Seasonal Prediction Initialization with the GMAO Ocean EnKF

Christian Keppenne[1], Michele Rienecker[2], Jossy Jacob[3] & Robin Kovach[4]

[1,3,4,]Science Applications International Corporation
San Diego, CA 92121

[2]Global Modeling and Assimilation Office
NASA Goddard Space Flight Center
Greenbelt, MD 20771

AMS 2007 Meeting, San Antonio, TX

[1]clk@gmao.gsfc.nasa.gov, [2]michele.rienecker@nasa.gov, [3]jjacob@gmao.gsfc.nasa.gov , [4]kovach@gmao.gsfc.nasa.gov

Objective: Transit from R&D status to a production application

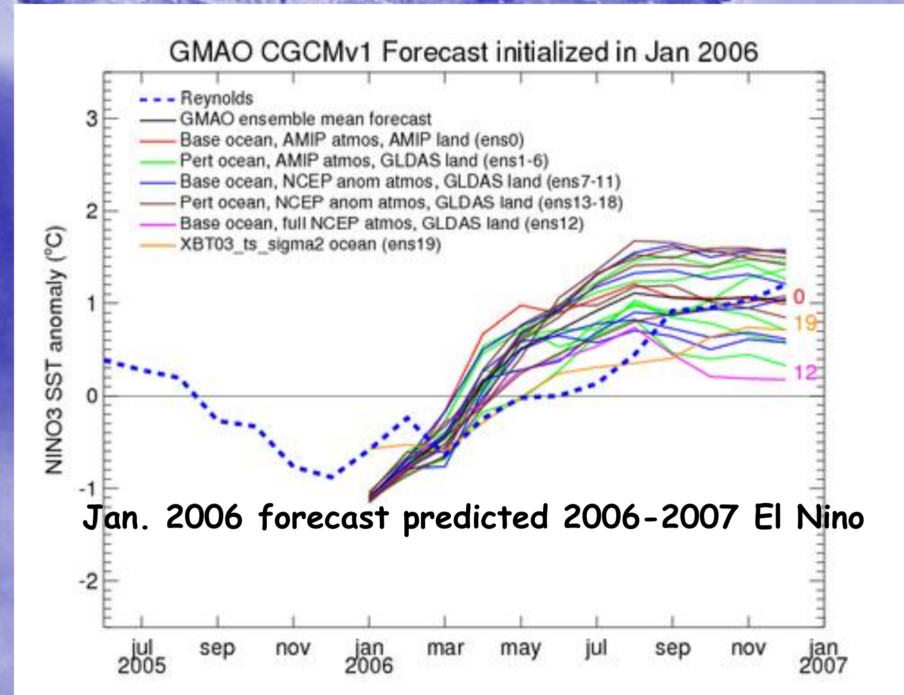
- Analysis speed up
- Good performance with very small ensembles
- Match the performance/outperform production ODAS

Outline

- GMAO coupled S/I forecasting system
- Acceleration of EnKF analysis
- Background covariance conditioning and consequences
- Impact of T+SSH assimilation on CGCM seasonal hindcast skill

GMAO CGCM coupled forecasting system

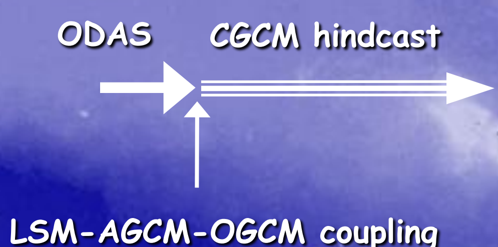
- *GMAO AGCM*: 90 x 144 x L34 (2° x 2.5° x L34)
- Poseidon v4 *OGCM*: 538 x 572 x L27 ((1/3° x 5/8° x L27)
- LSM: Mosaic catchment model



Poseidon v4 OGCM (Schopf and Loughe, 1995)

- Quasi-isopycnal vertical coordinate
- Prognostic variables are *h*, *t*, *s*, *u* and *v*
- Sea surface height (SSH) is diagnostic: $\eta = \Sigma_i \mathrm{buoyancy}(t_i, s_i) h_i / g$
- About 30 million prognostic variables at current resolution

Production ODAS: *OI of in situ T profiles with S(T) correction*



Ocean EnKF (MWR 130, 2951-2965, 2002; JMS 40-41, 363-380, 2003; NPG 12, 491-503, 2005)

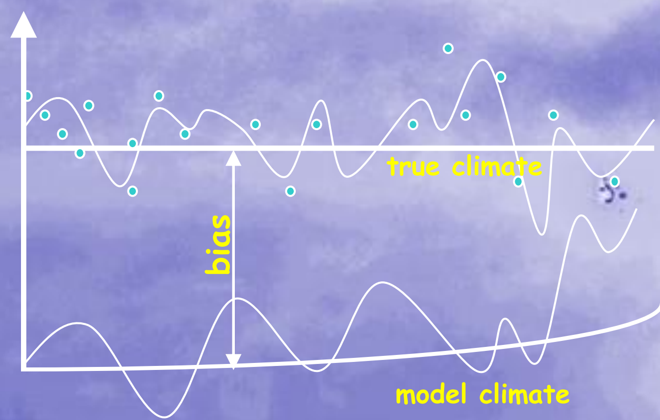
- Multivariate compactly supported background covariances: updates T, S, u & v
- Massively parallel: same domain decomposition used for model and analysis
- System noise representation: Model-error and forcing-error model
- Online bias correction used in SSH assimilation

Ocean EnKF (MWR 130, 2951-2965, 2002; JMS 40-41, 363-380, 2003; NPG 12, 491-503, 2005)

