# Lecture 4:
# 4D-Var Diagnostics

# Outline

- Posterior/analysis error estimates
- Consistency checks, hypothesis tests, degrees of freedom & information content
- Array modes
- Clipped analyses

# Posterior Error Estimates

# Posterior/Analysis Error Estimates

**Posterior error covariance:**

$$\mathbf{E}^{\mathbf{a}} = \left\langle (\mathbf{z_a} - \mathbf{z_t})(\mathbf{z_a} - \mathbf{z_t})^{\mathrm{T}} \right\rangle$$

$$= \left\langle (\mathbf{z_b} + \delta\mathbf{z_a} - \mathbf{z_t})(\mathbf{z_b} + \delta\mathbf{z_a} - \mathbf{z_t})^{\mathrm{T}} \right\rangle$$

$$= \left(\mathbf{I} - \mathbf{KG}\right)\mathbf{D}\left(\mathbf{I} - \mathbf{KG}\right)^{\mathrm{T}} + \mathbf{KRK}^{\mathrm{T}}$$

**(Caveat: $\mathbf{E}^{\mathbf{a}}$ correct *only if* D and R are correct)**

**But for dual Lanczos vector formulation:**

$$\mathbf{K} \approx \tilde{\mathbf{K}}_k = \mathbf{DG}^{\mathrm{T}}\mathbf{R}^{-1/2}\mathbf{V}_k\mathbf{T}_k^{-1}\mathbf{V}_k^{\mathrm{T}}\mathbf{R}^{-1/2}$$

**Lanczos vectors are orthonormal, and normal to subspace neglected by $\tilde{\mathbf{K}}_k$ :**

$$\mathbf{E}^{\mathbf{a}} \approx \tilde{\mathbf{E}}^{\mathbf{a}} = (\mathbf{I} - \tilde{\mathbf{K}}\mathbf{G})\mathbf{D}$$

# Posterior/Analysis Error Estimates

**Approx. posterior error covariance:**

$$\mathbf{E^a} \approx \tilde{\mathbf{E}}^{\mathbf{a}} = (\mathbf{I} - \tilde{\mathbf{K}}\mathbf{G})\mathbf{D}$$

$$\tilde{\mathbf{E}}^{\mathbf{a}} = \underbrace{\left(\mathbf{I} - \mathbf{D}\mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1/2}\mathbf{V}_k\mathbf{T}_k^{\mathrm{T}}\mathbf{V}_k^{\mathrm{T}}\mathbf{R}^{-1/2}\mathbf{G}\right)}\mathbf{D}$$

**Everything is available during inner-loops of 4D-PSAS and R4D-Var, at no extra cost**

$\tilde{\mathbf{E}}^{\mathbf{a}} \sim N_{\mathrm{model}} \times N_{\mathrm{model}}$ **HUGE!**

**Diagonal elements – *posterior* variances**

**define POSTERIOR_ERROR_I**

**Cross-covariance information from EOFs of** $\tilde{\mathbf{E}}^{\mathbf{a}}$

**define POSTERIOR_EOFS**

# Posterior/Analysis Error Estimates

**For the primal formulation:**

$$\mathbf{E^a} = \left(\mathbf{D}^{-1} + \mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{G}\right)^{-1}$$

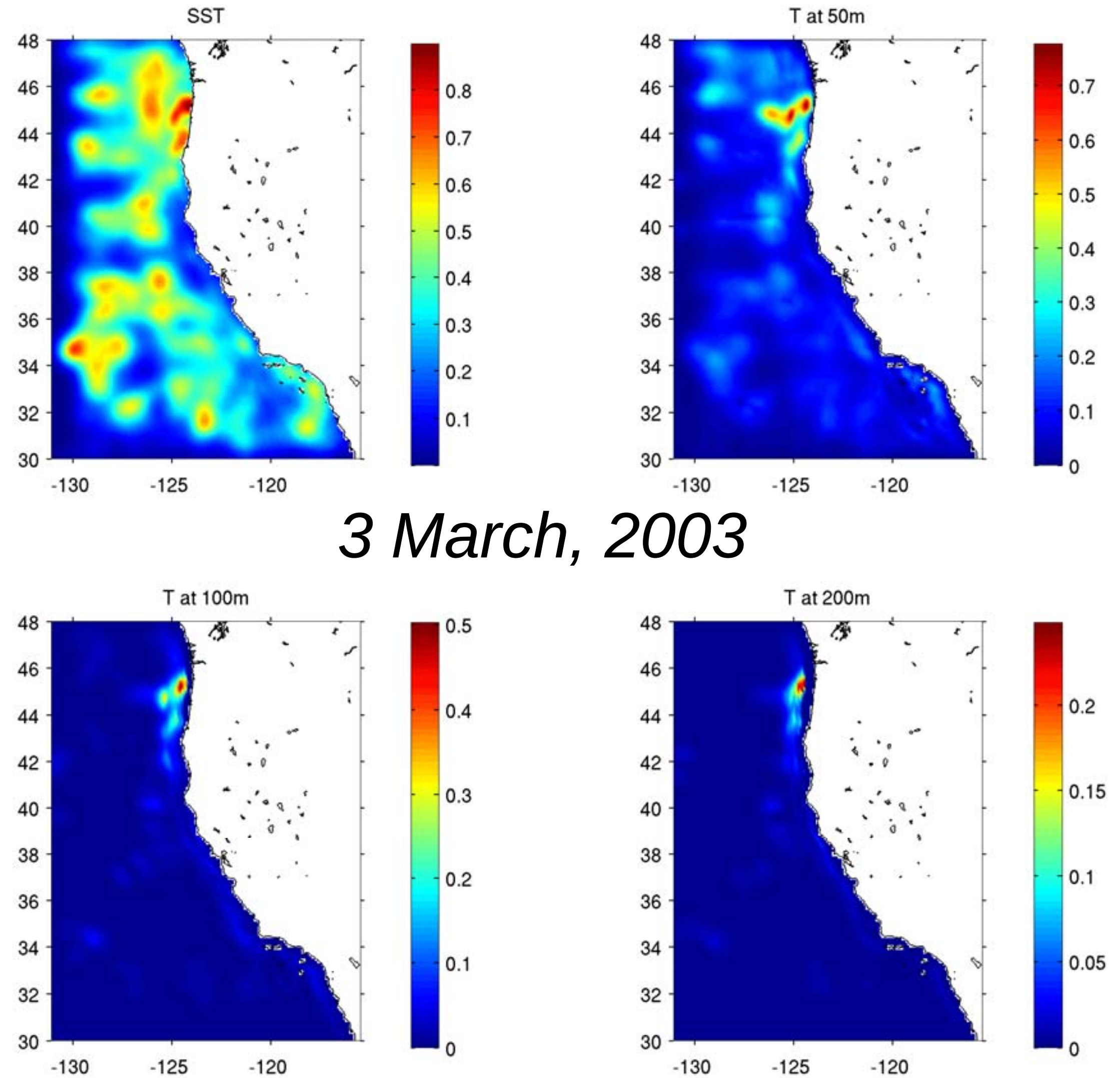
$$\mathbf{E^a} \approx \tilde{\mathbf{E}}^{\mathbf{a}} = \mathbf{D}^{1/2}\mathbf{V}_k\mathbf{T}_k^{-1}\mathbf{V}_k^{\mathrm{T}}\mathbf{D}^{1/2}$$

