Reduced-Rank Array Modes of the California Current Ocean Observing System

Andy Moore[1], Hernan Arango[2] & Chris Edwards[1]

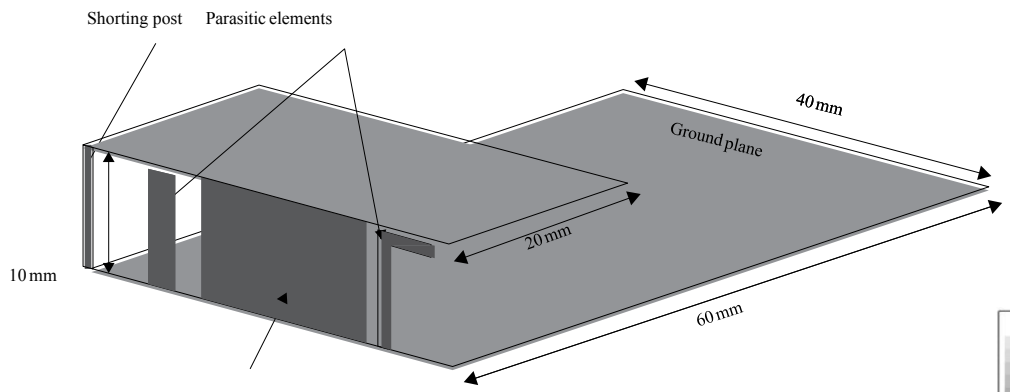
1: UC Santa Cruz

2: Rutgers University

Outline

- Array modes – an idea borrowed from antenna theory
- Reduced-rank Array Modes (RAMs)
- RAMs of the California Current observing system
- Overfitting of DA to the data – a useful stopping criteria
- How effective are the observations in constraining the background errors?

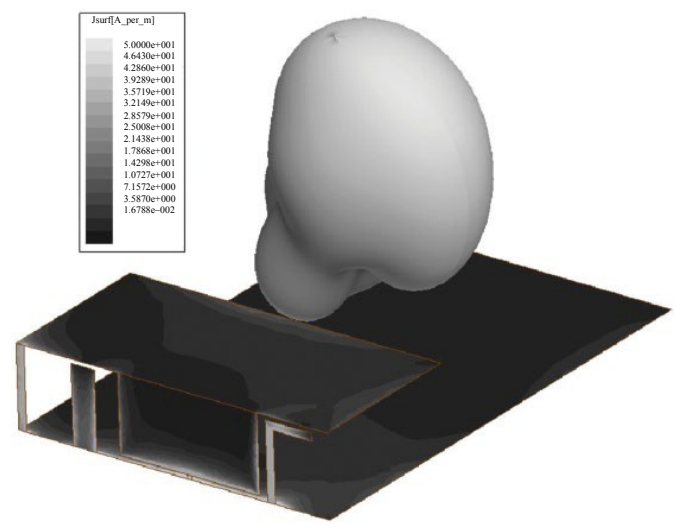
Characteristic Modes of an Antenna



Geometry of a multiband planar inverted-F antenna.

Figures from Chen and Wang (2015: *Characteristic Modes: Theory and Applications in Antenna Engineering*)

Resonance $\propto 1/|\lambda-\omega|$



Oceanographic applications:

- **Bennett (1985)**
- **Le Hénaff et al (2009)**

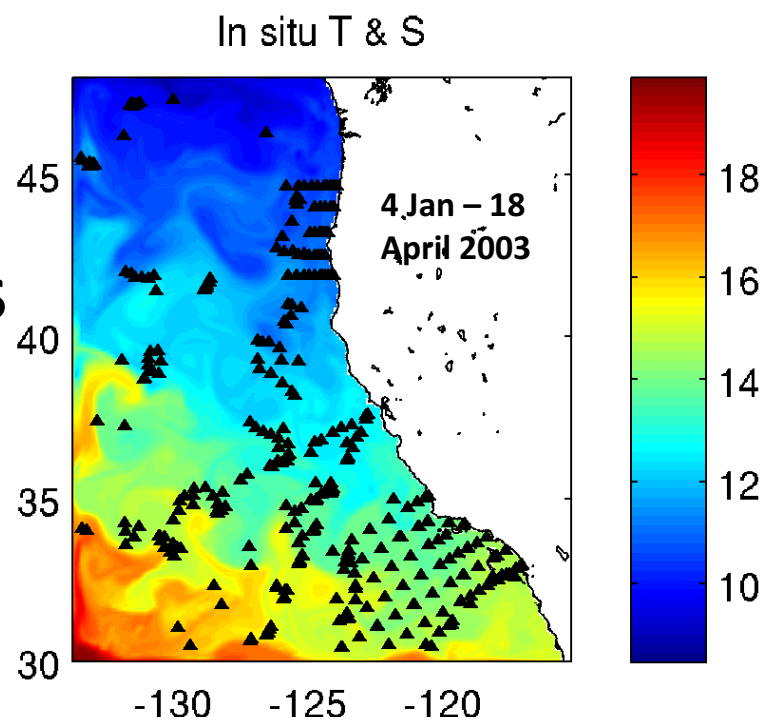
Response $\propto 1/\lambda$

Current distribution and radiation pattern for a characteristic mode with a resonant frequency of 4.1GHz

Array Modes: Assessing the Efficiency of the CA Current Observing System

- **How well does the CA Current observing system "observe" the circulation given our prior hypotheses about the errors?**

ROMS CCS domain



CCS observing system:

- **Satellite SST – daily (AVHRR, AMSR, MODIS)**
- **Aviso gridded SSH - daily**
- ***In situ* T & S profiles**

The Analysis Equation

- Recall the analysis equation:

$$\mathbf{x}_a = \mathbf{x}_b + \boxed{\mathbf{B}\mathbf{G}^{\mathrm{T}}\left(\mathbf{G}\mathbf{B}\mathbf{G}^{\mathrm{T}} + \mathbf{R}\right)^{-1}\left(\mathbf{y} - H(\mathbf{x}_b)\right)}$$

analysis

background

adjoint obs operator

background error cov matrix

obs error cov matrix

obs

obs operator

Analysis increment

- Solved iteratively in ROMS 4D-Var system using the Lanczos formulation of the **B**-preconditioned RPCG.

(Reduced Rank) Array Modes (RAMs)

In **ROMS 4D-Var** the analysis equation is solved using the Lanczos algorithm:

$\equiv(\mathbf{R}^{-1}\mathbf{GBG}^{\mathrm{T}}+\mathbf{I})^{-1}$ *m*=# of inner-loops

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{BG}^{\mathrm{T}}\boxed{\mathbf{V}_{\mathrm{m}}\mathbf{T}_{\mathrm{m}}^{-1}\mathbf{V}_{\mathrm{m}}^{\mathrm{T}}\mathbf{GBG}^{\mathrm{T}}}\mathbf{R}^{-1}(\mathbf{y}-H(\mathbf{x}_b))$$

EOFs of $(\mathbf{R}^{-1}\mathbf{GBG}^{\mathrm{T}}+\mathbf{I})$

This can be rewritten as: $\mathbf{V}_{\mathrm{m}}^{\mathrm{T}}\mathbf{GBG}^{\mathrm{T}}\mathbf{V}_{\mathrm{m}} = \mathbf{I}_{\mathrm{m}}$

$$\mathbf{x}_a = \mathbf{x}_b + \sum_{i=1}^{m}\alpha_i\Psi_i \quad \text{where} \quad \Psi_i = \mathbf{BG}^{\mathrm{T}}\boxed{\mathbf{V}_{\mathrm{m}}\mathbf{u}_i}$$ are the *"RAMs"*

