ROMS 4D-Var, Observation Impact and Observation Sensitivity

Andy Moore
Dept. of Ocean Sciences
University of California, Santa Cruz

Reverend Thomas Bayes
(1702-1761)

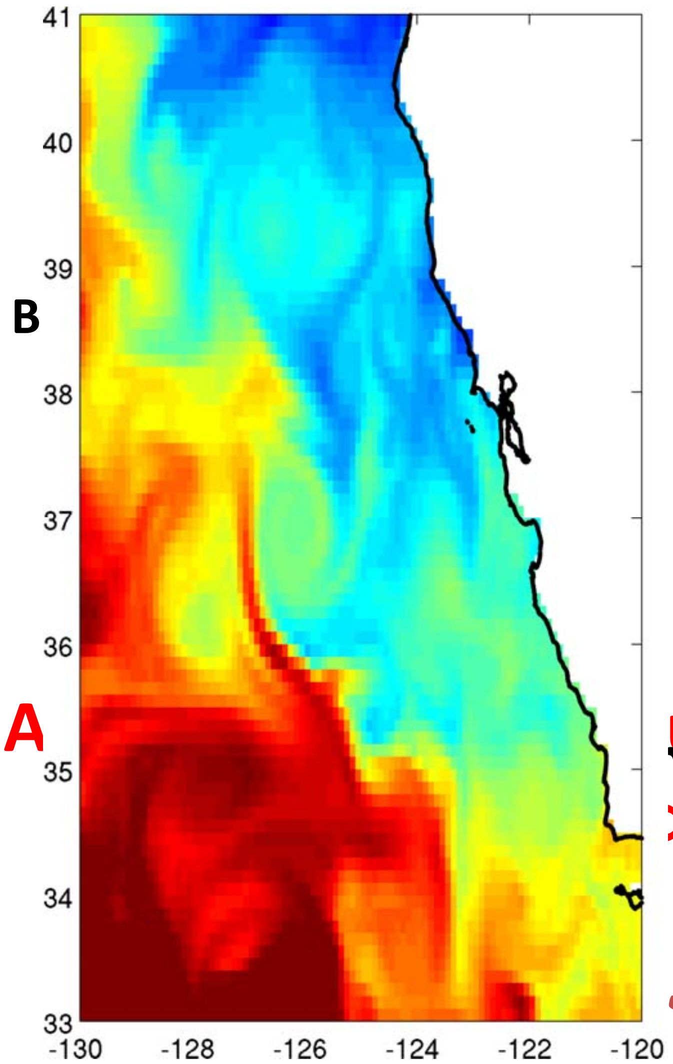
Data Assimilation

Model

Observations

$b_b(t)$, B

+



A

ut

n

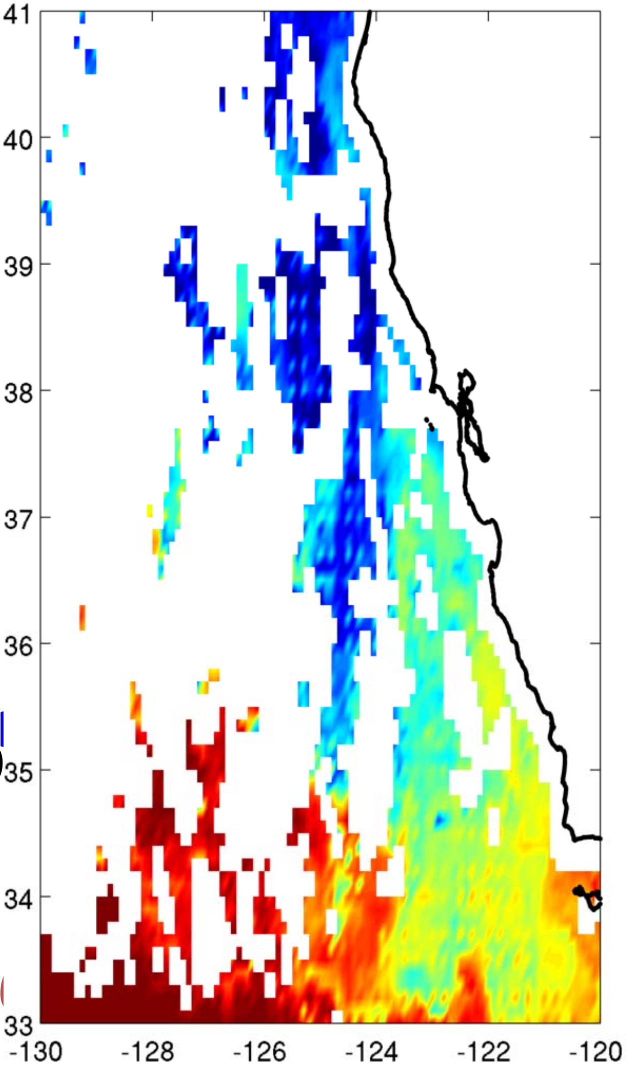
d

steri

D

I

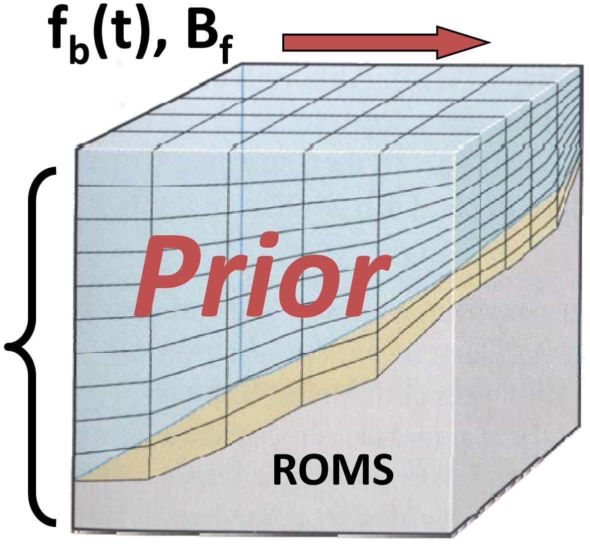
f



Data Assimilation

Model

$f_b(t)$, B_f

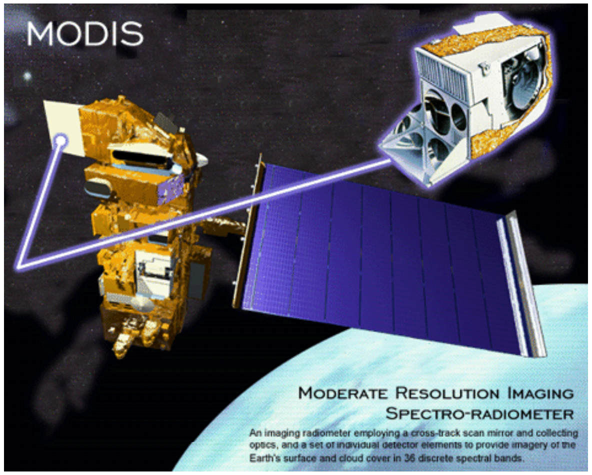


$b_b(t)$, B_b

$x_b(0)$, 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:**

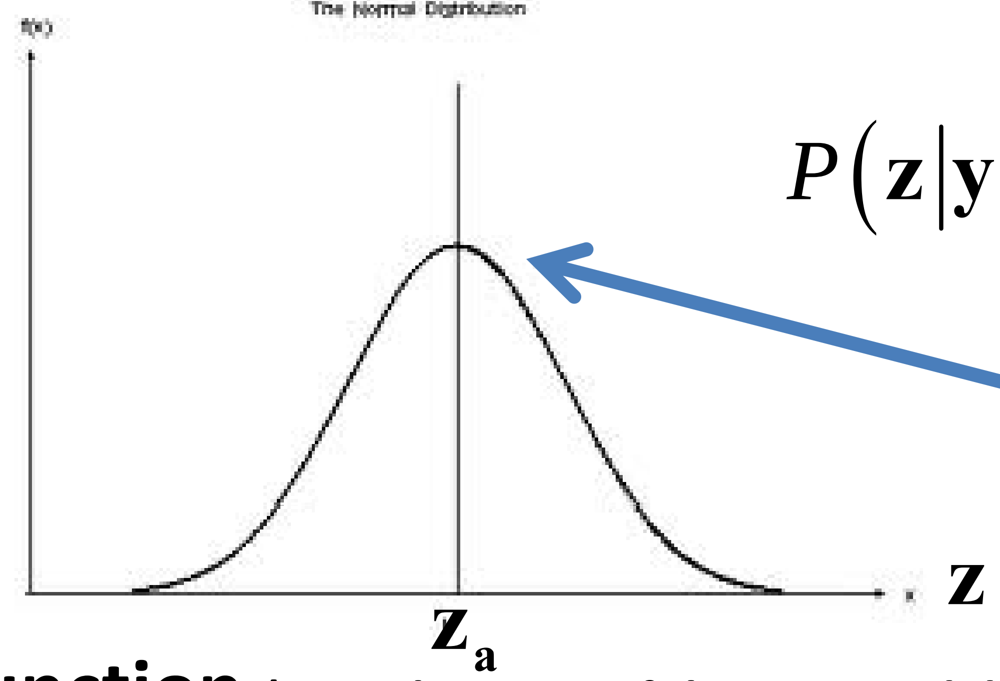
