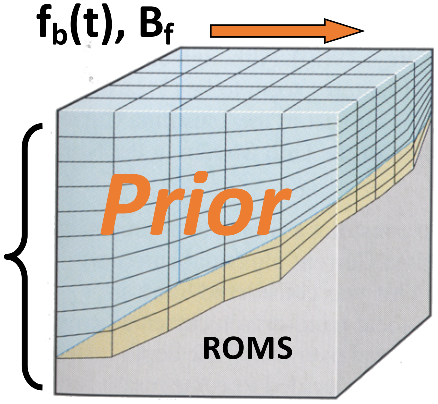
Lecture 5:
Array Modes

Data Assimilation

Model



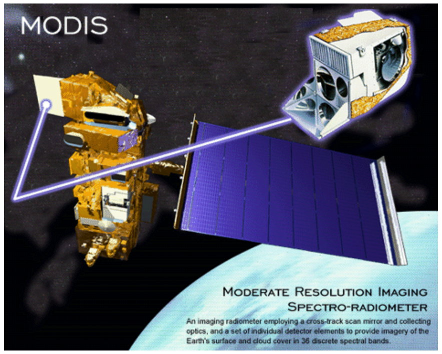
$f_b(t)$, B_f

$b_b(t)$, B_b

$x_b(0)$, B_x

+

Observations



An imaging radiometer employing a cross-track scan mirror and collecting optics, and a set of individual detector elements to provide imagery of the Earth's surface and cloud cover in 36 discrete spectral bands.

The control vector:

$$\mathbf{z} = \begin{pmatrix} \mathbf{x}(0) \\ \mathbf{f} \\ \mathbf{b} \end{pmatrix}$$

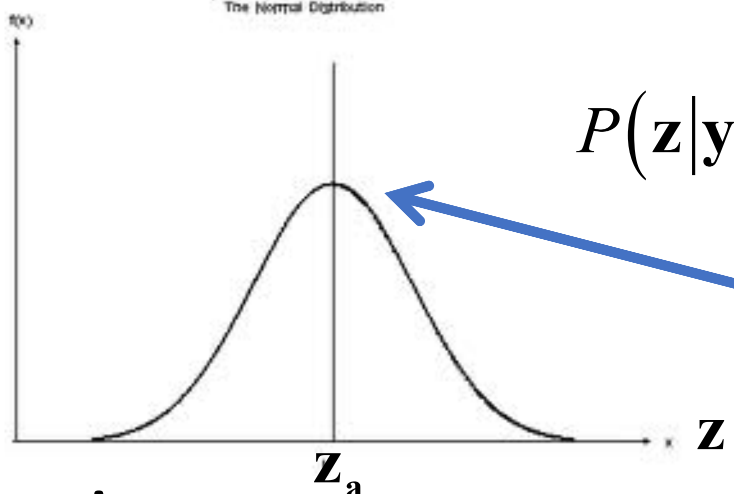
***Prior* error covariance:**

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{\mathbf{x}} & & \\ & \mathbf{B}_{\mathbf{f}} & \\ & & \mathbf{B}_{\mathbf{b}} \end{pmatrix}$$

Maximum Likelihood Estimate & 4D-Var

The Reverend Bayes would have said:

Probability $P(\mathbf{z}|\mathbf{y})$



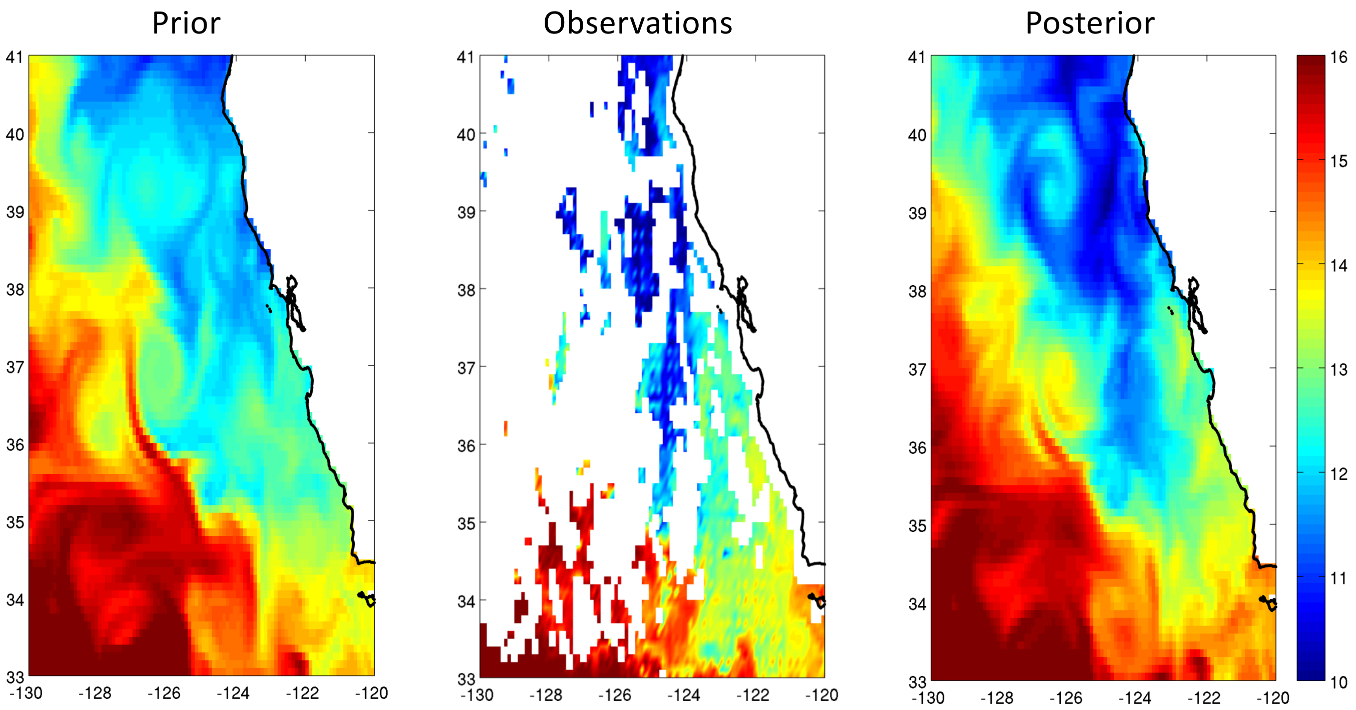
$$P(\mathbf{z}|\mathbf{y}) \propto \exp(-J)$$

Maximize $P(\mathbf{z}|\mathbf{y})$ by minimizing J using variational calculus

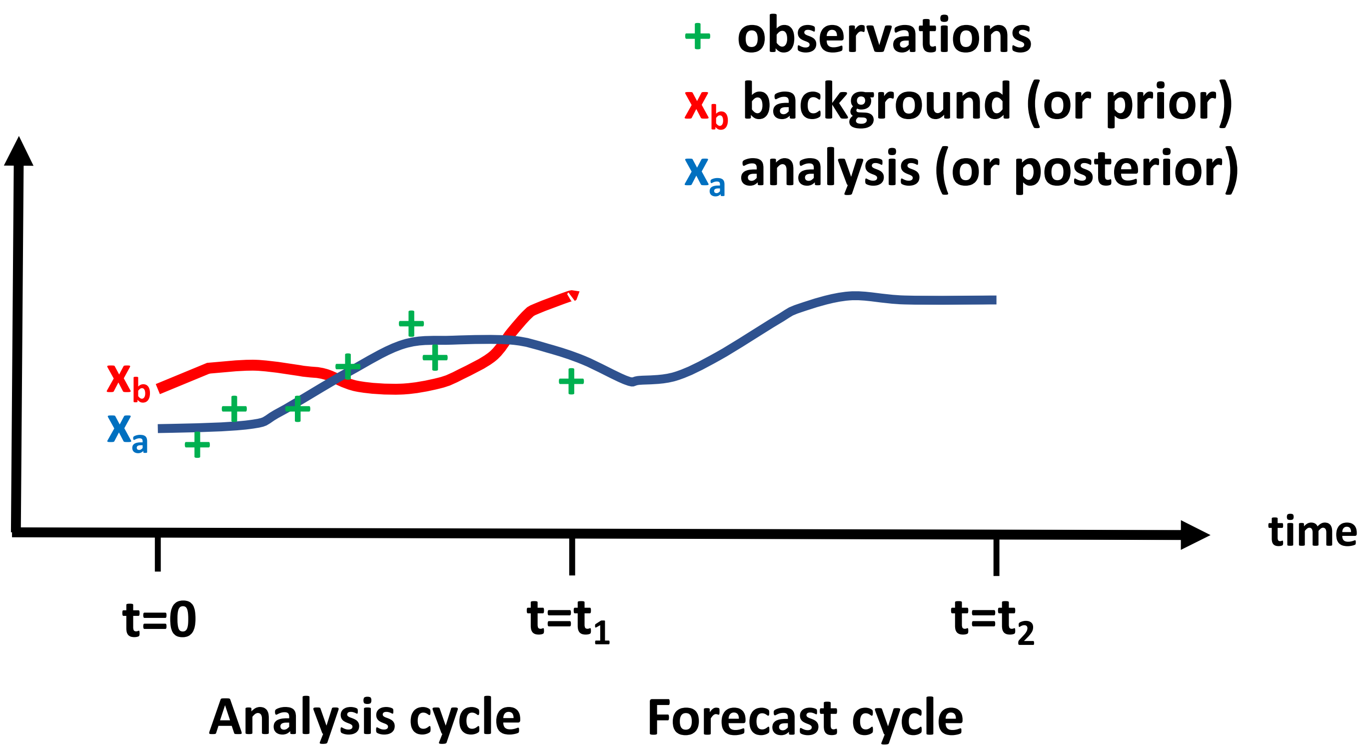
The cost function (a combination of the prior and data distributions):

$$J = \left(\mathbf{z}-\underset{\substack{\uparrow\\ \text{Prior}}}{\mathbf{z}_\mathrm{b}}\right)^T \underset{\substack{\uparrow\\ \text{Prior}\\ \text{error}\\ \text{cov.}}}{\mathbf{B}^{-1}} \left(\mathbf{z}-\mathbf{z}_\mathrm{b}\right) + \left(\underset{\substack{\uparrow\\ \text{Obs}}}{\mathbf{y}} - \underset{\substack{\uparrow\\ \text{Obs}\\ \text{operator}}}{G}(\mathbf{z})\right)^T \underset{\substack{\uparrow\\ \text{Obs}\\ \text{error}\\ \text{cov.}}}{\mathbf{R}^{-1}} \left(\mathbf{y}-G(\mathbf{z})\right)$$

Sea Surface Temperature, Jan. 2010



4-dimensional variational (4D-Var) data assimilation



Notation & Nomenclature

$$\mathbf{x} = \begin{bmatrix} \mathbf{T} \\ \mathbf{S} \\ \mathbf{u} \\ \mathbf{v} \\ \zeta \end{bmatrix}$$

State vector

$$\mathbf{z} = \begin{bmatrix} \mathbf{x}(0) \\ \mathbf{f}(t) \\ \mathbf{b}(t) \\ \boldsymbol{\eta}(t) \end{bmatrix}$$

Control vector

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ \vdots \\ y_N \end{bmatrix}$$

Observation vector

Prior

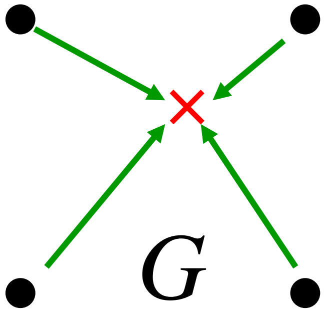
$$\mathbf{d} = \left(\mathbf{y} - G(\mathbf{x}^{\mathrm{b}})\right)$$

Innovation vector

η(t) = Correction for model error

η(t)=0 : Strong constraint

η(t)≠0 : Weak constraint



Observation operator

In 4D-Var, *G* includes NL ROMS

Array Modes: Assessing the Efficacy of the Observing System

- We have explored how the observations impact different aspects of the 4D-var circulation estimates and ensuing forecasts.
- However, we have not yet established how effective the observing system is at "observing" the circulation given our prior hypotheses about the system.
- Recall the analysis equation:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{B}\underbrace{\mathbf{G}^{\mathrm{T}}(\mathbf{G}\mathbf{B}\mathbf{G}^{\mathrm{T}} + \mathbf{R})^{-1}(\mathbf{y} - H(\mathbf{x}_b))}_{\mathbf{w}}$$

$$= \mathbf{x}_b + \mathbf{B}\mathbf{w}$$

- So the increment $\mathbf{x}_a - \mathbf{x}_b$ lies *entirely* in the space spanned by **B**.