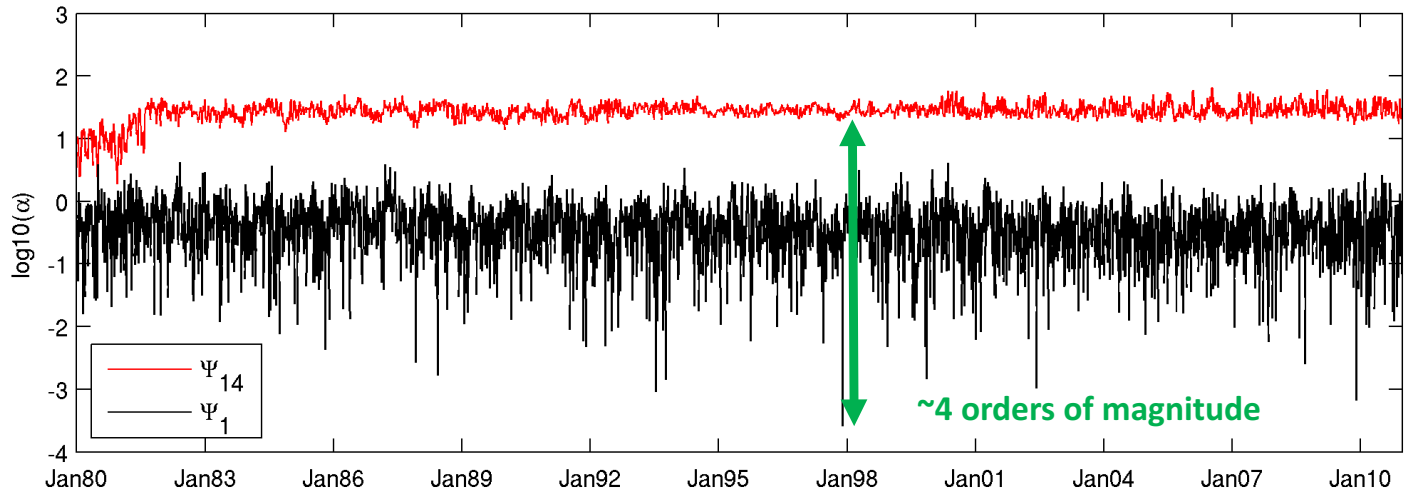
$$\boxed{\alpha_i = \lambda^{-1}\mathbf{u}_i^{\mathrm{T}}\mathbf{V}_{\mathrm{m}}^{\mathrm{T}}\mathbf{GBG}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y}-H(\mathbf{x}_b))}$$

Array mode amplitudes depend on the obs values, *y*

$(\lambda_i, \mathbf{u}_i)$ are the eigenpairs of $\mathbf{T}_{\mathrm{m}}$

NOTE: The array modes depend *ONLY* on the obs locations, and *NOT* the obs values