$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_x & & \\ & \mathbf{B}_f & \\ & & \mathbf{B}_b \end{pmatrix}$$

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*(**z**|**y**) by minimizing *J* using variational calculus

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

$$J = (\mathbf{z}-\mathbf{z}_\mathbf{b})^T \mathbf{B}^{-1} (\mathbf{z}-\mathbf{z}_\mathbf{b}) + (\mathbf{y}-G(\mathbf{z}))^T \mathbf{R}^{-1} (\mathbf{y}-G(\mathbf{z}))$$

Prior

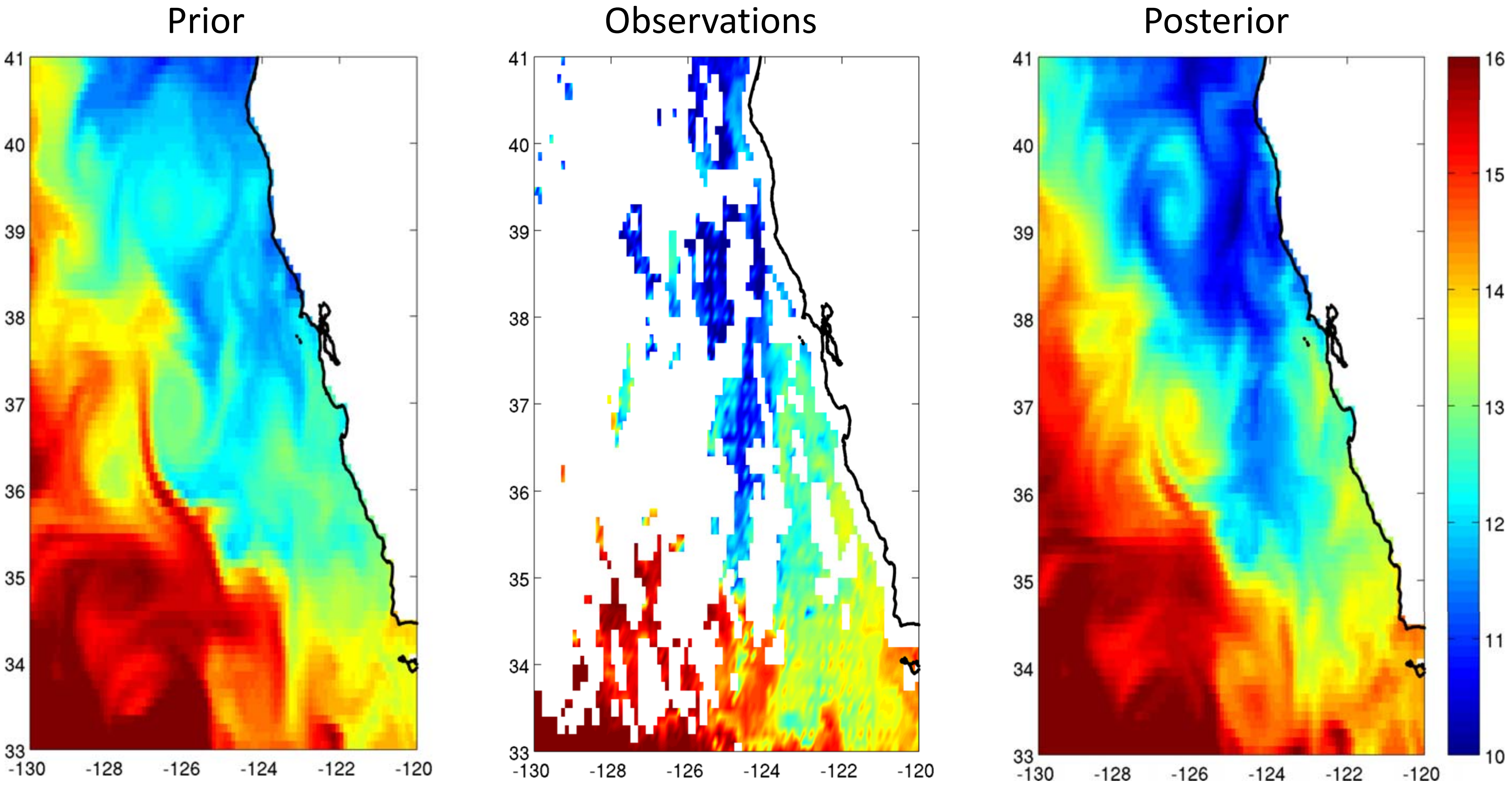
Prior error cov.

Obs

Obs operator

Obs error cov.