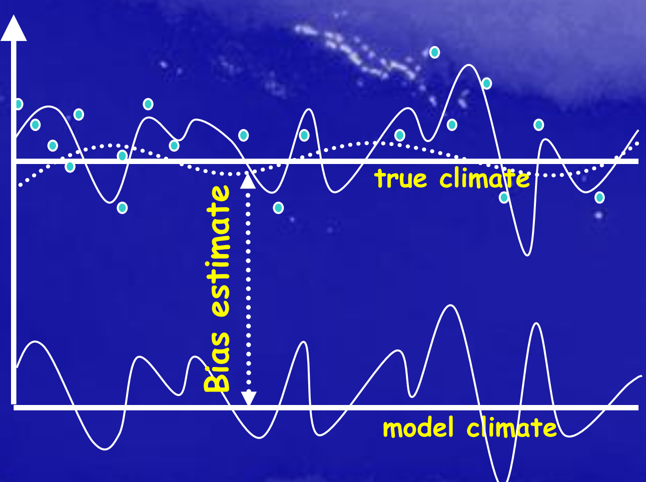
- Multivariate compactly supported background covariances: updates T, S, u & v
- Massively parallel: same domain decomposition used for model and analysis
- System noise representation: Model-error and forcing-error model
- Online bias correction used in SSH assimilation

T/P altimetry data are anomalies
Hence 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)

Compactly supported EnKF (bias estimation omitted)

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

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

$$\boldsymbol{HP}^{\mathrm{f}}\boldsymbol{H}^{\mathrm{T}} = \frac{1}{\mathrm{n}-1}\boldsymbol{SS}^{\mathrm{T}}, \qquad (1c)$$

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

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

Analysis speed up

- Costliest step of analysis algorithm is (1e)
    - Perform the matrix-vector multiplications S^T ($C_l \circ a_i$) on coarse grid
    - Bilinearly imterpolating to the fine grid
    - Perform every other operation on fine grid

Coarsening ratio	1 (no coarsening)	4	16	64
Wallclock time (S) (with 641s ensemble integration)	544 (1185)	152 (793)	43 (684)	16 (657)
RMS OMA (m)	0.05631	0.05634	0.05639	0.05647
Performance loss (%)		0.15	0.4	0.8

Table 1. Wallclock time and RMS OMA difference for one SSH analysis cycle with 17 ensemble members on 128 SGI® Altix™ CPUs. The parenthesized times include the 2-day 17-member ensemble integration preceding the analysis.

Compactly supported EnKF (bias estimation omitted)

Time filter: $$\mathbf{x}^{\Phi}_{i,k} = (1-\beta)\mathbf{x}^{\Phi}_{i,k-1} + \beta \mathbf{x}^{f}_{i,k} \qquad , i = 1,\ldots,n,$$

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

$$\mathbf{S} = \{\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n\} = \{\mathbf{H}(\Phi(\mathbf{x}^{f}_1 - \bar{\mathbf{x}}^{f})), \mathbf{H}(\Phi(\mathbf{x}^{f}_2 - \bar{\mathbf{x}}^{f})), \ldots, \mathbf{H}(\Phi(\mathbf{x}^{f}_n - \bar{\mathbf{x}}^{f}))\}, \qquad (1b)$$

Spatial filtering operator $$\mathbf{HP}^{f}\mathbf{H}^{T} = \frac{1}{n-1}\mathbf{SS}^{T}, \qquad (1c)$$

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

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

Improving the performance for small ensembles

Spatio-temporal filtering of background-error covariances

- Temporal filter applied to X^f integration (exponential moving average)
- Spatial filter applied to (X^f -<X>) deviations (Gaussian filter)

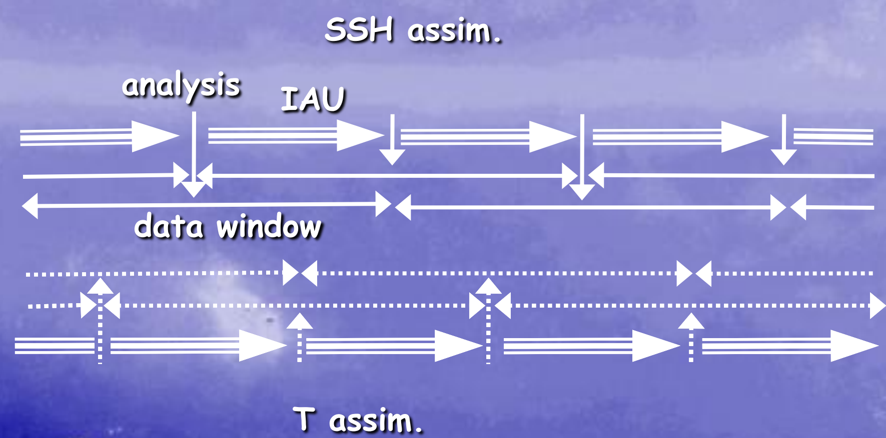
Effect of time filtering

β	1 (no time filter)	0.01	0.005	0.0025	0.00125
RMS OMF (K)	1.174	1.165	1.161	1.164	1.177

Table 2. RMS OMF difference for T as a function of the time-filtering parameter β in 30 day 17-member EnKF runs assimilating TAO and XBT observations (model timestep: 1200s).