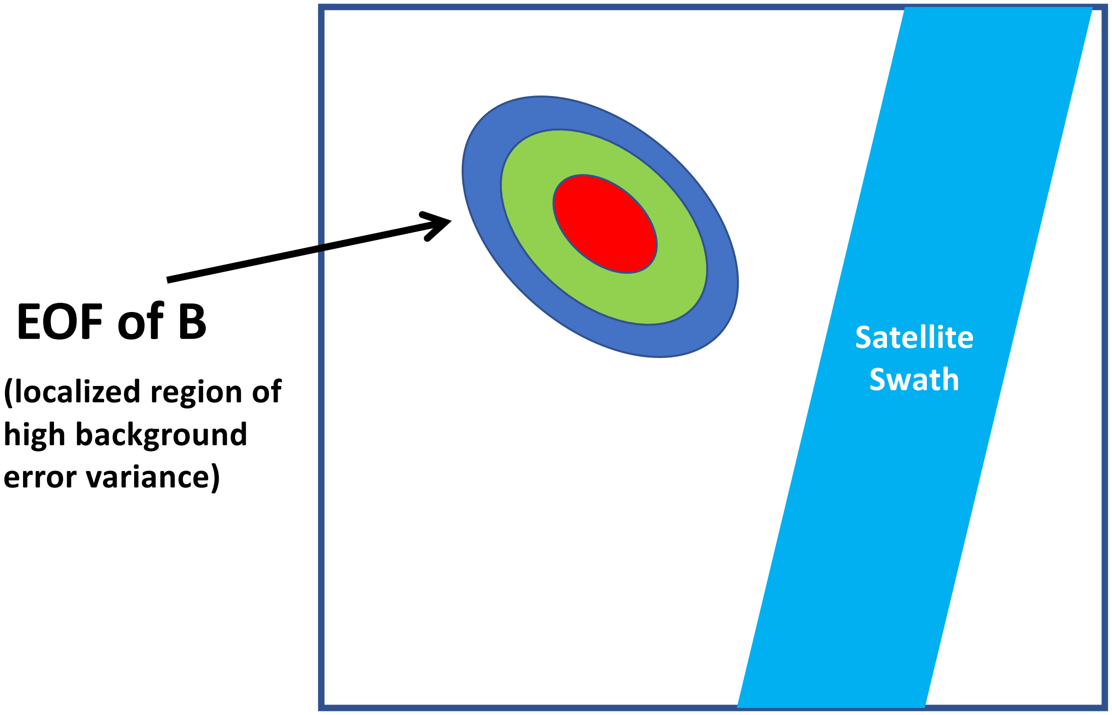
The Importance of the Background Error Covariance Matrix

$$\underset{\substack{\uparrow\\ \text{analysis}}}{\mathbf{x}_{\mathrm{a}}} = \underset{\substack{\uparrow\\ \text{background}}}{\mathbf{x}_{\mathrm{b}}} + \underbrace{\boxed{\mathbf{B}\underset{\substack{\uparrow\\ \text{adjoint}\\ \text{obs operator}}}{\mathbf{G}}{}^{\mathrm{T}}\left(\mathbf{G}\underset{\substack{\uparrow\\ \text{background}\\ \text{error cov}\\ \text{matrix}}}{\mathbf{B}}\mathbf{G}^{\mathrm{T}}+\underset{\substack{\uparrow\\ \text{obs error}\\ \text{cov matrix}}}{\mathbf{R}}\right)^{-1}\left(\underset{\substack{\uparrow\\ \text{obs}}}{\mathbf{y}}-\underset{\substack{\uparrow\\ \text{obs operator}}}{H}(\mathbf{x}_{\mathrm{b}})\right)}}_{\text{Analysis increment}}$$

The analysis increment “lives” in the space spanned by B !!!

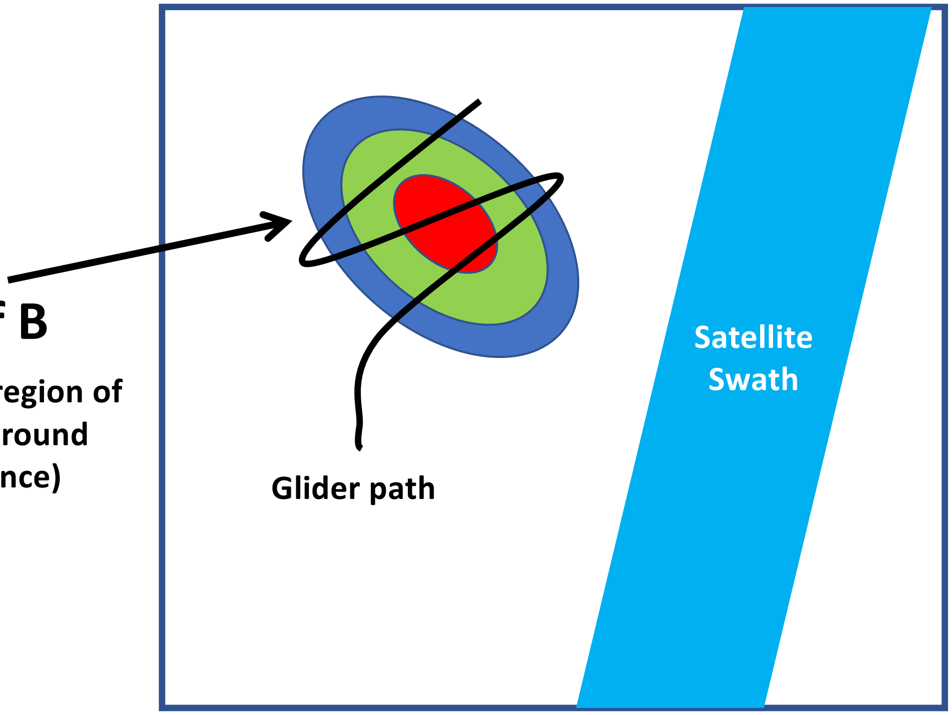
Therefore, to reduce errors in x_b, the observing system must effectively observe (directly via G or indirectly via G^T) the dominant EOFs of B.

An Illustrative Example



The satellite swath does not directly (G) or indirectly (G^T) observe the region of elevated background error variance associated with the EOF of B, so errors in this regions will not be corrected during data assimilation by the satellite observations.

An Illustrative Example



EOF of B

(localized region of high background error variance)

The glider path does directly observe the region of high error background error variance associated with the EOF of B, so errors in this regions will be corrected during data assimilation by the glider observations.

Eigenvectors

We will be concerned with two different sets of eigenvectors:

1. The EOFs of **B:** $\mathbf{B} = \mathbf{E}\Lambda\mathbf{E}^{\mathrm{T}}$ (**More specifically the EOFs of** $\mathbf{C} = \Phi\Pi\Phi^{\mathrm{T}}$ **where** $\mathbf{B} = \Sigma\mathbf{C}\Sigma^{\mathrm{T}}$ **)**
   These tell us about the space in which the increments live.
2. The eigenvectors of the inverse stabilized representer matrix:

$$(\mathbf{G}\mathbf{B}\mathbf{G}^{\mathrm{T}} + \mathbf{R})^{-1}$$

If this is poorly conditioned, then the increment will be dominated by the eigenvectors of (**GBG$^{\mathrm{T}}$+R**) with the *smallest* eigenvalues.

In some sense, it is the juxtaposition of these two sets of eigenvectors that determines the efficacy of the observing system.

Array Modes

Recall that the analysis equation is solved using the Lanczos vectors:

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

This can be rewritten as:

$$\mathbf{x}_a = \mathbf{x}_b + \sum_{i=1}^{m} \alpha_i \Psi_i$$ where $\Psi_i = \mathbf{B}\mathbf{G}^{\mathrm{T}}\mathbf{V}_{\mathrm{m}}\mathbf{u}_i$ are the “array modes” (Bennett, 1985)

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

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

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

Array Modes

- The array modes are a set of generally non-orthogonal basis functions that depend *only* on the obs locations.
- The contribution of each Ψ_i to the increment $\mathbf{x}_a$-$\mathbf{x}_b$ (i.e. the amplitude α_i) depends on the obs values.
- Each Ψ_i is associated with an eigenpair (λ_i,$\mathbf{u}_i$).
- The number of arrays modes equals the number of inner-loops
- Bennett (1985) refers to the array modes as “interpolation patterns.”
- The amplitude α_i depends on $(\lambda_i)^{-1}$, so Ψ_1 represents the most “stable” interpolation pattern wrt changes in the obs values.
- Ψ_m is the least stable, and may represent a significant source of unphysical noise.

