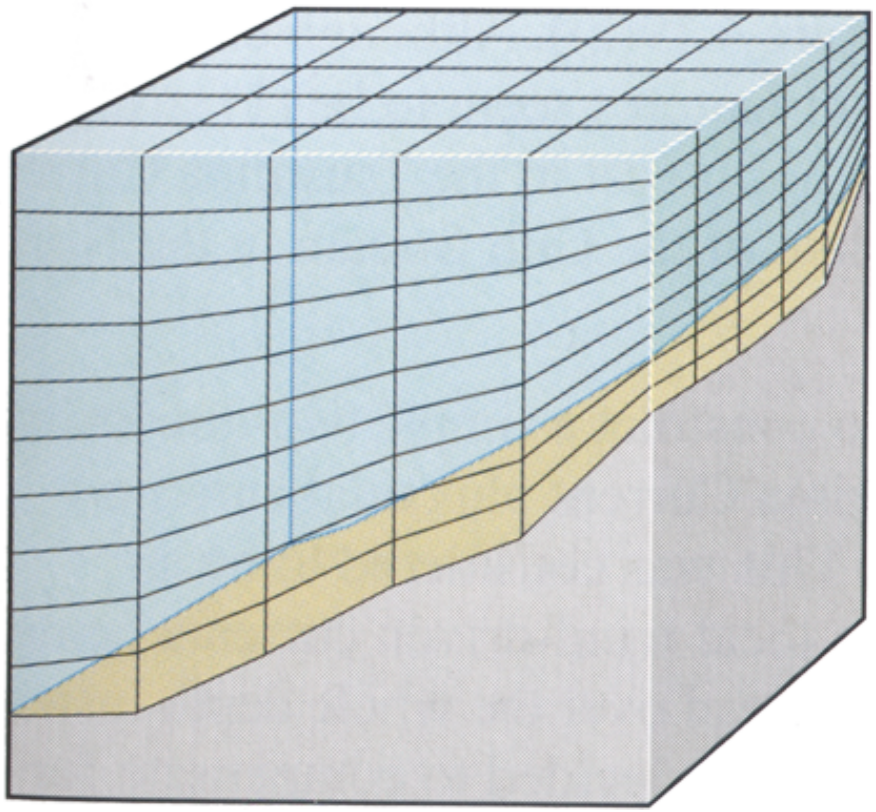
The Regional Ocean Modeling System (ROMS) 4D-Var Assimilation Systems applied to the California Current System

Andy Moore, Gregoire Broquet, Hernan Arango, Chris Edwards, Brian Powell, Milena Veneziani and Javier Zavala-Garay

Research Supported by: NSF, ONR, NOPP

ROMS

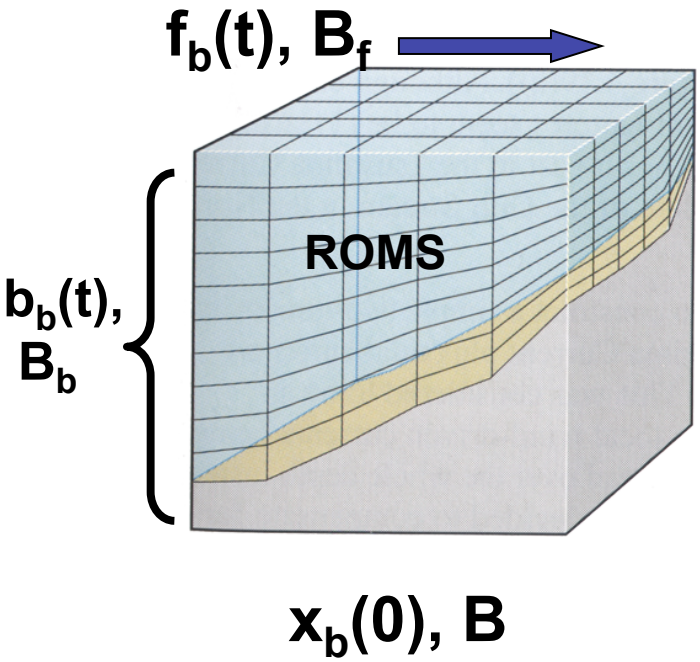


- Primitive equation, hydrostatic
- Terrain-following vertical coordinates
- Orthogonal curvilinear grids
- Open boundaries
- Comprehensive suite of numerical and physical options
- Community model

ROMS 4D-Var

- Incremental (linearize about a prior) (Courtier et al, 1994)
- Primal & dual formulations (Courtier 1997)
- Primal – Incremental 4-Var (I4D-Var)
- Dual – PSAS (4D-PSAS) & indirect representer (R4D-Var) (Da Silva et al, 1995; Egbert et al, 1994)
- Strong and weak (dual only) constraint
- Preconditioned, Lanczos formulation of conjugate gradient (Lorenc, 2003; Tshimanga et al, 2008; Fisher, 1997)
- Diffusion operator model for prior covariances (Derber & Bouttier, 1999; Weaver & Courtier, 2001)
- Multivariate balance for prior covariance (Weaver et al, 2005)
- Physical and ecosystem components
- Parallel

The Incremental Control Vector



$$\mathbf{z} = (\delta \mathbf{x}^T(0), \ \boldsymbol{\eta}_b^T(t), \ \boldsymbol{\eta}_f^T(t), \ \boldsymbol{\varepsilon}(t))^T$$

initial condition increment

boundary condition increment

forcing increment

corrections for model error

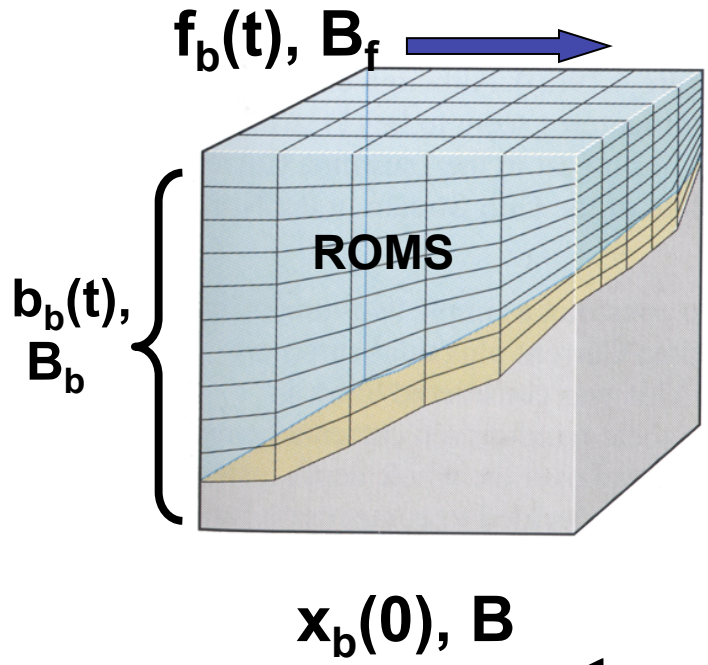
State vector: $\mathbf{x}=(T,S,u,v,\zeta)^T$

Controls: initial conditions, $\mathbf{x}(0)$

surface forcing, $\mathbf{f}(t)$

boundary conditions, $\mathbf{b}(t)$

The Incremental Control Vector



$$\mathbf{z} = \left(\delta \mathbf{x}^T(0),\ \boldsymbol{\eta}_b^T(t),\ \boldsymbol{\eta}_f^T(t),\ \boldsymbol{\varepsilon}(t)\right)^T$$

initial condition increment

boundary condition increment

forcing increment

corrections for model error

Cost/Penalty: $$J = \frac{1}{2}\mathbf{z}^T\mathbf{D}^{-1}\mathbf{z} + \frac{1}{2}\left(\mathbf{G}\mathbf{z} - \mathbf{d}\right)^T \mathbf{R}^{-1}\left(\mathbf{G}\mathbf{z} - \mathbf{d}\right)$$

$$\mathbf{D} = \underbrace{\mathrm{diag}(\mathbf{B}, \mathbf{B}_b, \mathbf{B}_f, \mathbf{Q})}_{\text{Prior error covariance}}$$