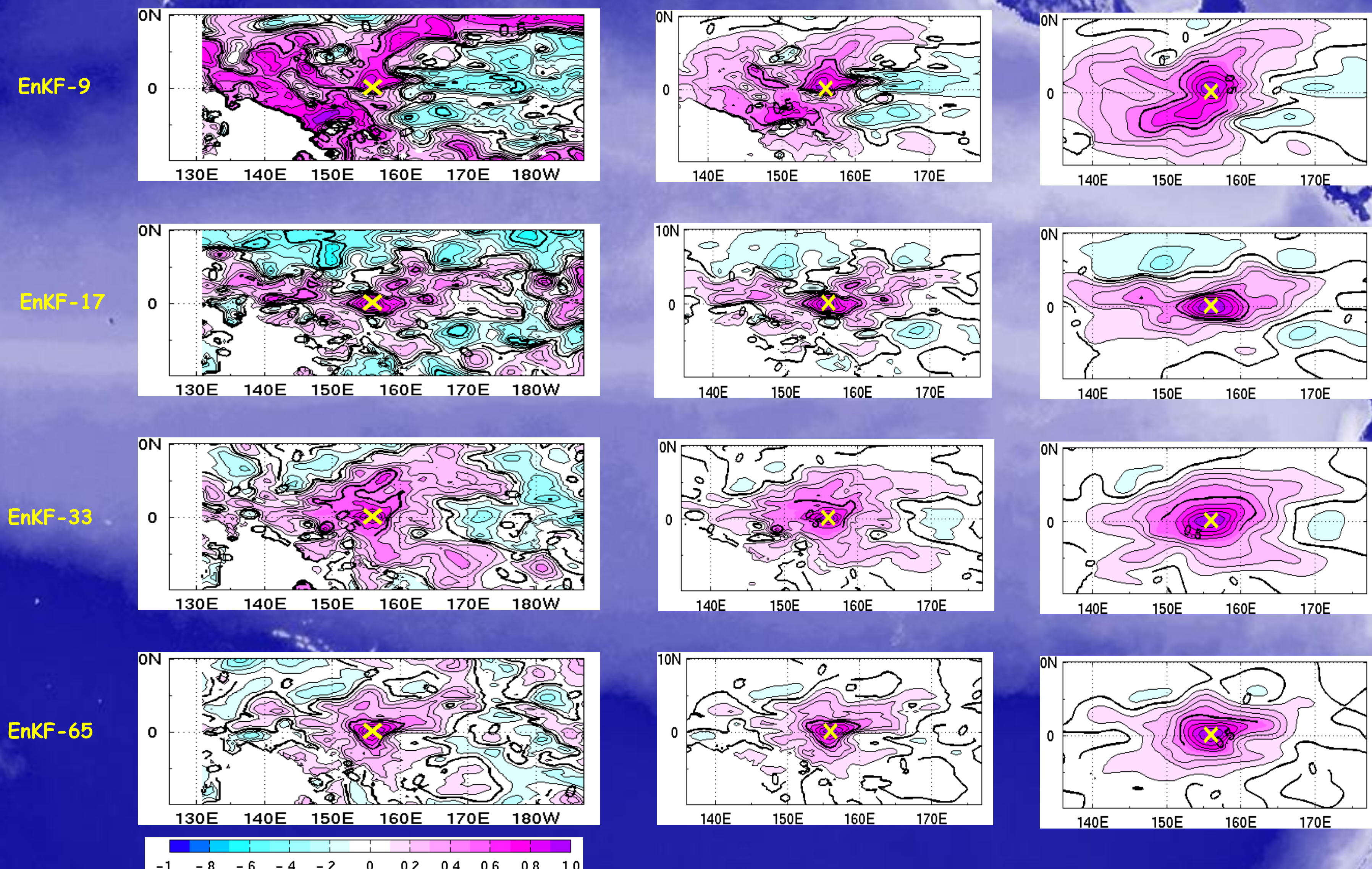
EnKF runs

- Assimilate T/P SSH anomalies + TAO & XBT temperature profiles 1/1/2001-12/31/2001
- Online bias estimation in SSH assimilation
- 4 runs: 9, 17, 33 & 65 member EnKF
- Compare with
  - no-assimilation control
  - Production ODAS (temperature OI + S(T) correction)



Effect of spatial filtering

Example of marginal Kalman gain: T obs @(0n,156E,150m) on 12/31/01
horizontal section through <T',T'> covariances



Unfiltered, not compactly supported

Unfiltered, compactly supported

Filtered, compactly supported

Effect of spatial filtering

Example of marginal Kalman gain: T obs @(0n,156E,150m) on 12/31/01
horizontal section through <T',T'> covariances

	EnKF-9	EnKF-17	EnKF-33
Unfiltered, globally supported (UGS)	0.36	0.46	0.67
Unfiltered, compactly supported (UCS)	0.51	0.58	0.75
Filtered, compactly supported (FCS)	0.63	0.70	0.77

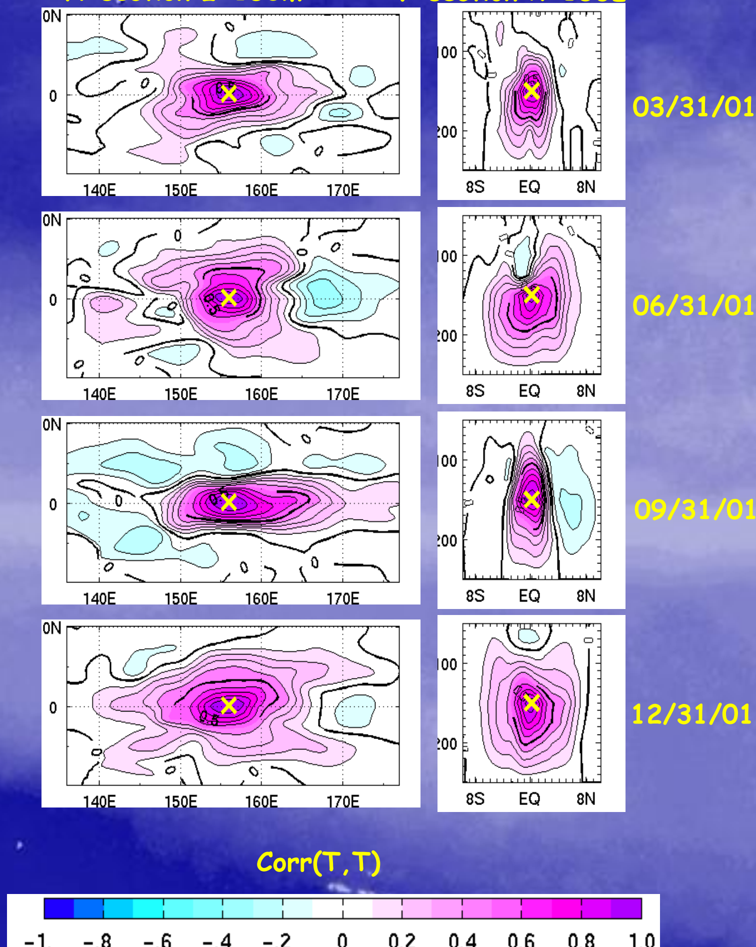
Table 3. Correlation of the horizontal section through the unfiltered globally supported (UGS) marginal gain for *T* in the EnKF-65 run with the corresponding UGS, unfiltered compactly supported (UCS) and filtered compactly supported (FCS) horizontal sections through the corresponding marginal gain in the EnKF-9, EnKF-17 and EnKF-33 runs.