**Straightforward but not yet implemented in ROMS I4D-Var (due to large I/O requirements)**

# Expected Posterior Error

**R4D-Var**
**1 outer-loop**
**100 inner-loops**

**3-10 March, 2003**
**(10km, 42 levels)**



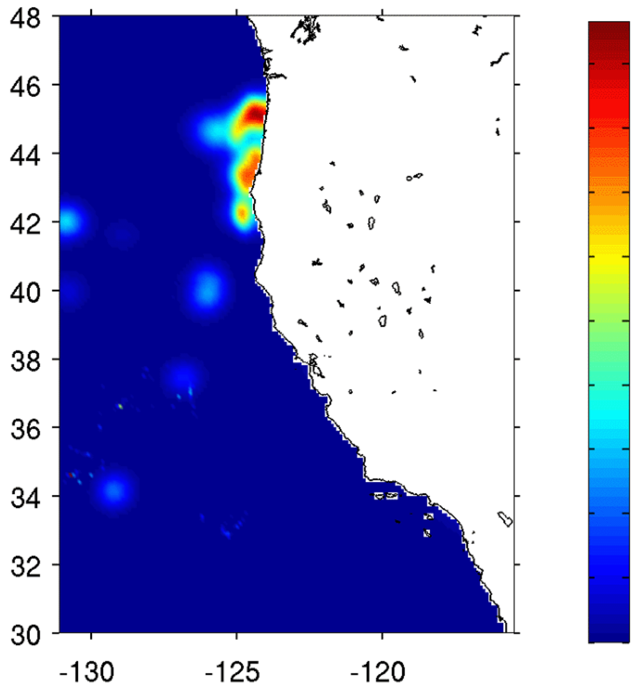

*3 March, 2003*

$\sigma_b$ = *prior error std*

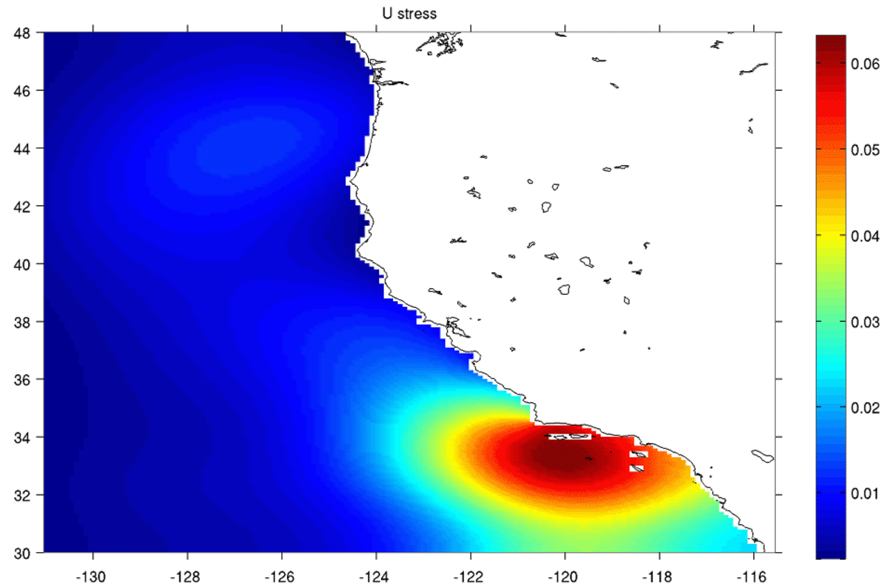
$\sigma_a$ = *posterior error std*

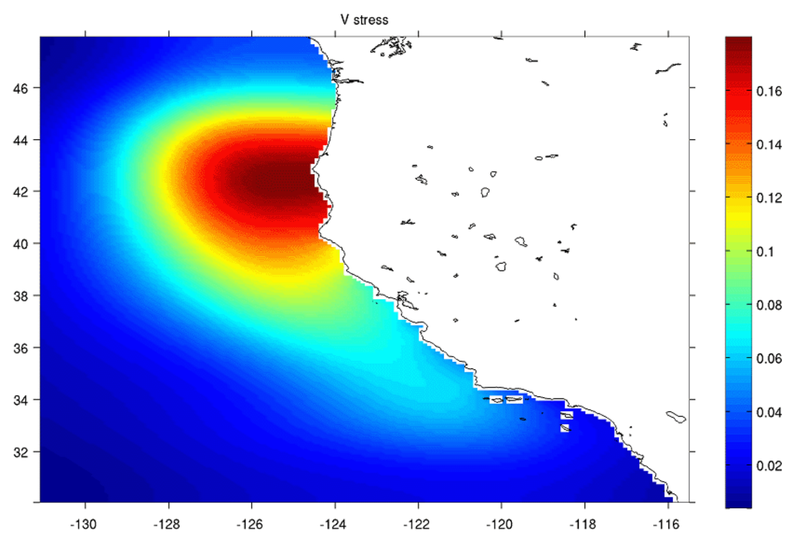
$$(\sigma_b - \sigma_a) / \sigma_b$$

SSS

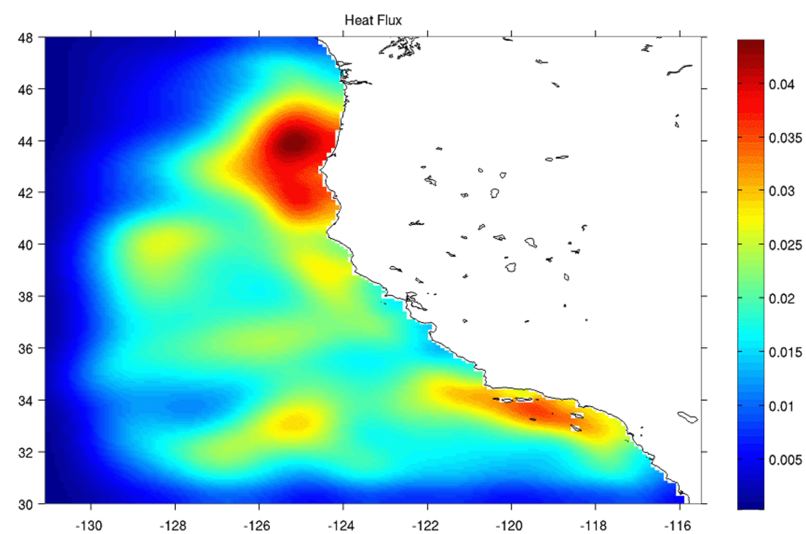



Zonal wind stress



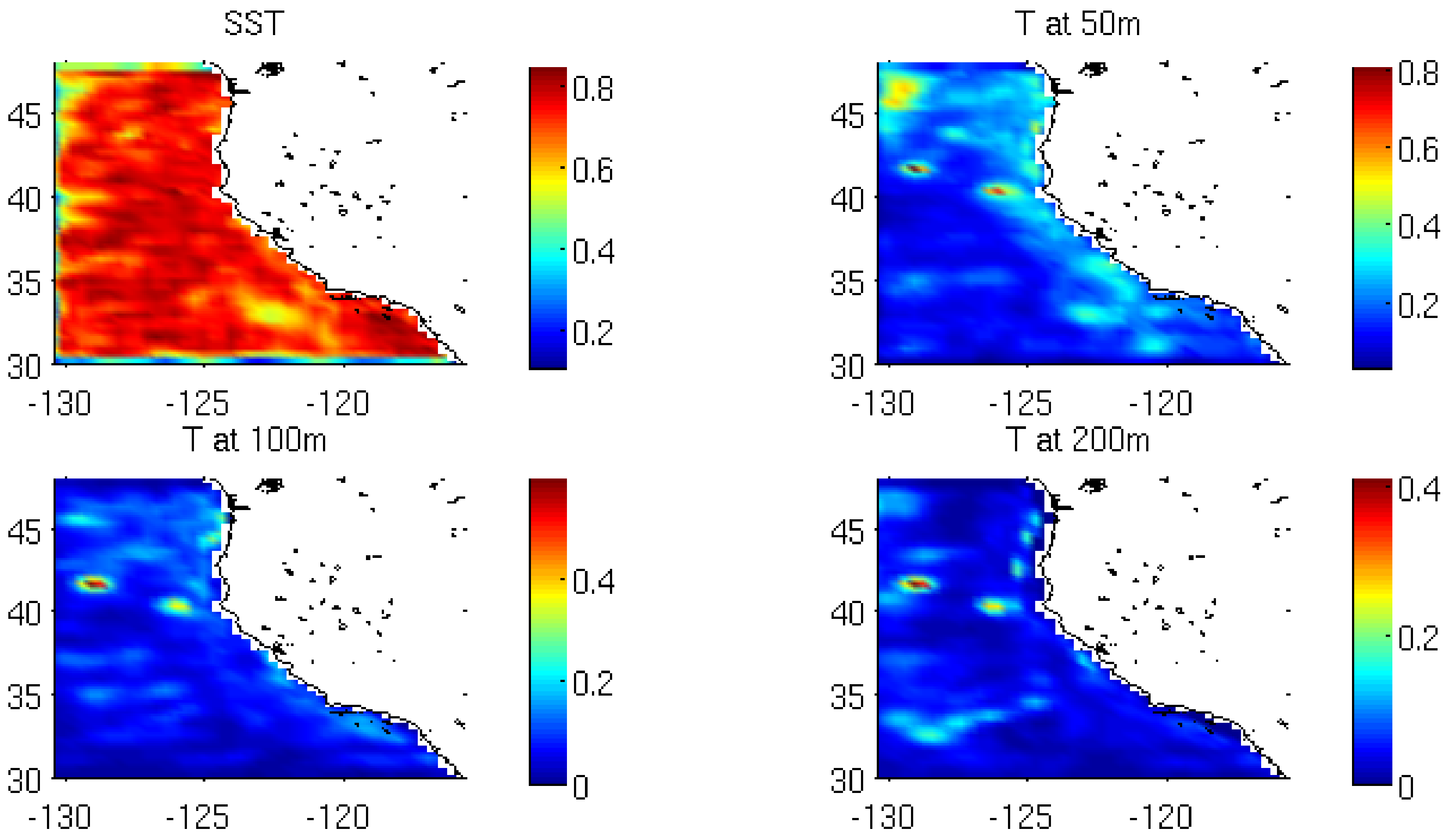




Meridional wind stress




Heat Flux

# Cautionary Note about E<sup>a</sup> Estimates



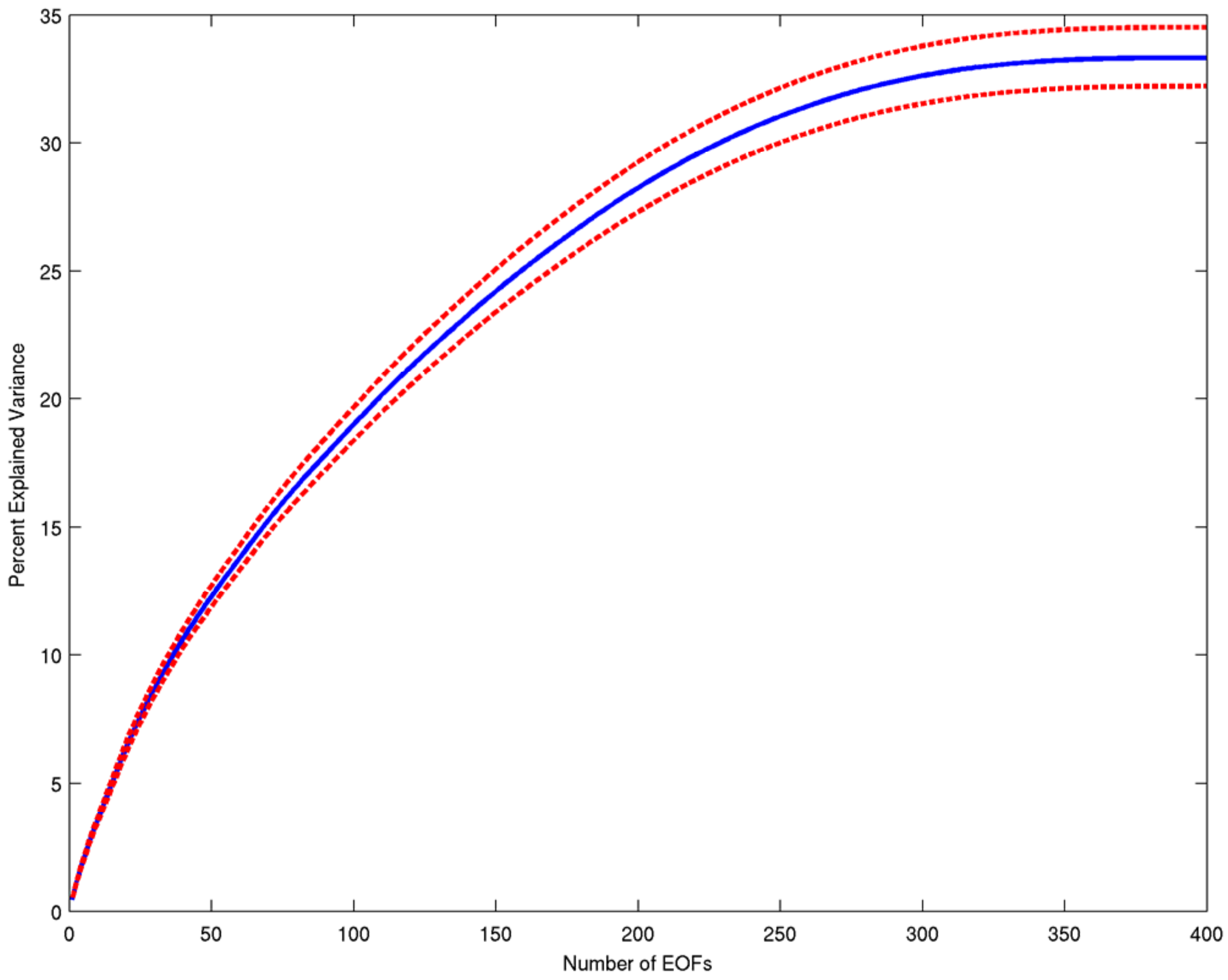

**800 Lanczos vectors vs 100 Lanczos vectors (30 km)**

$$(\sigma_{800}^2 - \sigma_{100}^2) / \sigma_{100}^2$$