(Reduced Rank) Array Modes (RAMs)

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

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

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

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

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

$$\mathbf{x}_a = \mathbf{x}_b + \sum_{i=1}^{m} \alpha_i \Psi_i \quad \text{where} \quad \Psi_i = \mathbf{B}\mathbf{G}^{\mathrm{T}}\boxed{\mathbf{V}_{\mathrm{m}}\mathbf{u}_i}$$

are the *"RAMs"*

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

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

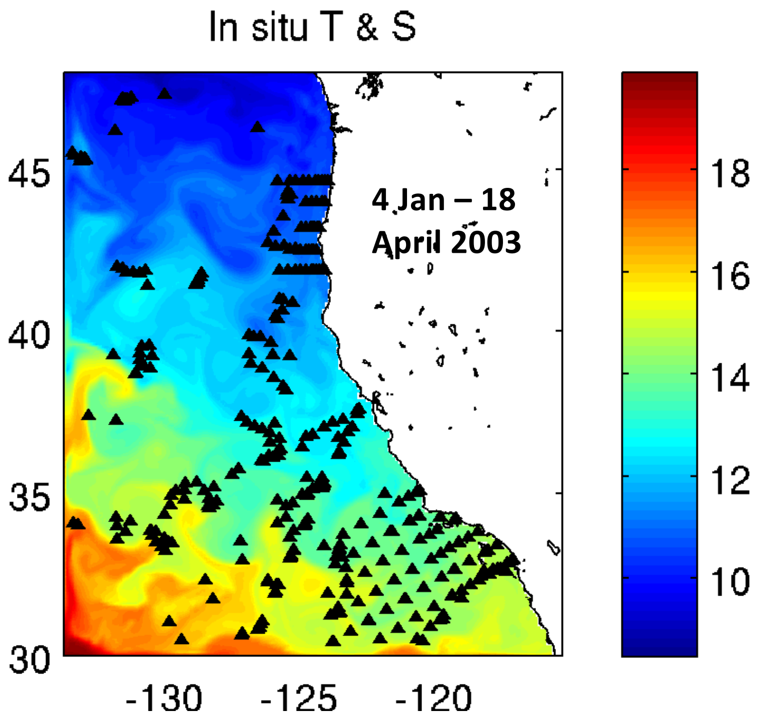
$(\lambda_i, \mathbf{u}_i)$ are the eigenpairs of $\mathbf{T}_{\mathrm{m}}$

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

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

- **How well does the California Current observing system “observe” the circulation given our prior hypotheses about the errors?**

ROMS CCS domain



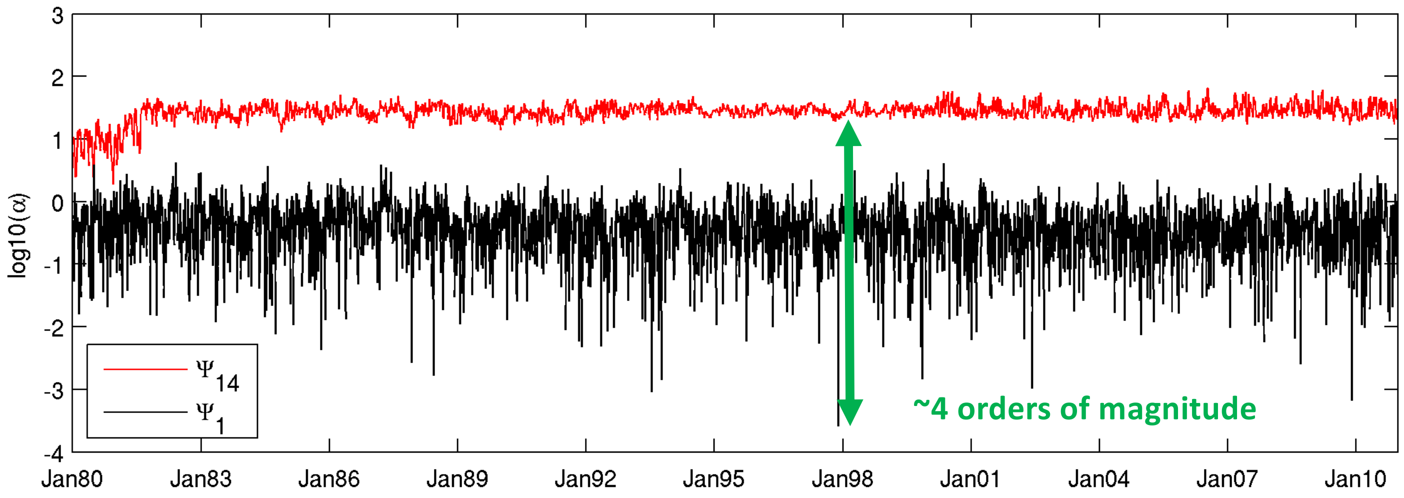
In situ T & S

4 Jan – 18 April 2003

CCS observing system:

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

RAM amplitudes of the California Current system

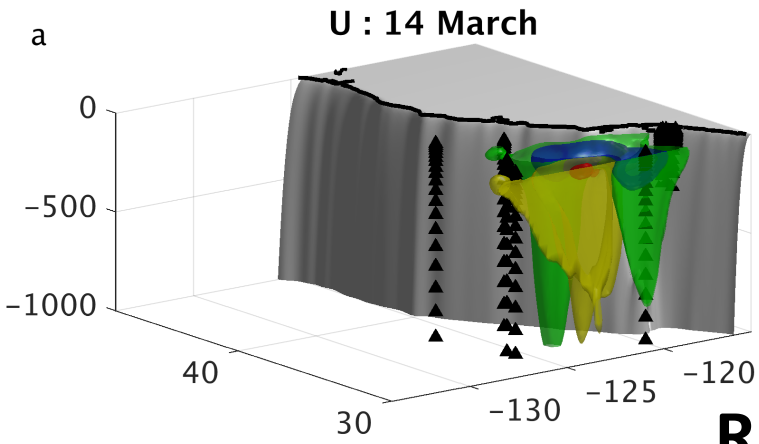


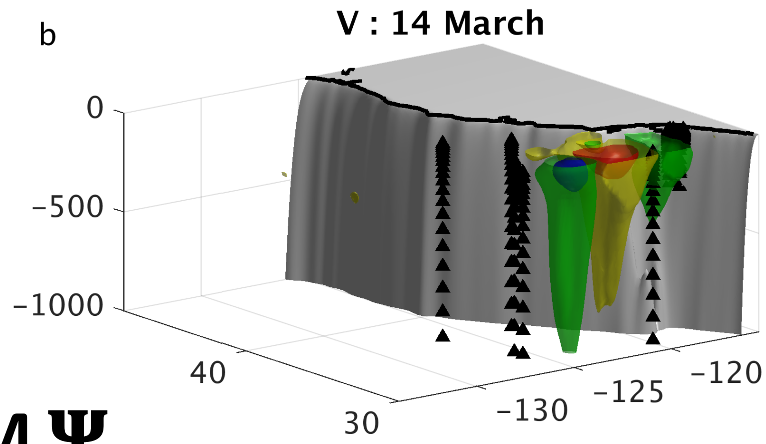
~4 orders of magnitude

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

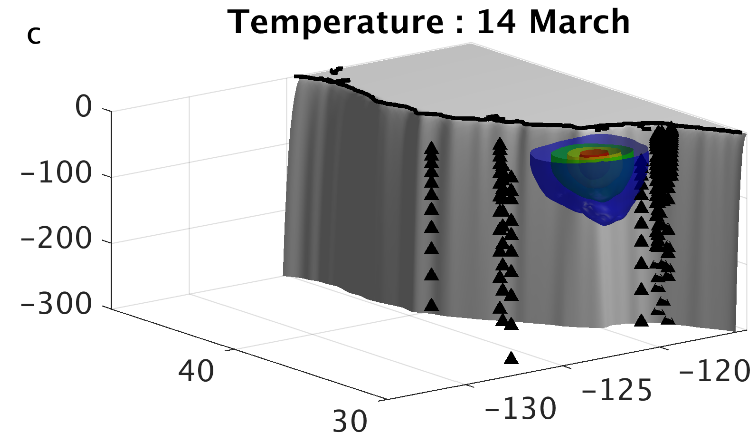
14 March, 2001

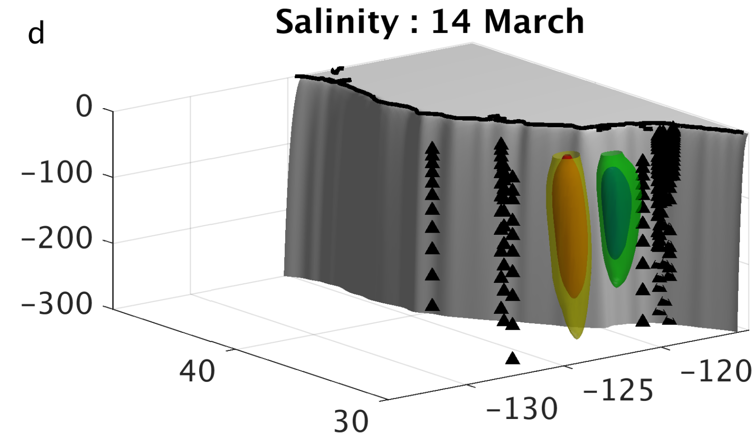
The sub-space activated by the obs & d.f.s





RAM Ψ_1



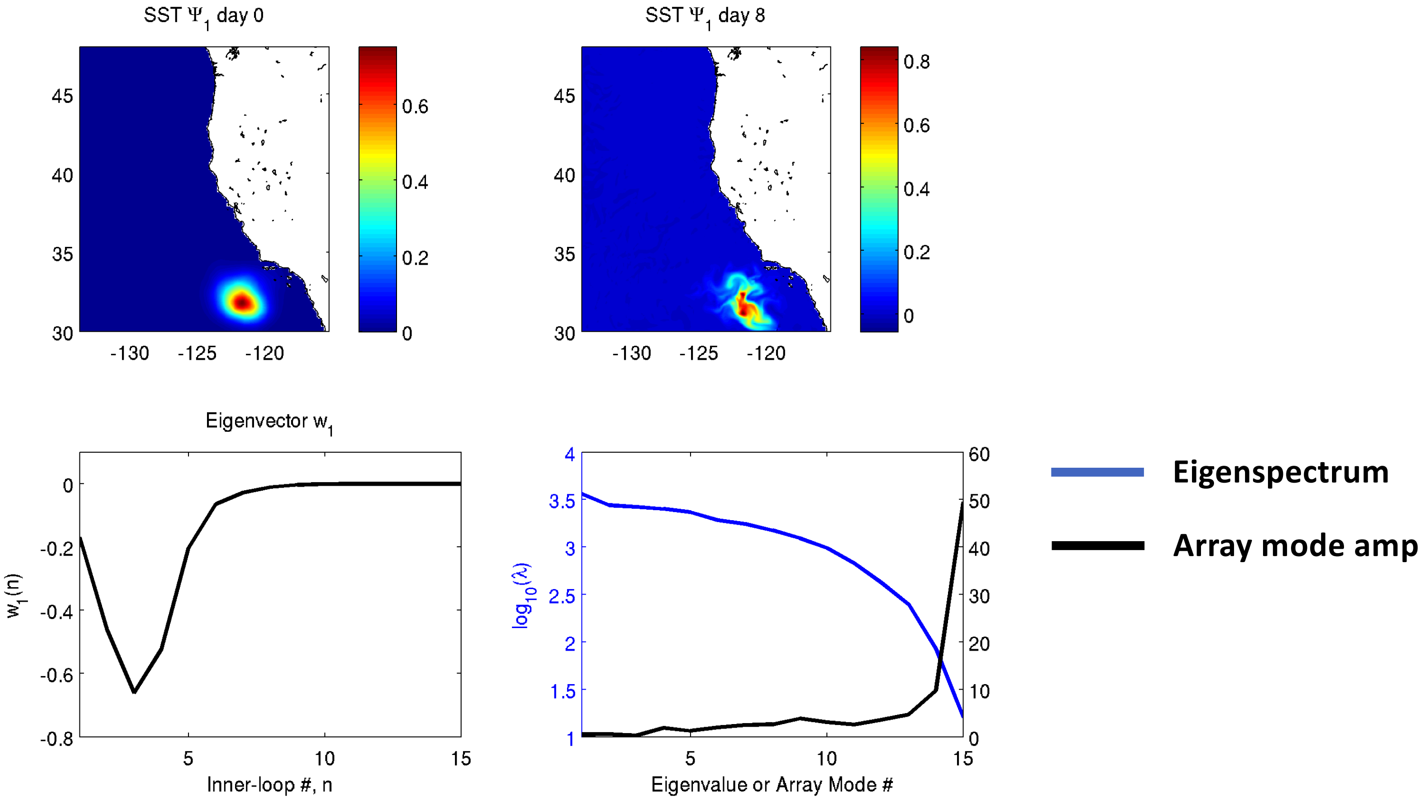


RAM #1

Time evolution:

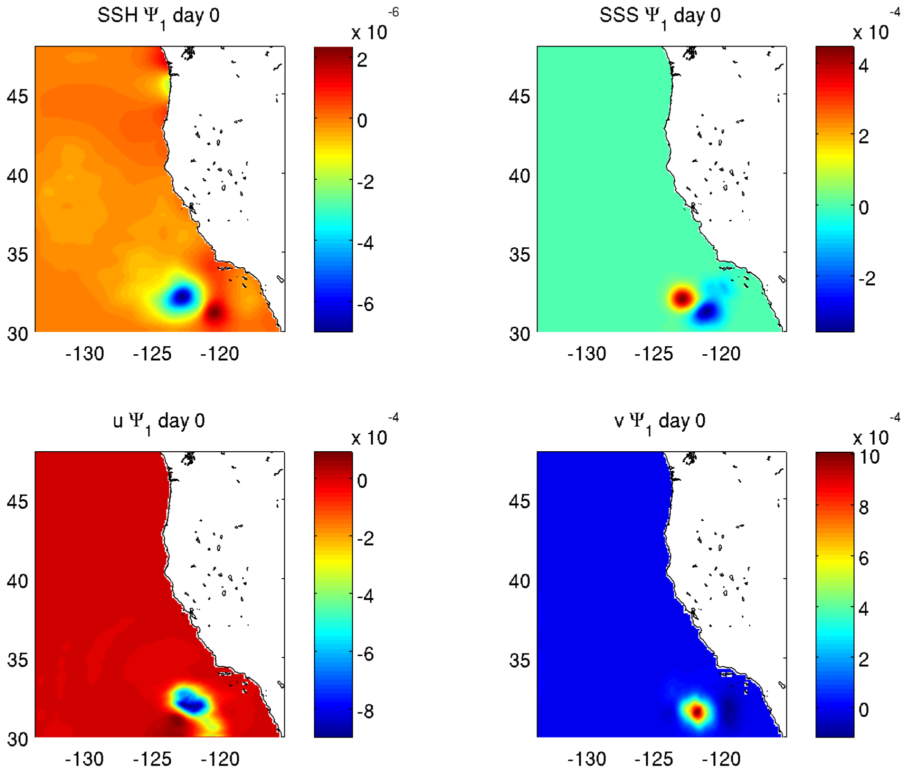
$$\Psi_i(t) = \mathbf{M}\Psi_i$$

15-23 March, 2001

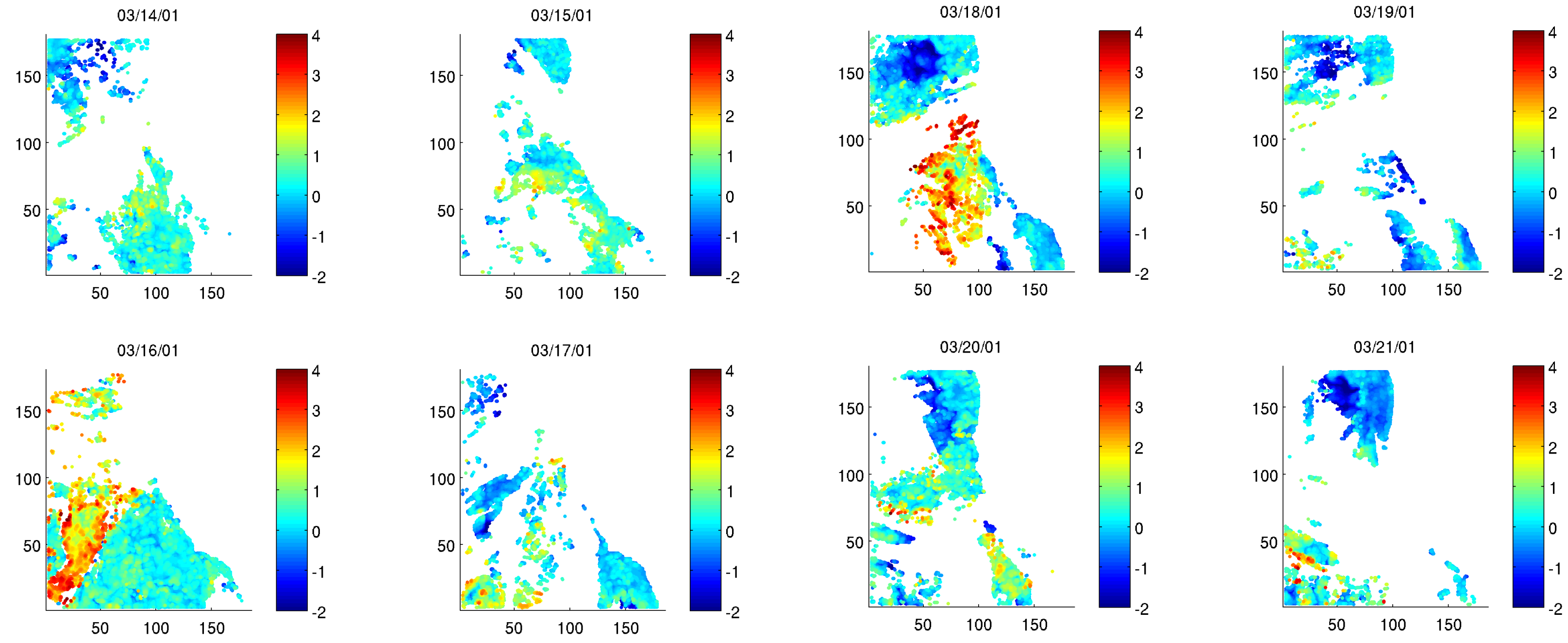


RAM #1

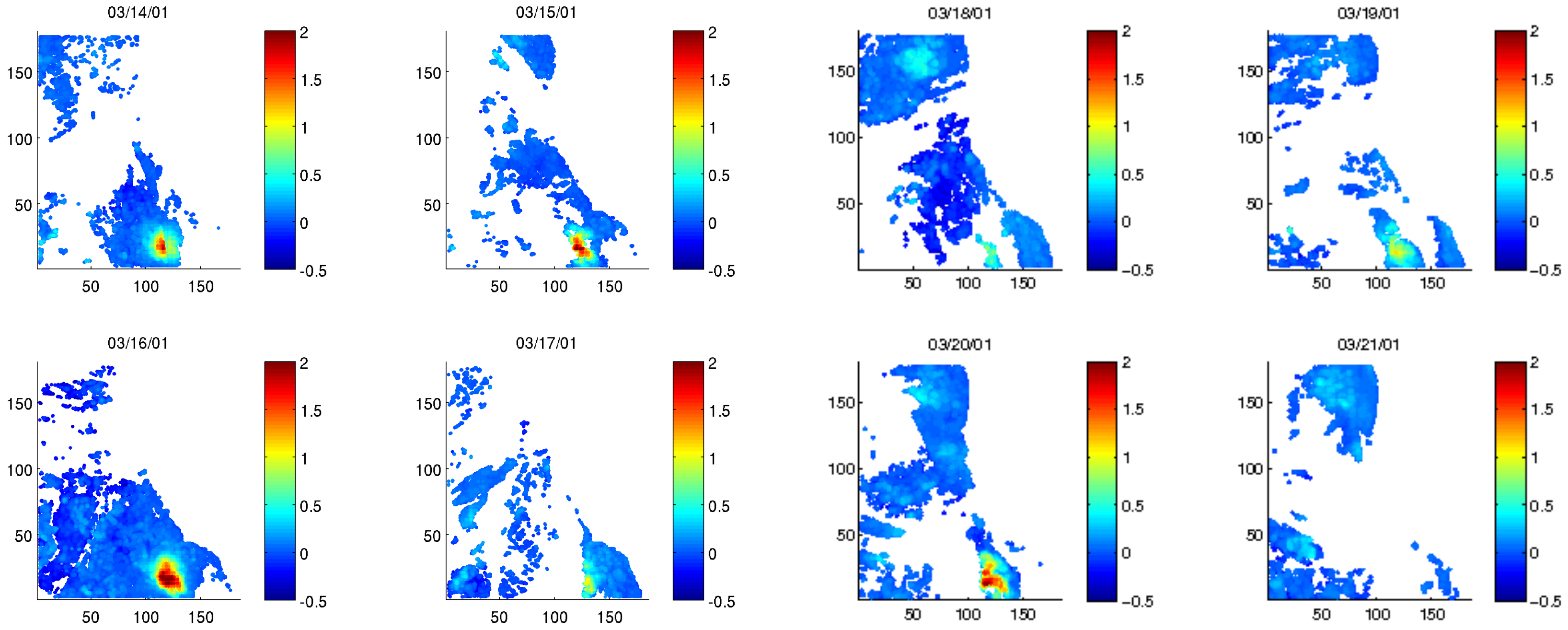
15-23 March, 2001



SST observations: 15-23 March, 2001



$$\Psi_1 = \mathbf{B}\mathbf{G}^{T}\mathbf{V}_m\mathbf{u}_1$$

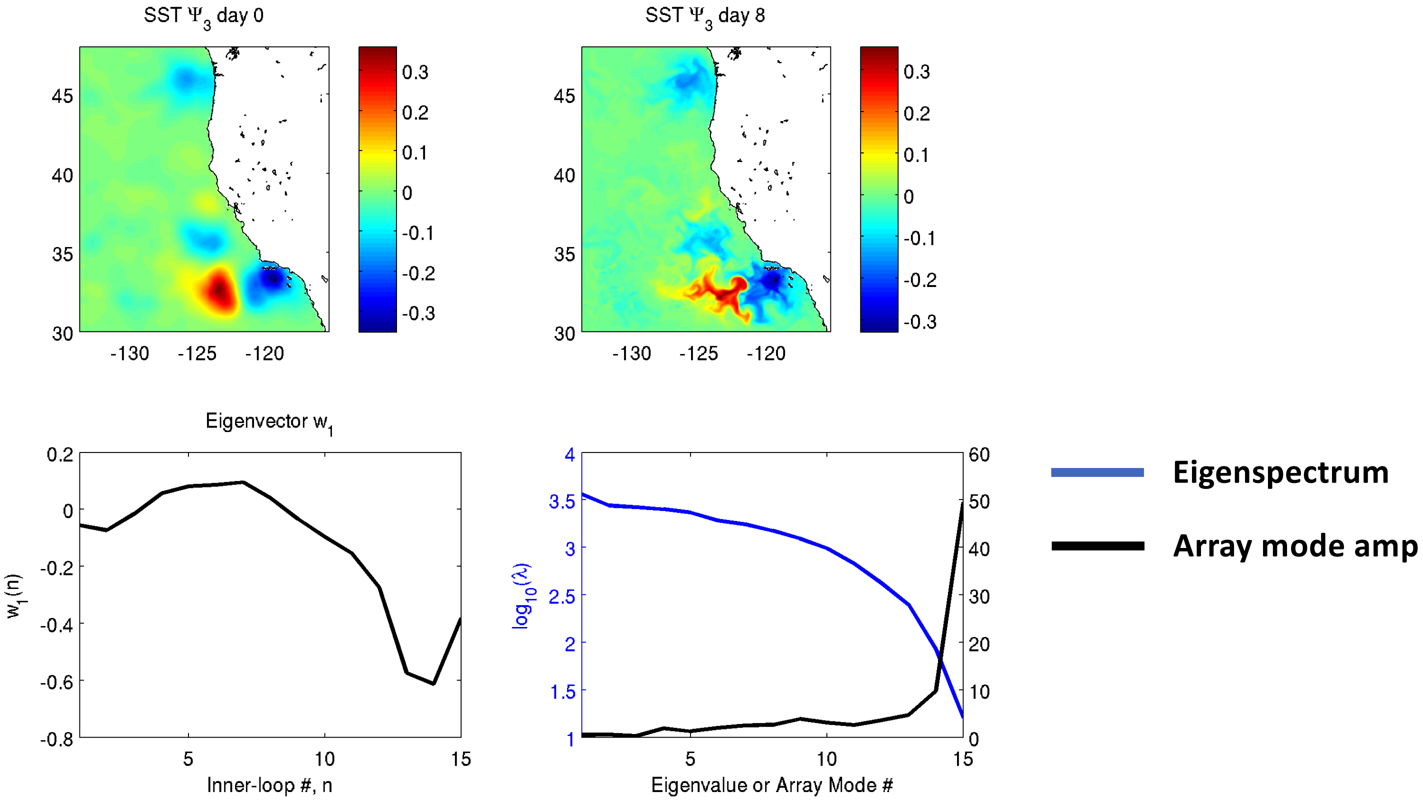


G^T maps these fields into state space

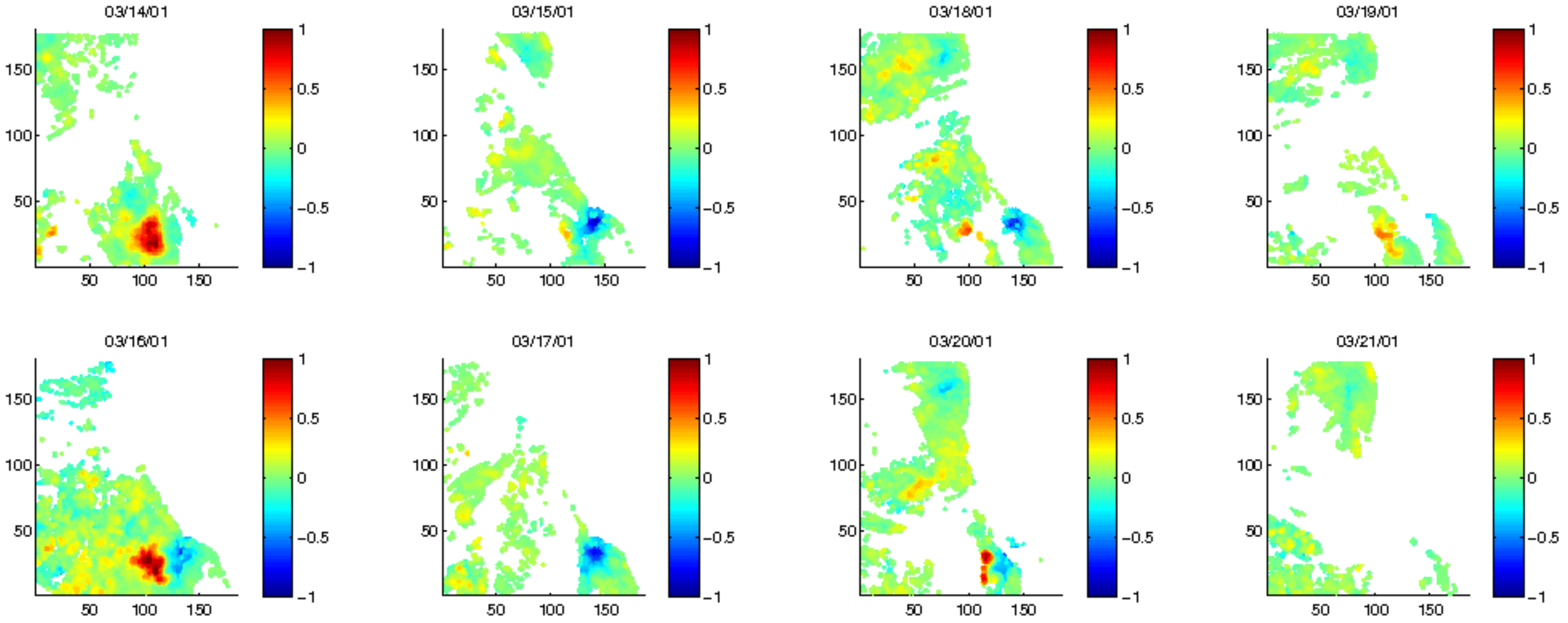
RAM #3

$$\Psi_i(t) = \mathbf{M}\Psi_i$$

15-23 March, 2001



$$\Psi_3 = \mathbf{B}\mathbf{G}^T\mathbf{V}_m\mathbf{u}_3$$