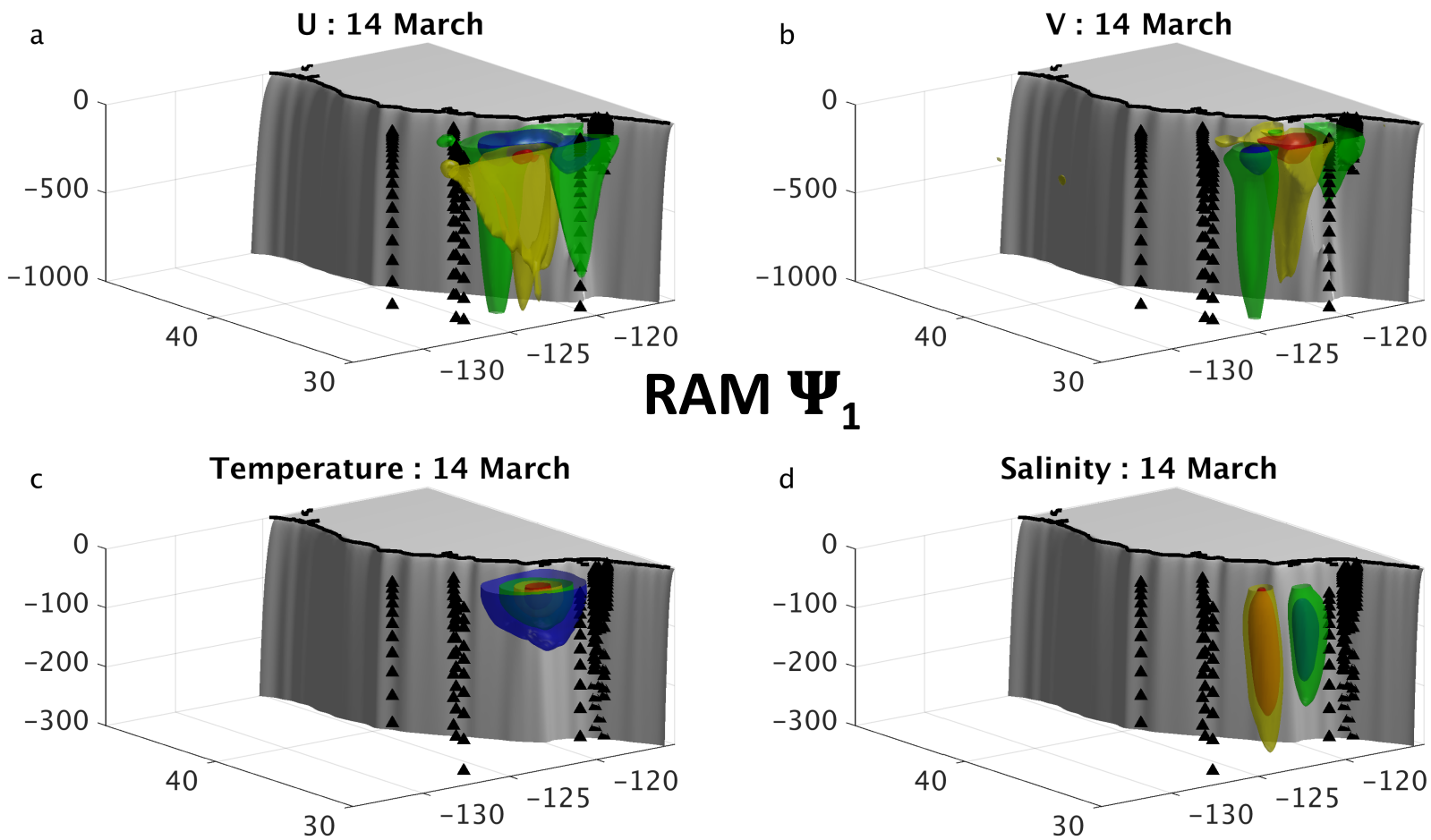
RAM amplitudes of the California Current system



31 year sequence of 4D-Var analyses (1980-2010)
8 day overlapping windows
Obs assimilated: SST, SSH, in situ T and S
1 outer-loop, 14 inner-loops

