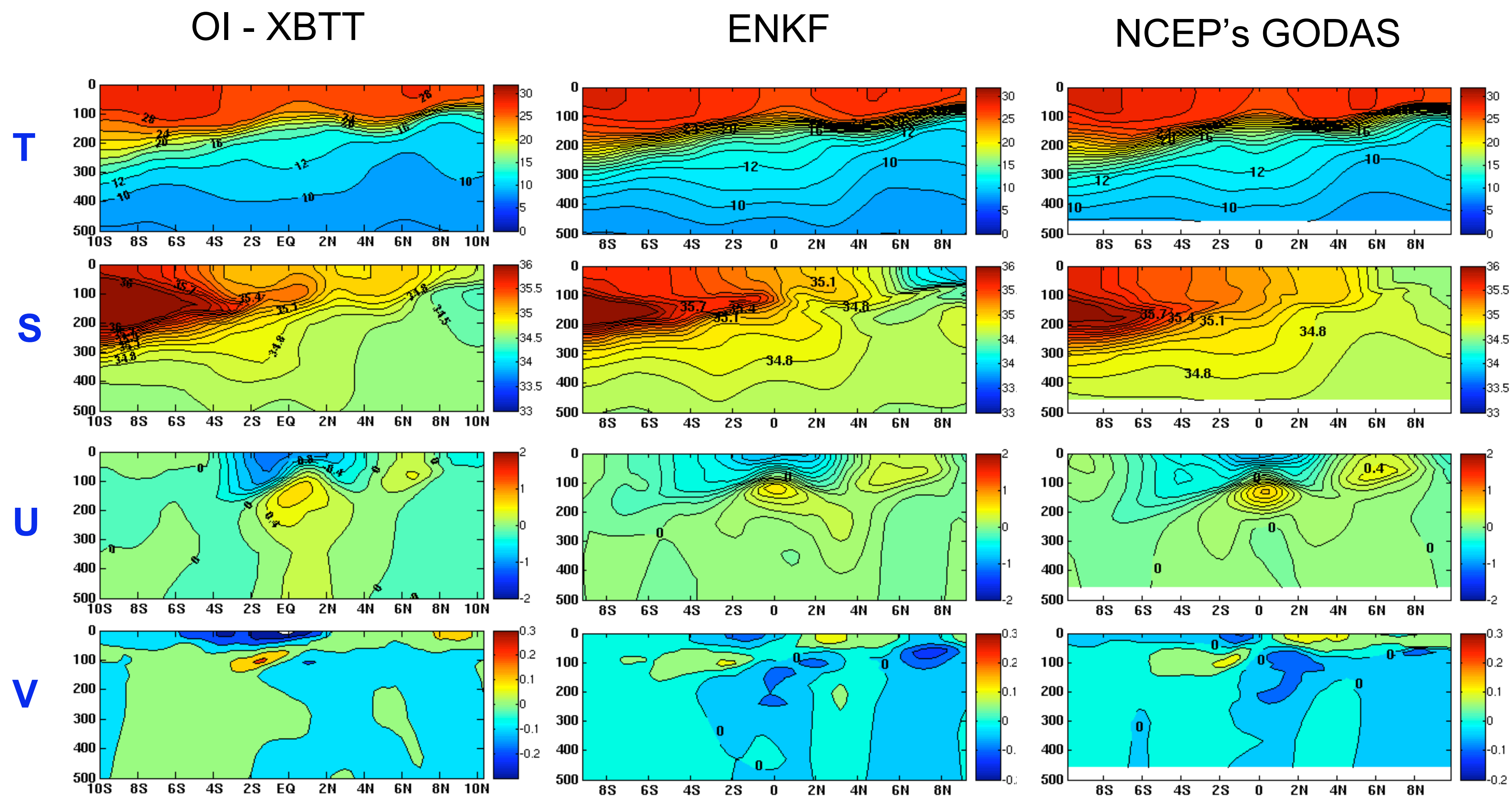




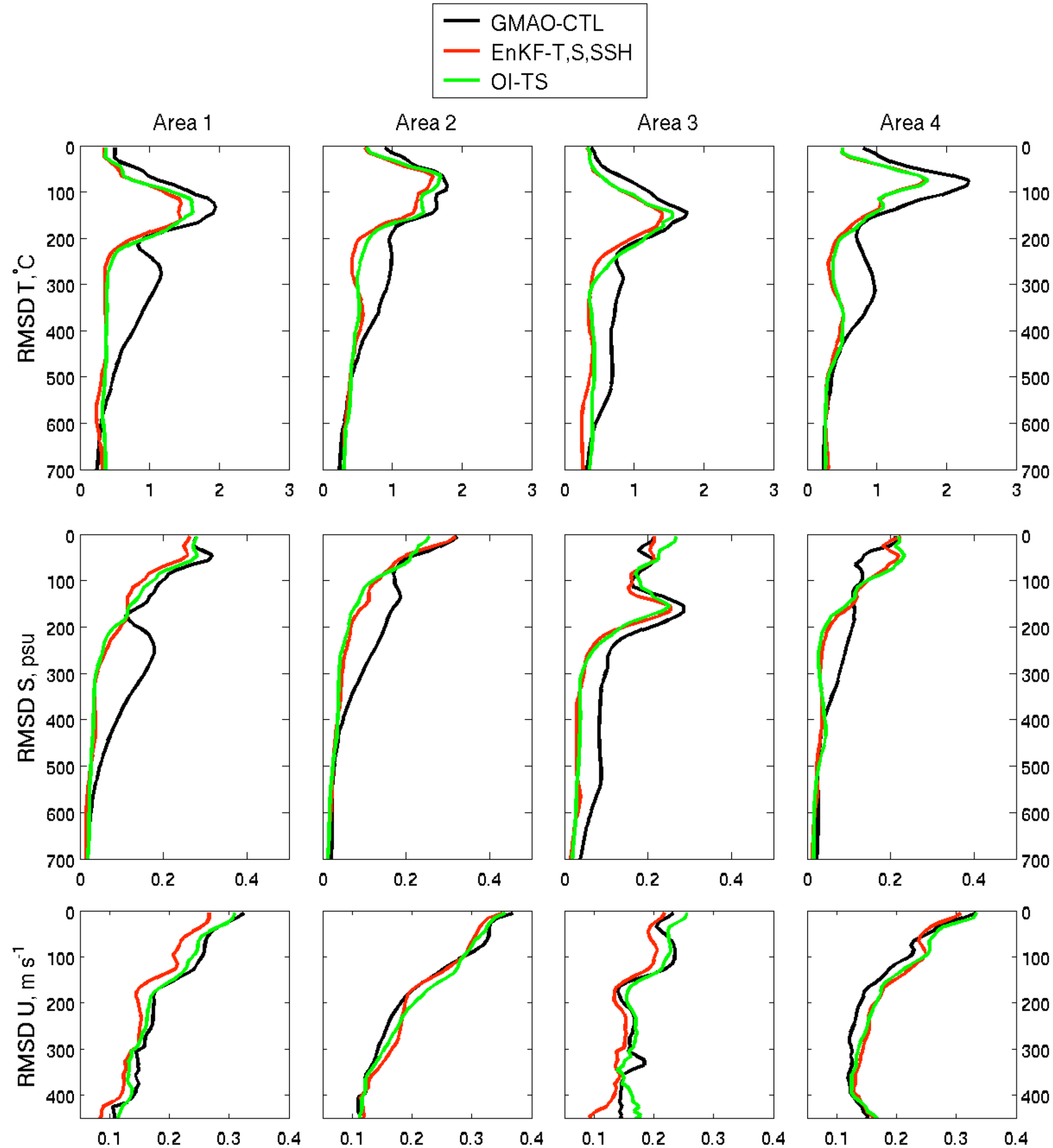
Ocean climate for June 2007 along the equatorial Pacific