Sea Surface Temperature, Jan. 2010



Notation & Nomenclature

$$\mathbf{x} = \begin{bmatrix} \mathbf{T} \\ \mathbf{S} \\ \mathbf{u} \\ \mathbf{v} \\ \zeta \end{bmatrix} \qquad \mathbf{z} = \begin{bmatrix} \mathbf{x}(0) \\ \mathbf{f}(t) \\ \mathbf{b}(t) \\ \boldsymbol{\eta}(t) \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ \vdots \\ y_N \end{bmatrix}$$

State vector

Control vector

Observation vector

Prior

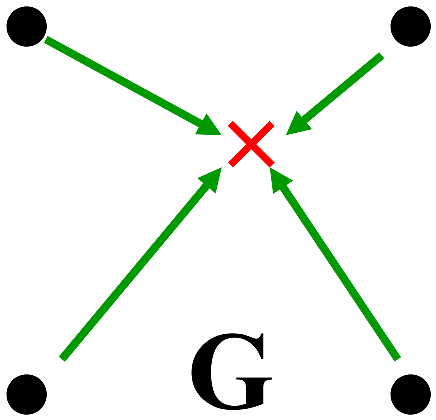
$$\mathbf{d} = \left(\mathbf{y} - \mathbf{G}\mathbf{x}^{\mathrm{b}}\right)$$

Innovation vector

η(t) = Correction for model error

η(t)=0 : Strong constraint

η(t)≠0 : Weak constraint



Observation matrix

The Linear Optimal Estimate

Analysis: $$\mathbf{z_a} = \mathbf{z_b} + \mathbf{Kd}$$

Gain (dual):

$$\mathbf{K} = \mathbf{BG}^{\mathrm{T}}(\mathbf{GBG}^{\mathrm{T}} + \mathbf{R})^{-1}$$

Gain (primal):

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

Regional Ocean Modeling System (ROMS) 4D-Var

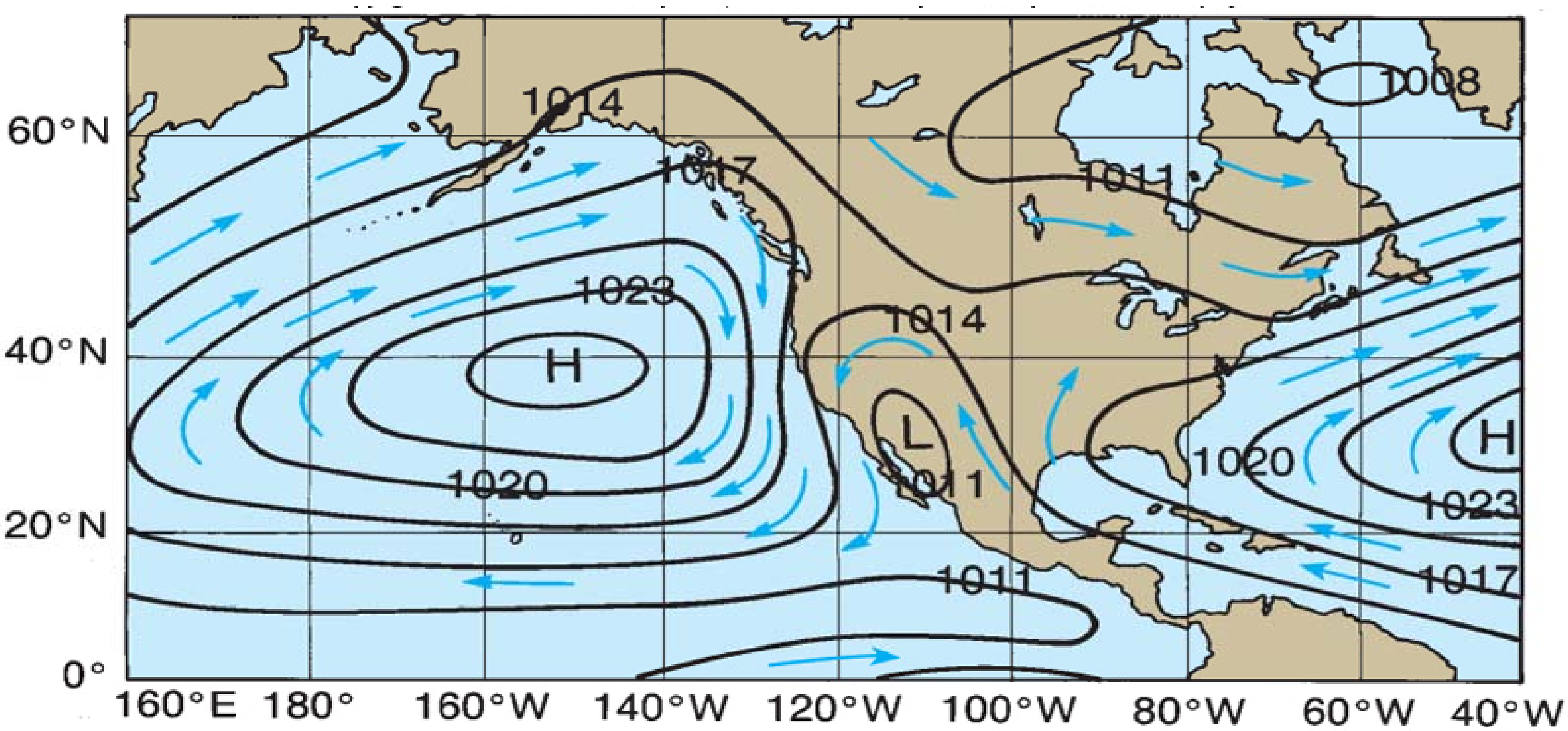
- **Incremental (linearized about a prior)** (Courtier et al, 1994)
- **Control vector: initial conditions, surface forcing, boundary conditions.**
- **Primal & dual formulations** (Courtier 1997)
- **Primal – Incremental 4-Var (I4D-Var)**
- **Dual – Lanczos-augmented RPCG & indirect representer (R4D-Var)** (Egbert et al, 1994; Gürol et al, 2014)
- **Strong and weak (dual only) constraint**
- **Preconditioned, Lanczos formulation of conjugate gradient** (Lorenc, 2003; Tshimanga et al, 2008; Fisher, 1997)
- **2nd-level preconditioning for multiple outer-loops**
- **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** (Song et al, 2012)