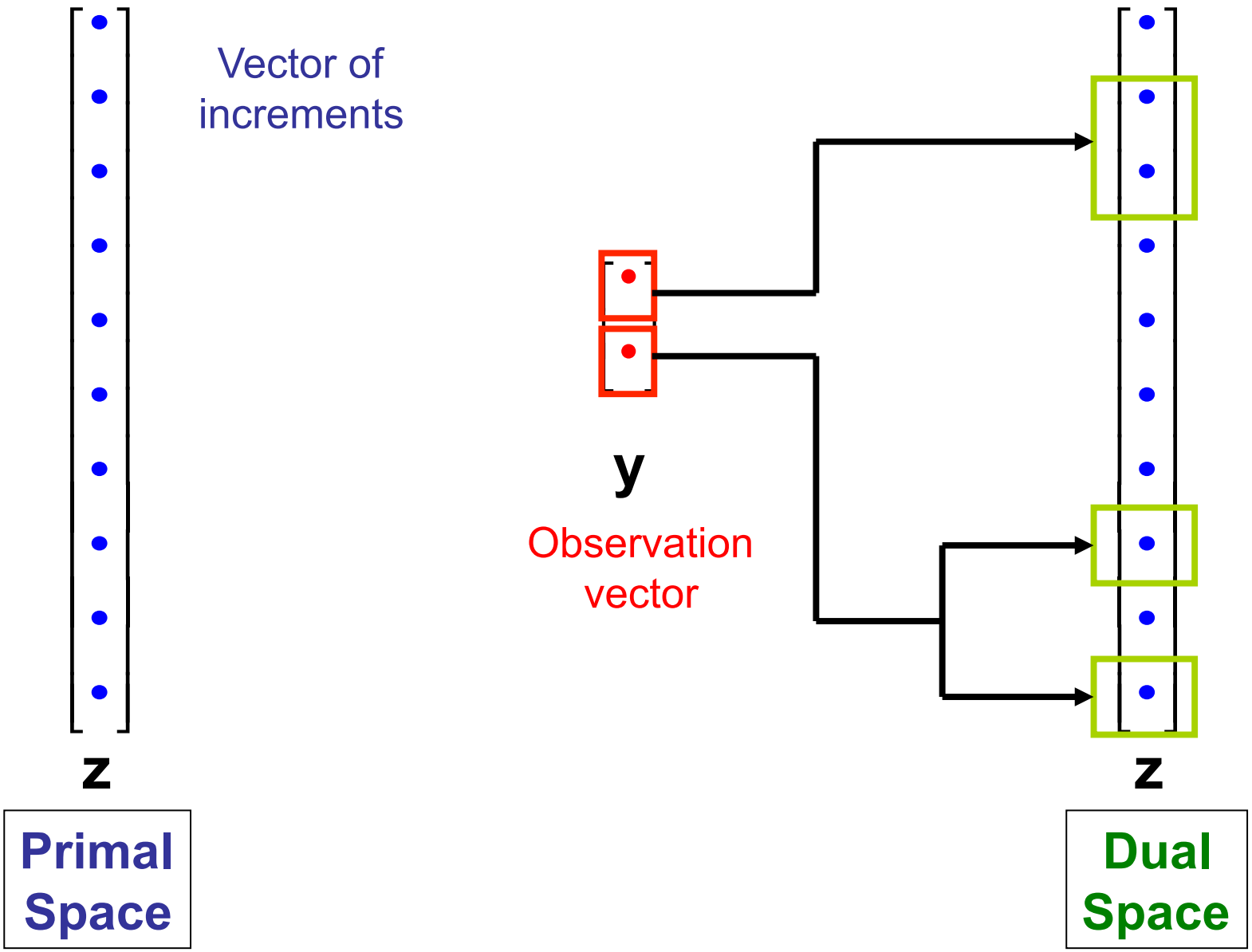
Tangent Linear Model

Obs Error Cov.

Innovation

$$\mathbf{d} = \mathbf{y} - \mathbf{H}\mathbf{x}_b$$

Primal vs Dual Formulation



Primal vs Dual Formulation

ROMS has primal and dual 4D-Var options
& strong vs weak constraint

Analysis: $$\mathbf{z}^{a} = \mathbf{z}^{b} + \mathbf{K}\mathbf{d}$$

Gain (primal): $$\mathbf{K} = (\mathbf{D}^{-1} + \mathbf{G}^{T}\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{G}^{T}\mathbf{R}^{-1}$$

Gain (dual): $$\mathbf{K} = \mathbf{D}\mathbf{G}^{T}(\mathbf{G}\mathbf{D}\mathbf{G}^{T} + \mathbf{R})^{-1}$$

Primal vs Dual Formulation

$$N_{state} \times N_{state}$$

Analysis: $$\mathbf{z}^{\mathbf{a}} = \mathbf{z}^{\mathbf{b}} + \mathbf{K}\mathbf{d}$$

Gain (primal): $$\mathbf{K} = \boxed{(\mathbf{D}^{-1} + \mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{G})^{-1}}\,\mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1}$$

Gain (dual): $$\mathbf{K} = \mathbf{D}\mathbf{G}^{\mathrm{T}}\boxed{(\mathbf{G}\mathbf{D}\mathbf{G}^{\mathrm{T}} + \mathbf{R})^{-1}}$$

$$N_{obs} \times N_{obs}$$

$$\boxed{N_{obs} \ll N_{state}}$$

Primal vs Dual Formulation

The Lanczos formulation of the CG algorithm lends considerable utility to 4D-Var:

Analysis: $$\mathbf{z}^{\mathbf{a}} = \mathbf{z}^{\mathbf{b}} + \tilde{\mathbf{K}}\mathbf{d}$$

Gain (primal): $$\tilde{\mathbf{K}} = \mathbf{V}_{\mathbf{p}}\mathbf{T}_{\mathbf{p}}^{-1}\mathbf{V}_{\mathbf{p}}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1}$$

Gain (dual): $$\tilde{\mathbf{K}} = \mathbf{D}\mathbf{G}^{\mathrm{T}}\mathbf{V}_{\mathbf{d}}\mathbf{T}_{\mathbf{d}}^{-1}\mathbf{V}_{\mathbf{d}}^{\mathrm{T}}$$

$\mathbf{V}_{\mathbf{p}}$ = matrix of Lanczos vectors of the Hessian

$\mathbf{V}_{\mathbf{d}}$ = matrix of dual Lanczos vectors

Primal vs Dual Formulation

The Lanczos formulation of the CG algorithm lends considerable utility to 4D-Var:

Analysis: $$\mathbf{z}^{\mathbf{a}} = \mathbf{z}^{\mathbf{b}} + \tilde{\mathbf{K}}\mathbf{d}$$

Gain (primal): $$\tilde{\mathbf{K}} = \mathbf{V}_{\mathbf{p}}\mathbf{T}_{\mathbf{p}}^{-1}\mathbf{V}_{\mathbf{p}}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1}$$

Gain (dual): $$\tilde{\mathbf{K}} = \mathbf{D}\mathbf{G}^{\mathrm{T}}\mathbf{V}_{\mathbf{d}}\mathbf{T}_{\mathbf{d}}^{-1}\mathbf{V}_{\mathbf{d}}^{\mathrm{T}}$$

- Posterior (analysis) error variance (dual)
- Posterior (analysis) error EOFs (dual)
- Observation impact (primal & dual)
- Observation sensitivity (dual)

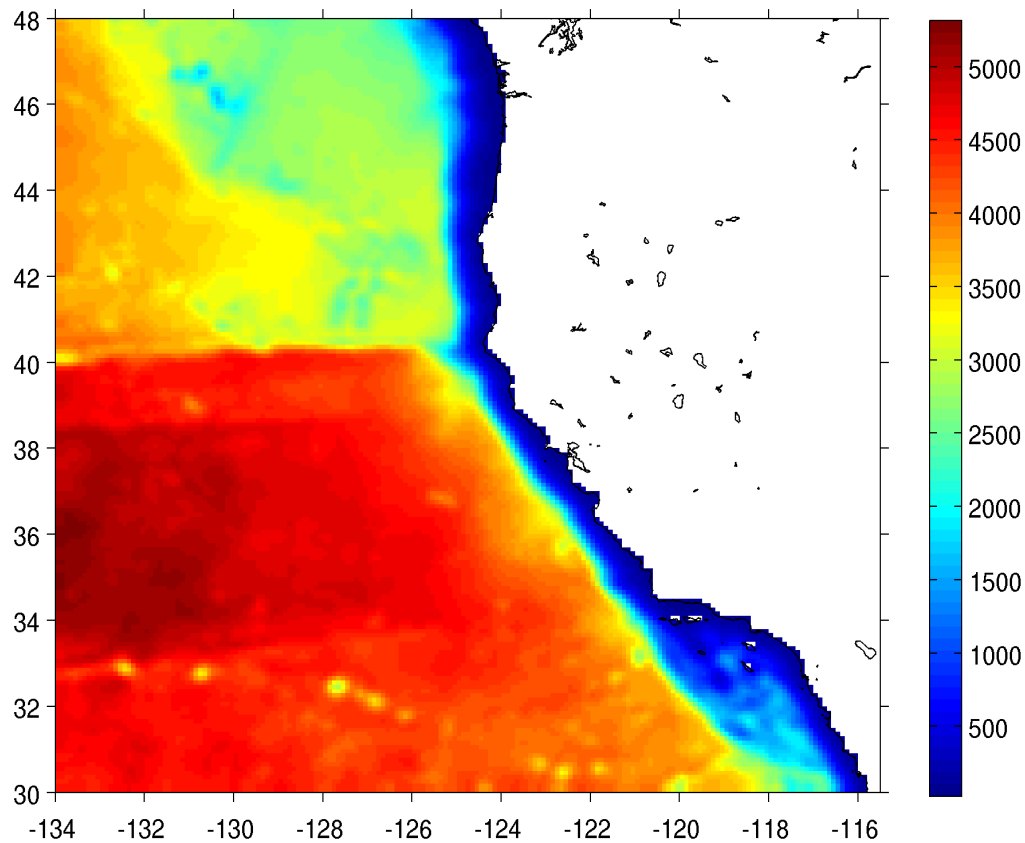
ROMS: California Current System (CCS)