**EOF spectrum is quite "flat"**

# Expected Posterior Error EOFs

**Flat spectrum - perhaps due to preconditioning which clusters eigenvalues of ($R^{1/2}GDG^{T}R^{1/2}+I$) around 1.**



# Consistency checks, hypothesis tests, degrees of freedom & information content

# Consistency Checks in Obs Space

**Statistics of the innovation vectors following Desroziers et al (2005):**

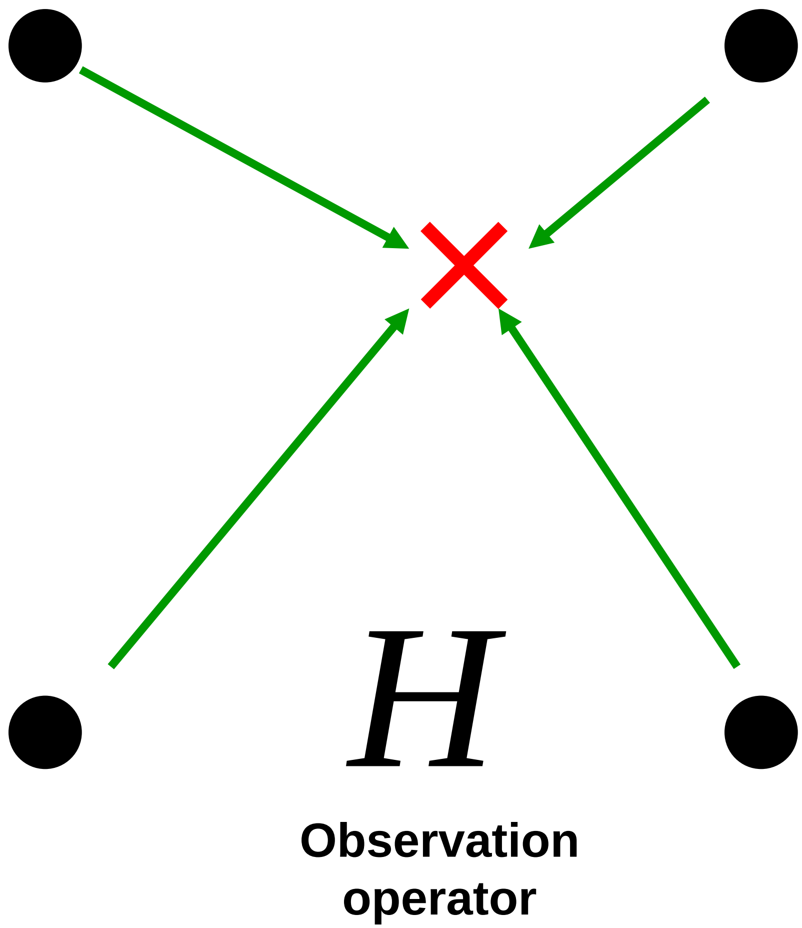
$$\mathbf{d} = (\mathbf{y} - H(\mathbf{x_b}))$$

$$\mathbf{d_a^o} = (\mathbf{y} - H(\mathbf{x_a}))$$

$$\mathbf{d_b^a} = (H(\mathbf{x_a}) - H(\mathbf{x_b}))$$

$$\tilde{\sigma}_b^2 = (\mathbf{d_b^a})^{\mathrm{T}}\mathbf{d} / p$$

$$\tilde{\sigma}_o^2 = (\mathbf{d_a^o})^{\mathrm{T}}\mathbf{d} / p$$




Compare $\tilde{\sigma}_o$ with $\sigma_{\mathrm{o}}$ & $\tilde{\sigma}_b$ with $\sigma_{\mathrm{b}}$