ROMS 4D-Var Diagnostic Tools

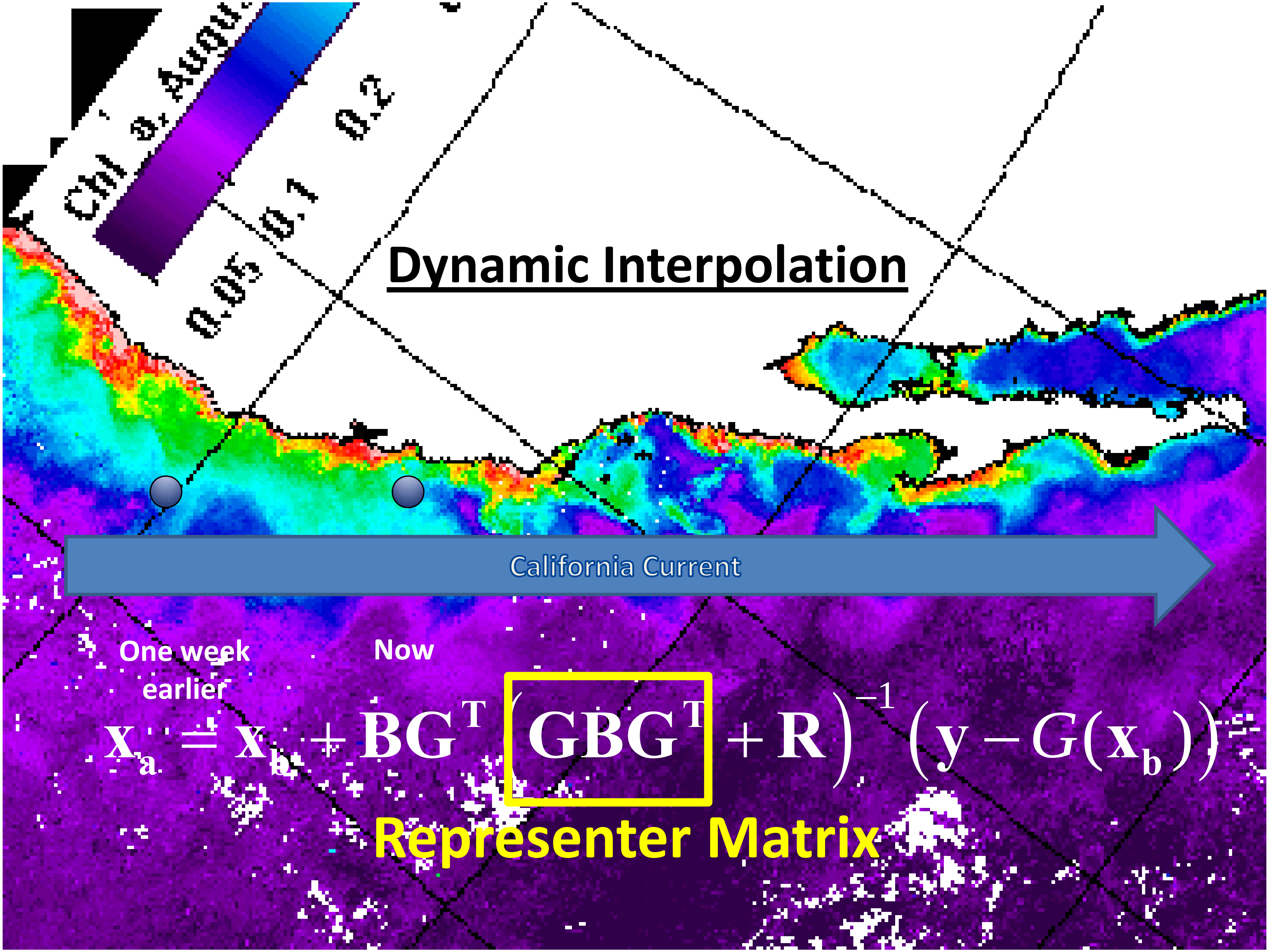
- **Observation impact** (Langland and Baker, 2004; Errico 2007)
- **Observation sensitivity – adjoint of 4D-Var**

  (Gelaro et al, 2004)
- **Singular value decomposition** (Barkmeijer et al, 1998; Moore et al., 2004, 2009)
- **Expected errors** (Moore et al., 2012; Smith et al., 2015)