14 March, 2001

The sub-space activated by the obs & d.f.s

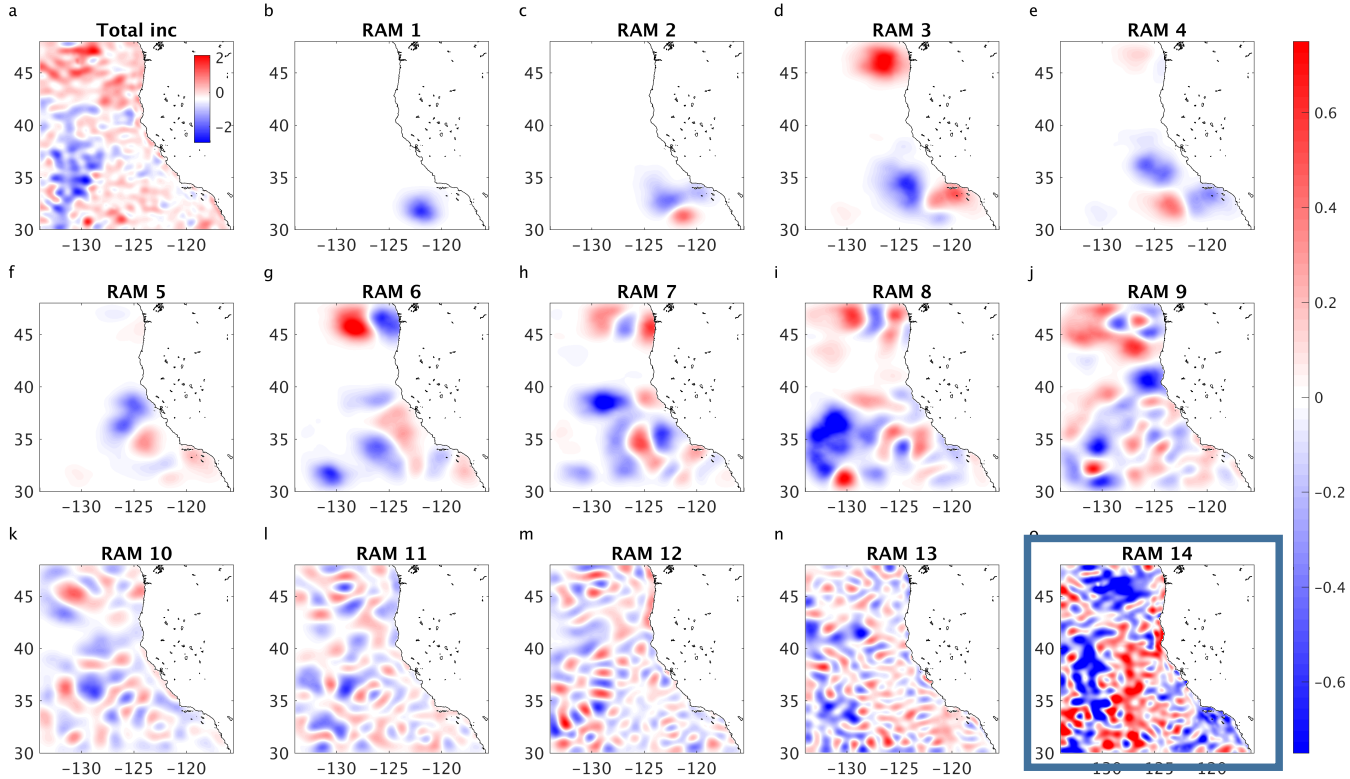


RAM Ψ_1

Overfitting of the Model to the Observations

Contribution of each array mode to the SST increment on 14 March 2001: recall *m=14 inner-loops*.

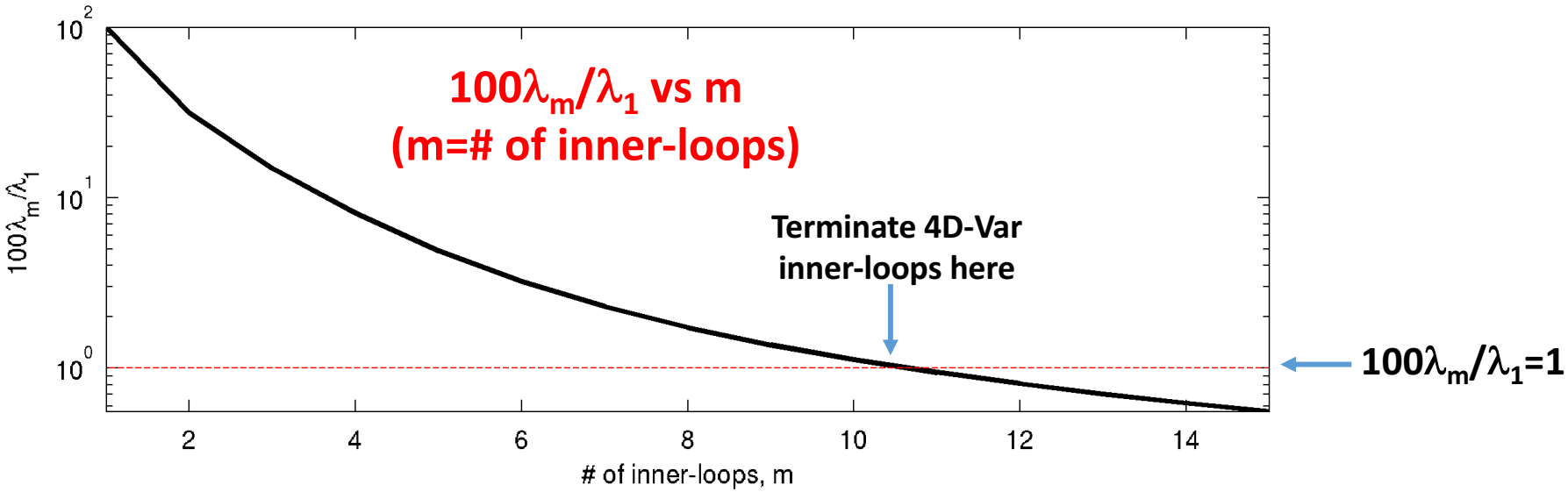
The Bennett and McIntosh (1984) "1% rule": Discard array modes with eigenvalues 1% or less of the max value.