COAMPS forcing $f_b(t)$, B_f

ECCO open boundary conditions $b_b(t)$, B_b

$x_b(0)$, B

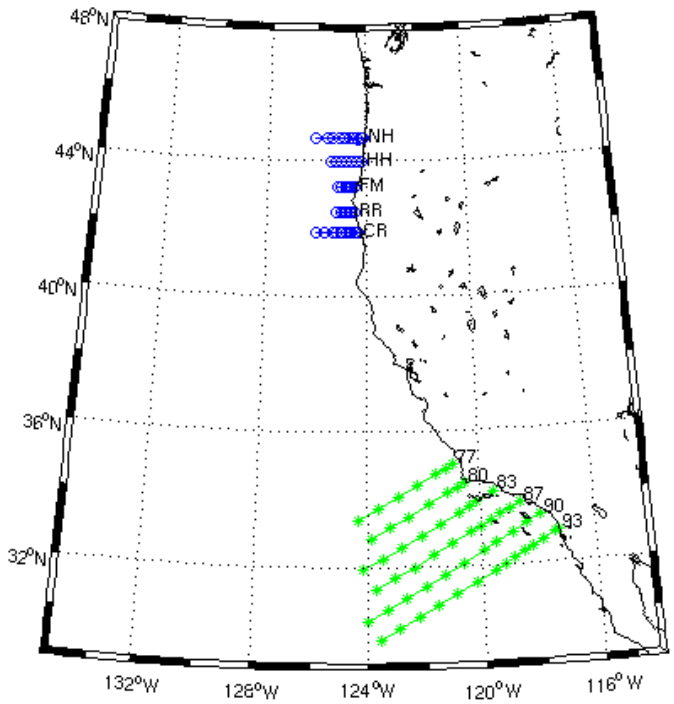
Previous assimilation cycle



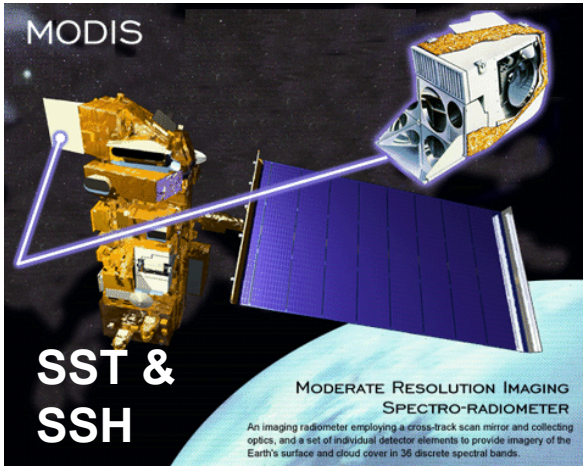
30km, 10 km & 3 km grids, 30- 42 levels

Veneziani et al (2009)
Broquet et al (2009)

Observations (y)

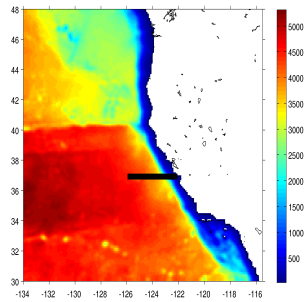


CalCOFI & GLOBEC/LTOP



Ingleby and Huddleston (2007)

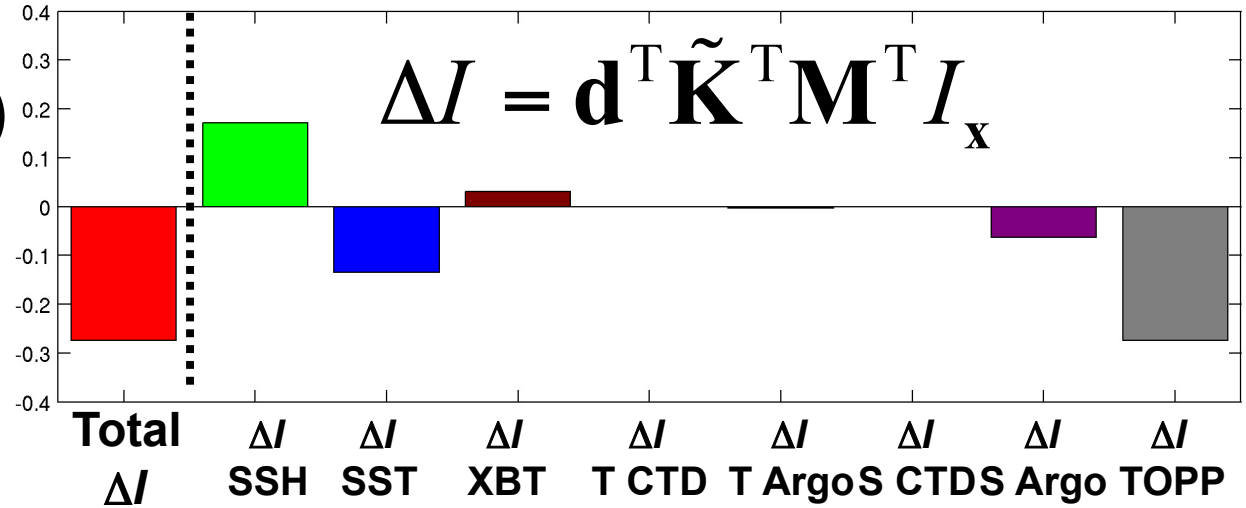
Observation Impact
(37N transport, 0-500m)



$$\Delta I = \mathbf{d}^{\mathrm{T}} \tilde{\mathbf{K}}^{\mathrm{T}} \mathbf{M}^{\mathrm{T}} I_{\mathbf{x}}$$

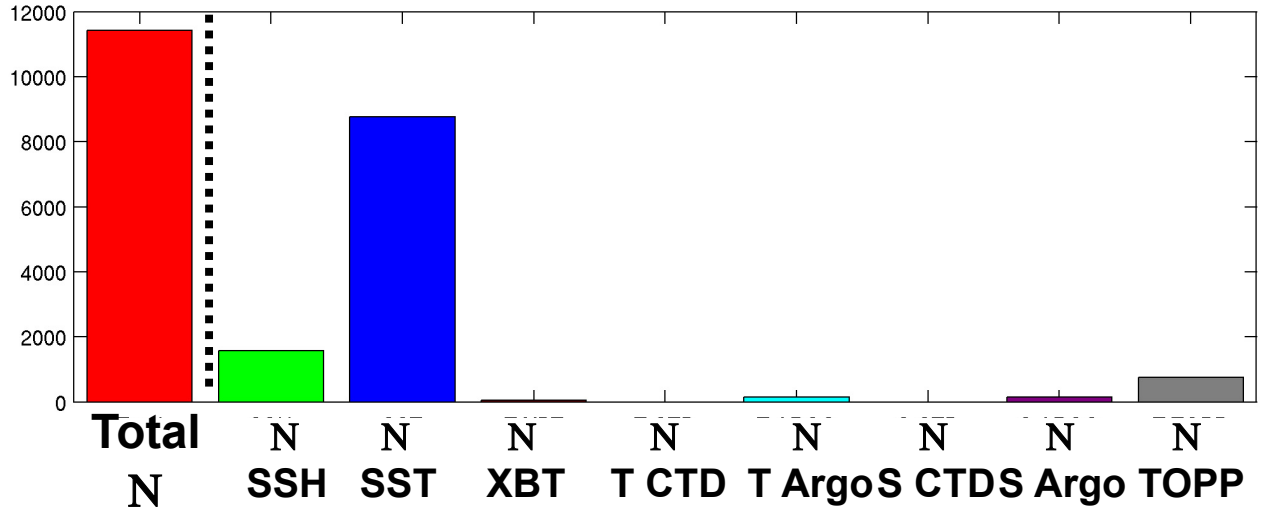
ΔI (Sv)

Transport Increment

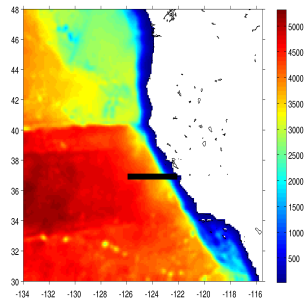


7 day assim cycle (strong)

N_{obs}



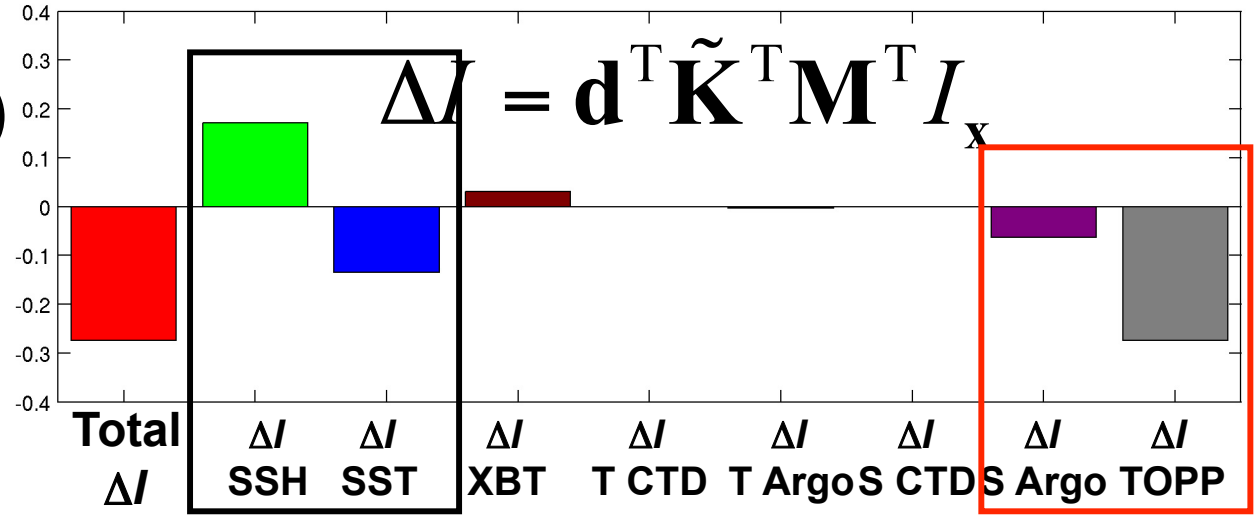
Observation Impact
(37N transport, 0-500m)



$$\Delta I = \mathbf{d}^{\mathrm{T}} \tilde{\mathbf{K}}^{\mathrm{T}} \mathbf{M}^{\mathrm{T}} I_{\mathrm{x}}$$