G^T maps these fields into state space

Array Modes

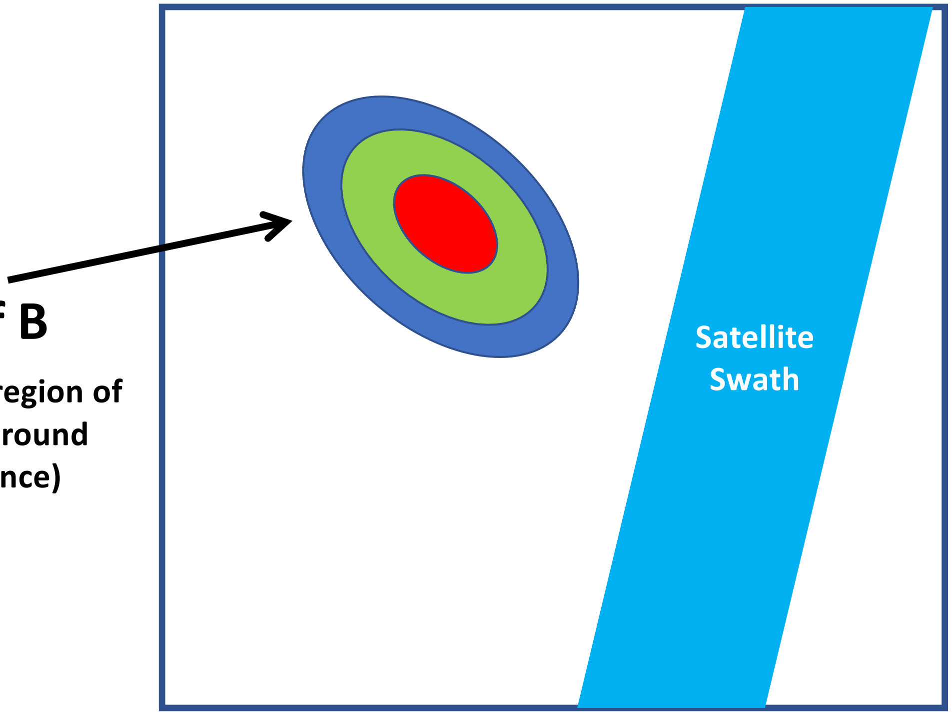
Recall the definition of an array mode: $\Psi_i = \mathbf{B}\mathbf{G}^{\mathrm{T}}\mathbf{V}_{\mathrm{m}}\mathbf{u}_i$

B can be expressed in terms of its EOFS: $\mathbf{B} = \mathbf{E}\boldsymbol{\Lambda}\mathbf{E}^{\mathrm{T}}$

So the array modes are linear combinations of the EOFs of **B**

In which case, if $\mathbf{G}^{\mathrm{T}}\mathbf{V}_{\mathrm{m}}\mathbf{u}_{\mathrm{i}}$ does not project onto a particular EOF of **B**, then that EOF will not be resolved by the array modes.