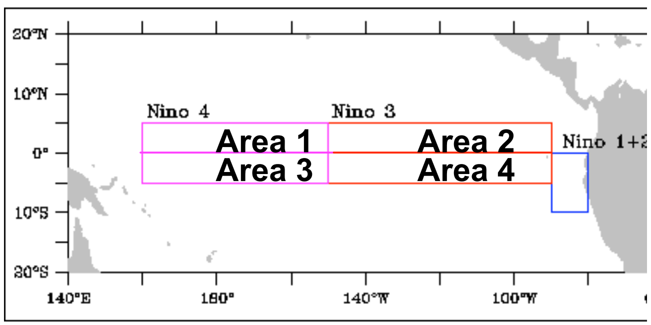


Ocean climate for June 2007 along 155°W

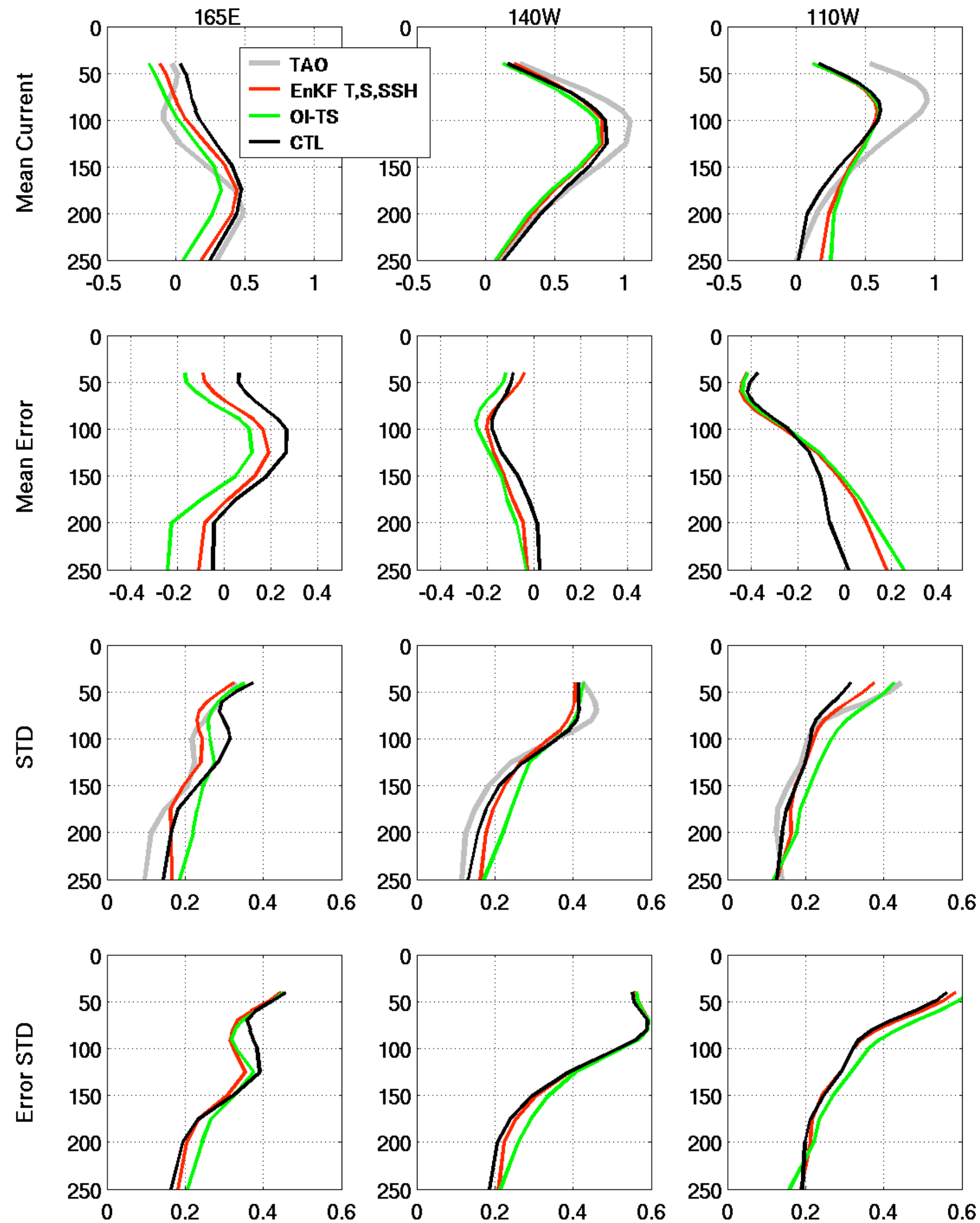


Independent Validation
RMSD of analysis c.f.
TAO servicing cruise CTDs
1994-1998

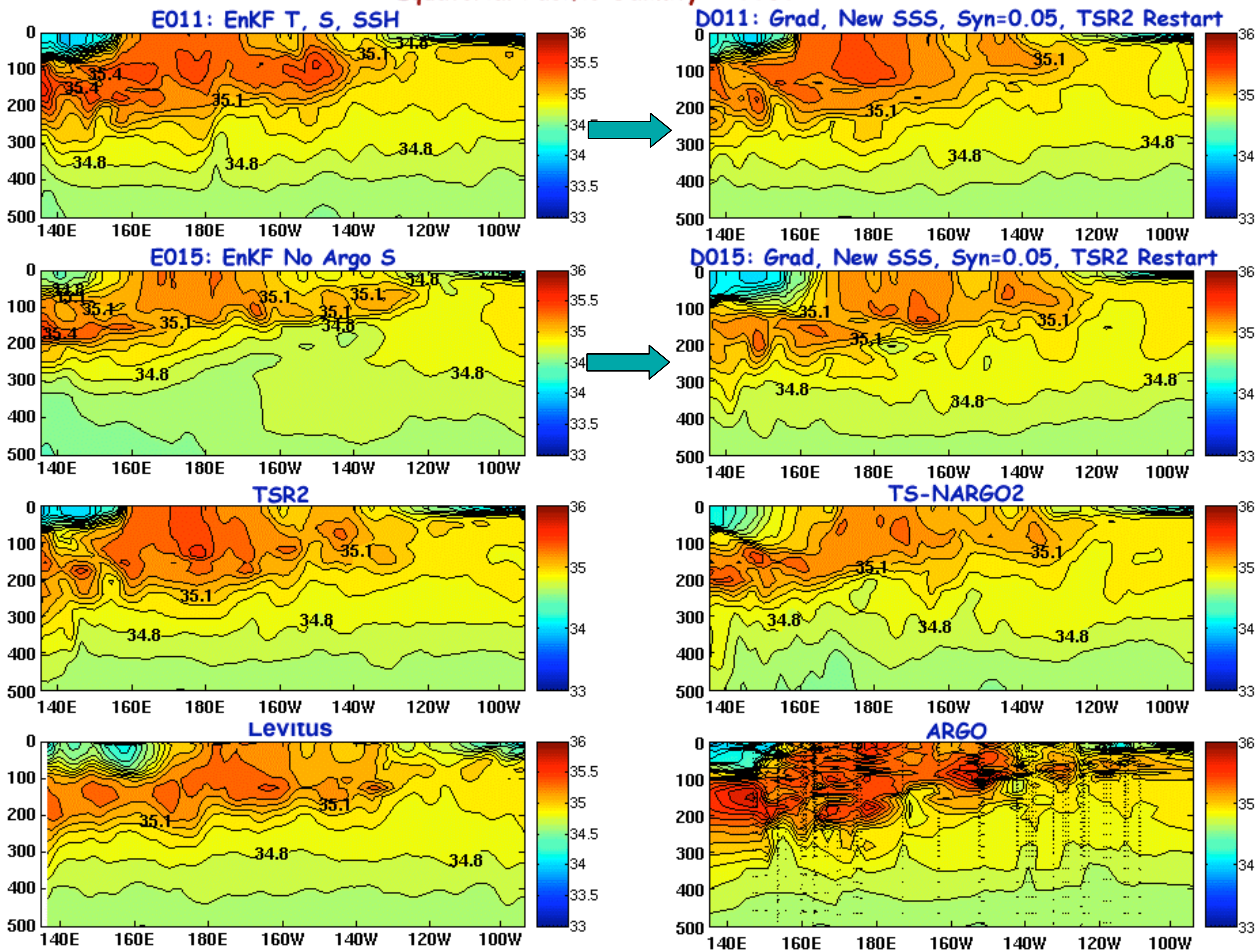




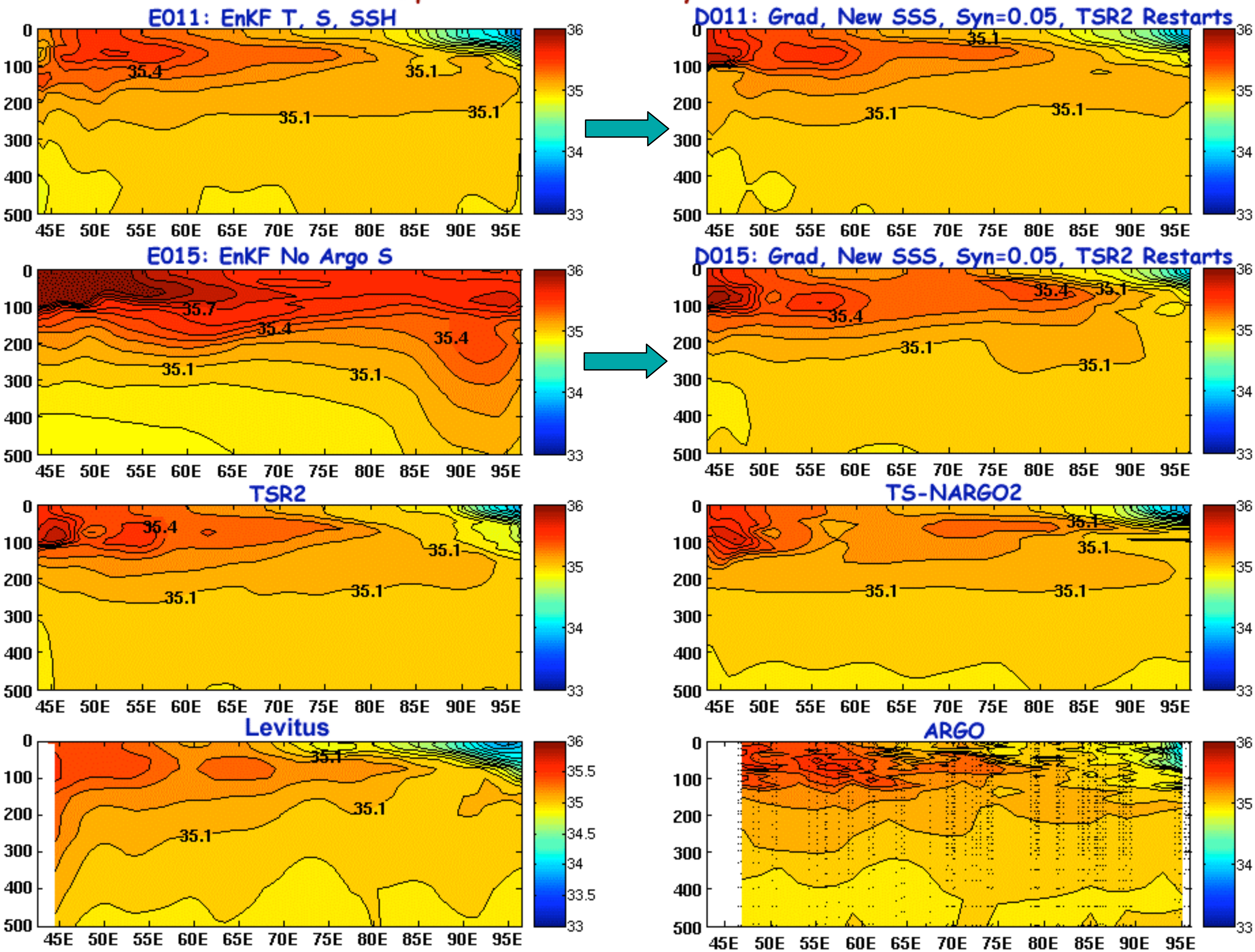
Independent Validation
RMSD of analysis c.f.
TAO ADCP zonal currents
1993-2006