- Filtering increases correlation of Kalman gain with corresponding raw (no filter, globally supported) gain from EnKF-65 run

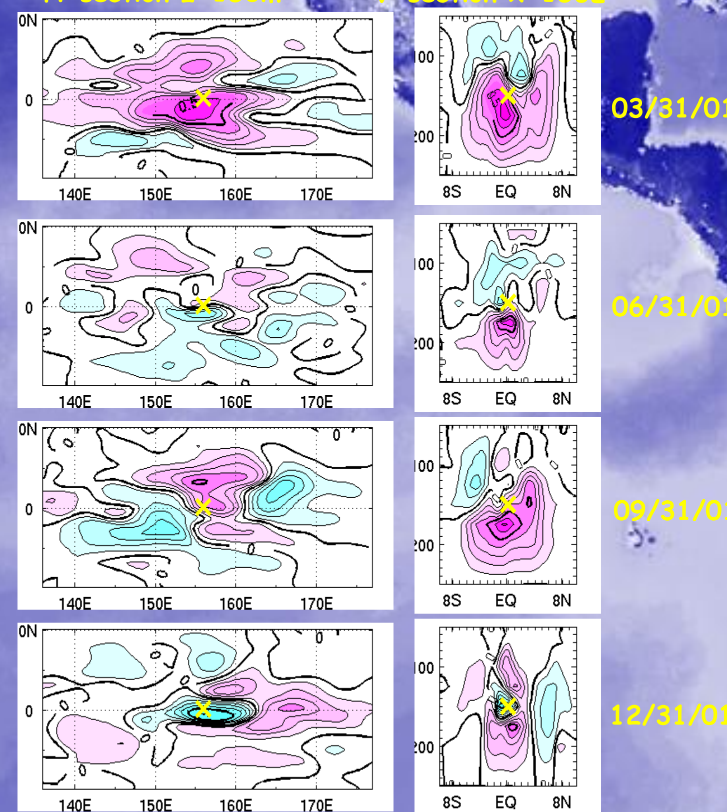
- Especially for very small ensembles, the filter "simulates" the covariances one would get from a larger ensemble

Temporal evolution of Kalman gain for T obs.

EnKF-33: filter
Schur(C,P) @(0N,156E,150m)