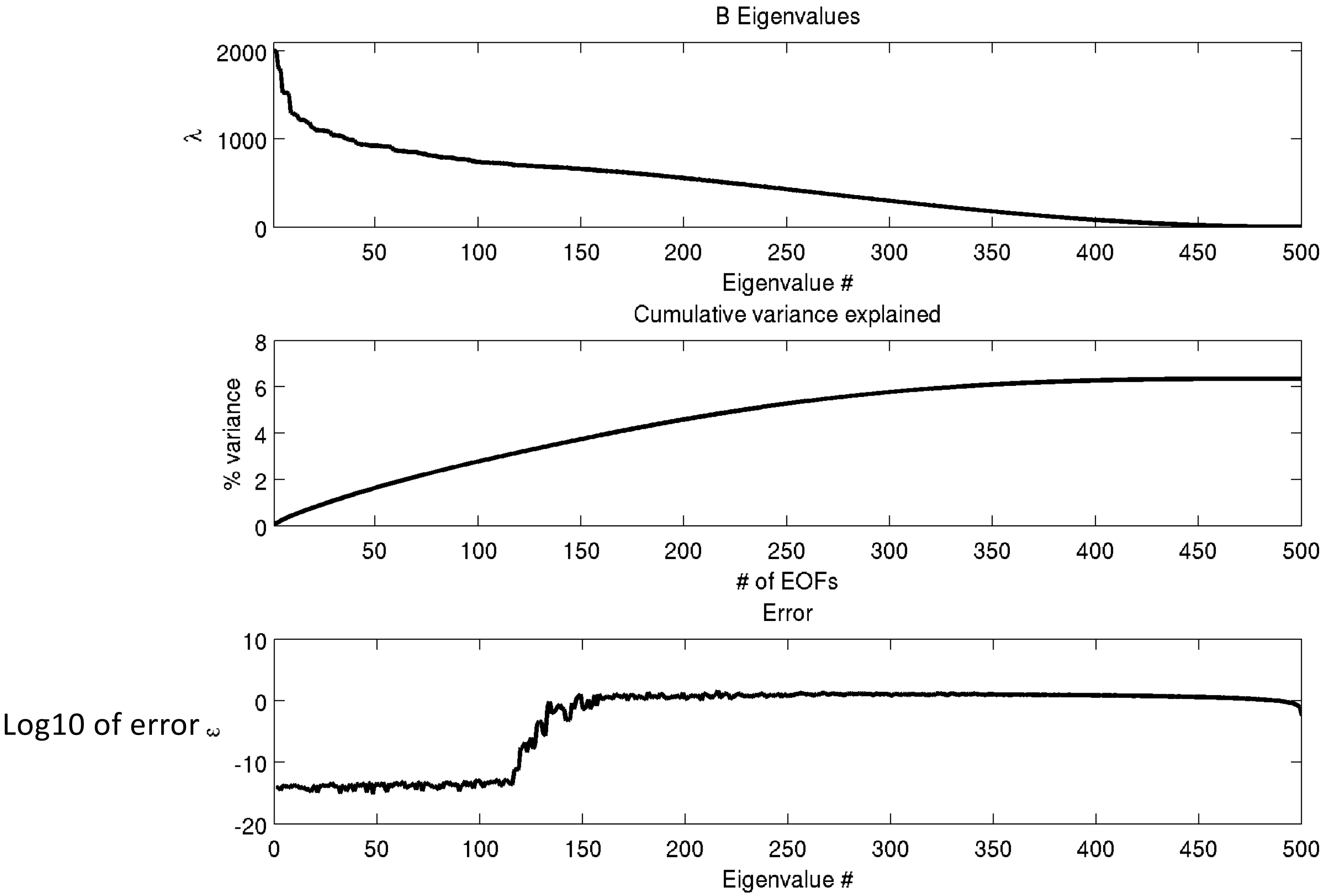
Recall the Illustrative Example



EOF of B

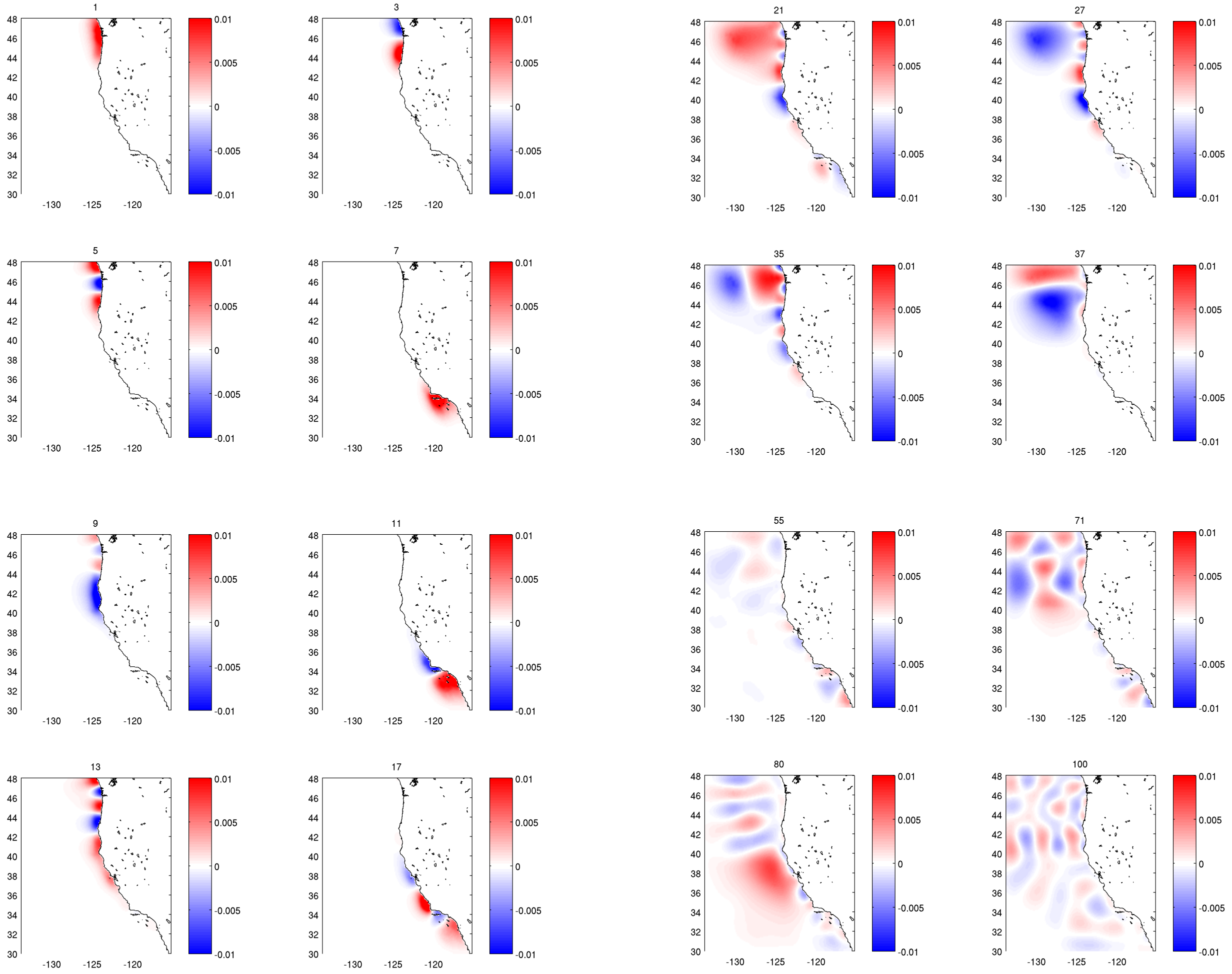
(localized region of
high background
error variance)

Do the array modes ”overlap” the EOFs?

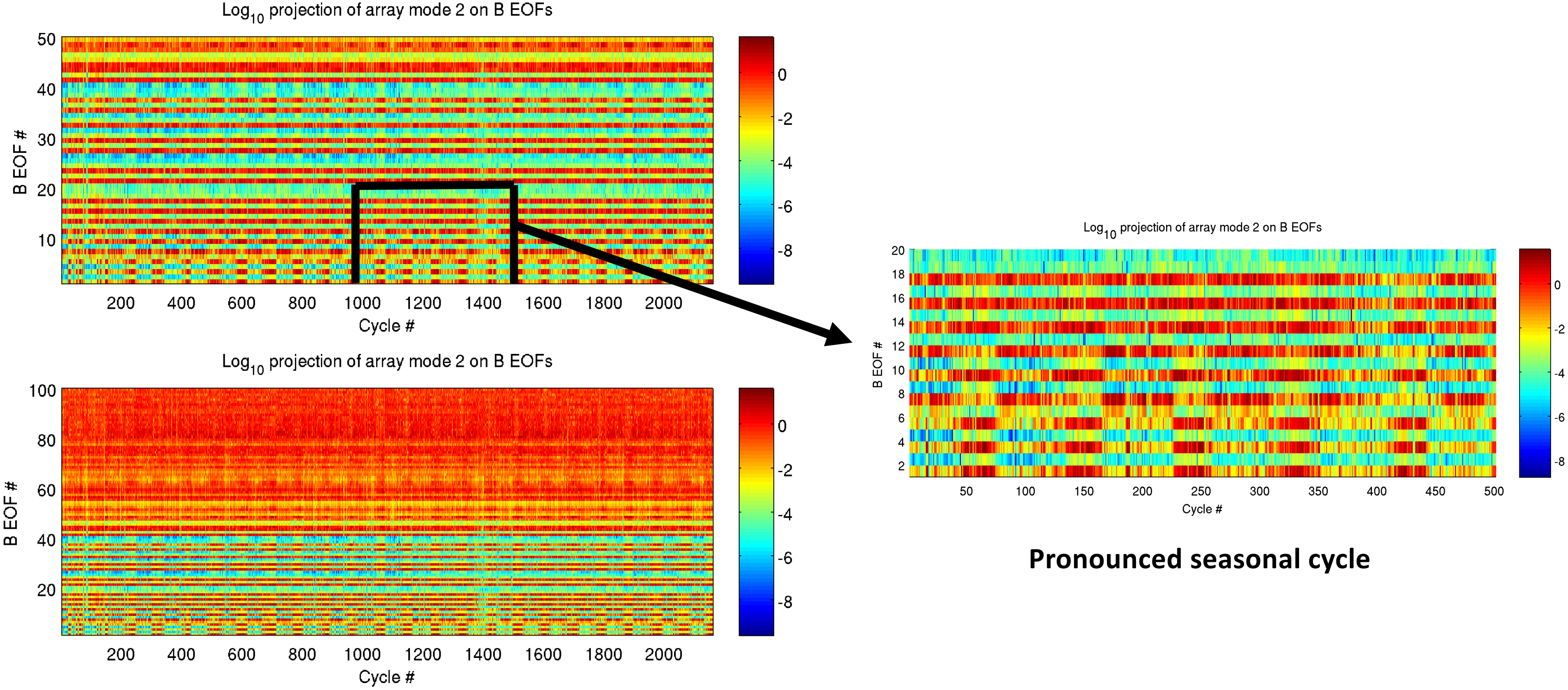

B Eigenvalues
λ
2000
1000
0
50
100
150
200
250
300
350
400
450
500
Eigenvalue #
Cumulative variance explained
% variance
8
6
4
2
0
50
100
150
200
250
300
350
400
450
500
# of EOFs
Error
Log10 of error ε
10
0
-10
-20
0
50
100
150
200
250
300
350
400
450
500
Eigenvalue #