Equatorial Pacific Salinity: 200602

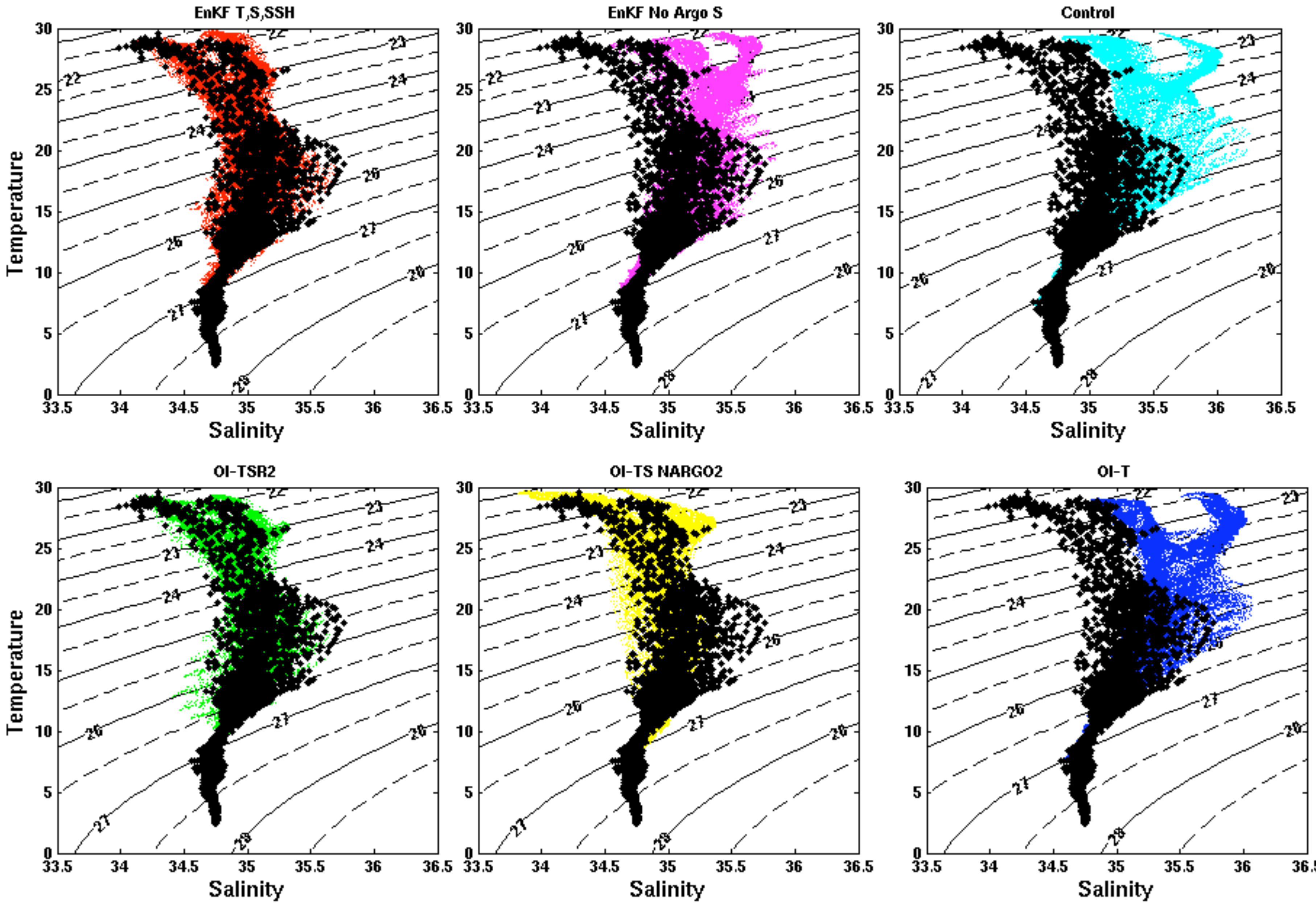


Equatorial Indian Salinity: 200602



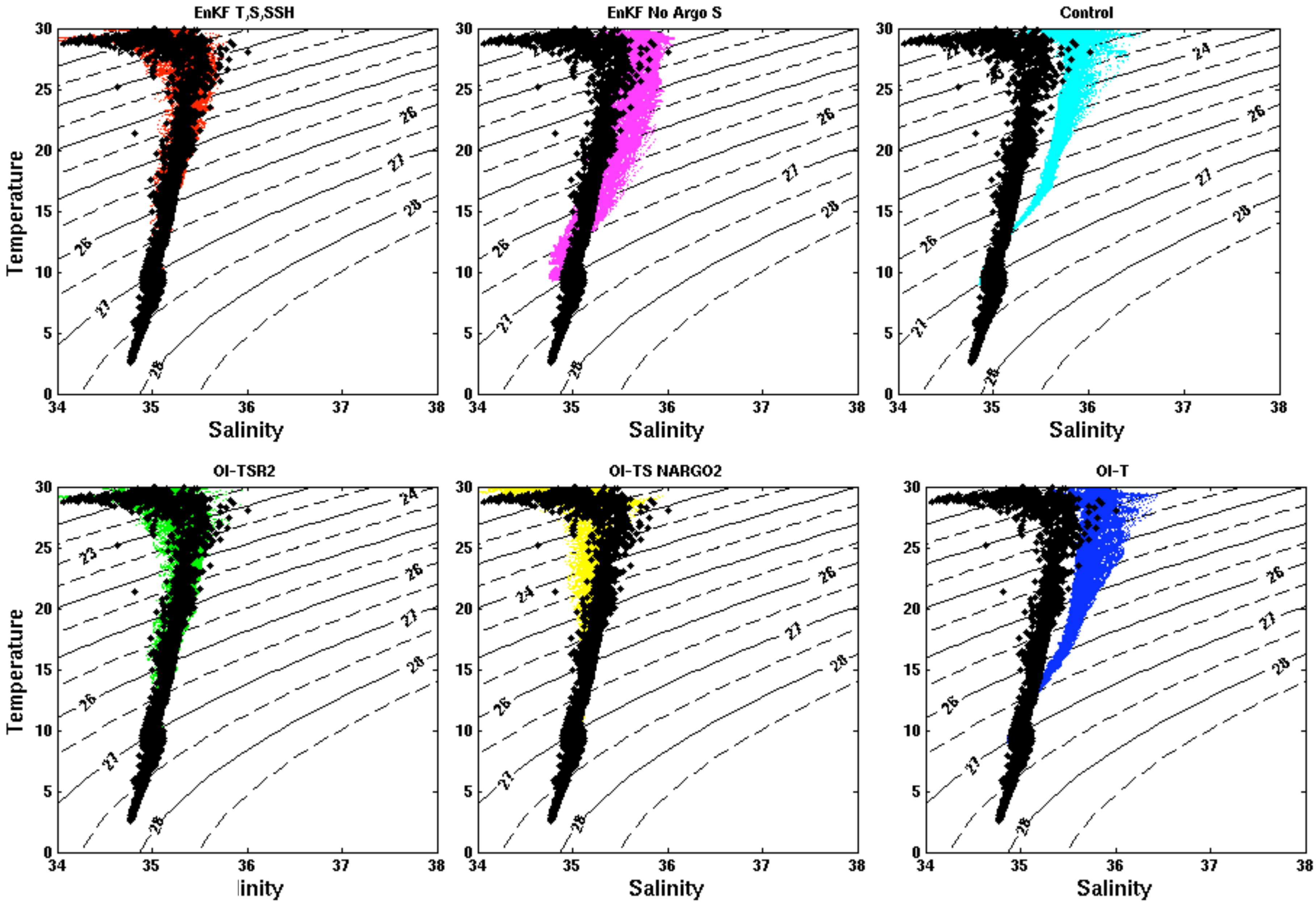
Temperature-Salinity Diagrams (Density Contours in kg/m^3)

Central Indian Ocean : 70E-75E, 10S-15S (Black Dots are Argo T-S Values)



Temperature-Salinity Diagrams (Density Contours in kg/m^3)

Equatorial Indian Ocean : 70E-75E, 2.5S-2.5N (Black Dots are Argo T-S Values)