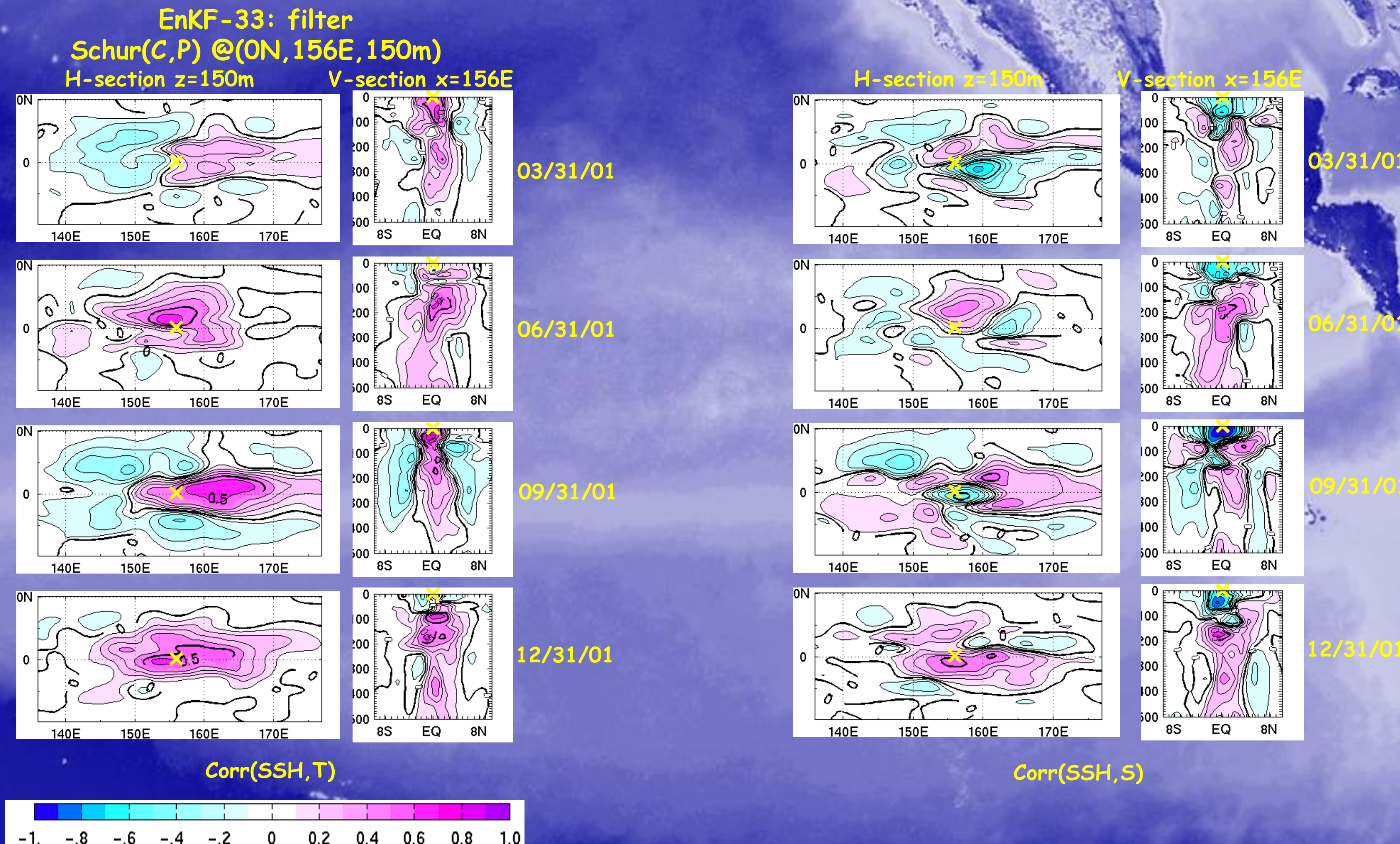


Corr(T,T)

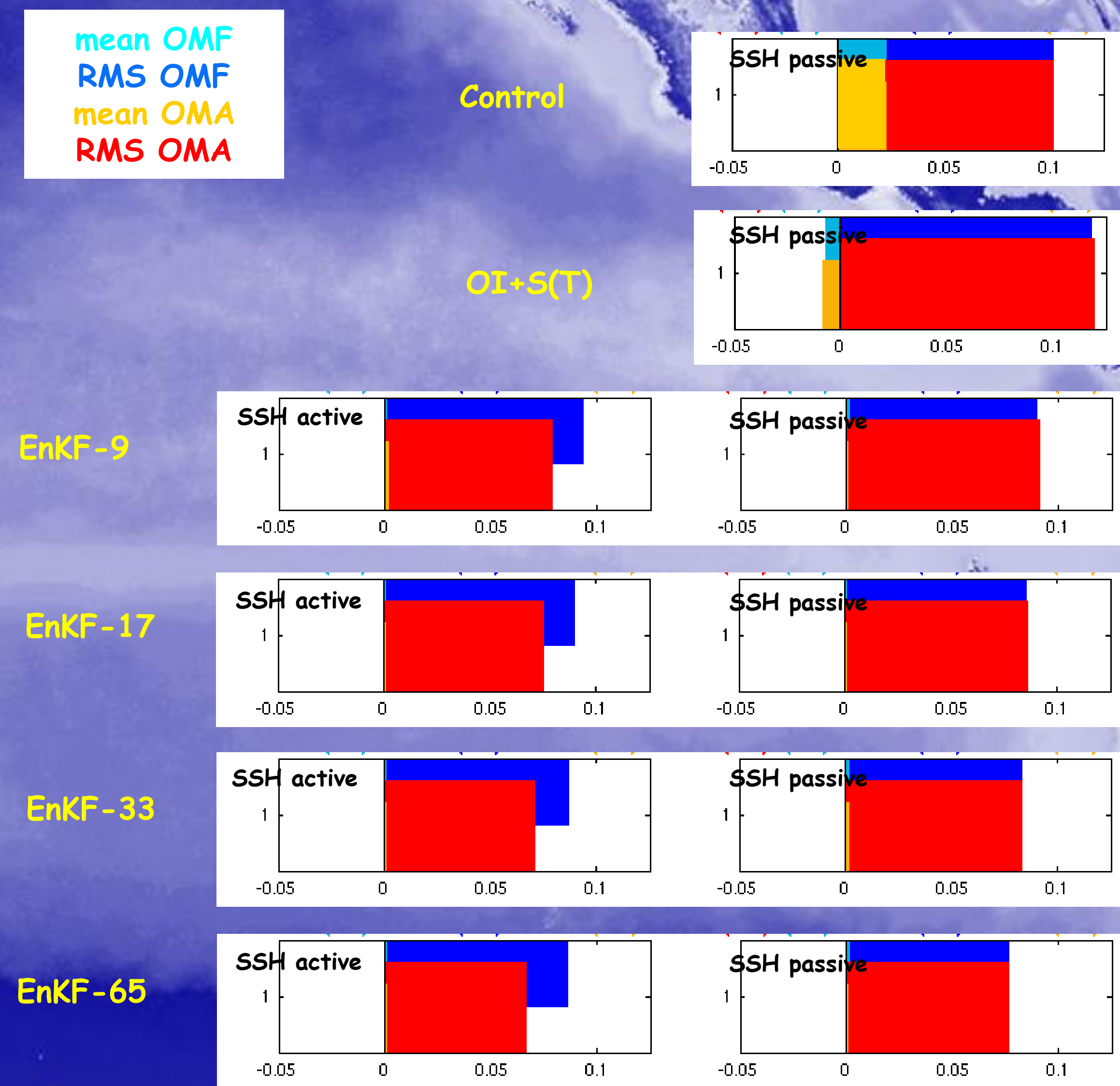


Corr(T,S)

Temporal evolution of Kalman gain for SSH obs.