# Consistency Checks in Obs Space

**Statistics of the innovation vectors following Desroziers et al (2005):**

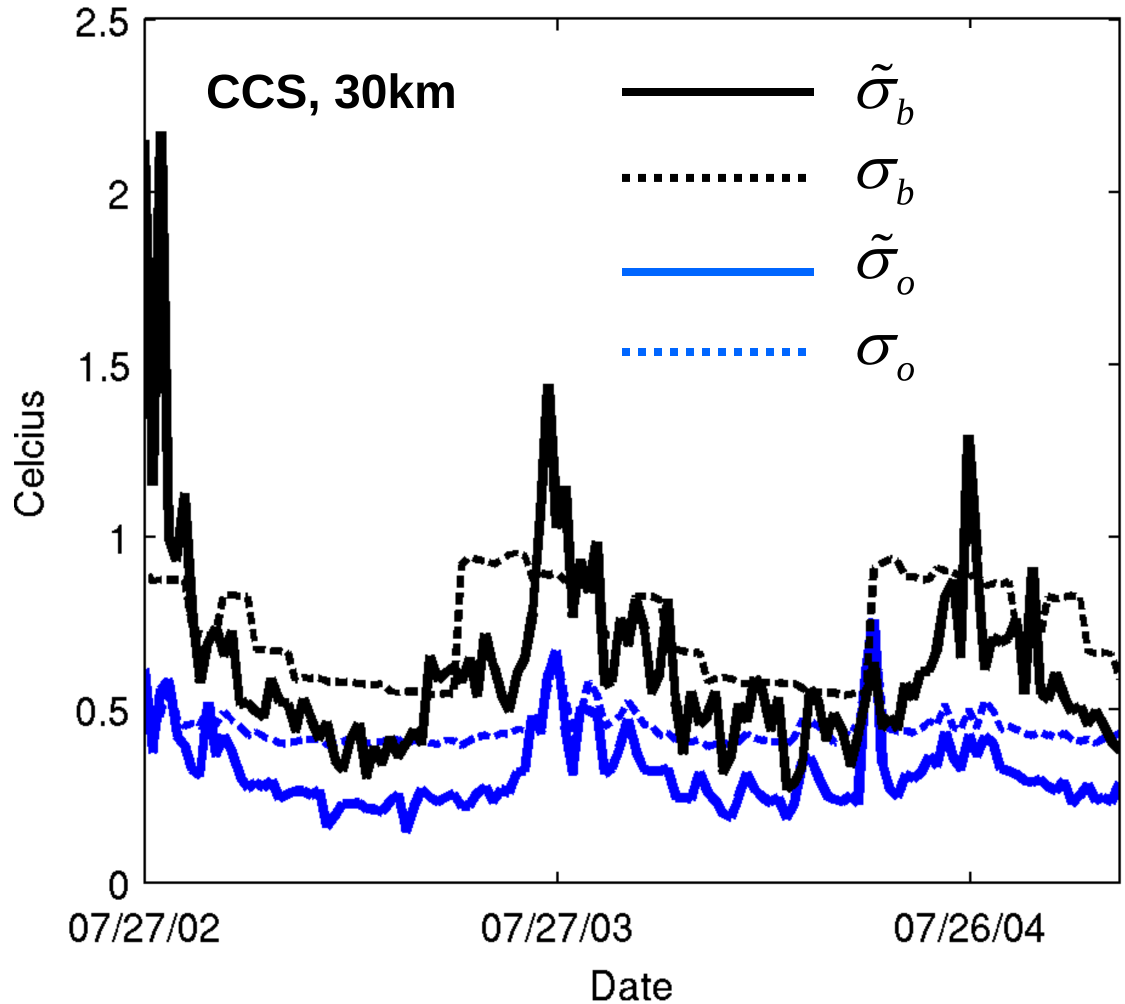
$$\mathbf{d} = (\mathbf{y} - H(\mathbf{x_b}))$$

$$\mathbf{d_a^o} = (\mathbf{y} - H(\mathbf{x_a}))$$

$$\mathbf{d_b^a} = (H(\mathbf{x_a}) - H(\mathbf{x_b}))$$

$$\tilde{\sigma}_b^2 = (\mathbf{d_b^a})^{\mathrm{T}} \mathbf{d} / p$$

$$\tilde{\sigma}_o^2 = (\mathbf{d_a^o})^{\mathrm{T}} \mathbf{d} / p$$




Compare $\tilde{\sigma}_o$ with $\sigma_o$ & $\tilde{\sigma}_b$ with $\sigma_b$

# Hypothesis Tests &Degrees of Freedom

**Recall that the optimal increments minimize:**

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

$$J_b \qquad\qquad\qquad J_o$$

**Theoretical min:** $J_{\mathrm{min}} = N_{\mathrm{obs}}/2$

**No. of dof in obs** $\longrightarrow$ $(J_b)_{\mathrm{min}} = \mathrm{Tr}(\mathbf{KG})/2$

**No. of dof in *prior*** $\longrightarrow$ $(J_o)_{\mathrm{min}} = \left(N_{\mathrm{obs}} - \mathrm{Tr}(\mathbf{KG})\right)/2$

**"dof" – degrees of freedom**

**(Bennett et al, 1993; Cardinali et al, 2004; Desroziers et al., 2009)**

# Degrees of Freedom & Information Content

**Degrees of freedom in the obs:**

$$(J_b)_{\min} = \mathrm{Tr}(\mathbf{KG})/2$$

**But for** $m = N_{obs}$**:**

$$\mathbf{K} = \mathbf{DG}^{\mathrm{T}}\mathbf{R}^{-1/2}\mathbf{V}_m\mathbf{T}_m^{-1}\mathbf{V}_m^{\mathrm{T}}\mathbf{R}^{-1/2}$$

**So:**

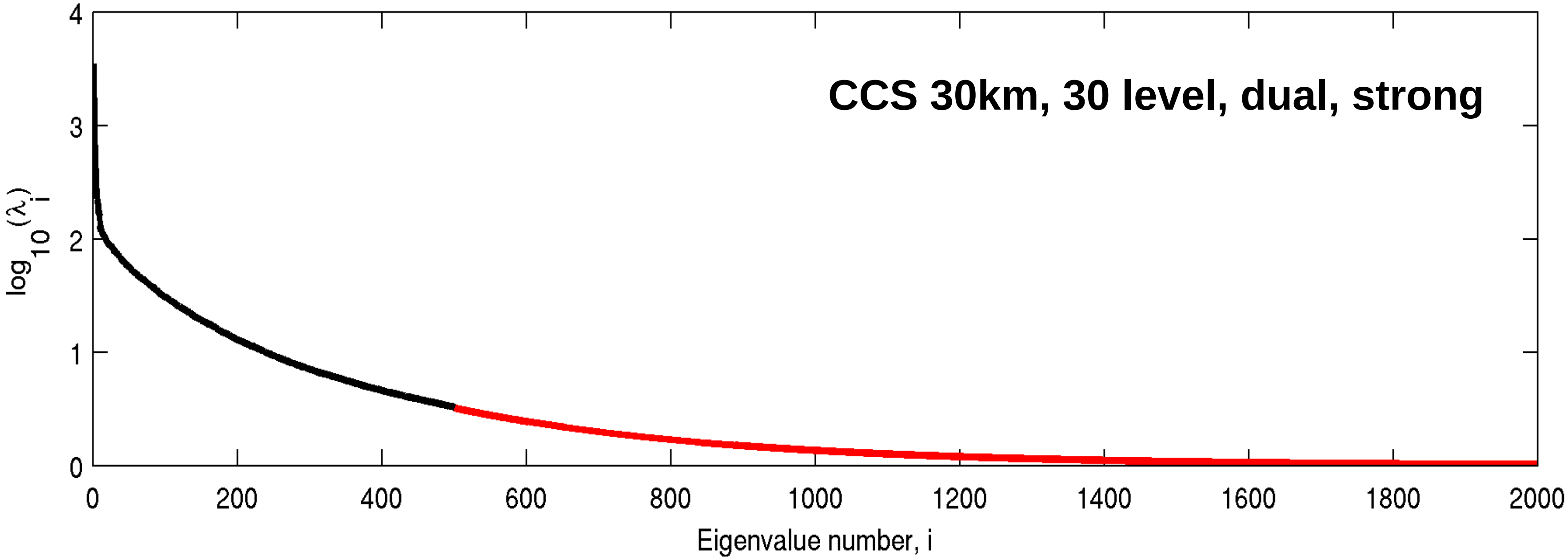
$$(J_b)_{\min} = \frac{1}{2}(N_{obs} - \mathrm{Tr}(\mathbf{T}_m^{-1}))$$

$$= \frac{1}{2}\left(N_{obs} - \sum_{i=1}^{N_{obs}} \boxed{\lambda_i^{-1}}\right)$$

Eigenvalues of $\mathbf{T}_m^{-1}$

# Degrees of Freedom & Information Content

(LhessianEV=T, s4dvar.in)