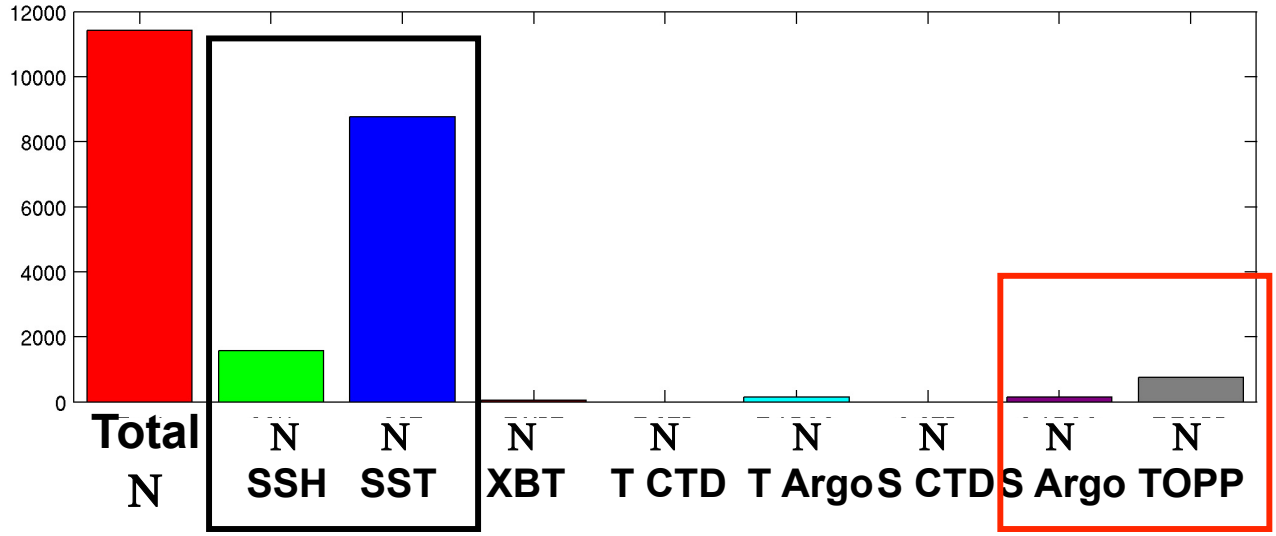
ΔI (Sv)

Transport Increment

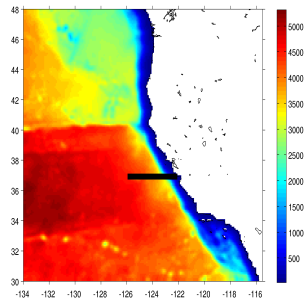


7 day assim Cycle (strong)

N_{obs}

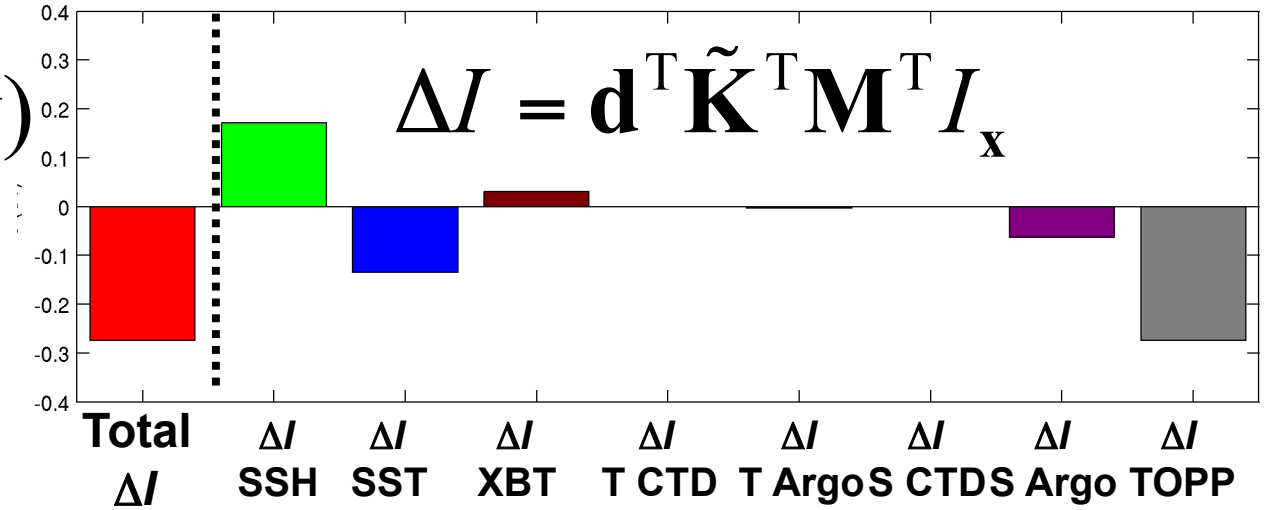


Obs Impact vs Obs Sensitivity



ΔI (Sv)

Transport Increment

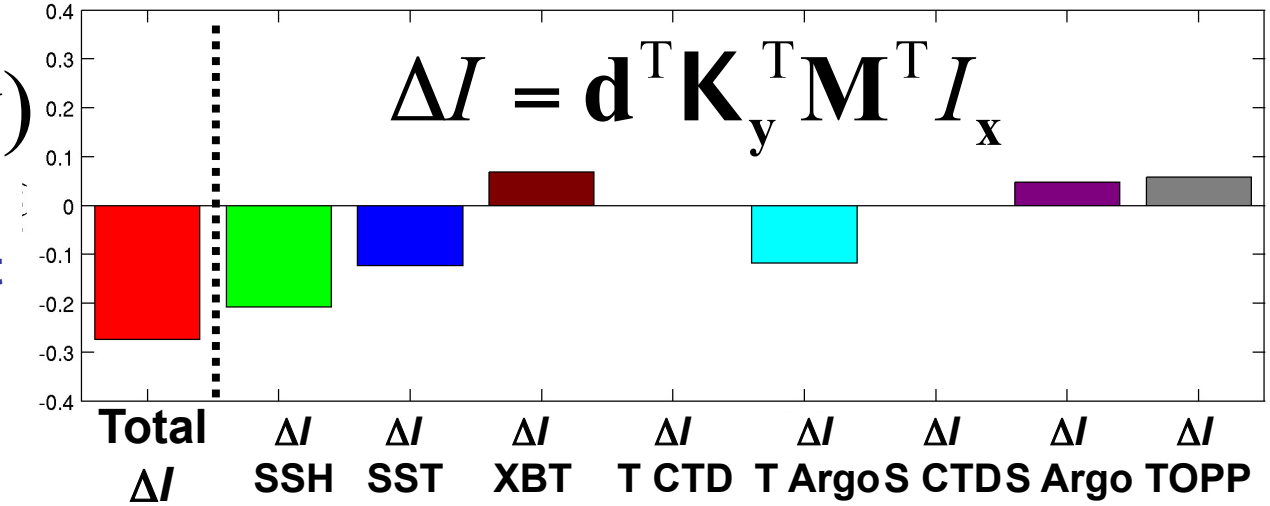


Impact

7 day assim cycle (strong)

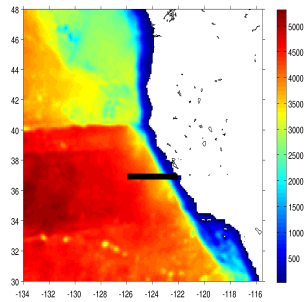
ΔI (Sv)

Transport Increment



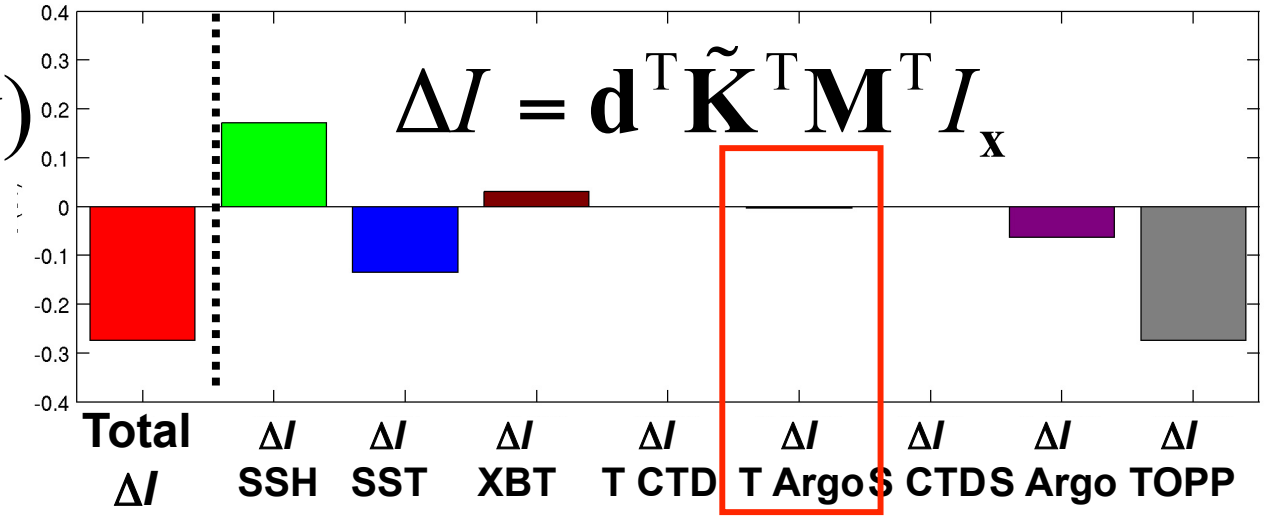
Sensitivity: based on adjoint of 4D-Var

Obs Impact vs Obs Sensitivity



ΔI (Sv)