SST

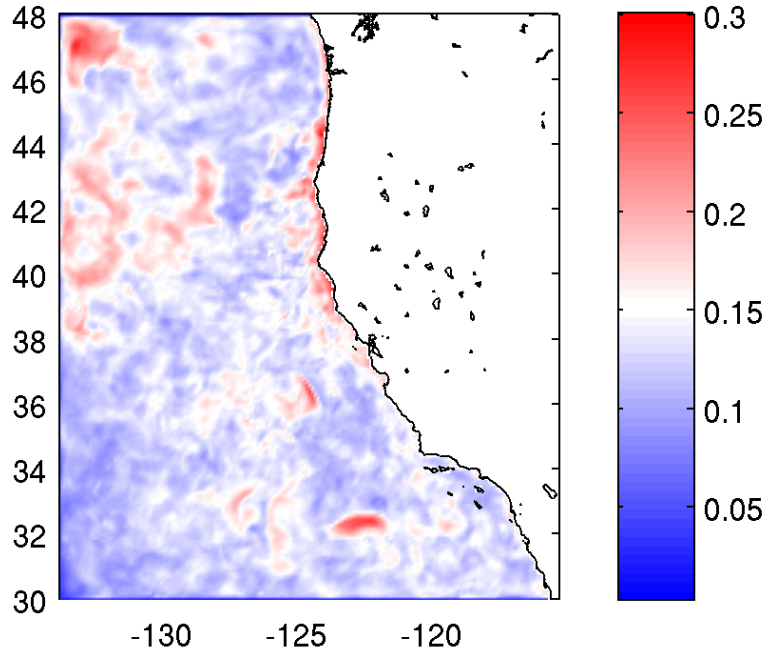
Reject based on 1% rule

The "1% Rule" to avoid overfitting to the Observations



Average eigenvalue ratio for all cycles vs number of inner-loops employed: (suggests we should use no more than 10 inner-loops to prevent over-fitting of the observations)

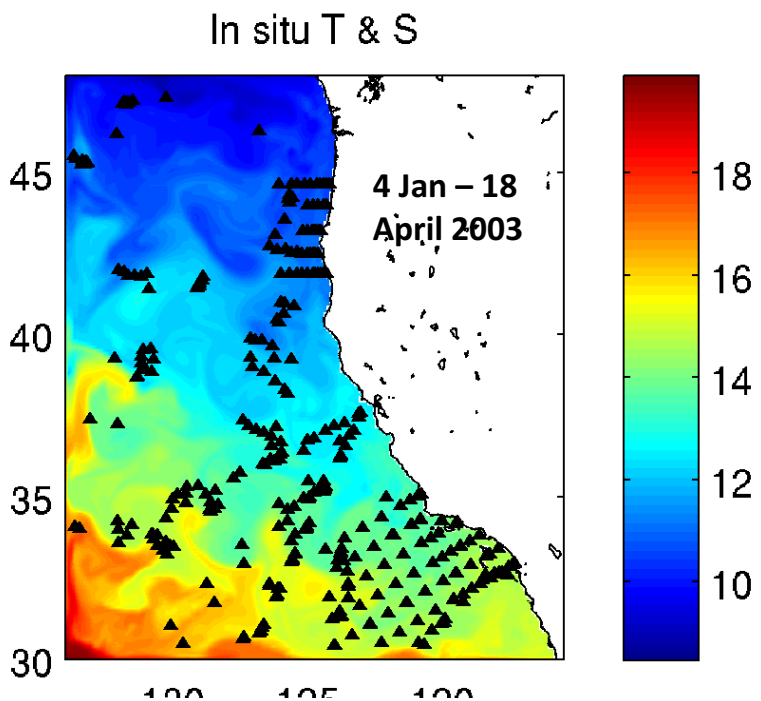
RMS SST difference (14 inner-loops minus 10 inner-loops) for 2001