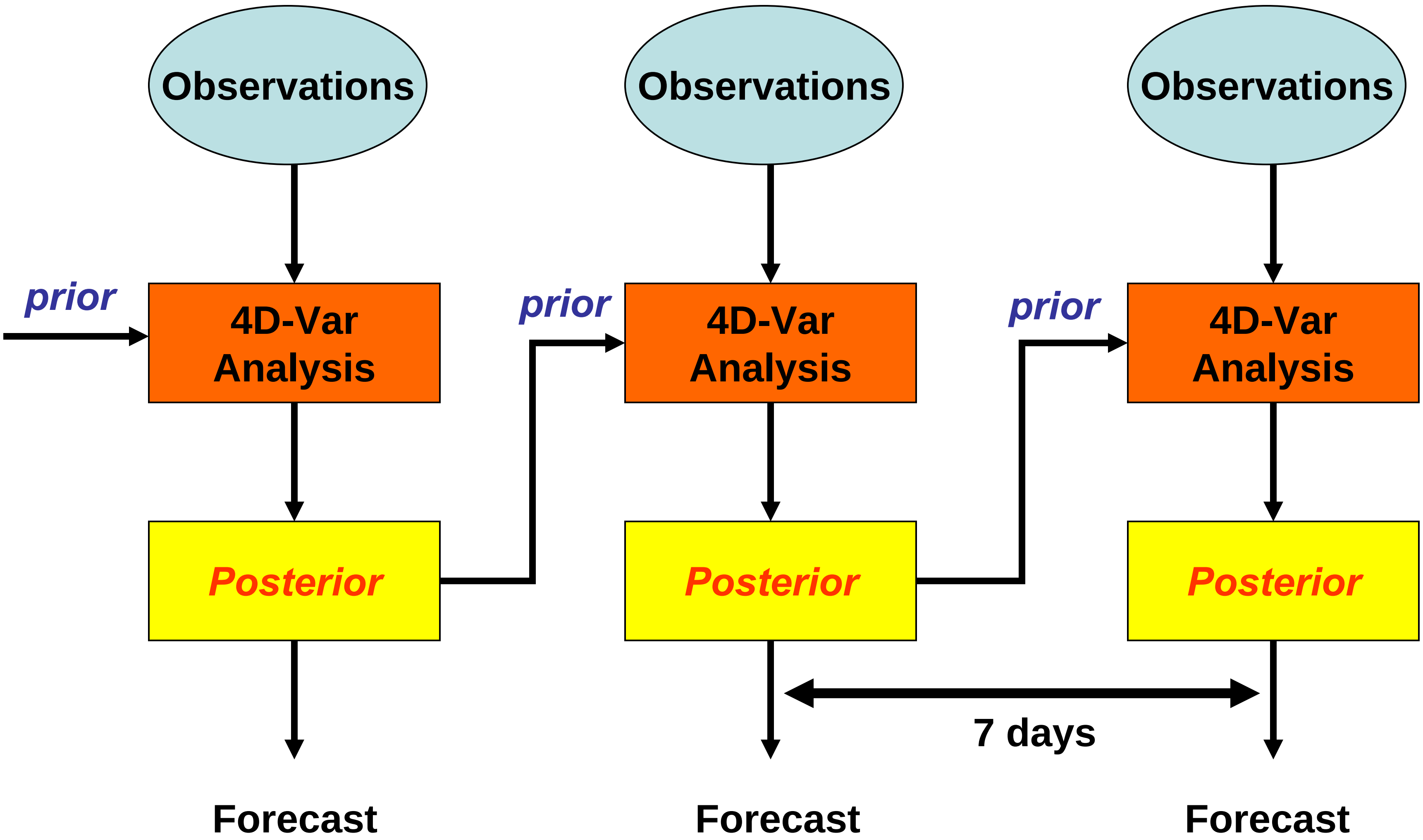



Computed directly during R4D-Var

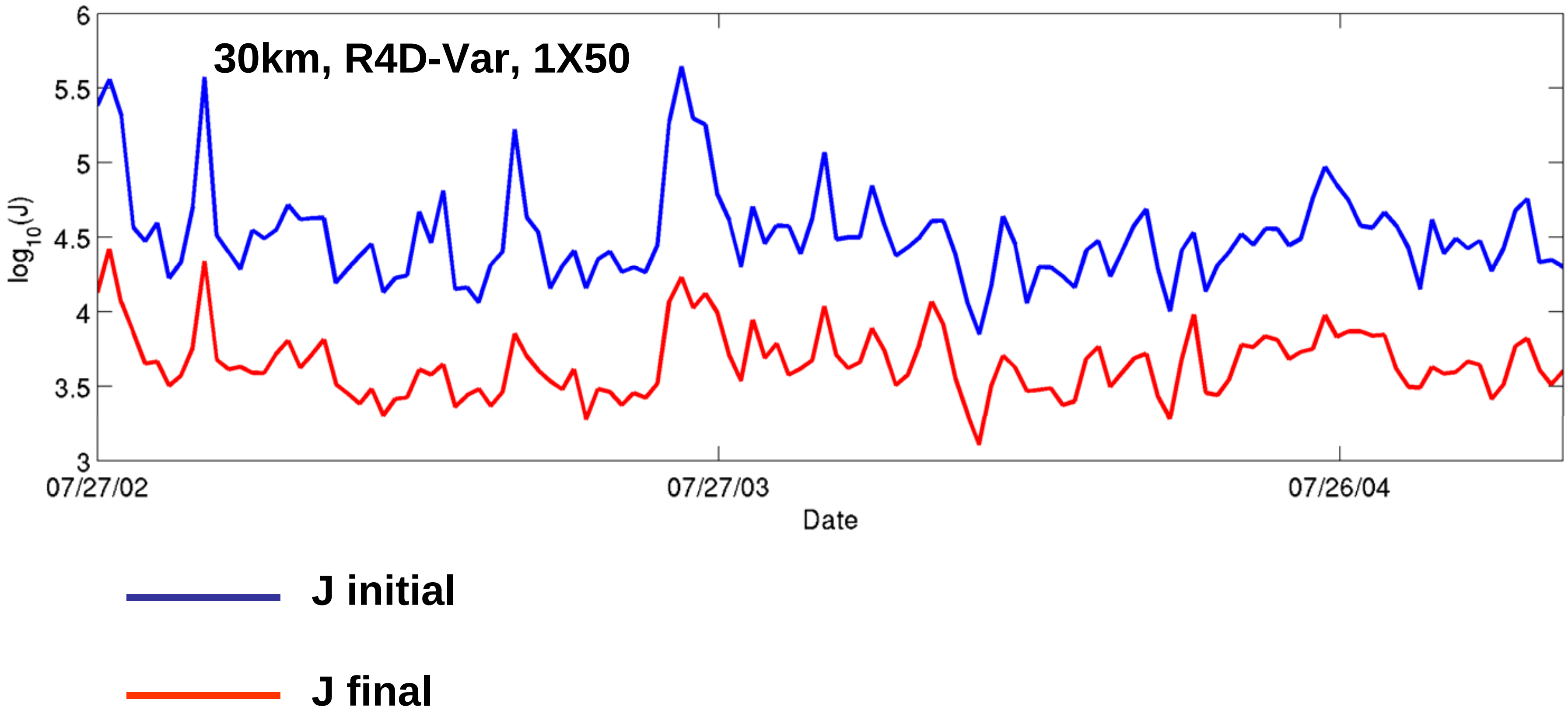
Computed from a curve fit

$$\log_{10}(\lambda_i) = ae^{-bi},\ a = 1.684,\ b = 2.7 \times 10^{-3}$$

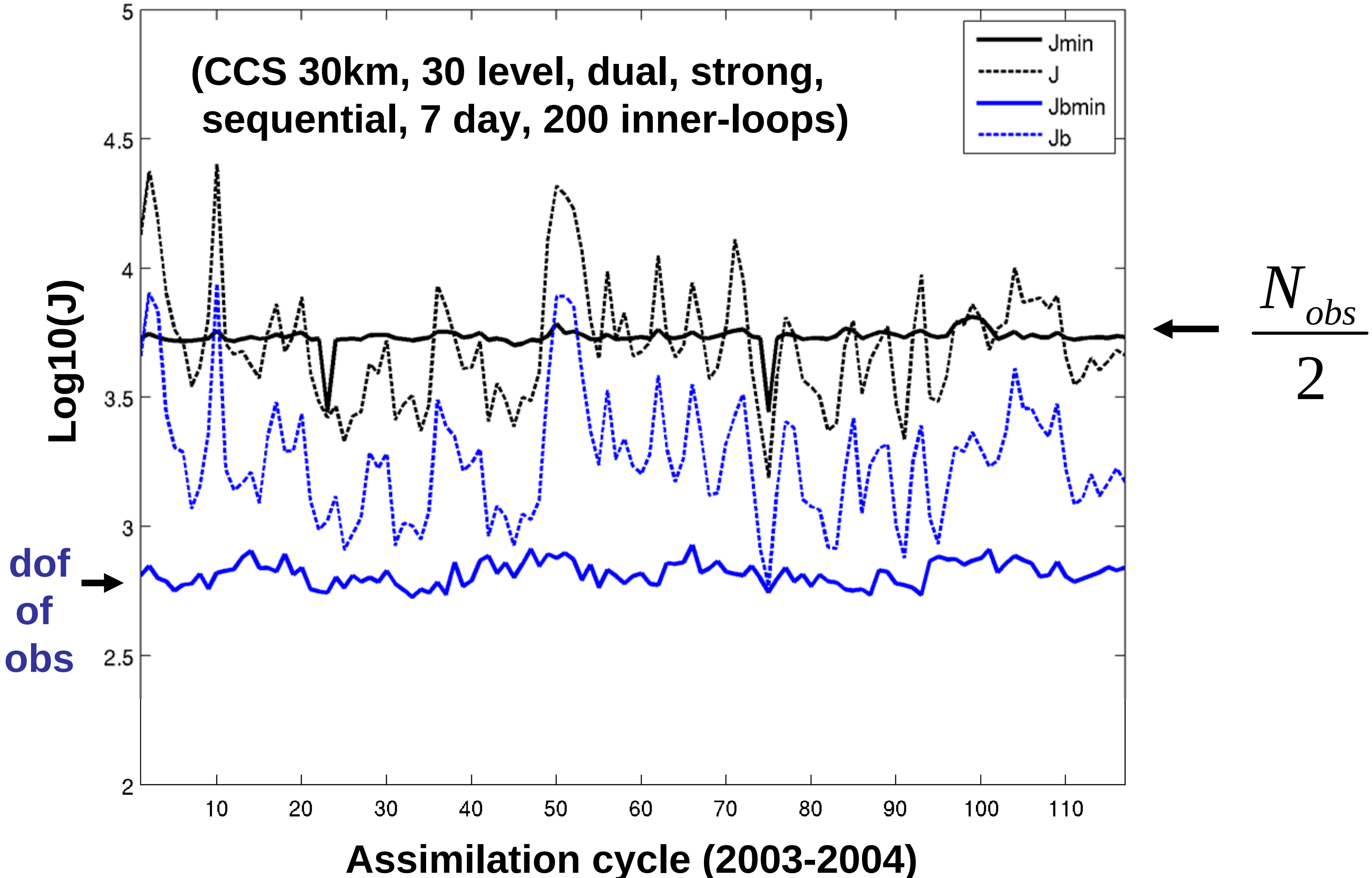
# Sequential 4D-Var with 30km CCS ROMS



# Sequential 4D-Var CCS ROMS



# Degrees of Freedom & Information Content




- **Less than 10% of <u>all</u> observations provide independent info**
- **LOTS OF REDUNDANCY!**
- **$J_b > (J_b)_{min}$ and indicates over fitting to the obs**
- **$J \neq J_{min}$ and indicates that *prior* hypotheses are incorrect**

# Array Modes

**The eigenvalues $\lambda_i$ of $T_m$ are also the eigenvalues of $(R^{1/2}GDG^TR^{1/2}+I)$, the preconditioned "stabilized representer matrix."**

**Consider the eigenpairs:** $\left(\lambda_i, \hat{\mathbf{w}}_i\right)$

**Following Bennett (1985):**

$$\delta \mathbf{x}_{\mathbf{a}}(t) = \mathcal{R}(t)\mathbf{R}^{-1/2}\mathbf{W}\mathbf{\Lambda}^{-1}\mathbf{W}^{\mathrm{T}}\mathbf{R}^{-1/2}\mathbf{d}$$