Transport Increment

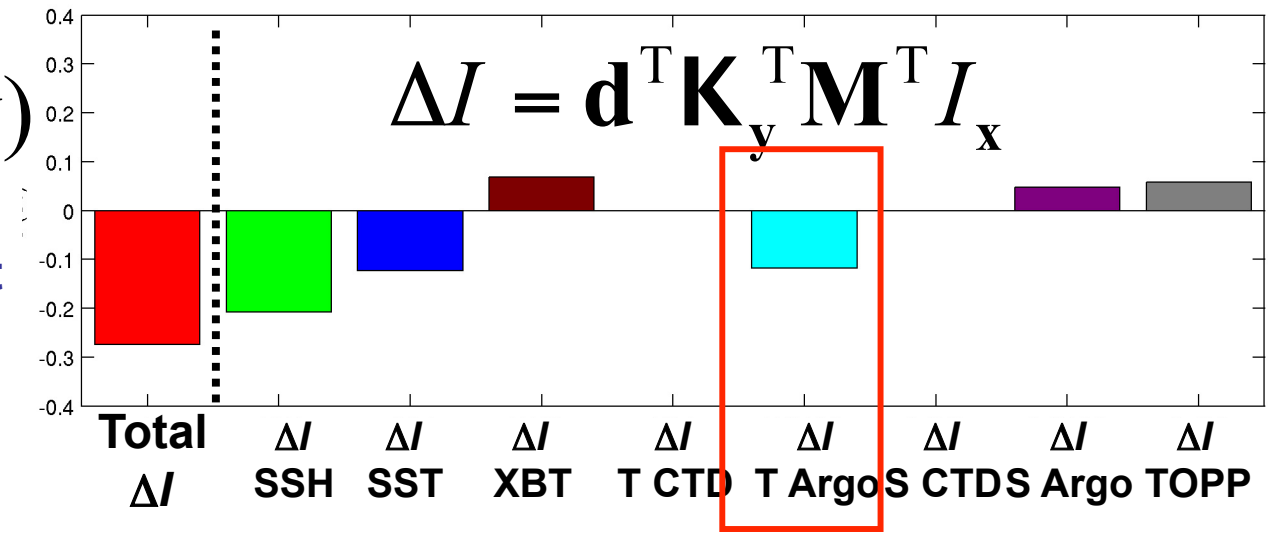


Impact

7 day assim cycle (strong)

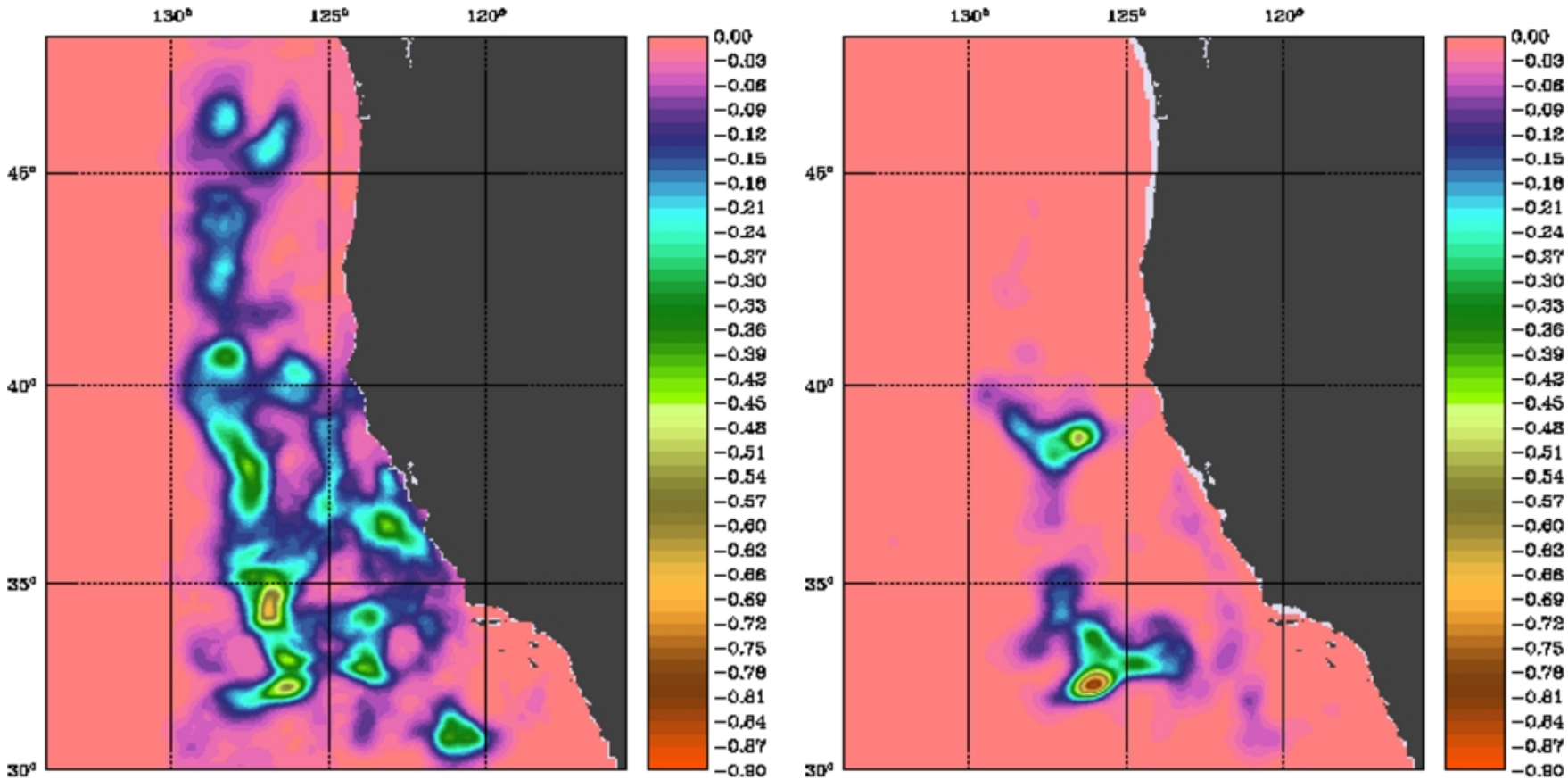
ΔI (Sv)