The California Current



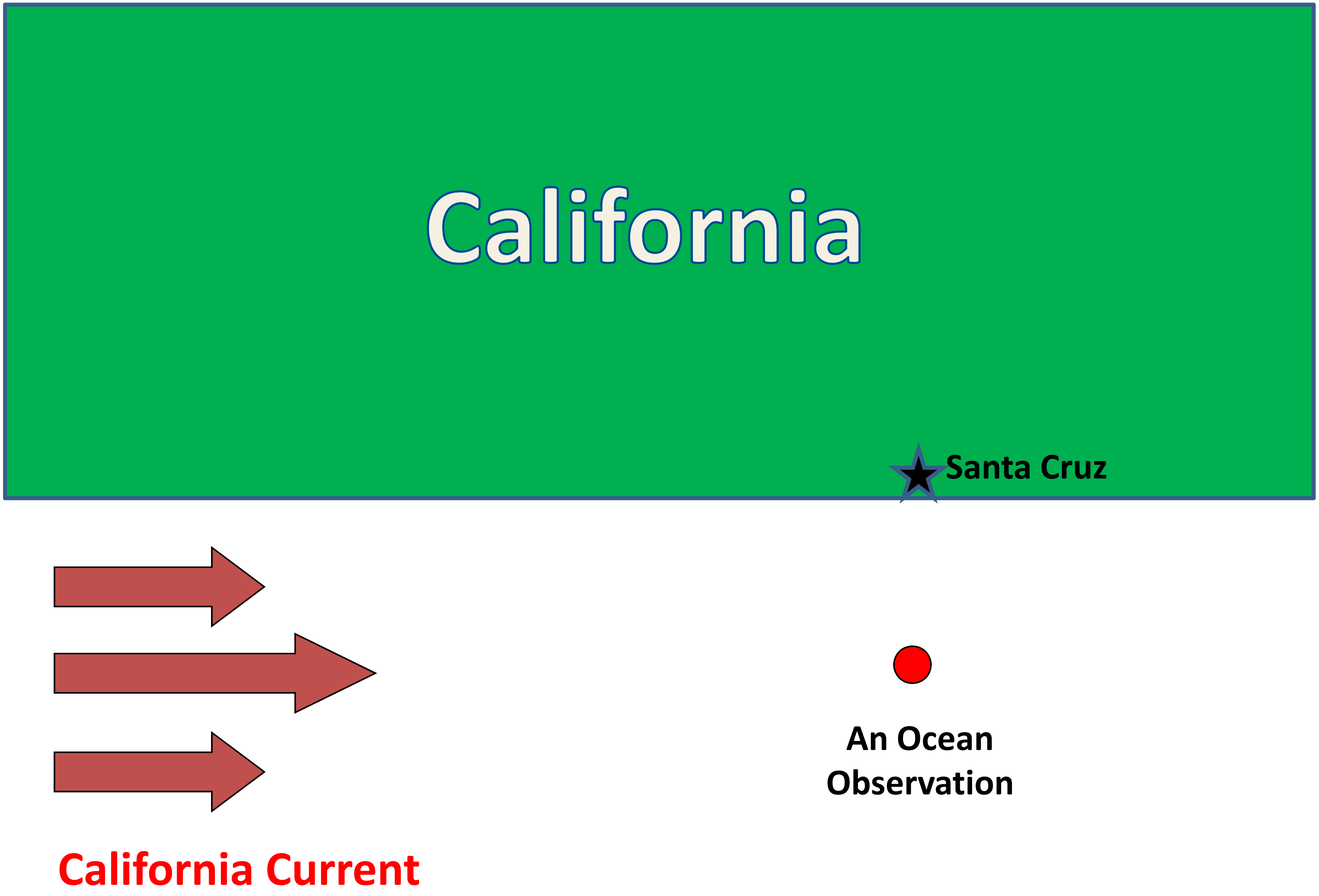
Dynamic Interpolation



$$x_a = x_b + BG^T\left(\boxed{GBG^T} + R\right)^{-1}\left(y - G(x_b)\right)$$

Representer Matrix

Dynamic Interpolation



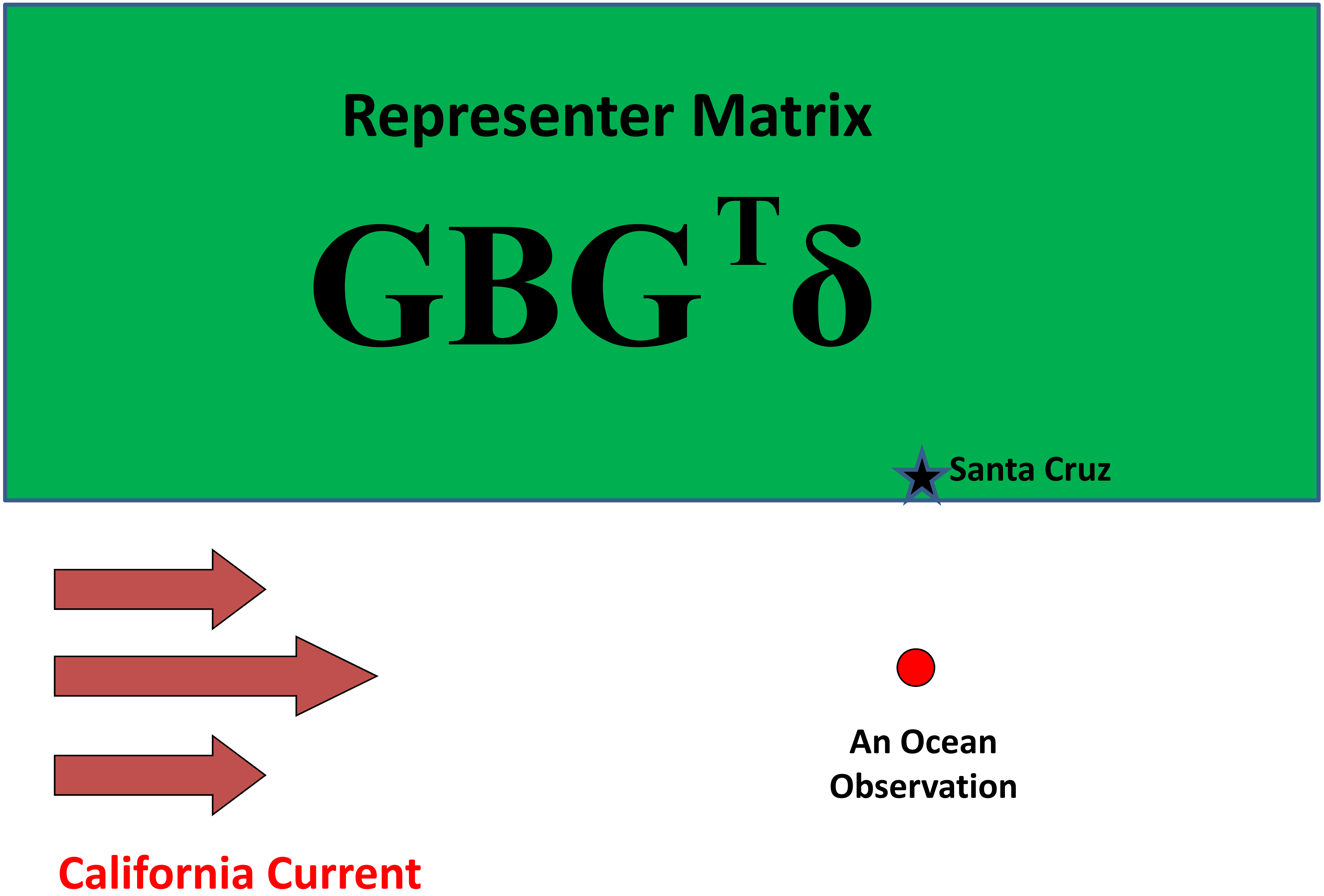
California

Santa Cruz

An Ocean Observation

California Current

Dynamic Interpolation