Forecast skill (ACC) from CGCMv1
Heat content anomaly in upper 300m
1993-2006

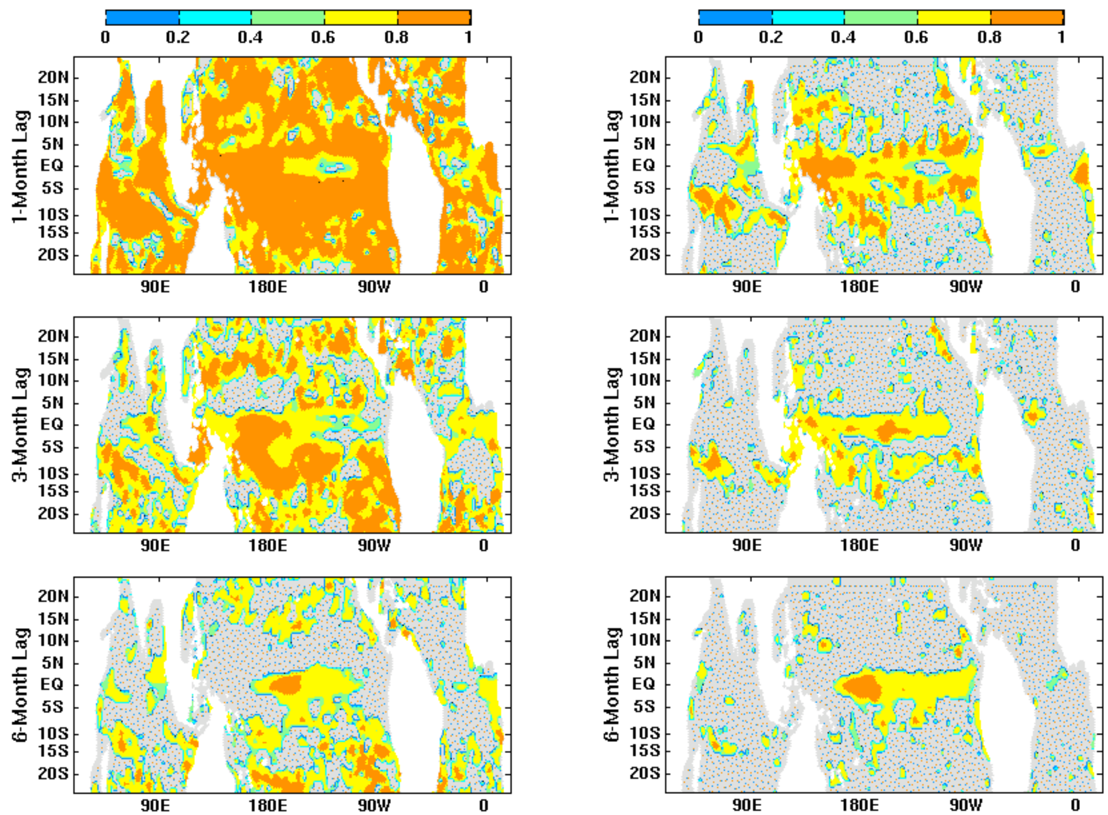
EnKF

OI-TS

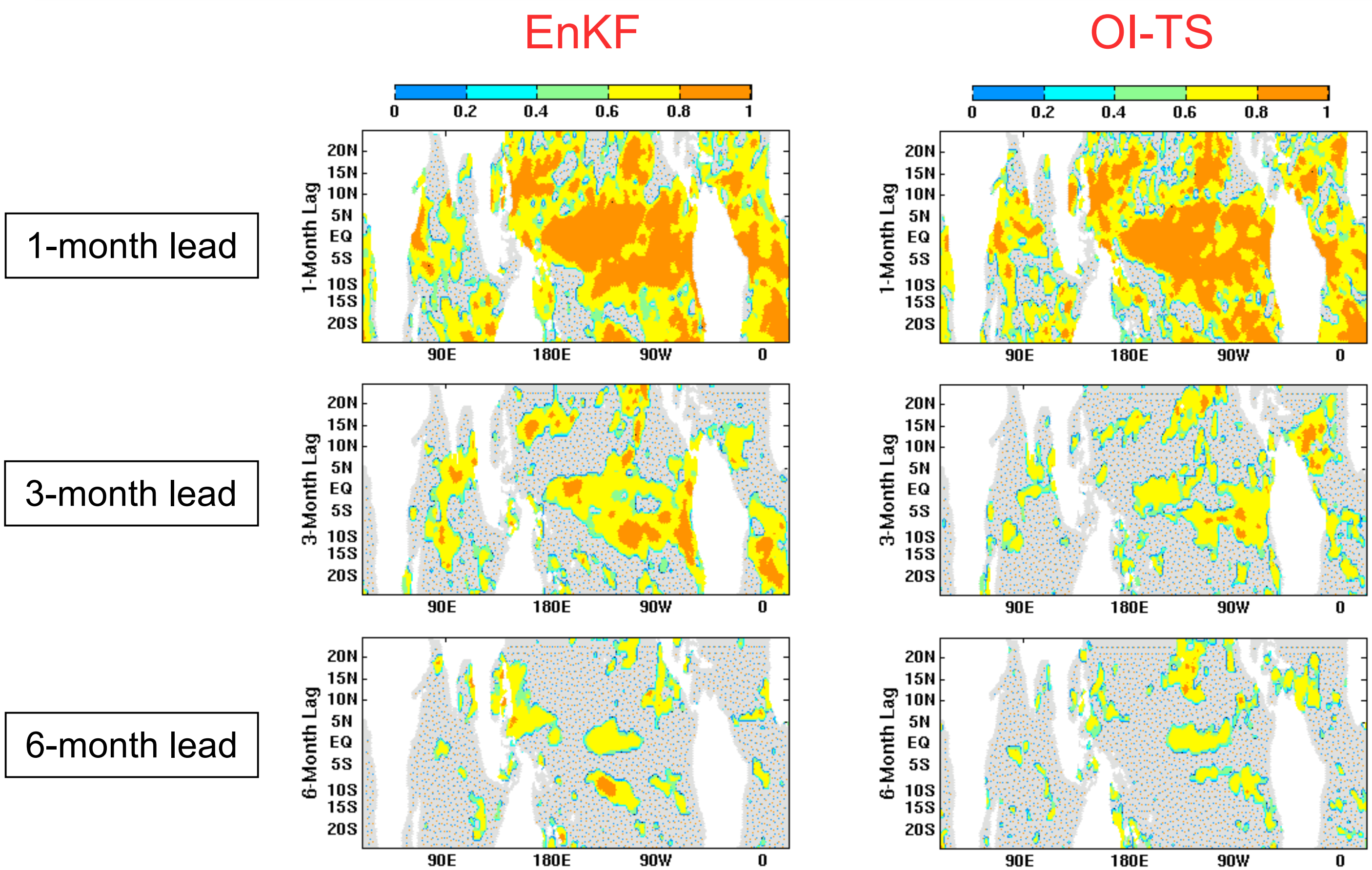
1-month lead

3-month lead

6-month lead

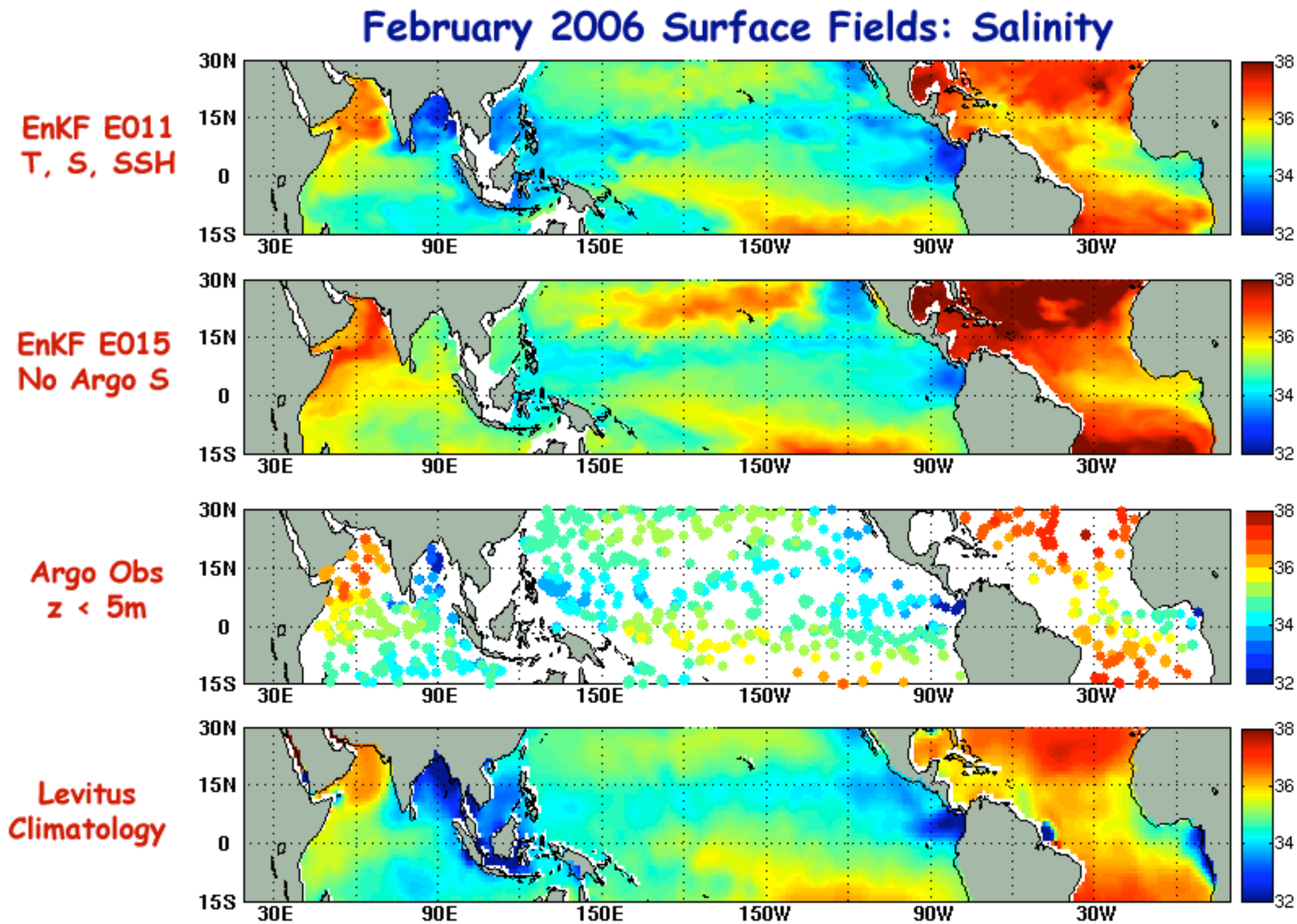


Forecast skill (ACC) from CGCMv1
SST anomaly
1993-2006



The impact of Argo - preparing for Aquarius

Christian Keppenne and Robin Kovach



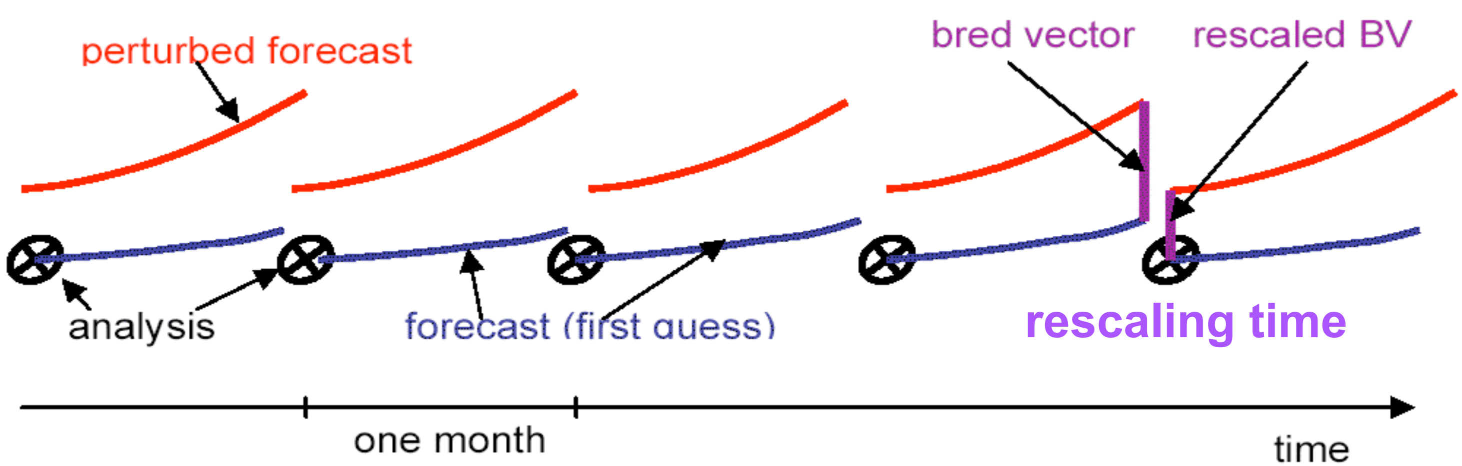
Augmenting covariance estimates with information from bred vectors

Shu-Chih Yang
Christian Keppenne, Eugenia Kalnay

Background

- Coupled breeding technique is designed to capture the growing errors related to slow-varying coupled instabilities, like ENSO-related growing errors.
- Breeding is a nonlinear approach and tightly related to the Ensemble Kalman Filter.
- Coupled breeding is implemented in the NASA/GMAO coupled general circulation model (CGCM). The applications of bred vectors (BVs) are explored for the purpose of improving couple forecasting:
  - use BVs as the initial ensemble perturbations of the ensemble forecast system for ENSO prediction
  - Augment the background error covariances in ocean data assimilation system with the structure of BVs.

Breeding in the GMAO coupled GCM



- **NASA/GMAO coupled GCM (Poseidon+ NSIPP-1 AGCM)**
- **Bred vectors** : Differences between the control forecast and perturbed run
- Coupled breeding cycle needs to choose physically meaningful breeding parameters in order to choose the type of instability

Coupled Bred vectors

- 4 different rescaling norms are chosen to measure the coupled atmosphere-ocean instability (10% of Climate variability, **rescale every month**)

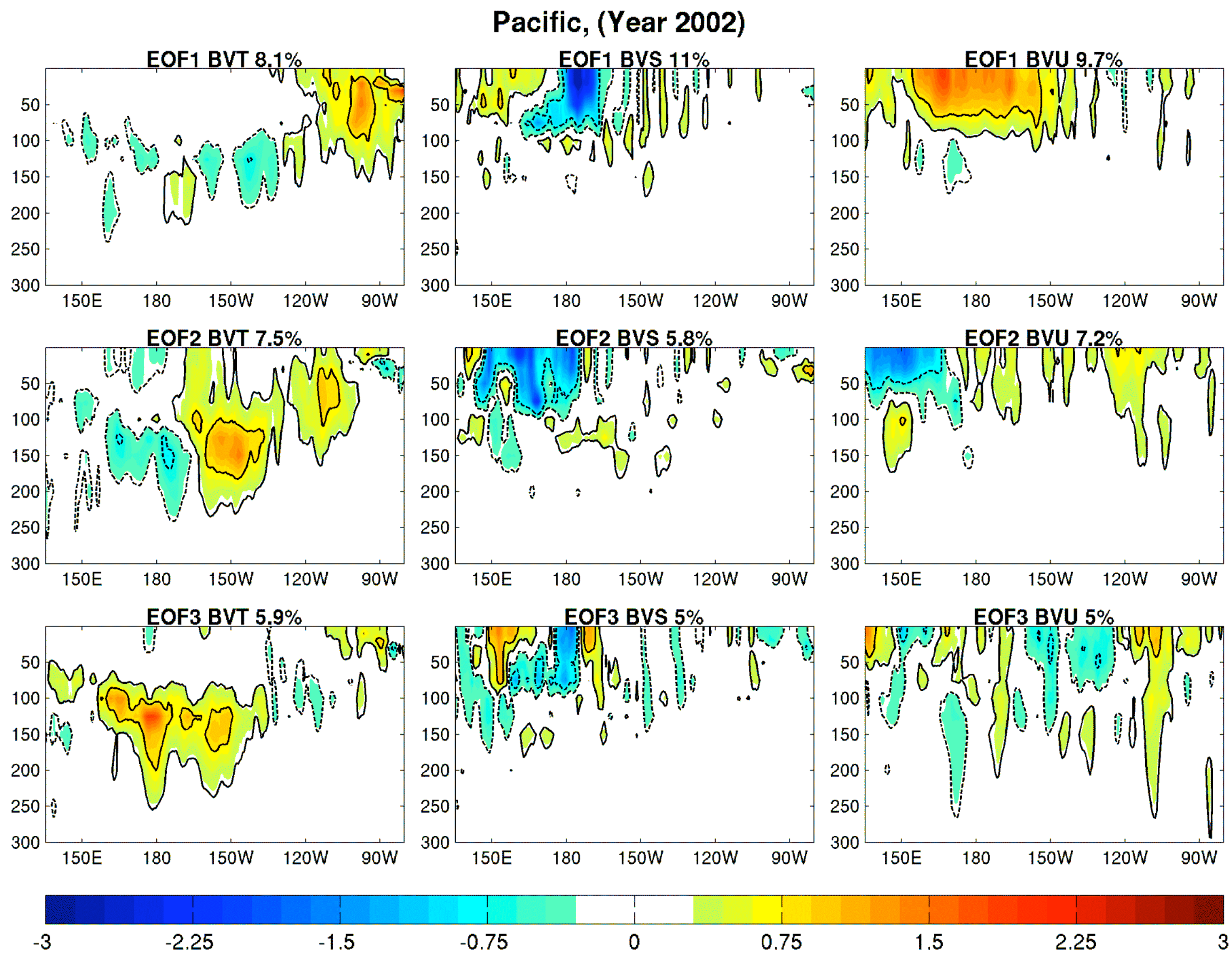
  1. **$|SST_{BV}|=0.1°C$** (in 150°W~90°W, 5°S~5°N)
  2. **$|D20_{BV}|=1.5$ m** (in 160°E~140°W, 2.5°S~2.5°N)
  3. **$|[u'_{BV},h'_{BV}]|=6.5\times10^{-3}$** (in 130°E-80°W, 5°S~5°N)