Transport Increment



Sensitivity: based on adjoint of 4D-Var

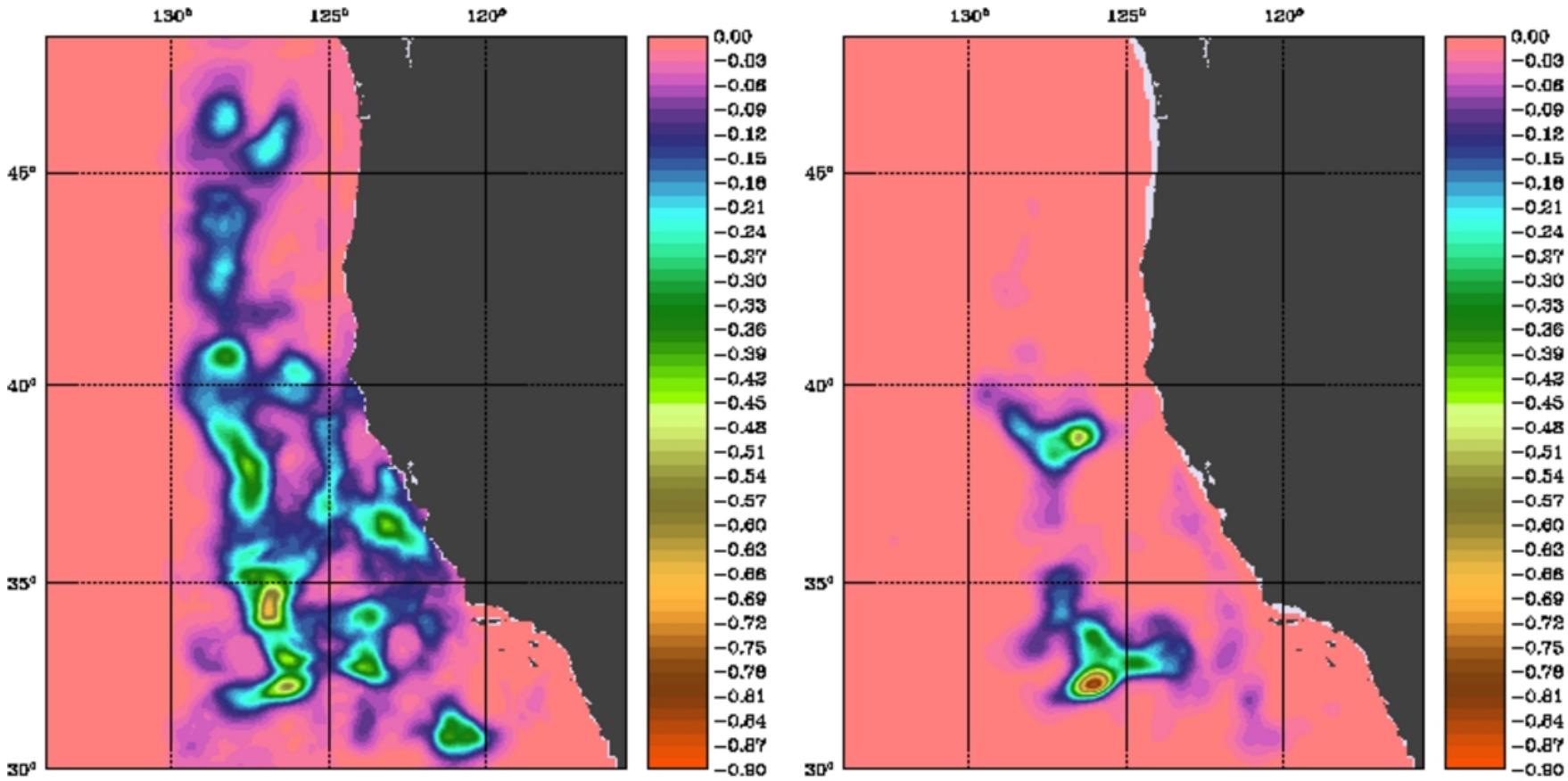
Expected Posterior Error



SST **75m**

$$\mathbf{E}^{\mathbf{a}} = (\mathbf{I} - \tilde{\mathbf{K}}\mathbf{G})\mathbf{D}(\mathbf{I} - \tilde{\mathbf{K}}\mathbf{G})^{\mathrm{T}} + \tilde{\mathbf{K}}\mathbf{R}\tilde{\mathbf{K}}^{\mathrm{T}}$$

Expected Posterior Error

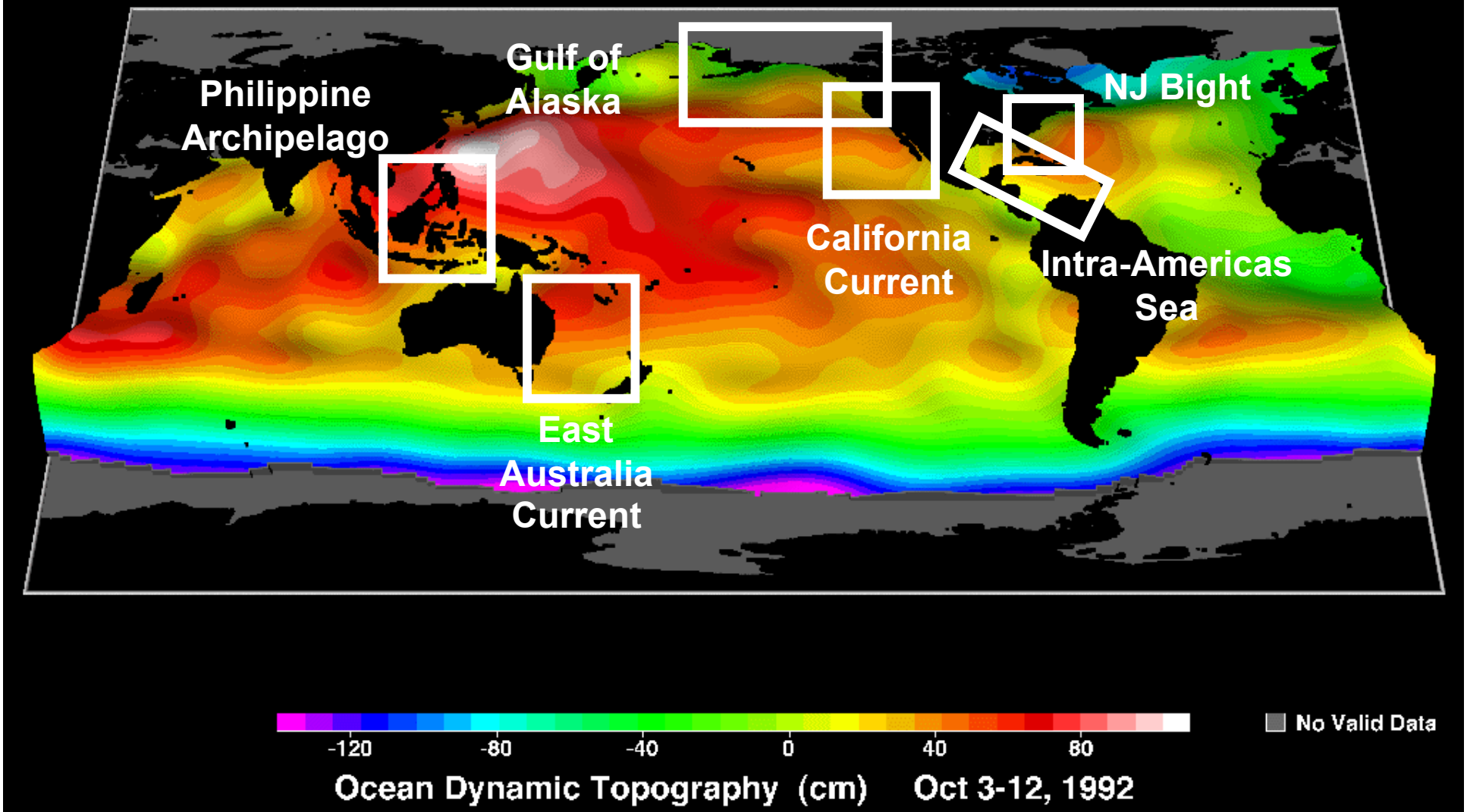


SST **75m**

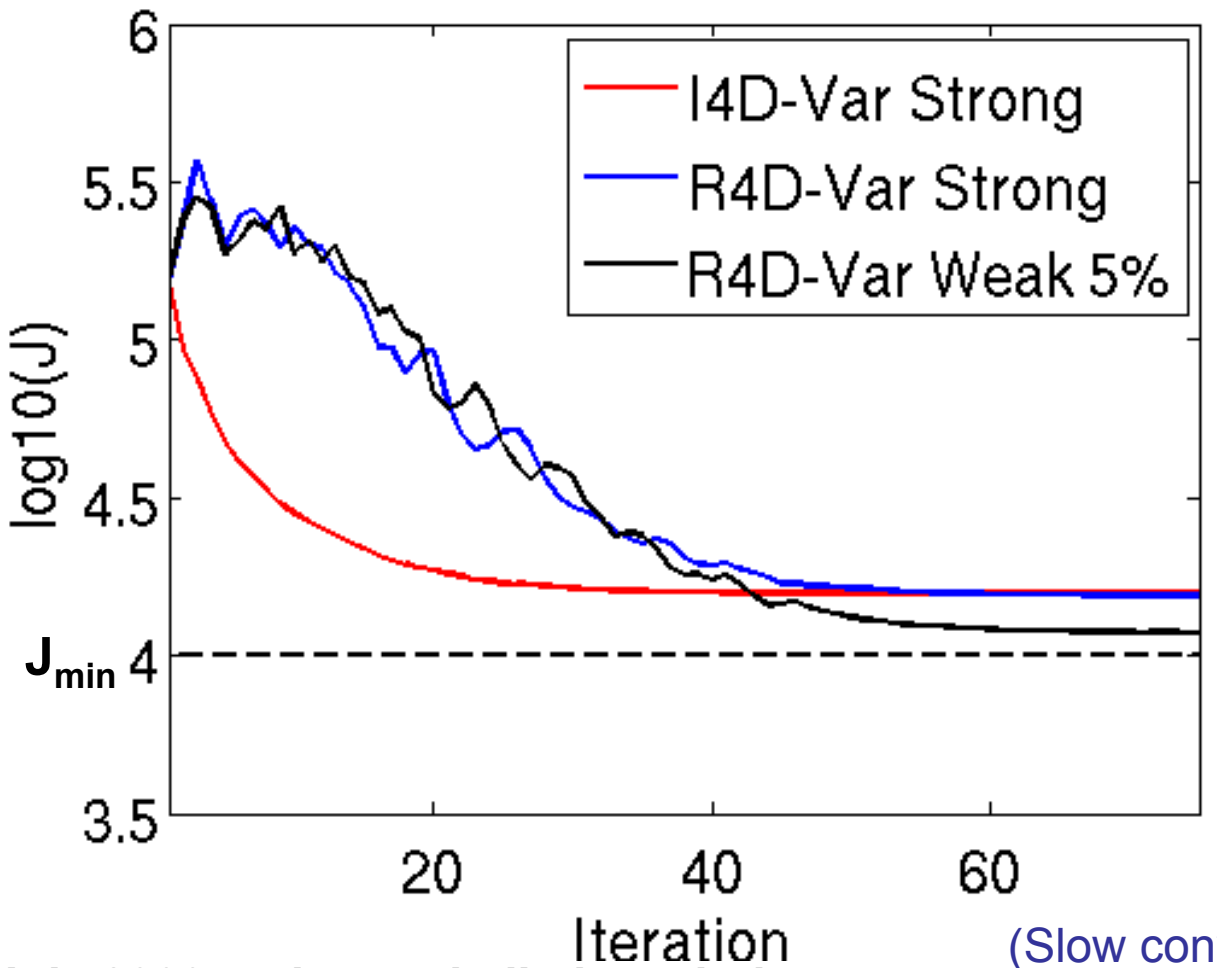
$$\mathrm{diag}(\mathbf{E}^{\mathbf{a}} - \mathbf{D})$$