Representer Matrix

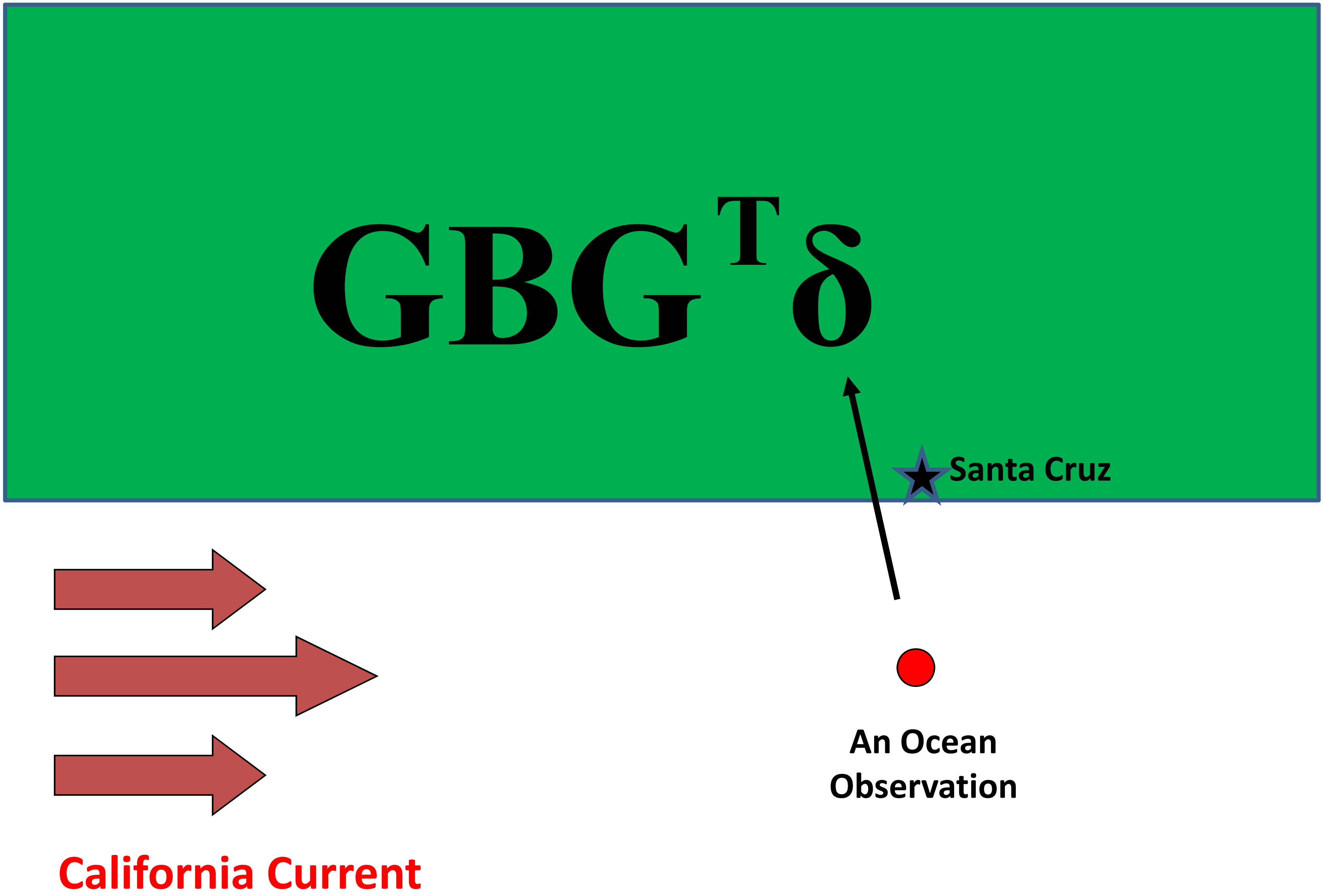
$$GBG^T\delta$$

Santa Cruz

An Ocean Observation

California Current

Dynamic Interpolation



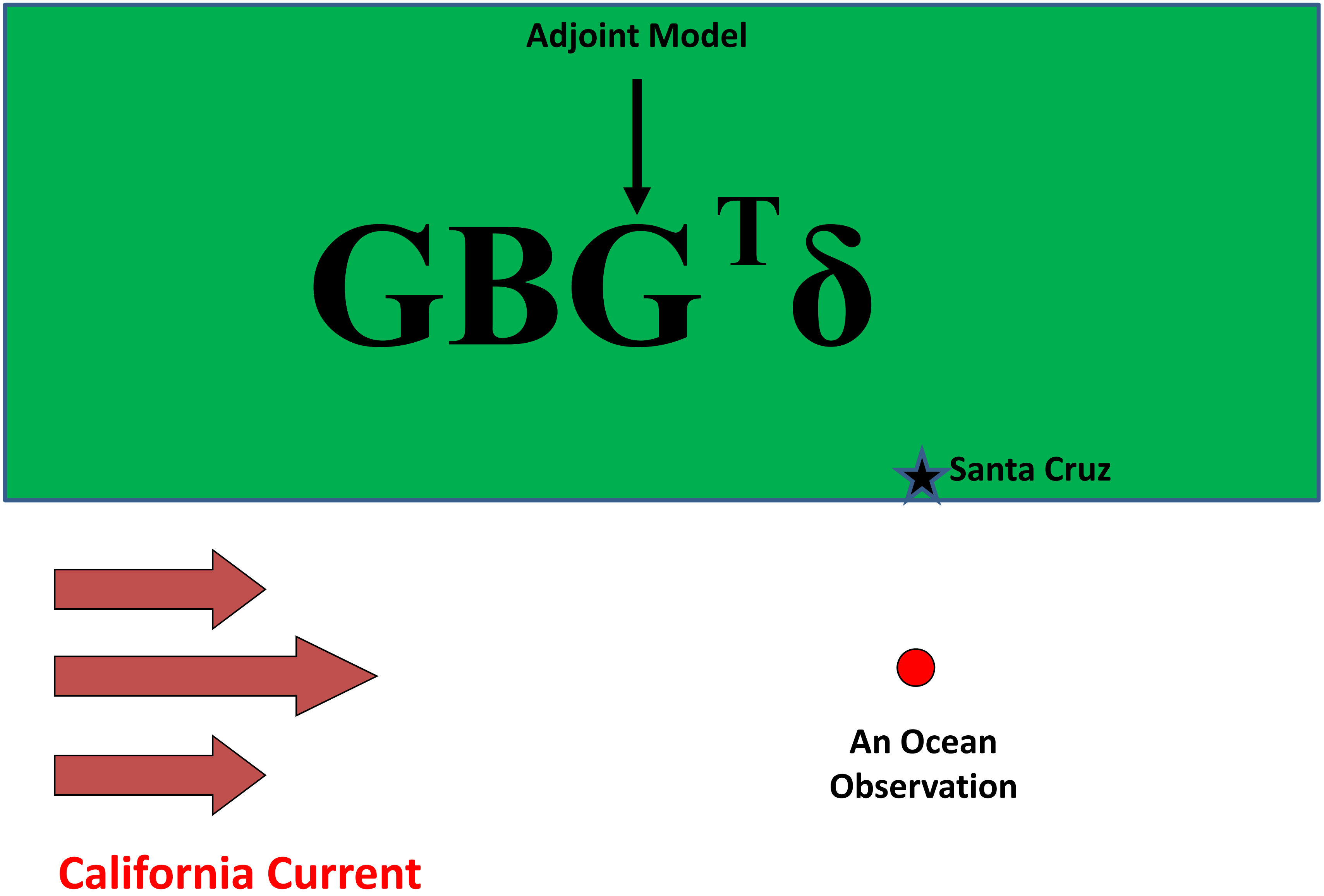
$GBG^{T}\delta$

Santa Cruz

An Ocean Observation

California Current

Dynamic Interpolation



Adjoint Model

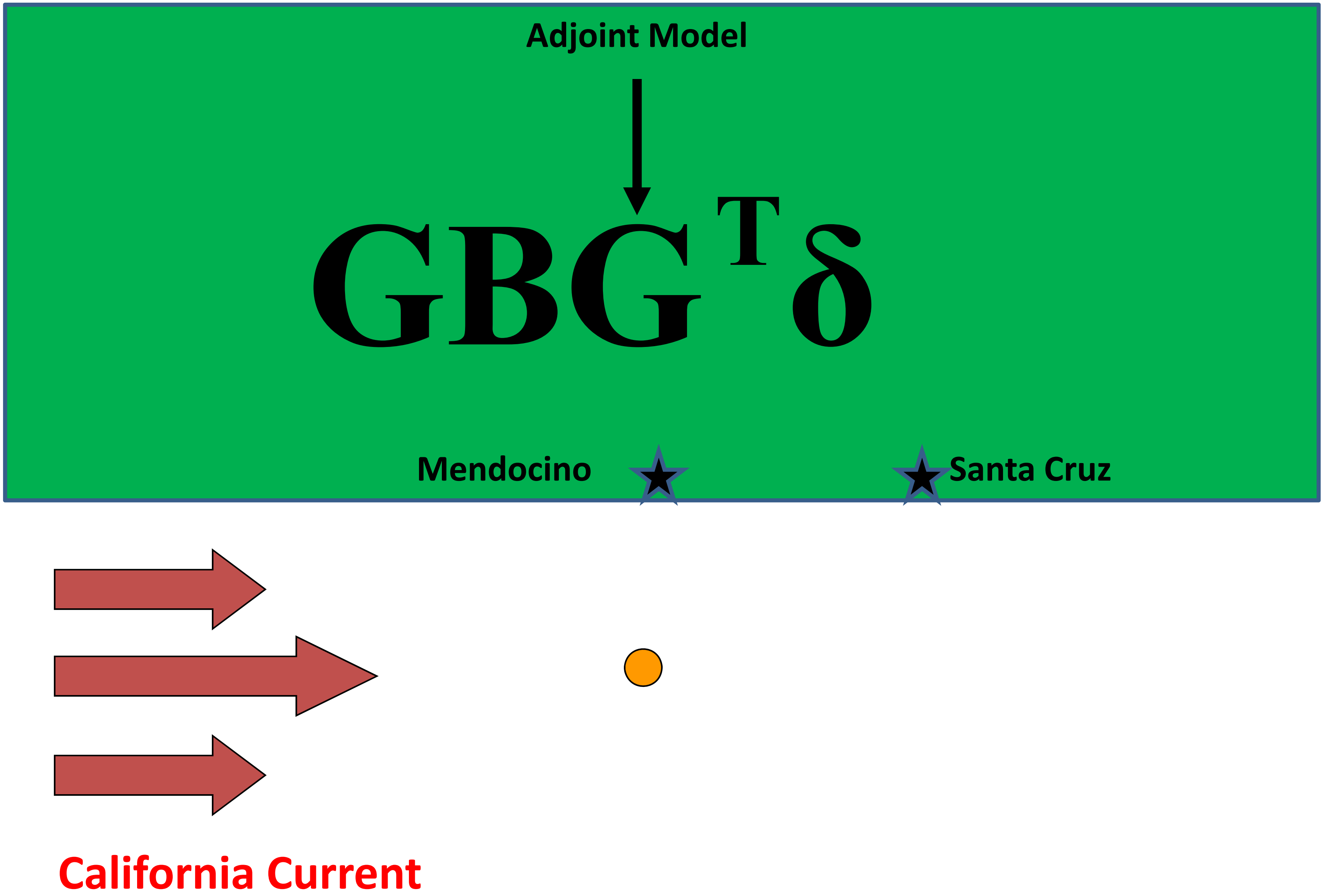