     >> the first 4 long wave modes (Kelvin+3 Rossby waves)
  4. **$|[u_{BV}\tau_{xc}+u_c\tau_{xBV}]|=0.1$** (in 130°E-80°W, 5°S~5°N)

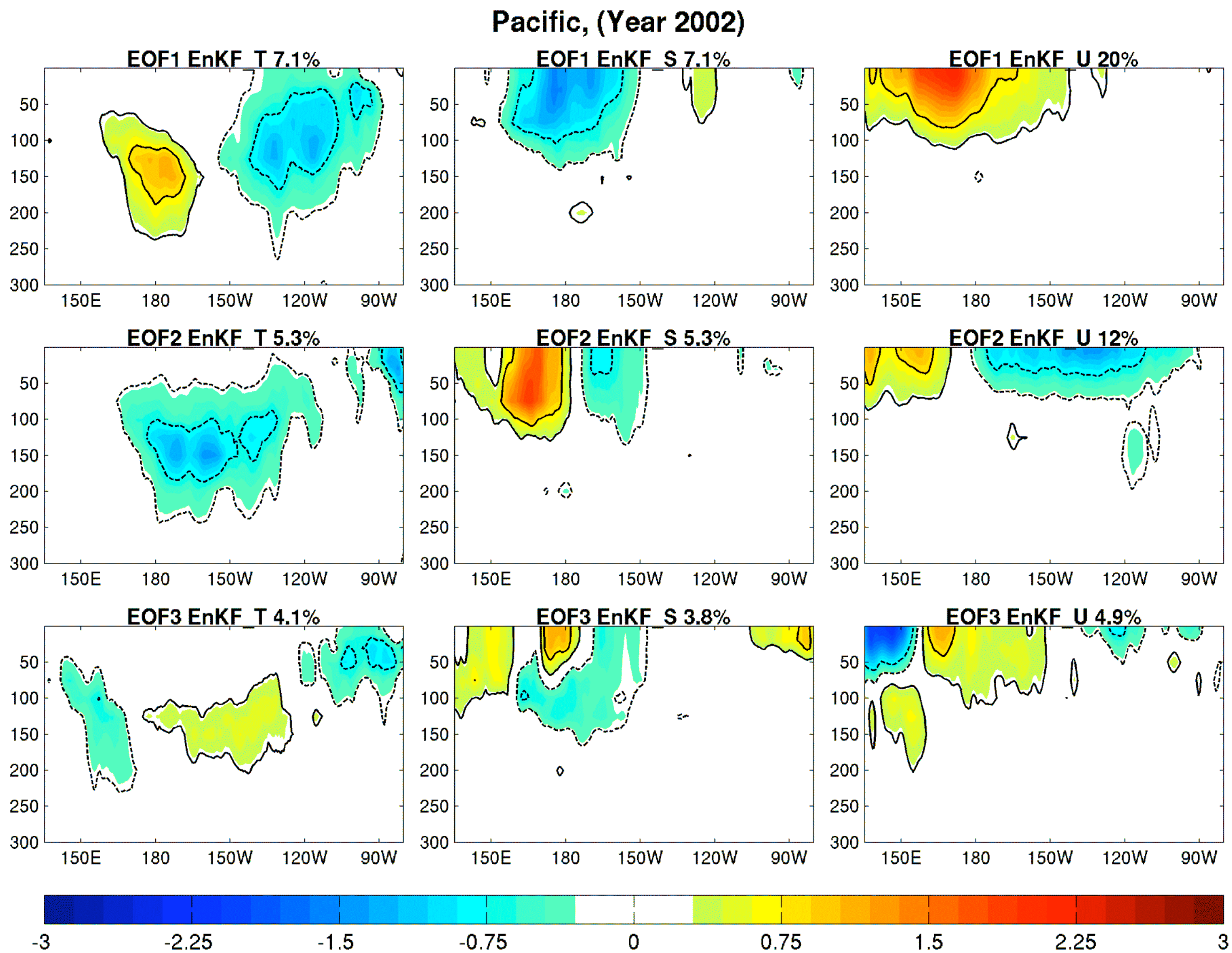
     >> work done on the ocean by the atmosphere (Goddard and Philander, 1999)

- Initial conditions for CGCM:
  - Ocean analysis (T, S assimilated with optimal interpolation scheme) + AMIP restart
  - 4 pairs of ± coupled BVs are centered at this initial condition

Dominant growing modes from BVs in Pacific



Dominant growing modes from EnKF in Pacific



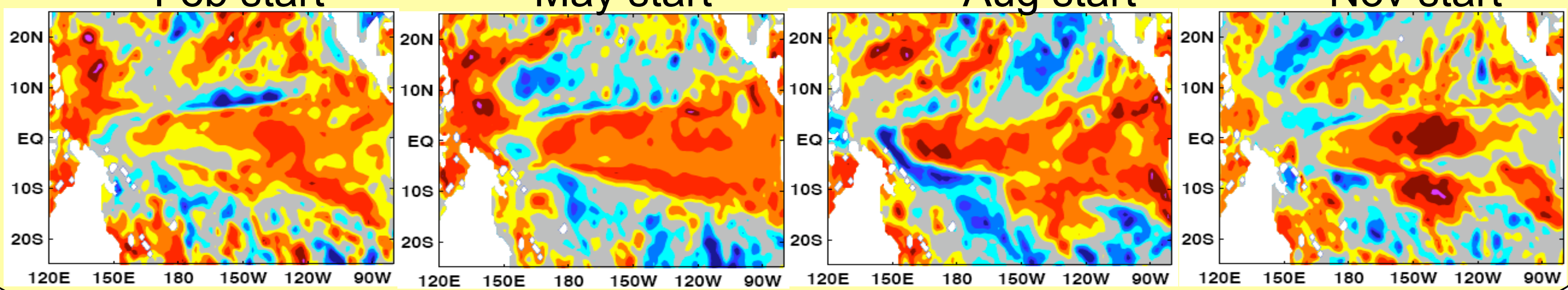
Ensemble forecasts initialized from 4 ±BVs

pattern correlation: SSTA vs. Reynolds SSTA at 9-month lead time (1993~2002)

4 BV ensemble mean has higher skill than control

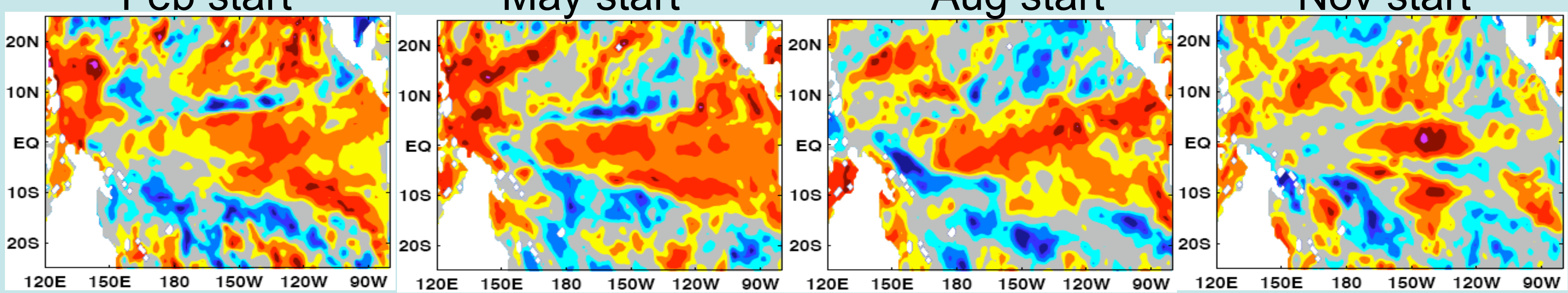
BVs ensemble mean

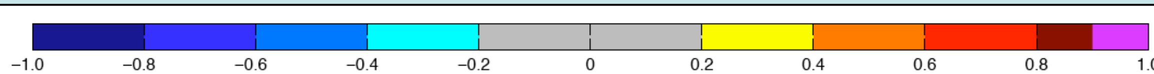
Feb start | May start | Aug start | Nov start



Control

Feb start | May start | Aug start | Nov start





Ensemble-based covariance in hybrid-OI scheme

$\mathbf{P}_f = (1-\alpha)\,\mathbf{P}_{\mathrm{OI}} + \alpha\,\mathbf{P}_f^0$

$\mathbf{P}_f$: the background error covariance

$\mathbf{P}_f^0$: Ensemble-based background error covariance

$\mathbf{P}_{\mathrm{control}}$: Gaussian covariance ($x_s$=20º, y_s=5º, z_s=100m)

α : the hybrid coefficient

$\mathbf{x}^a - \mathbf{x}^f = \mathbf{K}[\mathbf{y} - H(\mathbf{x}^f)] = \mathbf{Kd}$ (analysis increment)

$$\begin{aligned}
\mathbf{Kd} &= \mathbf{P}_f\mathbf{H}^T[\mathbf{H}\mathbf{P}_f\mathbf{H}^T + \mathbf{R}]^{-1}\mathbf{d} \\
&= (\mathbf{P}_f^0 + \mathbf{P}_{contr.})\mathbf{H}^T[\mathbf{H}(\mathbf{P}_f^0 + \mathbf{P}_{contr,})\mathbf{H}^T + \mathbf{R}]^{-1}\mathbf{d} \\
&= \mathbf{P}_f^0\mathbf{H}^T[\mathbf{H}(\mathbf{P}_f^0 + \mathbf{P}_{contr.})\mathbf{H}^T + \mathbf{R}]^{-1}\mathbf{d} + \mathbf{P}_{contr.}\mathbf{H}^T[\mathbf{H}(\mathbf{P}_f^0 + \mathbf{P}_{contr.})\mathbf{H}^T + \mathbf{R}]^{-1}\mathbf{d}
\end{aligned}$$

$\boldsymbol{d}$: the difference between forecast and observations (innovation vector)

α =0 : Fully $\mathbf{P}_{\mathrm{control}}$, approximate to Univariate OI

Assimilation experiment setup

Observations	TAO, XBT, ARGO, Pirata
Assimilation interval	4-day (Jan2002 ~ Dec2002)
Covariance localization for $\mathbf{P}_f^0$	x_s=8º, y_s=4º, z_s=100m
Gaussian horizontal filter for $\mathbf{P}_f^0$	x_f=4º, y_f=2º
Background error	σ_T=0.7°C, σ_S= 0.1psu

Experiments:

(1) only the Gaussian function (control)

- used as the benchmark

(2) P_f is based on 4 EOF modes

- EOFs are constructed from long and large ensemble runs

(3) P_f is based on 4 BVs (updated every month)

(4) P_f is based on 4 BVs (updated every 4 days by linear interpolation)