SSH OMF and OMA statistics

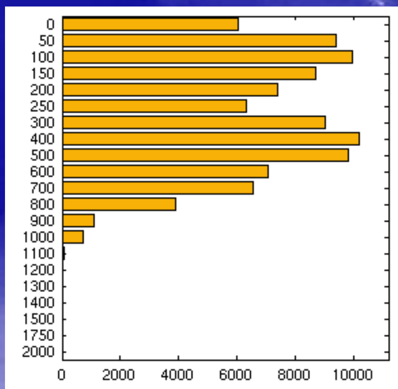


- Control is biased
- OI partly corrects SSH bias but worsens RMS OMF
- EnKF runs have no noticeable SSH bias

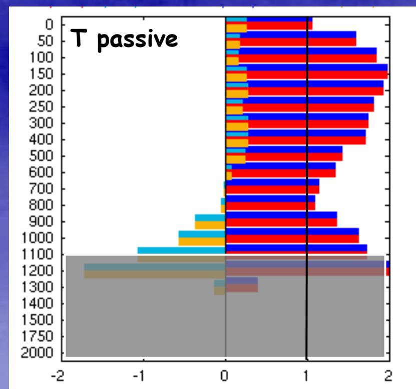
T OMF and OMA statistics

mean OMF
RMS OMF
mean OMA
RMS OMA

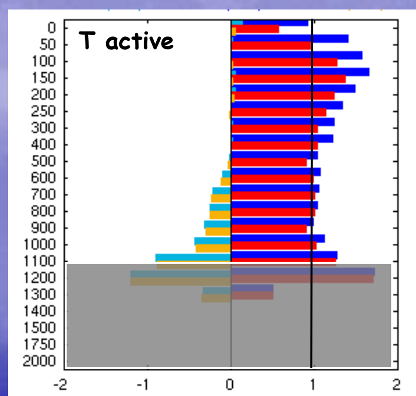
data count



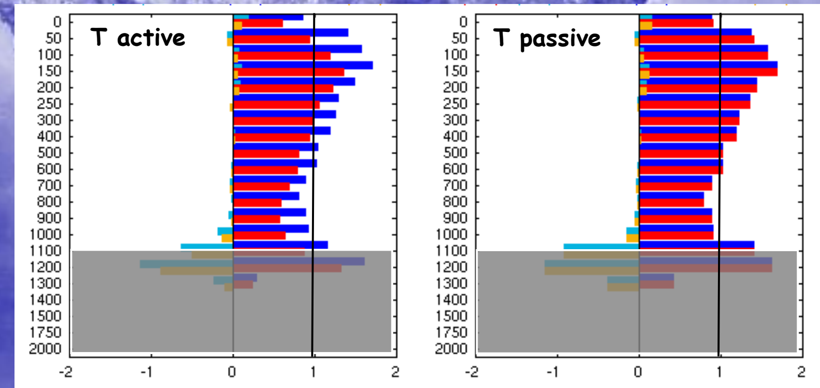
Control OMF stats



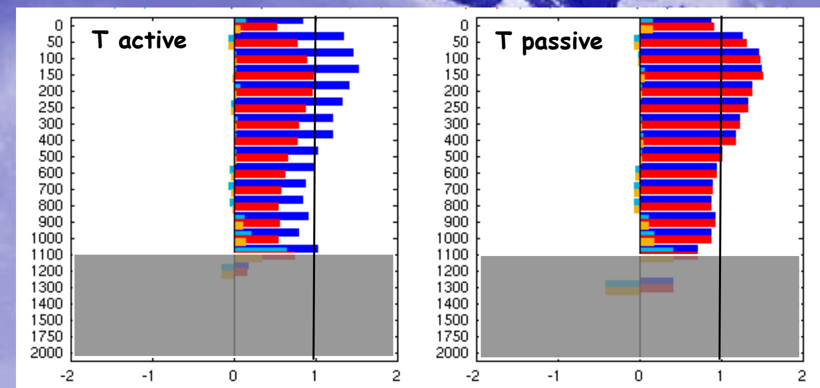
OI+S(T)