$$GBG^T\delta$$

Santa Cruz

An Ocean Observation

California Current

Dynamic Interpolation



Adjoint Model

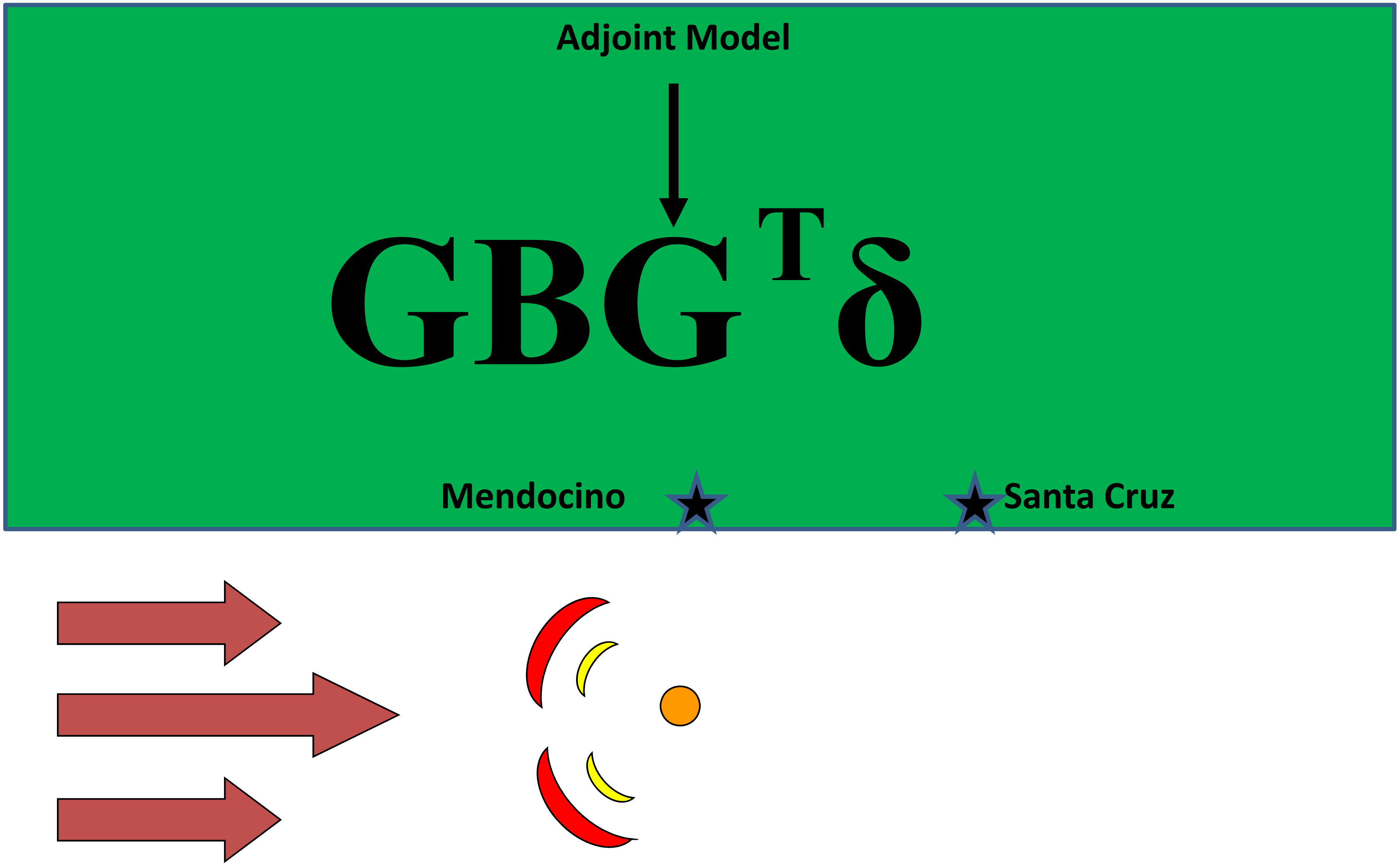
$$GBG^{T}\delta$$

Mendocino

Santa Cruz

California Current

Dynamic Interpolation



Adjoint Model

$$GBG^T\delta$$

Mendocino

Santa Cruz

California Current