RAMs: Assessing the Efficiency of the CA Current Observing System

- **How effective is the CA Current observing system at constraining the circulation given our prior hypotheses about the errors B and R?**

ROMS CCS domain



CCS observing system:

- **Satellite SST – daily (AVHRR, AMSR, MODIS)**
- **Aviso gridded SSH - daily**
- ***In situ* T & S profiles**

The Analysis Equation

- Recall the analysis equation:

$$\mathbf{x}_a = \mathbf{x}_b + \boxed{\mathbf{B}\mathbf{G}^T\left(\mathbf{G}\mathbf{B}\mathbf{G}^T + \mathbf{R}\right)^{-1}\left(\mathbf{y} - H(\mathbf{x}_b)\right)}$$

analysis

background

adjoint obs operator

background error cov matrix

obs error cov matrix

obs

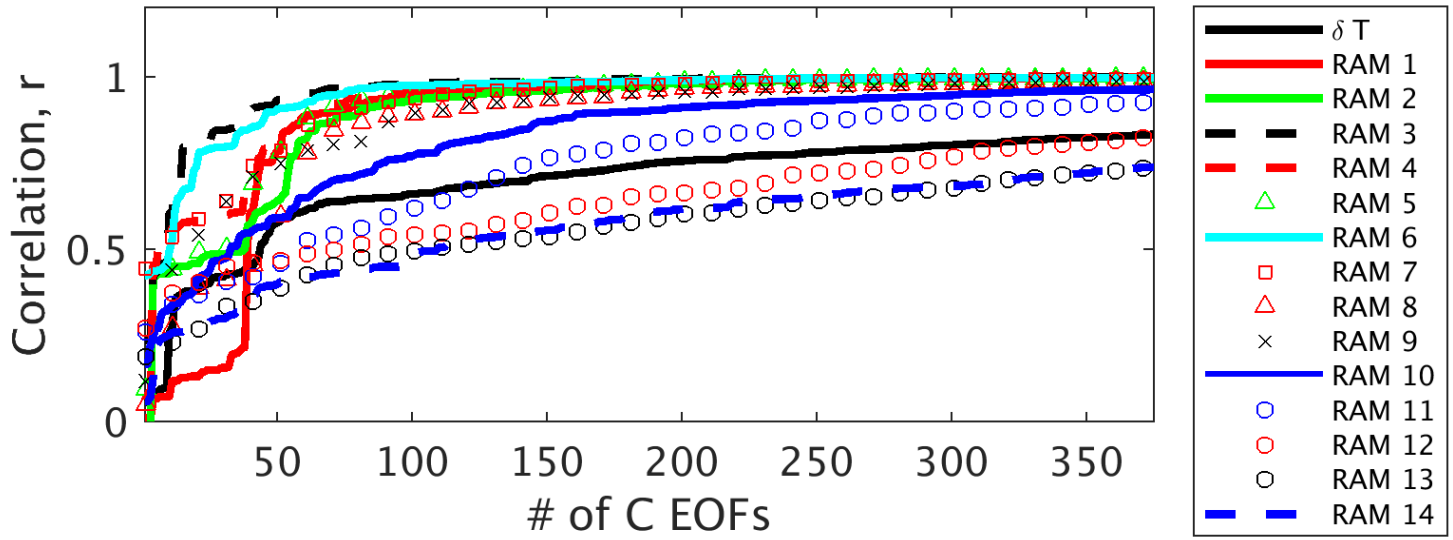
obs operator

Analysis increment

The analysis increment resides in the space spanned by B !!!

Therefore, to reduce errors in x_b, the observing system must effectively observe (directly via G or indirectly via G^T) the dominant EOFs of B.

Projection of RAMs on the EOFs of B



B is $O(10^6 \times 10^6)$

Number of EOFs required to recover RAMs << dim(**B**)

=> the observing system poorly constrains much of the space spanned by **B**

Summary & Conclusions

- The RAMs identify the sub-space informed by the observations during 4D-Var
- RAMs can be computed from archived 4D-Var output
- RAMs with small λ will amplify obs errors
- A useful practical stopping criteria identified for 4D-Var
- The current array observes the EOFs of **B** rather poorly
- Pseudo-observations could be used to optimize the observing system