Example EOFs of B

- **Flat spectrum**
- **Very small % variance explained by each**

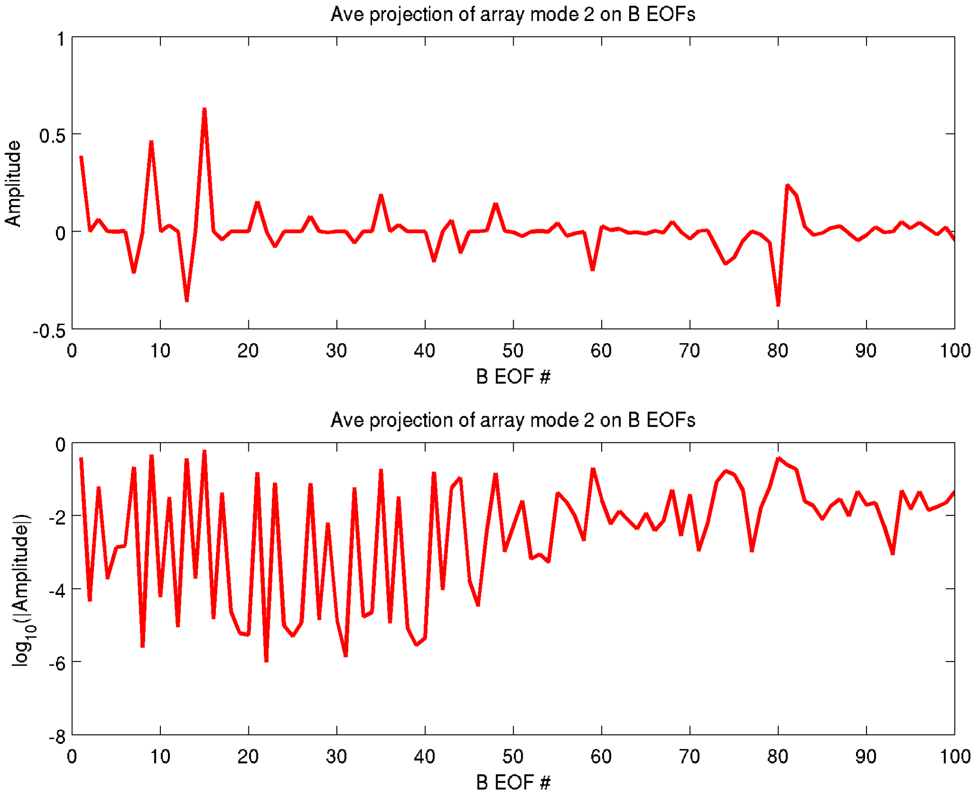


Projection of RAM 2 on B EOFs

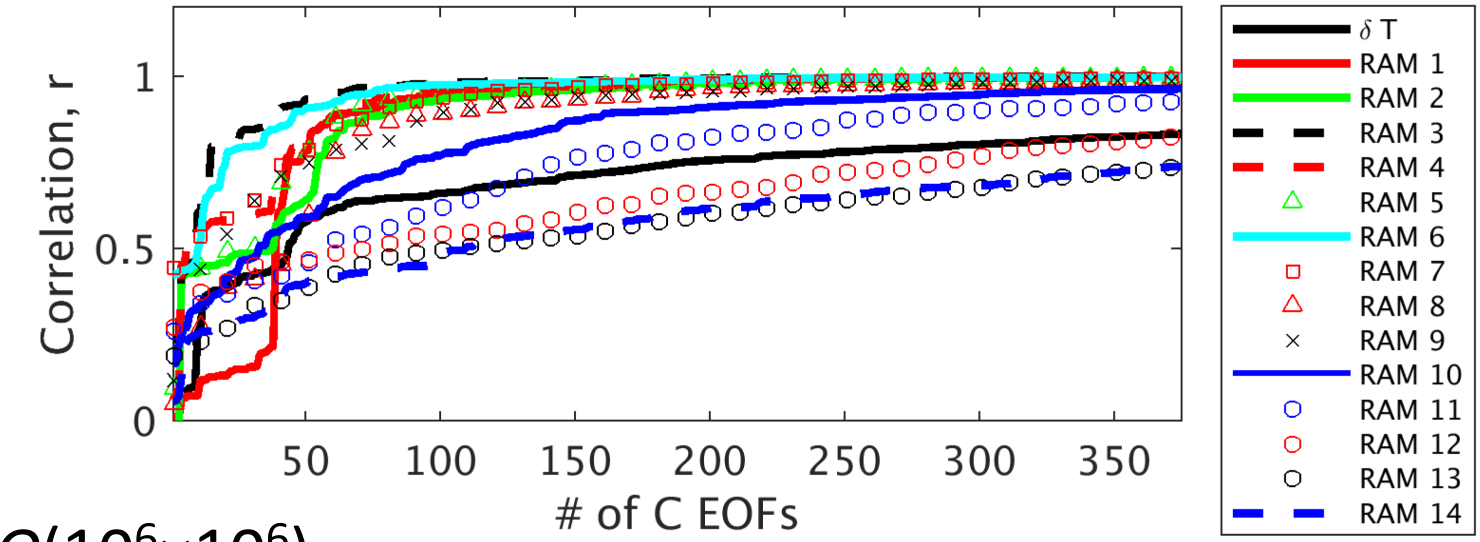


Pronounced seasonal cycle

Projection of RAM 2 on B EOFs



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 degrees of freedom span a small sub-space of **B**

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

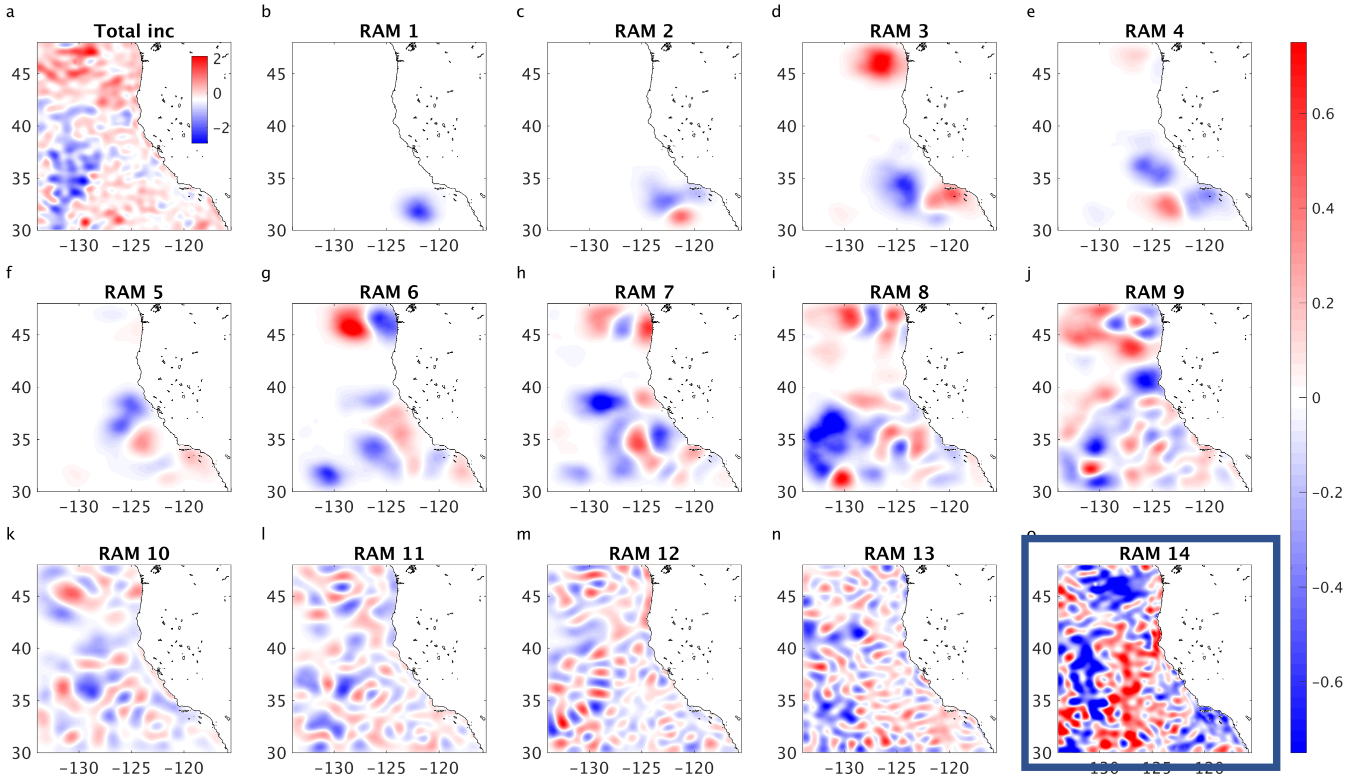
- **B** spectrum is very flat
- Each EOF explains a small % of background error variance
- Relatively few EOFs needed to recover array modes and increments (~0.002-0.02%)
- This suggests that the observations provide information about a very small part of the space occupied by **B**
- **Therefore most of the background error is unchanged by the obs**

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*.

(GBG^T+R)

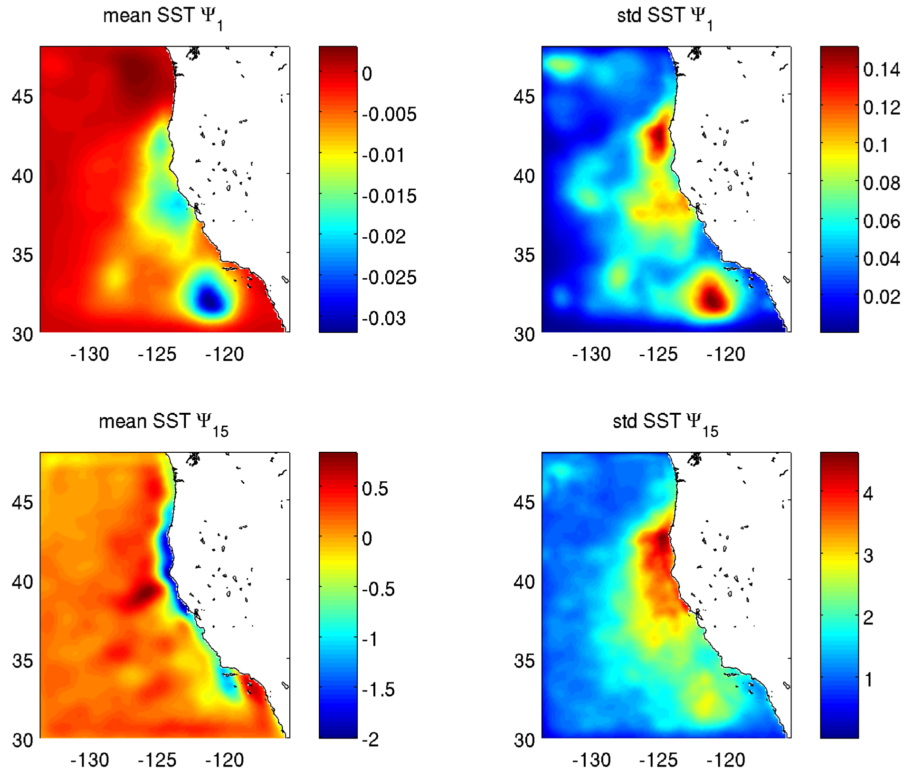
The Bennett and McIntosh (1984) “1% rule”: Discard array modes with eigenvalues 1% or less of the max value - more conservative.