$$= \sum_{i=1}^{N_{obs}} \lambda_i^{-1}\left(\hat{\mathbf{w}}_i^{\mathrm{T}}\mathbf{R}^{-1/2}\mathbf{d}\right)\mathbf{\Psi}_i(t)$$

**Analysis state-vector increment**

$$\mathbf{\Psi}_i(t) = \sum_{j=1}^{N_{obs}} w'_{ji}\mathbf{r}_j(t)$$

**"array modes"**

$$\mathbf{w}'_i = (w'_{ji}) = \mathbf{R}^{-1/2}\hat{\mathbf{w}}_i$$

# Array Modes

**State-vector analysis increment:**

$$\delta \mathbf{x_a}(t) = \sum_{i=1}^{N_{obs}} \boxed{\lambda_i^{-1}} \left( \hat{\mathbf{w}}_i^{\mathrm{T}} \mathbf{R}^{-1/2} \mathbf{d} \right) \mathbf{\Psi}_i(t)$$

**Largest eigenvalue $\lambda_1$ associated with $\Psi_1$(t), nominally contributes least to analysis increment.**

**Thus $\Psi_1$(t) represents the most stable component of $\delta\mathbf{x_a}$ with respect to changes in d.**

**Array modes can be readily computed after dual 4D-Var:**

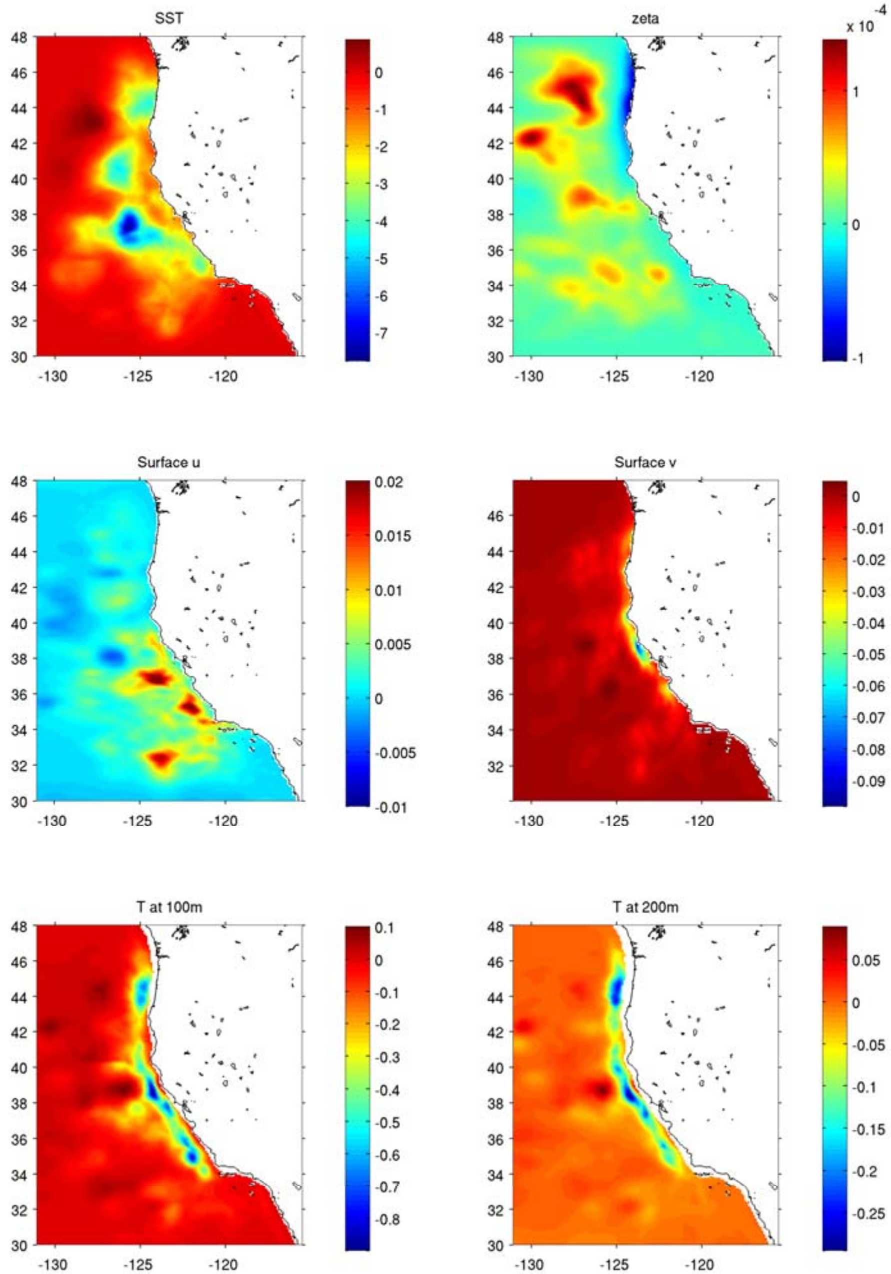
$$\mathbf{\Psi}_j(t) = \mathcal{M}_{\mathbf{b}}(t, t_0)\mathbf{D}\mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1/2}\hat{\mathbf{w}}_j$$