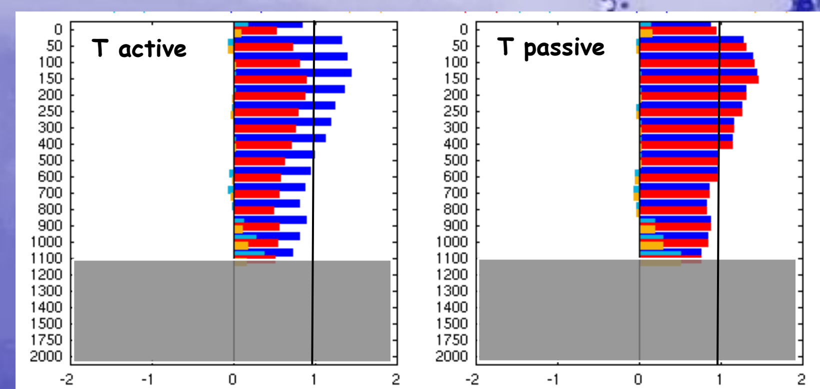
EnKF-9



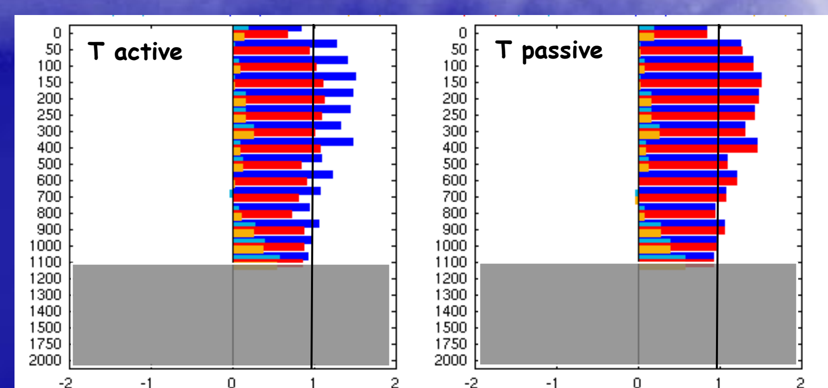
EnKF-17



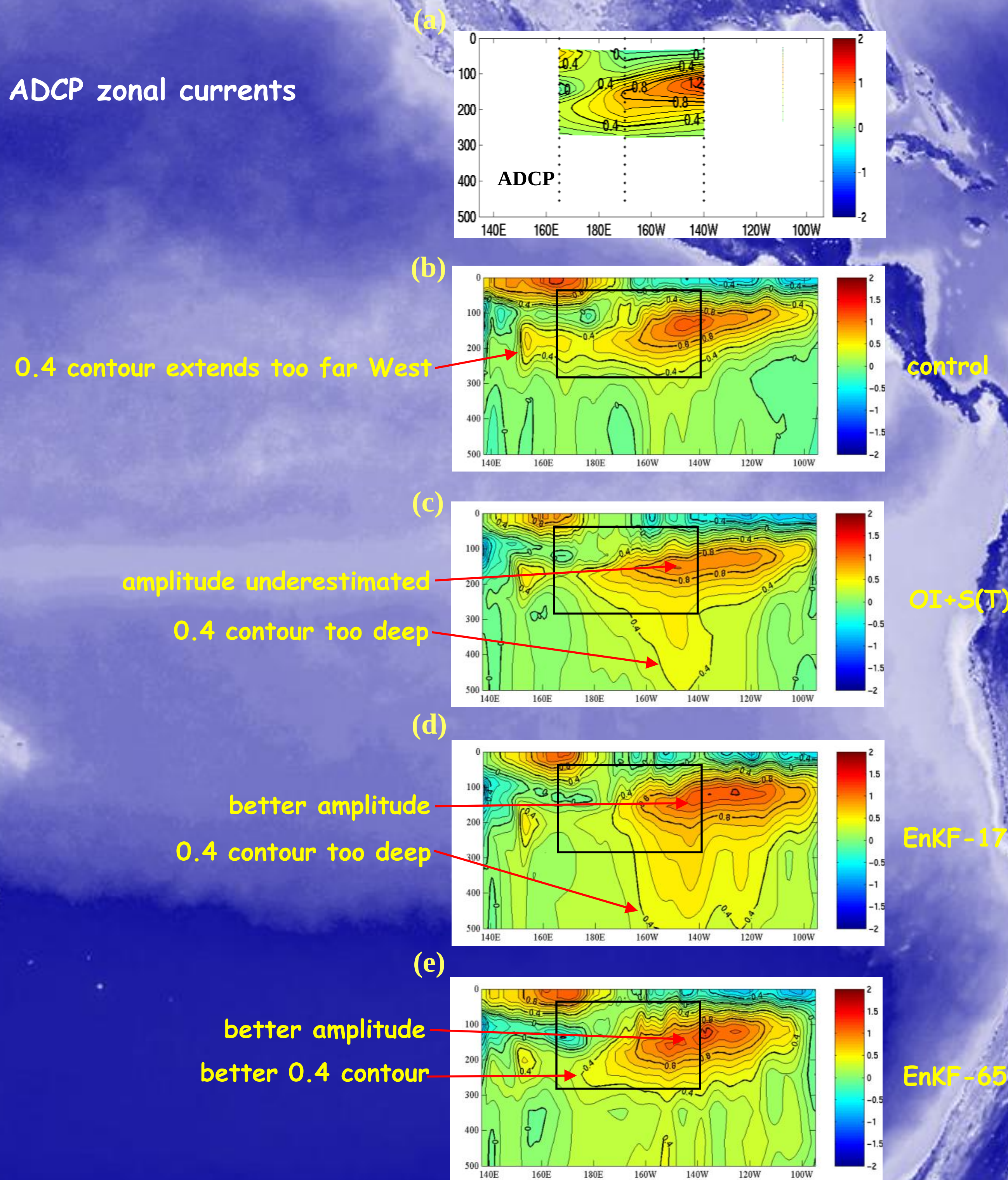
EnKF-33



EnKF-65



Validation with December 2001 ADCP zonal currents

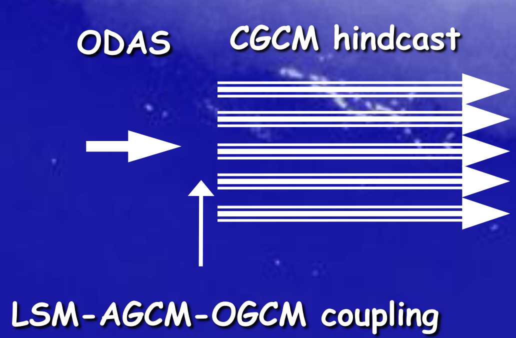


Impact of assimilation on CGCM hindcast skill

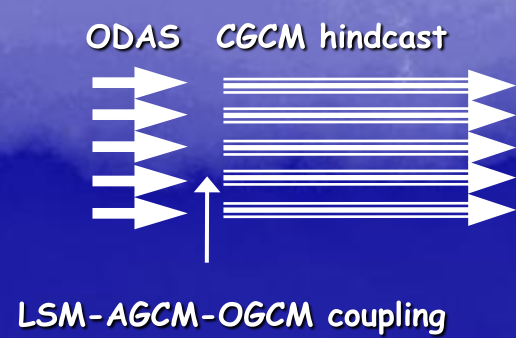
- 17-member EnKF
- Assimilate T + SSH for February-April of 1993-2002
- Couple OGCM to AGCM & LSM after running ODAS
- 12-month May start CGCM hindcasts initialized with ocean from EnKF runs (to save CPU time, CGCM hindcasts have only 5 ensemble member)