Posterior variance – Prior variance

Current Applications



Primal vs Dual Space
(I4D-Var vs 4D-PSAS vs R4D-Var)

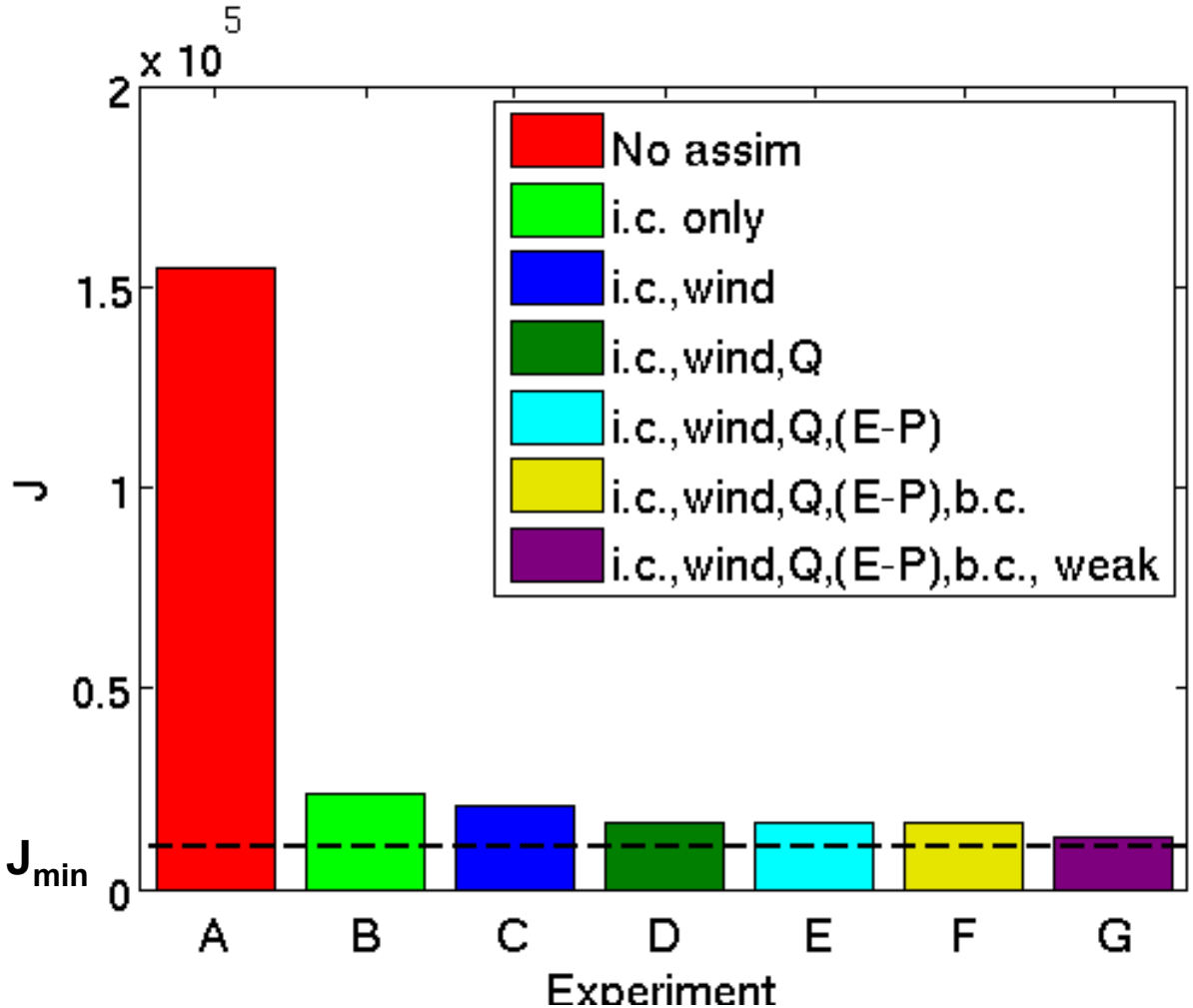


1 - 4 July, 2000

(1 outer, 75 inner,
L_h=50 km, L_v=30m,
Obs errors:
SSH 2 cm
SST 0.4 C
hydrographic 0.1 C
0.01psu)

July 2000: 4 day assimilation window
STRONG vs WEAK CONSTRAINT

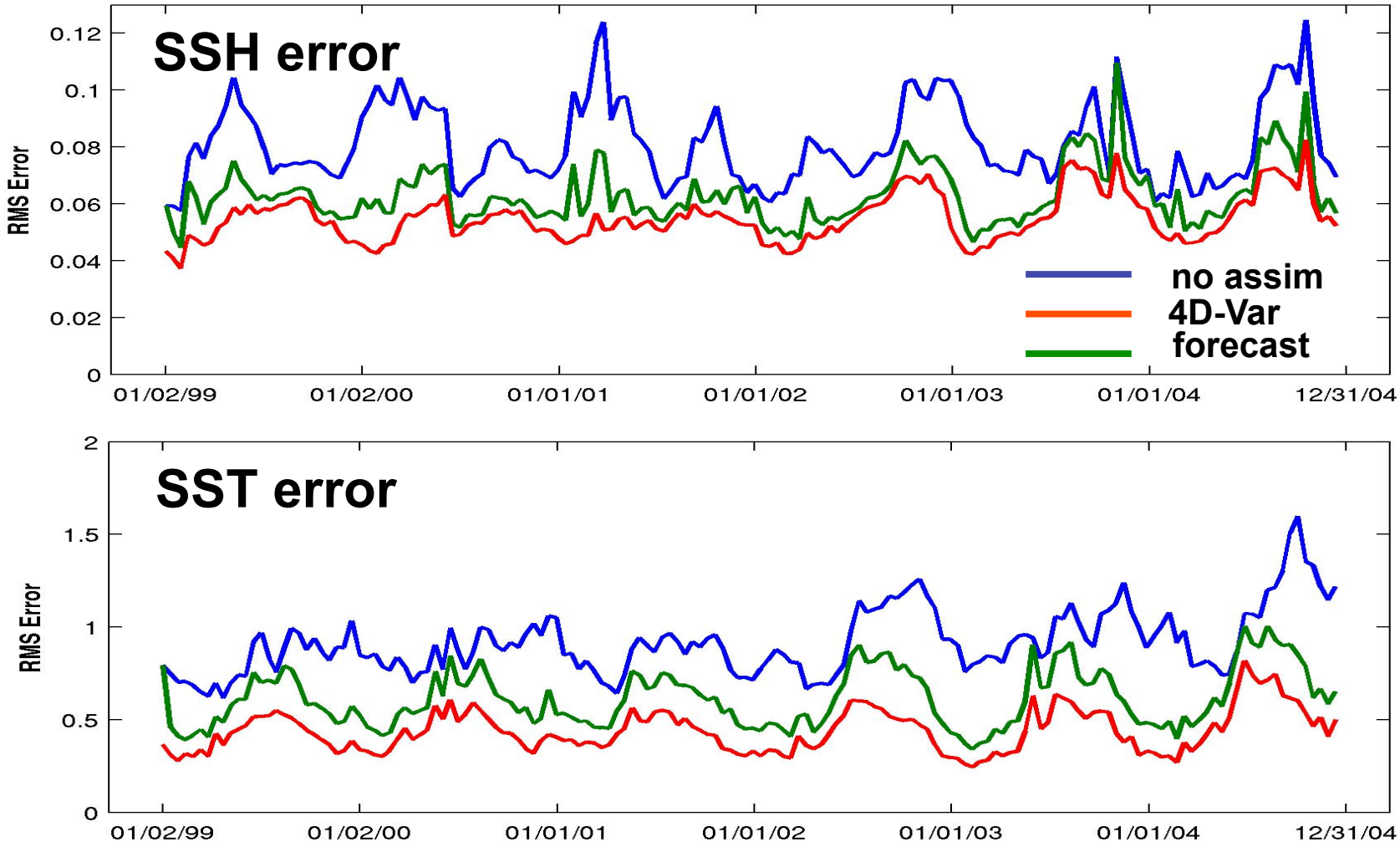
(Slow convergence of
dual form also noted by
El Akkraoui and Gauthier, 2009)



i.c. = initial condition
wind = wind stress
Q = heat flux
(E-P) = freshwater flux
b.c. = open boundary conditions

(R4DVAR: 1 outer-loop; 75 inner-loops)

ROMS 4D-Var in the CCS



ROMS 4D-Var Diagnostics

The Lanczos formulation of the CG algorithm lends considerable utility to 4D-Var:

Analysis: $$\mathbf{x}^{\mathbf{a}} = \mathbf{x}^{\mathbf{b}} + \mathbf{K}\mathbf{d}$$

Gain (primal): $$\mathbf{K} \approx \mathbf{V}_{\mathbf{p}} \mathbf{T}_{\mathbf{p}}^{-1} \mathbf{V}_{\mathbf{p}}^{\mathrm{T}} \mathbf{G}^{\mathrm{T}} \mathbf{R}^{-1}$$

Gain (dual): $$\mathbf{K} \approx \mathbf{D}\mathbf{G}^{\mathrm{T}} \mathbf{V}_{\mathbf{d}} \mathbf{T}_{\mathbf{d}}^{-1} \mathbf{V}_{\mathbf{d}}^{\mathrm{T}}$$