**(define ARRAY_MODES)**

**Example:**

**CCS, 10km**
**R4D-Var**
**3-7 March, 2003**

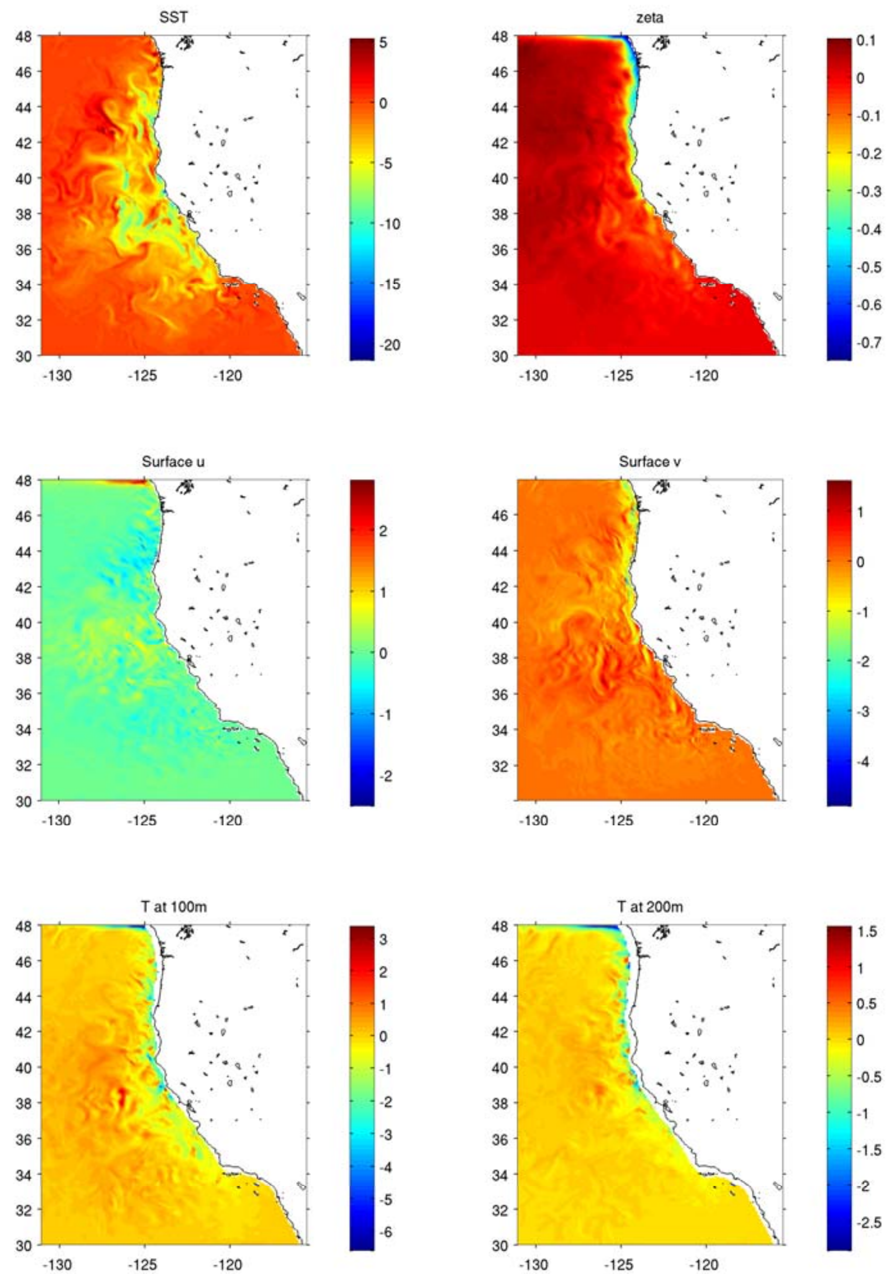
$\Psi_1$ 3 March



**<u>Example</u>:**

**CCS, 10km**
**R4D-Var**
**3-7 March, 2003**

$\Psi_1$ 7 March

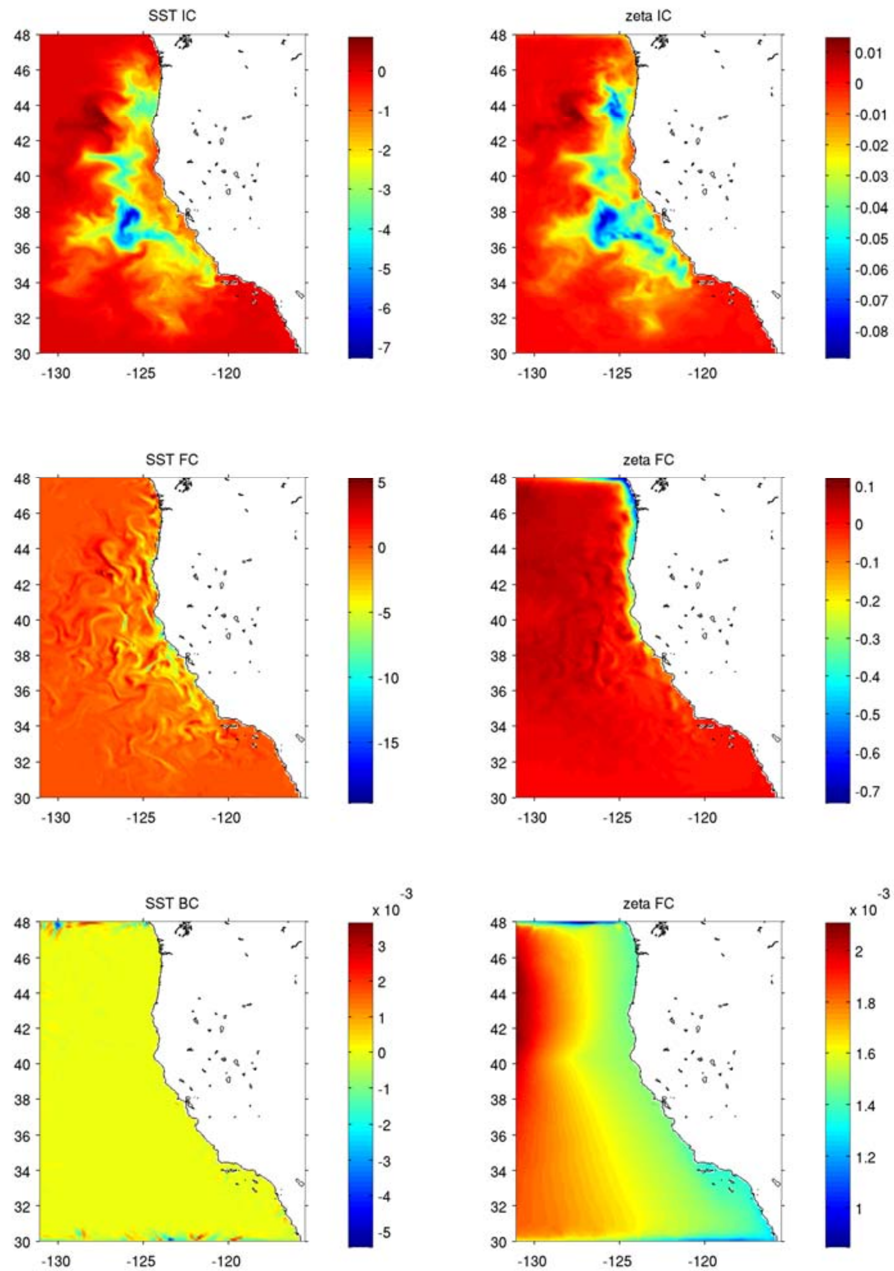


**Example**:

**CCS, 10km**
**R4D-Var**
**3-7 March, 2003**

# $\Psi_1$ 3 March

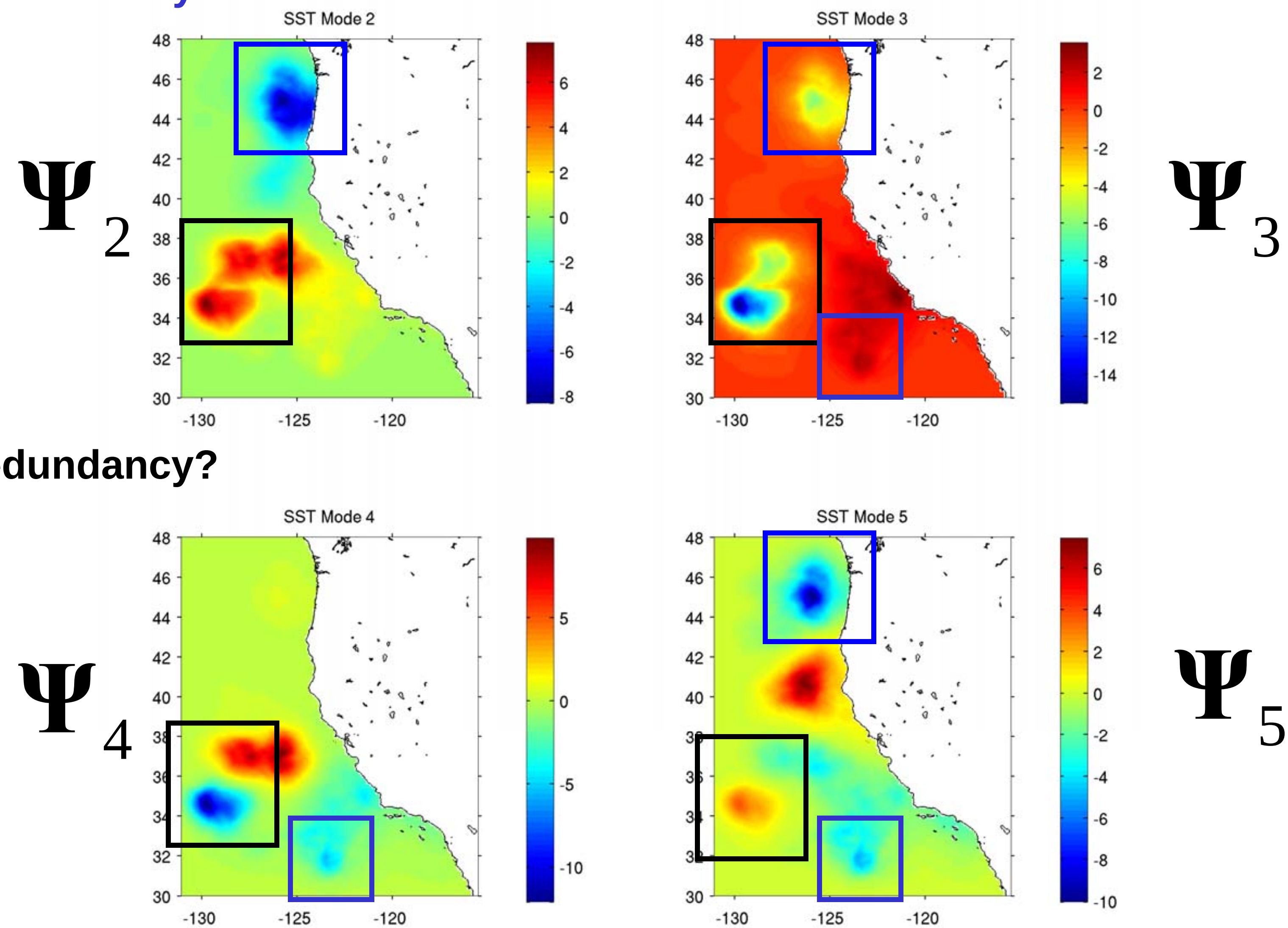
**Control vector contributions**



# Array Mode SST: 3 March

**Obs redundancy?**

**Obs redundancy?**




**Obs redundancy?**

# Clipped Analyses

**Following Bennett (1985)**

$$\delta \mathbf{x_a}(t) = \sum_{i=1}^{M} \lambda_i^{-1} \left( \hat{\mathbf{w}}_i^{\mathrm{T}} \mathbf{R}^{-1/2} \mathbf{d} \right) \boldsymbol{\Psi}_i(t)$$

**Truncate ("clip") the summation, discarding array modes $\Psi_i$ for $i>M$**

**$M$ is based on a criteria that reflects information content (*i.e.* $\lambda_i < \alpha \lambda_1$ for $i>M$).**

**$\alpha$ chosen so $\Psi_i$ have scales < model resolution for $i>M$**

# Clipping using Indirect Representer Algorithm

**Clipped analyses can be computed AFTER running R4D-Var or 4D-PSAS by exploiting the indirect representer algorithm according to:**

$$\delta \mathbf{x_a}(t) = \mathcal{M}_{\mathbf{b}}(t,t_0)\mathbf{D}\mathbf{G}^{\mathrm{T}}\sum_{i=1}^{M}\lambda_i^{-1}\left(\hat{\mathbf{w}}_i^{\mathrm{T}}\mathbf{R}^{-1/2}\mathbf{d}\right)\hat{\mathbf{w}}_i$$

**where $\mathcal{M}_{\mathbf{b}}(t,t_0)$ is TLROMS linearized about the *prior*/background.**

**(Moore et al, 2010b)**

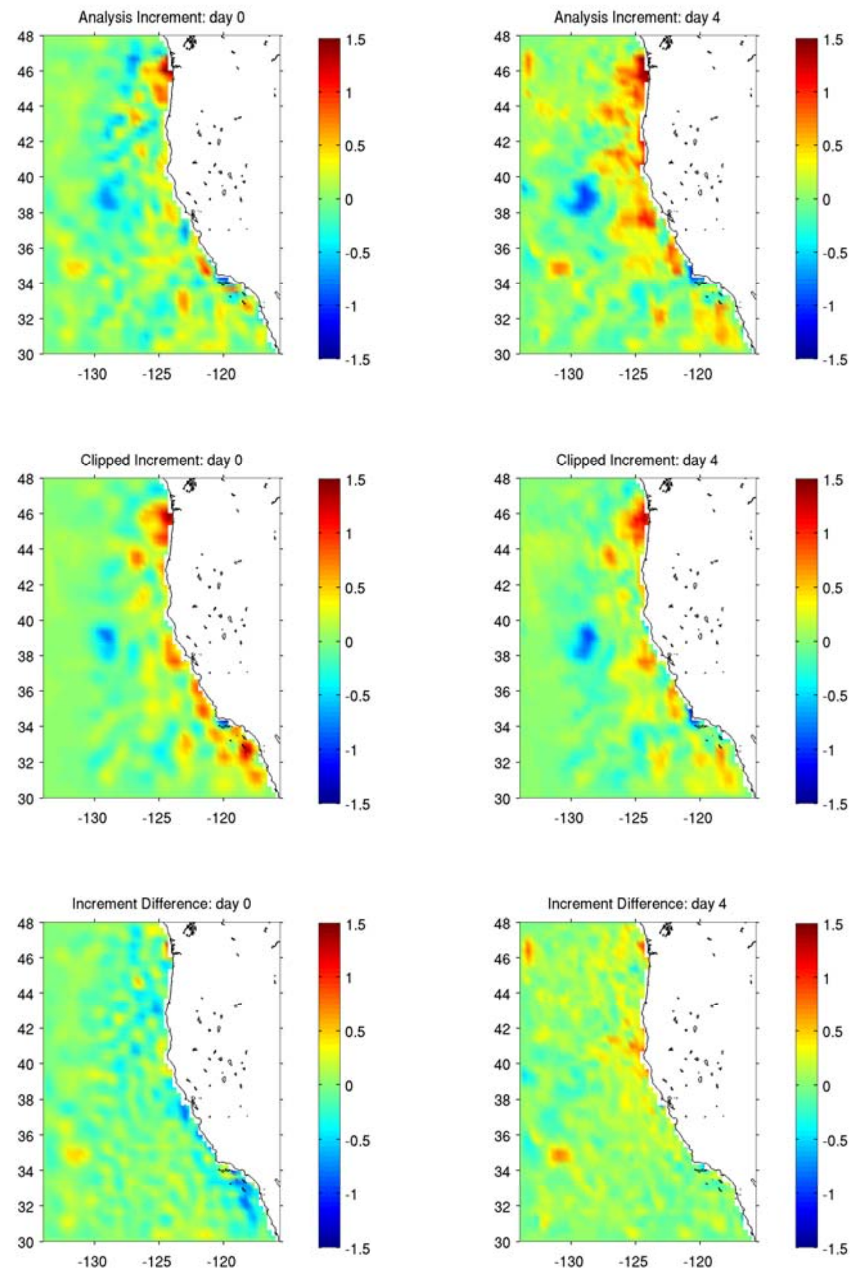
**<u>Example</u>:**
**CCS, 30km**
**R4D-Var**
**3-7 March, 2003**
**1 outer-loop**
**800 inner-loops**

**Clipped using "1% rule" of Bennett & McIntosh (1985)**

**Discard $\Psi_i$ for which**

$$\lambda_i < 0.01\lambda_1$$

$$M = 255$$



# Issues, Things to do, & Coming Soon

- *Posterior* error estimates available for 4D-PSAS & R4D-Var
  - available soon for I4D-Var
- *Posterior* error estimates are likely <u>over estimates</u> due to span of observation or model space by Lanczos vectors