- Assess impact of assimilation on SST hindcast skill
- Compare to history of production May-start hindcasts

T OI + S(T) initialization

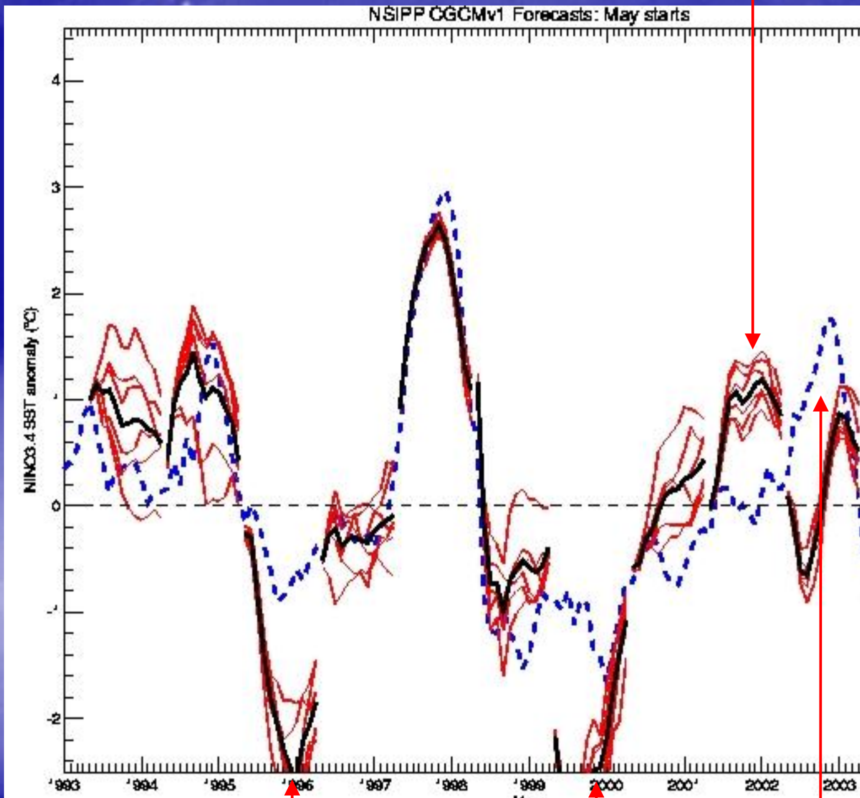


EnKF initialization



OI + S(T) initialization

False EN alert

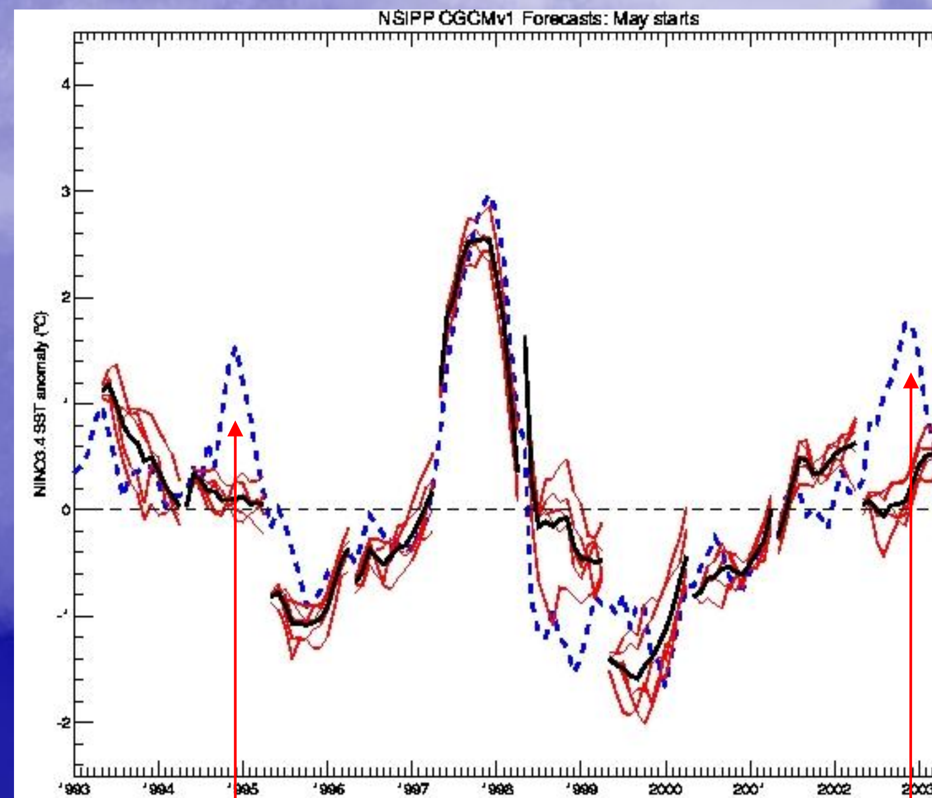


missed EN

false LN alert

false LN alert

EnKF-17 initialization



missed EN

missed EN

May start Niño-3.4 SST hindcasts

Conclusions

- Analysis speed up makes it practically feasible to use the EnKF in the production forecast initialization
- Small ensemble EnKF can outperform production system provided background covariances are appropriately preconditioned

Ongoing work

- More hindcast experiments with EnKF and Poseidon v4
- Multi-model multi-resolution EnKF with Poseidon v5/MOM 4 as part of GEOS-5 modeling system (ODAS-2)

Announcement:

- Please notice session A21 “Towards Operational Applications of Advanced Data Assimilation Methods” at the AGU Spring Meeting (May 22-25, Acapulco, Mexico)
- abstract deadline: March 1