$\mathbf{V}_{\mathbf{p}}$ = matrix of Lanczos vectors of the Hessian

$\mathbf{V}_{\mathbf{d}}$ = matrix of dual Lanczos vectors

ROMS 4D-Var Diagnostics

The Lanczos formulation of the CG algorithm lends considerable utility to 4D-Var:

Analysis: $$\mathbf{x}^{\mathbf{a}} = \mathbf{x}^{\mathbf{b}} + \mathbf{K}\mathbf{d}$$

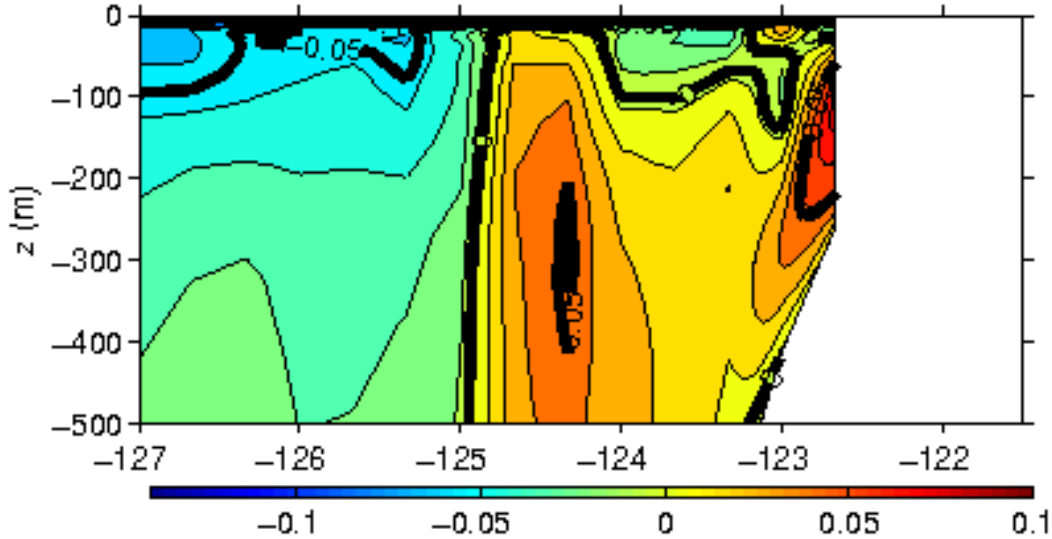
Gain (primal): $$\mathbf{K} \approx \mathbf{V}_{\mathbf{p}}\mathbf{T}_{\mathbf{p}}^{-1}\mathbf{V}_{\mathbf{p}}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1}$$

Gain (dual): $$\mathbf{K} \approx \mathbf{D}\mathbf{G}^{\mathrm{T}}\mathbf{V}_{\mathbf{d}}\mathbf{T}_{\mathbf{d}}^{-1}\mathbf{V}_{\mathbf{d}}^{\mathrm{T}}$$

- Posterior/Analysis error variance (dual)
- Posterior/Analysis error EOFs (dual)
- Observation impact (primal & dual)
- Observation sensitivity (dual)

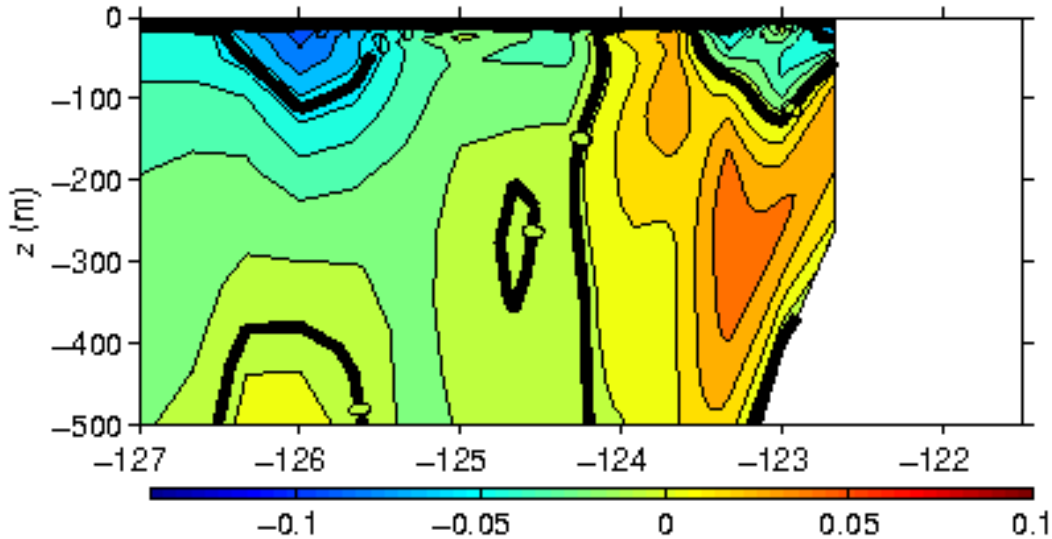
Assimilation impacts on CC

No assim



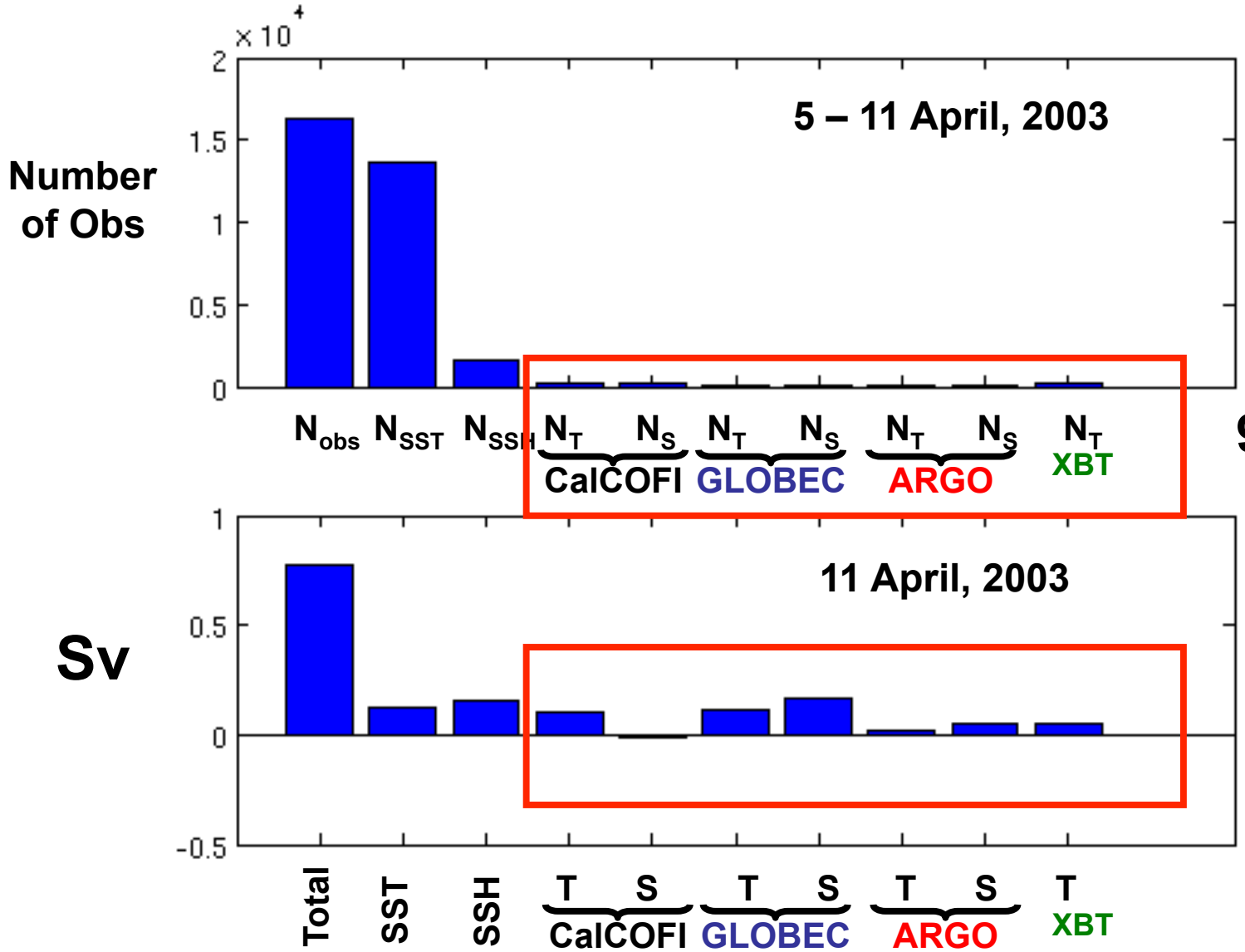
Time mean alongshore flow across 37N, 2000-2004 (30km)

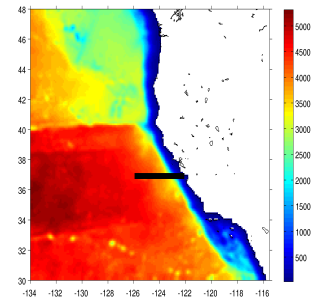
IS4D-Var



(Broquet et al, 2009)

Analysis Increment: upper 500 m Transport at 37N



980 obs (6%)

0.49Sv (63%)