Normalized Error covariance structure

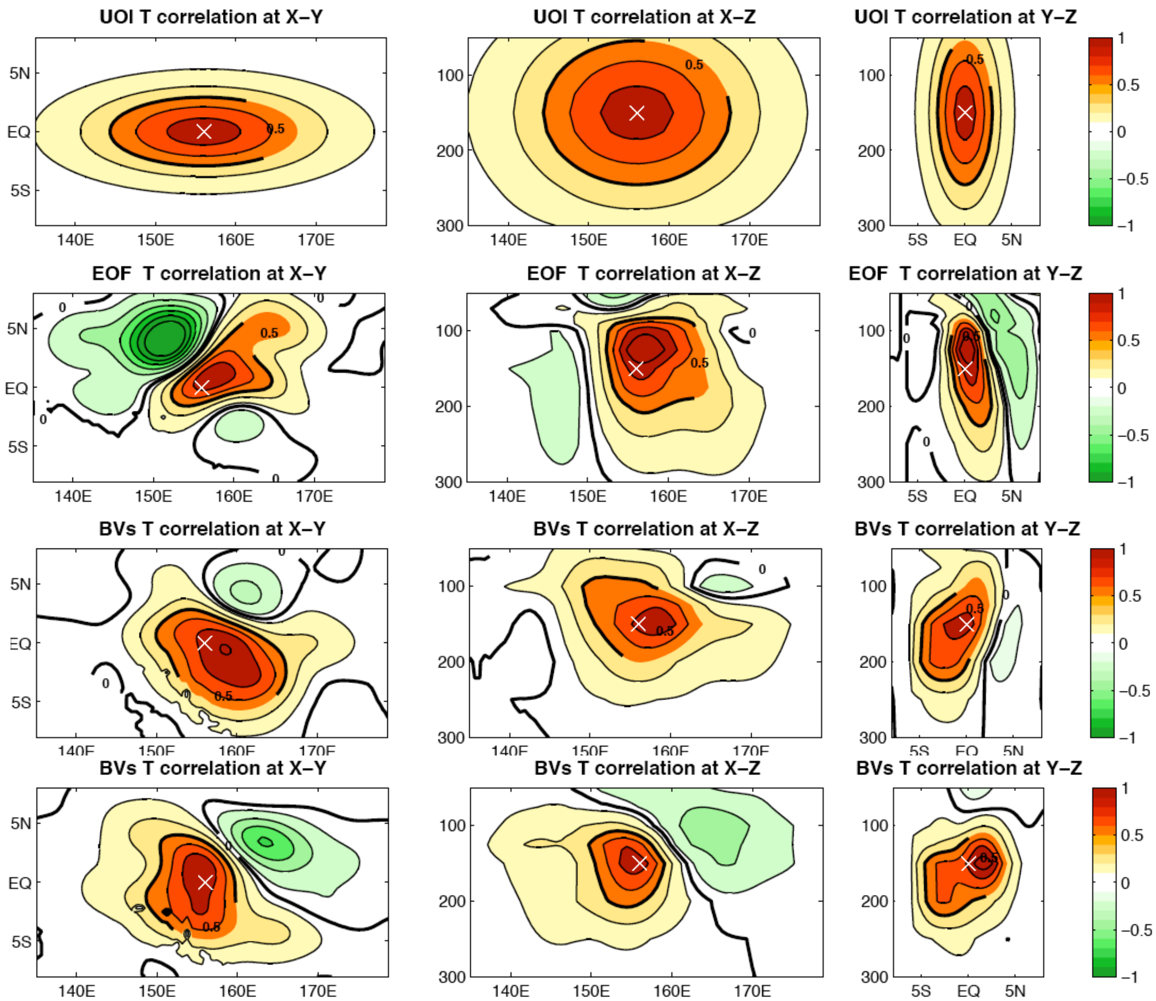
Temperature correlation of the location at (156°E, EQ, 150m)

control (UOI)

4EOF

4BVs

4BVs

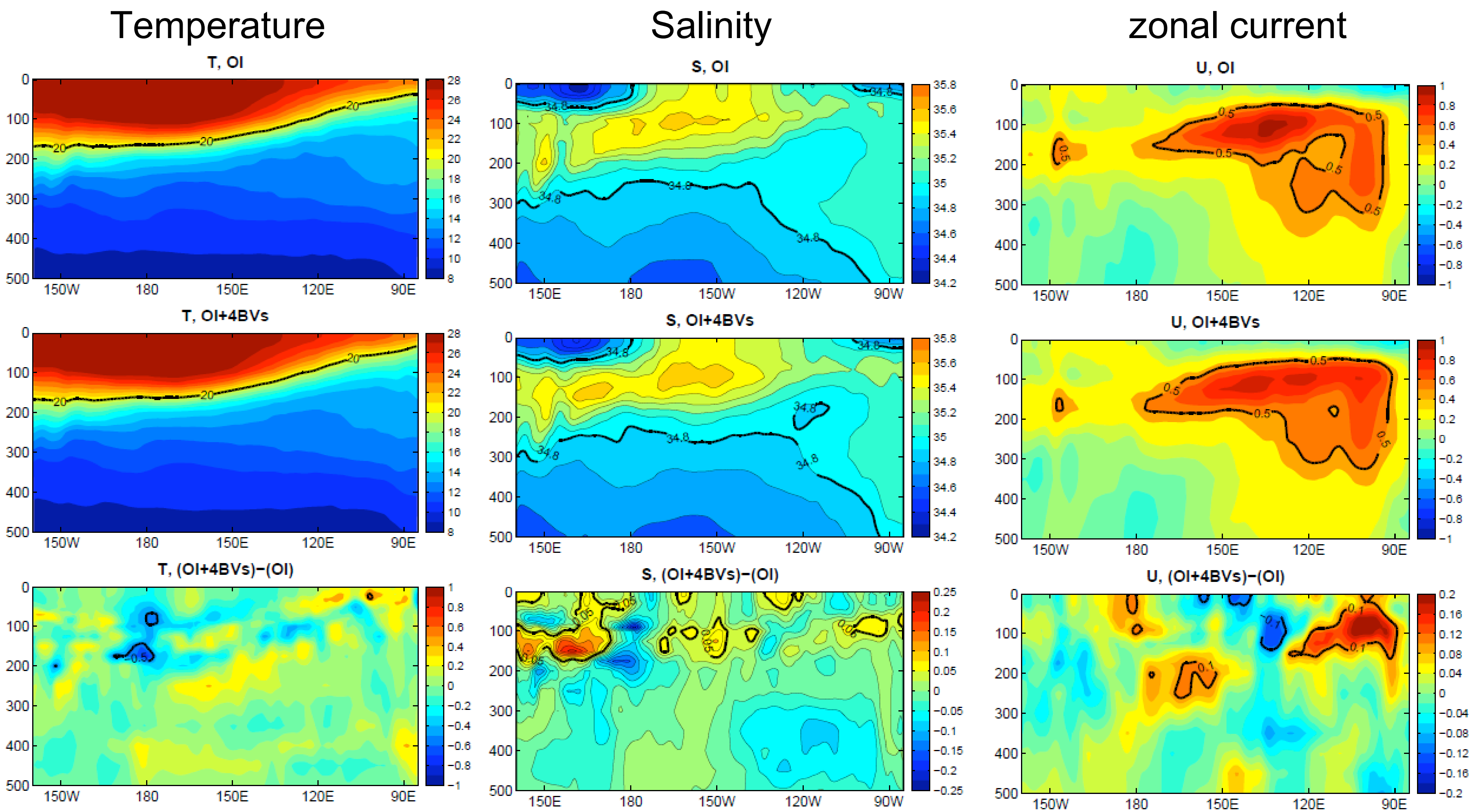


Fixed in time

Warm event

Cold event

Annual Mean 2002 structure in the equatorial Pacific

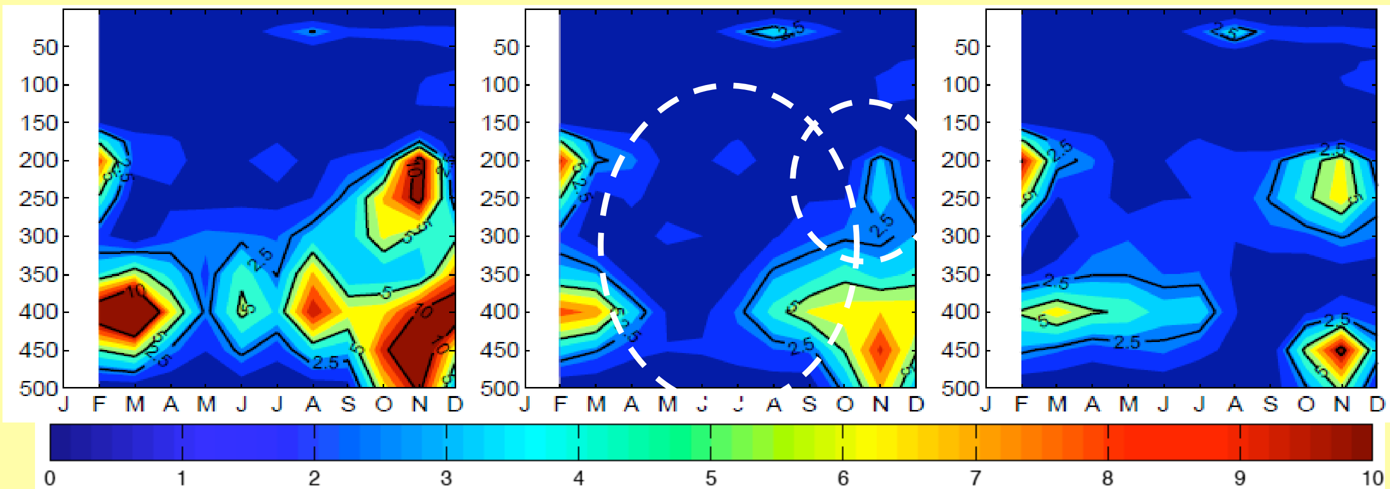


Comparisons with independent observations

Temp. observations from **G**lobal **T**emperature **S**alinity **P**rofile **P**rogram