- $(4D\text{-}Var)^T$ provides more reliable estimates (Lecture 5)

# Summary

- Posterior error estimates available via the utility afforded by the Lanczos algorithm.
- Posterior errors tend to be overestimates since $m \ll N_{obs}$
- Consistency checks and hypothesis tests are useful indicators of validity of *prior* hypotheses and info content.
- Array modes provide info about possible data redundancy.
- Clipping can remove unphysical features from analyses.

# References


- Bennett, A.F., 1985: Array design by inverse methods. *Prog. Oceanog*., **15**, 129-156.
- Bennett, A.F. and P.C. McIntosh, 1982: Open ocean modelling as an inverse problem: tidal theory. *J. Phys. Oceanogr*., **12**, 1004-1018.
- Bennett, A.F., L. Leslie, C. Hagelberg and P. Powers, 1993: Tropical cyclone prediction using a barotropic model initialized by a generalized inverse method. *Mon. Wea. Rev*., **121**, 1714-1729.
- Cardinali, C., S. Pezzulli and E. Andersson, 2004: Influence-matrix diagnostic of a data assimilation system. *Q. J. R. Meteorol. Soc*., **130**, 2767-2786.
- Desroziers, G., L. Berre, B. Chapnik and P. Poli, 2005: Diagnosis of observation, background and analysis-error statistics in observation space. *Q. J. R. Meteorol. Soc*., **131**, 3385-3396.
- Desroziers, G., L. Berre, V. Chabot and B. Chapnik, 2009: A posteriori diagnostics in an ensemble of perturbed analyses. *Mon. Wea. Rev*., **137**, 3420-3436.

# References


- Moore, A.M., H.G. Arango, G. Broquet, B.S. Powell, J. Zavala-Garay and A.T. Weaver, 2010a: The Regional Ocean Modeling System (ROMS) 4-dimensional data assimilation systems: Part I – System overview. *Ocean Modelling*, Submitted.
- Moore, A.M., H.G. Arango, G. Broquet, C.. Edwards, M. Veneziani, B.S. Powell, D. Foley, J. Doyle, D. Costa and P. Robinson, 2010b: The Regional Ocean Modeling System (ROMS) 4-dimensional data assimilation systems: Part II – Performance and application to the California Current System. *Ocean Modelling*, Submitted.