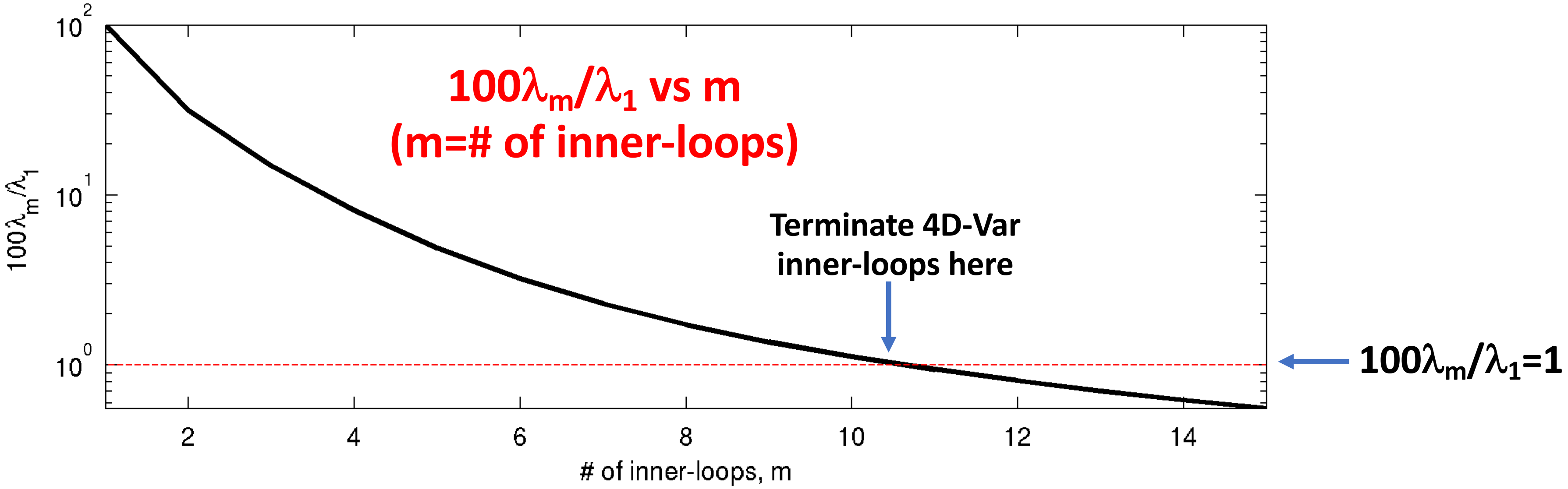
SST

Reject based on 1% rule

Mean and std of SST associated with array mode 1 and array mode 15 computed from ALL 4D-Var cycles, 1980-2010



The “1% Rule” to avoid overfitting to the Observations



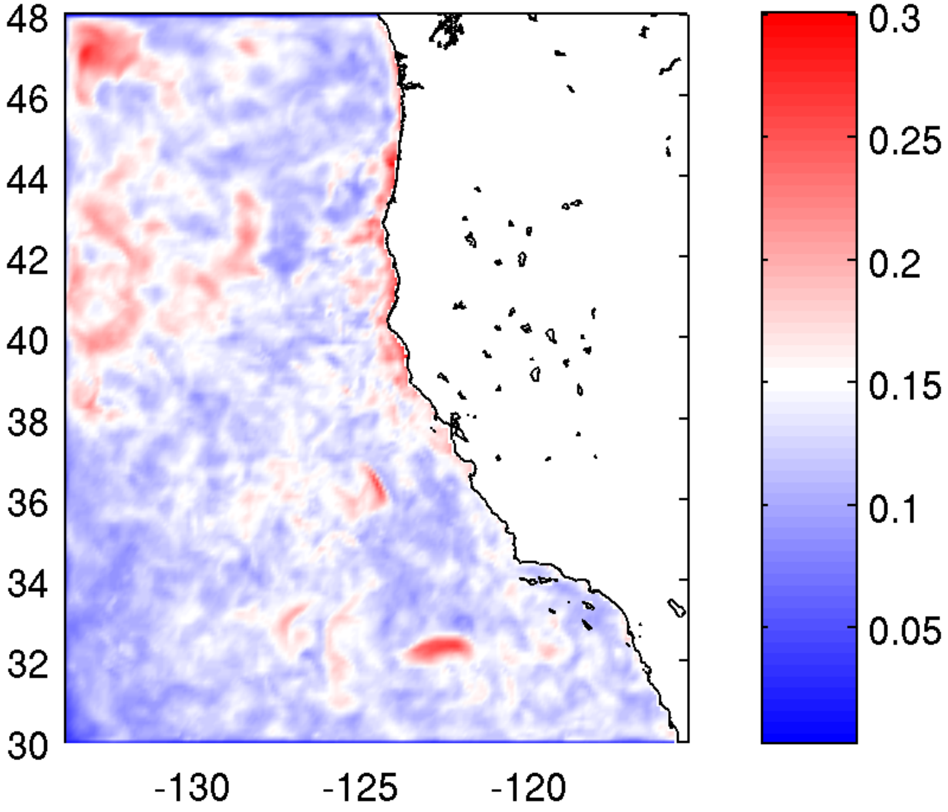
$100\lambda_m/\lambda_1$ vs m
(m=# of inner-loops)

Terminate 4D-Var
inner-loops here

$100\lambda_m/\lambda_1=1$

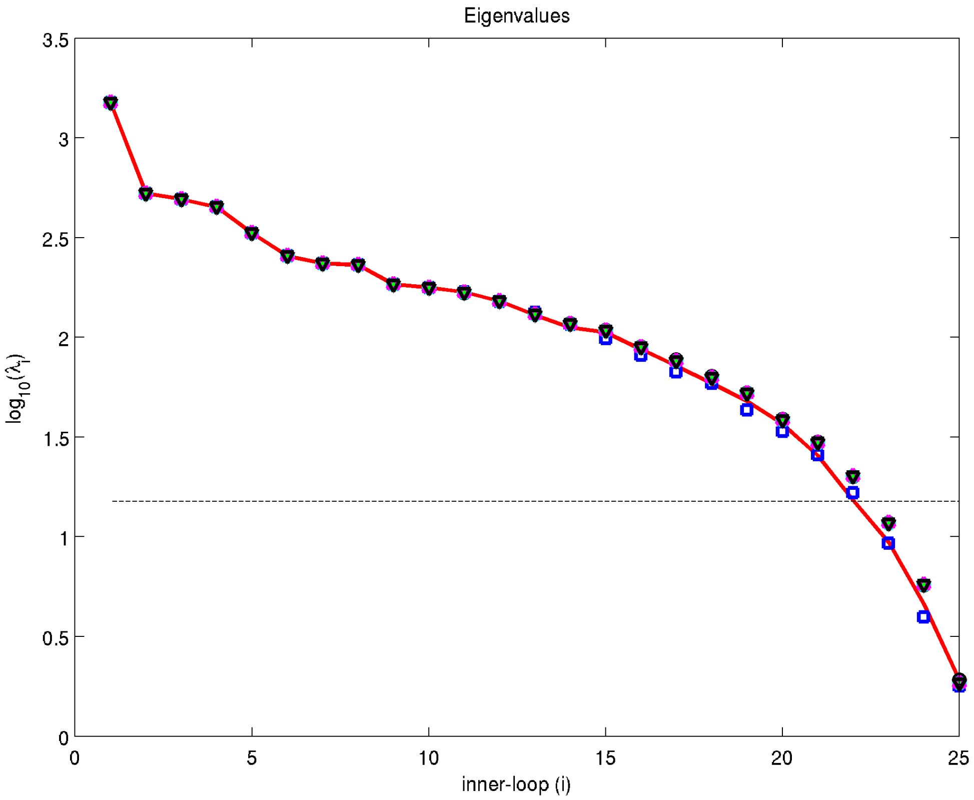
Average eigenvalue ratio for all cycles vs number of inner-loops employed: (suggests we should use no more than 10 inner-loops to prevent over-fitting of the observations)

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

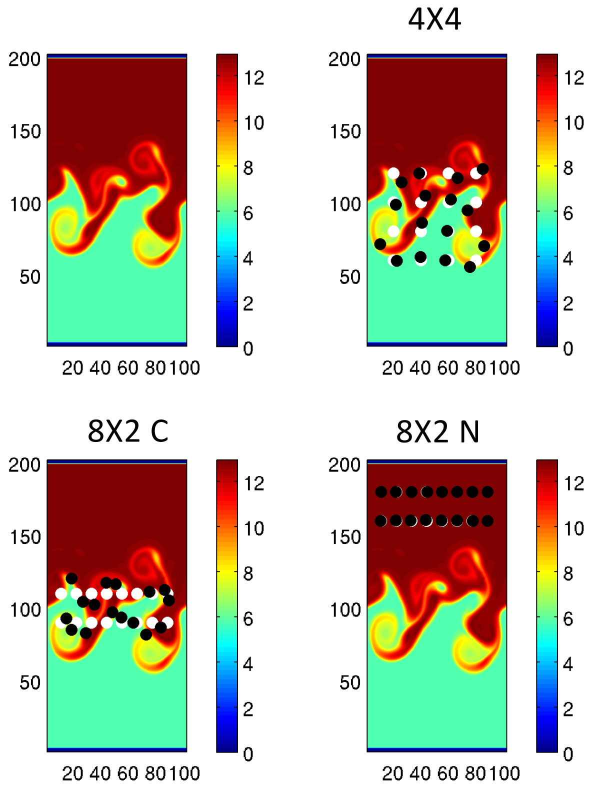


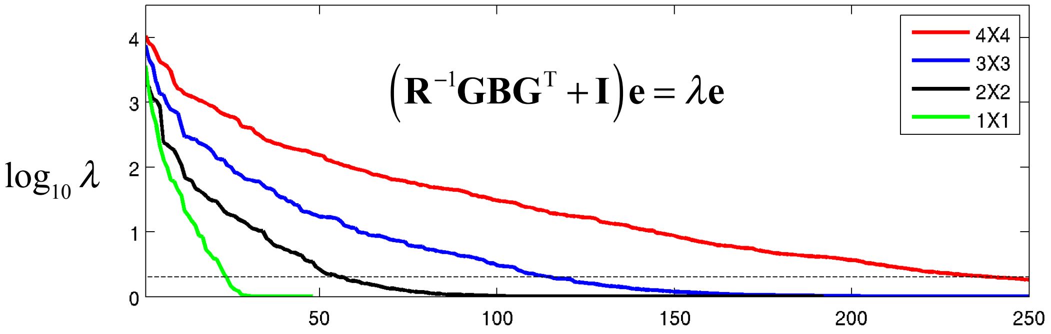
Eigenvalues of the preconditioned stabilized representer matrix

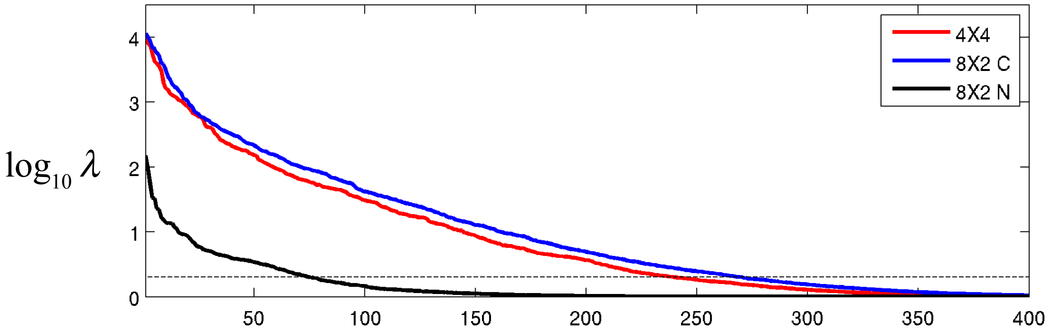
Computed from random innovations



Degrees of Freedom of the Observing Array







Array Modes

cpp options:

- ARRAY_MODES
- FORWARD_READ
- FORWARD_MIXING

Input files:

- FWDname – background circulation for ADROMS (ocean.in)
- Nvct - parameter to select required array mode (s4dvar.in)

Output files:

- TLMname - time evolution of the selected array mode (ocean.in)

Bibliography

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., 1985: Array design by inverse methods. *Prog. Oceanog*., **15**, 129-156.
Moore, A.M., H.G. Arango and C.A. Edwards, 2017: Reduced-Rank Array Modes of the California Current Observing System.
J. Geophys. Res. - Oceans, **122**, https://doi.org/10.1002/2017JC013172.