[GTSPP T profiles – monthly analysis] in Niño3 region

control 4BVs_4day 4EOF

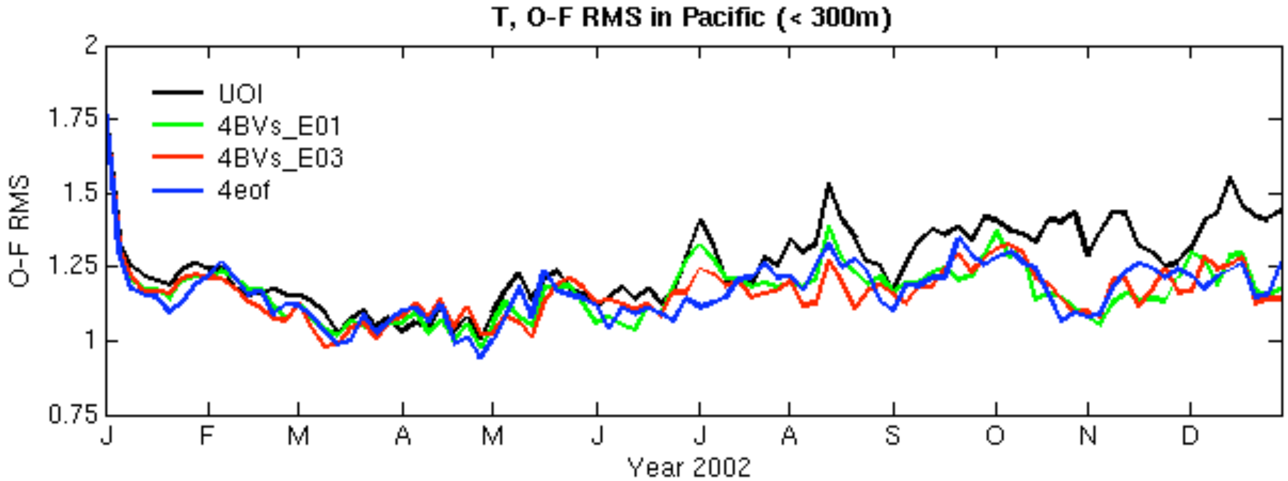


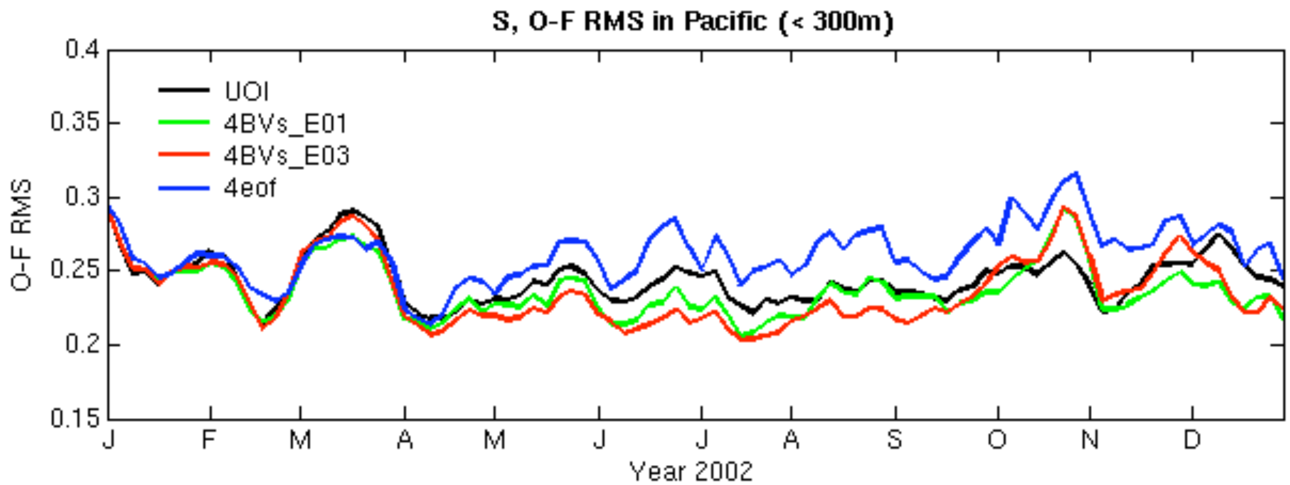
- Both the 4BV_4day and 4EOFs runs show improvement over the Control.
- The 4BV_4day run has positive impact on (i) summer season and (ii) the upper ocean of Nov&Dec.

RMS of Temp./Salin OMF in Pacific

E01: 4-day BVs

E03: monthly tendency BVs





Summary

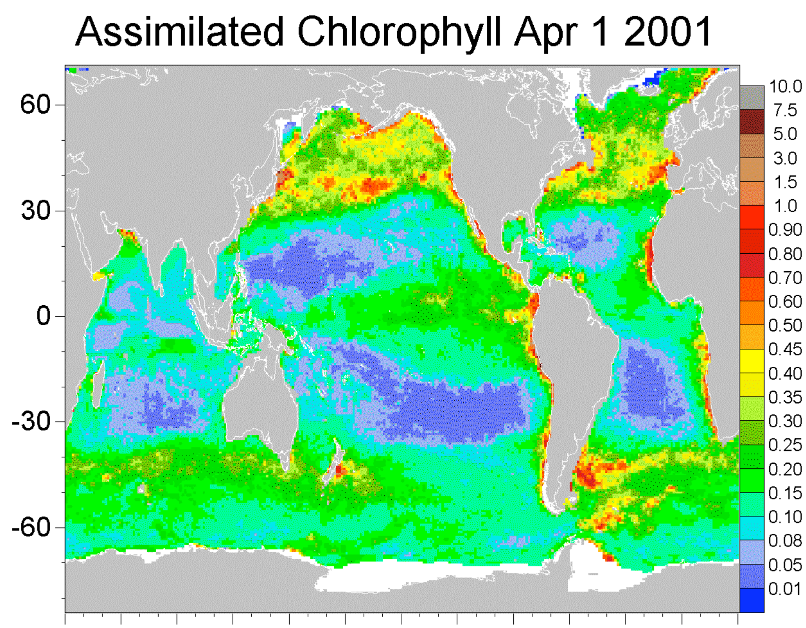
- Ensemble forecasts initialized from 4 coupled ±BVs have increased skill when starting from cold phase of the annual cycle.
- Augmenting the Gaussian background error covariance by 4BVs (a hybrid system) has positive impact when assimilating real T and S observations.
- The optimal hybrid weighting is 30-40% of the total background error covariance.
- Overall, between the two hybrid experiments, the one with the BVs applicable at the analysis time (BVs_4day) generates the better T and S analyses.
  - For T, the improvement over the control is seen in the tropical Pacific.
  - For S, the improvement is mainly located in the western Pacific during late spring to summer season.
  - BVs_4day carries the error structures most dynamically relevant to the slowly growing mode.

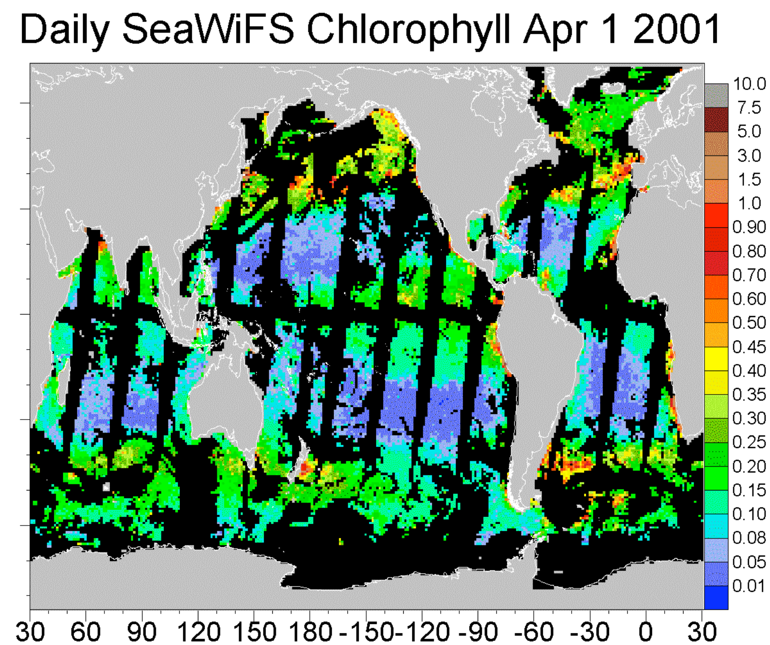
Ocean Color Assimilation

Watson Gregg and Lars Nerger

Ocean Color Data Assimilation complete, products available GES-DISC Giovanni (http://reason.gsfc.nasa.gov)

Goal: Consistent (climate) products from CZCS - MODIS

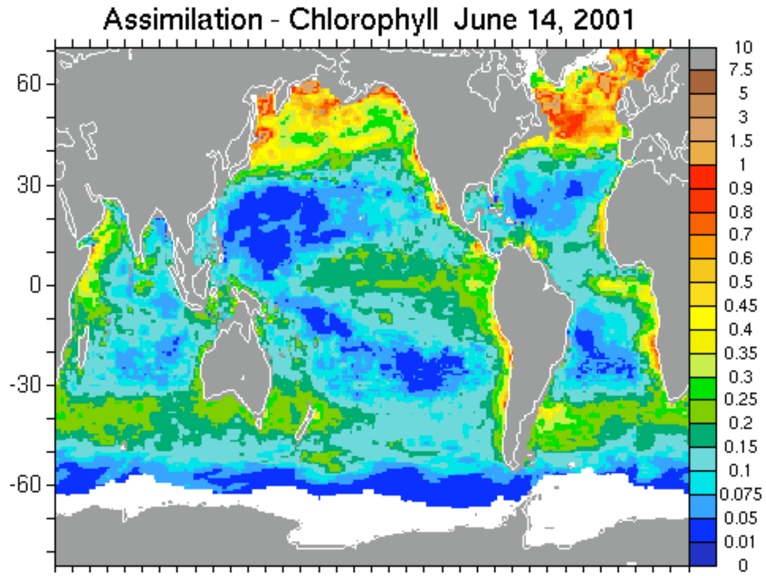


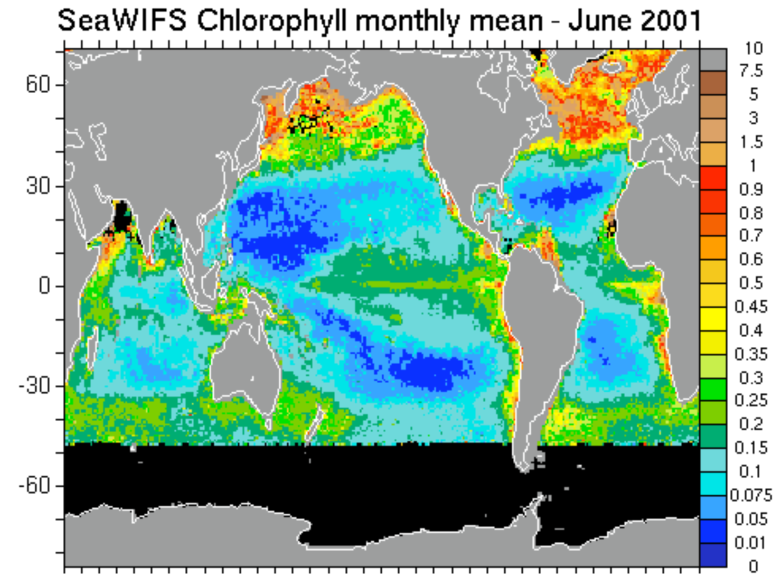


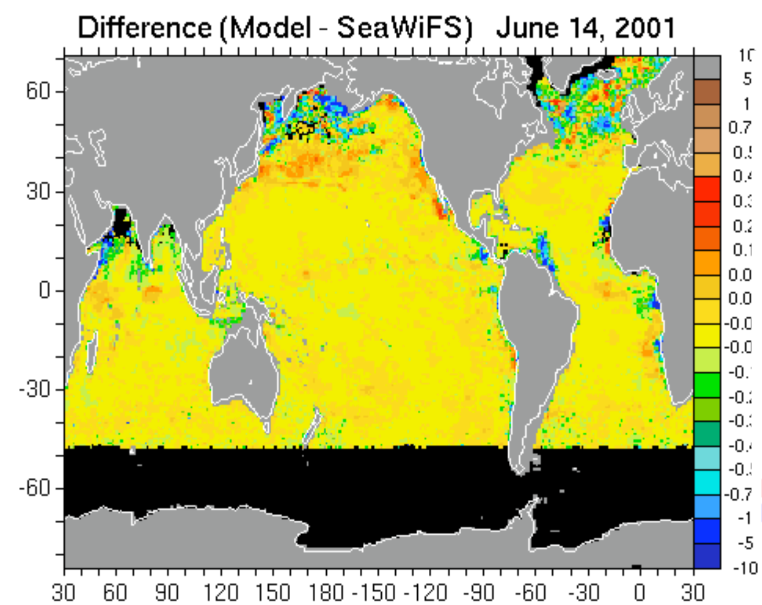
http://gmao.gsfc.nasa.gov/research/oceanbiology/

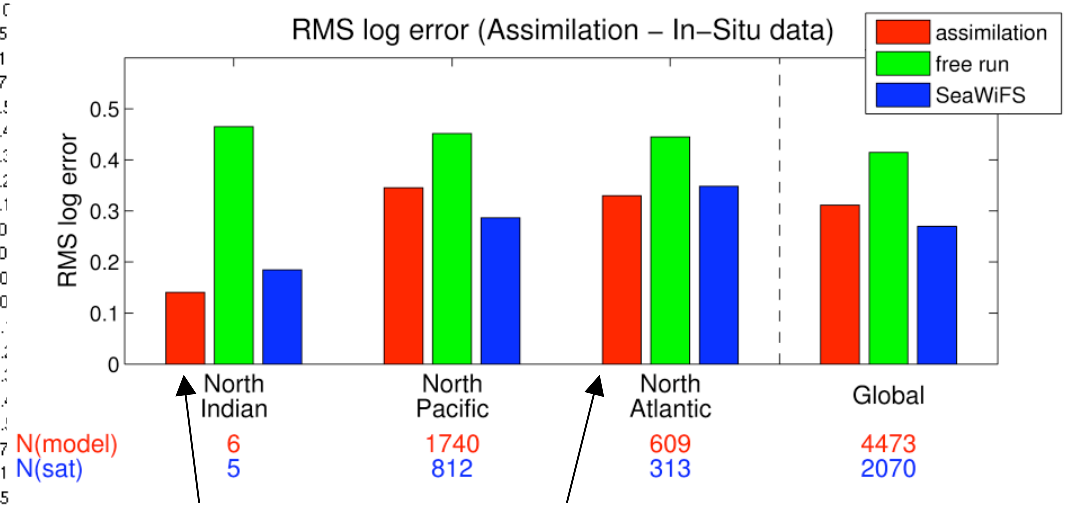
Constraining a Global Three-Dimensional Ocean Biogeochemical Model by SeaWiFS Ocean Chlorophyll Data Using a Local SEIK Filter

Lars Nerger, Watson W. Gregg





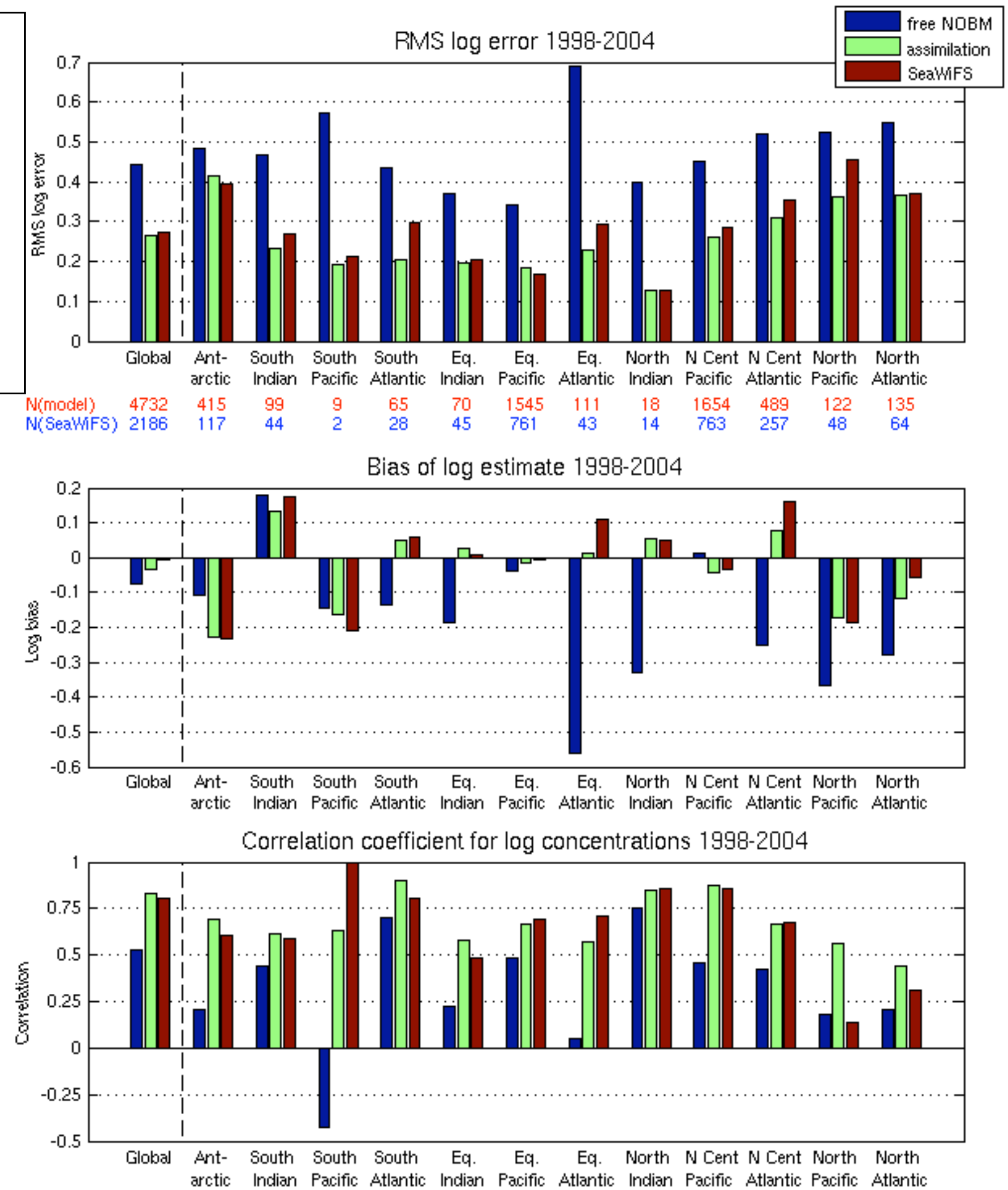




Smaller error than SeaWiFS

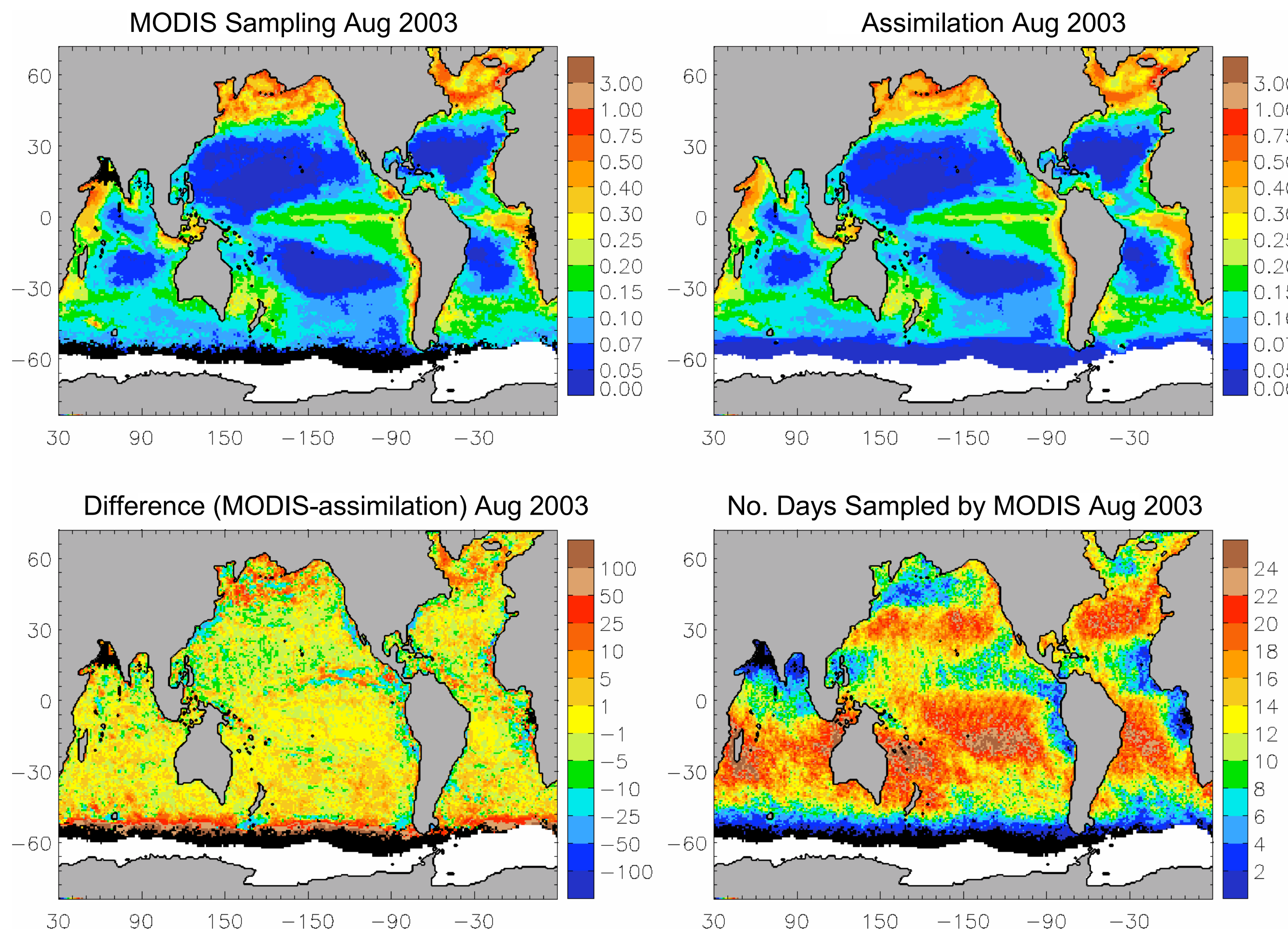
Comparison of the surface chlorophyll from free-running model, assimilation, and SeaWiFS with in situ data for 1998-2004: globally and separated over 12 major oceanographic basins.

Nerger and Gregg, 2007
J. Mar. Syst. (submitted)



Assimilation helps to identify sampling biases in MODIS ocean chlorophyll

Sampling biases in MODIS ocean chlorophyll were determined by "flying" the MODIS daily sampling over the complete daily coverage provided by data assimilation. The results showed that MODIS annual mean chlorophyll estimates are about 8% too high. Considering that the maximum interannual variability in the 10-year SeaWiFS record is about 3%, this sampling bias should be considered when making estimates of global chlorophyll.



Some References:

• Keppenne, C.L., et al. 2007: Error covariance modeling in the GMAO ocean assimilation Ensemble Kalman filter, *Mon. Wea. Rev. (to appear)*

• Keppenne, C.L., et al., 2005: Ensemble Kalman filter assimilation of temperature and altimeter data with bias correction and application to seasonal prediction. *Nonlin. Proc. Geophys.,* **12**, 491-503

• Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99**, 10,143-10,162.

• Houtekamer, P., and H. Mitchell, 2005: Ensemble Kalman filtering, *Q. J. Roy. Met. Soc.,* **131C**, 3269-3289.

• Yang, S.-C., et al., 2007: Bred vectors and tropical Pacific forecast errors in the NASA coupled general circulation model, *Mon. Wea. Rev. (to appear).*

• Yang, S-C, M. Cai, E. Kalnay, M. Rienecker, G. Yuan and Z. Toth, 2006: ENSO bred vectors in coupled ocean-atmosphere general circulation models. *J. Climate,* **19**, 1422-1436.

• Nerger, L. and W.W. Gregg, 2007: Improving Assimilation of SeaWiFS Data by the Application of Bias Correction with a Local SEIK Filter**,** *J. Mar. Syst.* (submitted).