Dynamic Interpolation

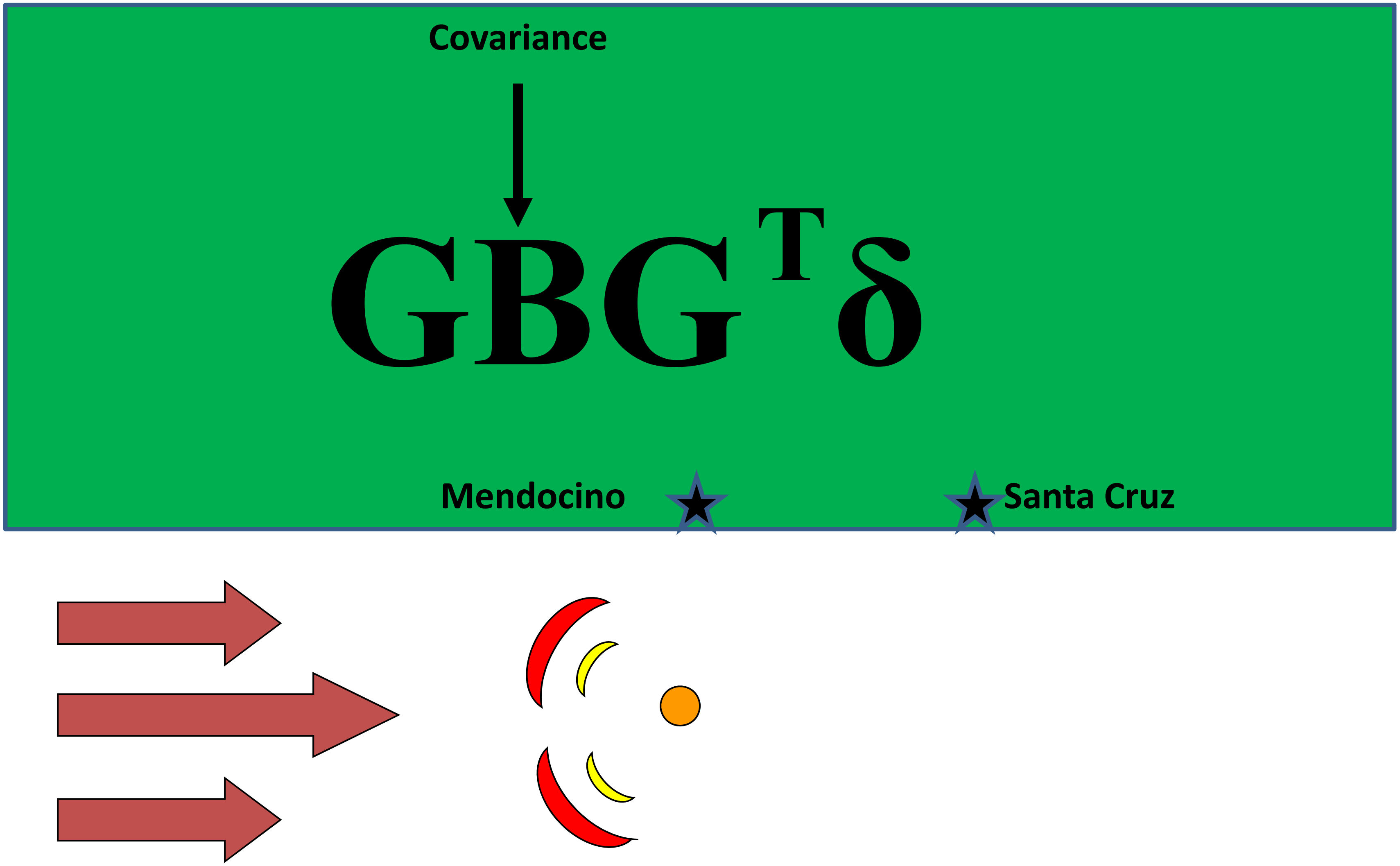
$$GB\overbrace{G^{T}\delta}^{\boxed{\text{Green's Function}}}$$

Mendocino

Santa Cruz

California Current

Dynamic Interpolation



Covariance

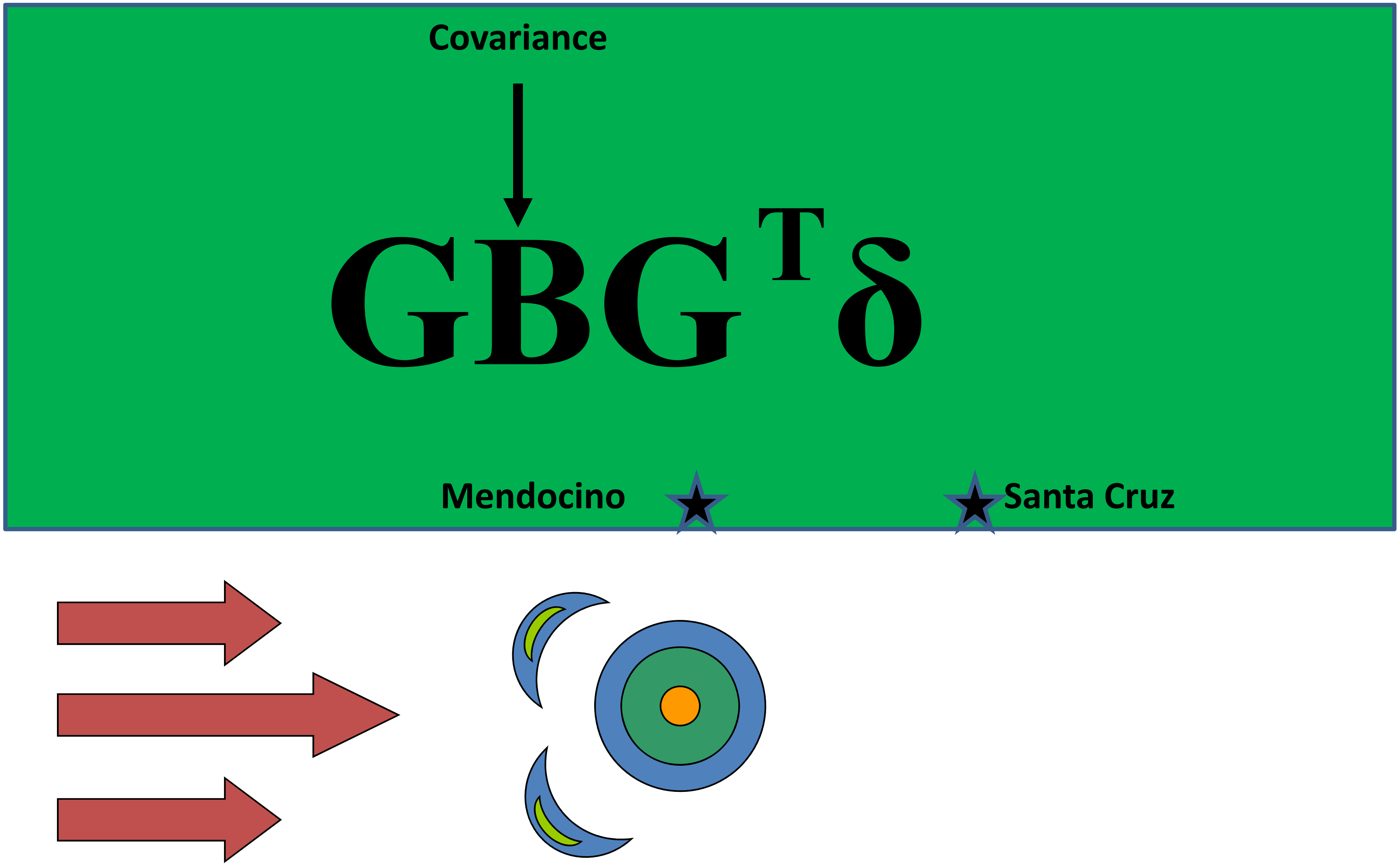
$$GBG^{T}\delta$$

Mendocino

Santa Cruz

California Current

Dynamic Interpolation



Covariance

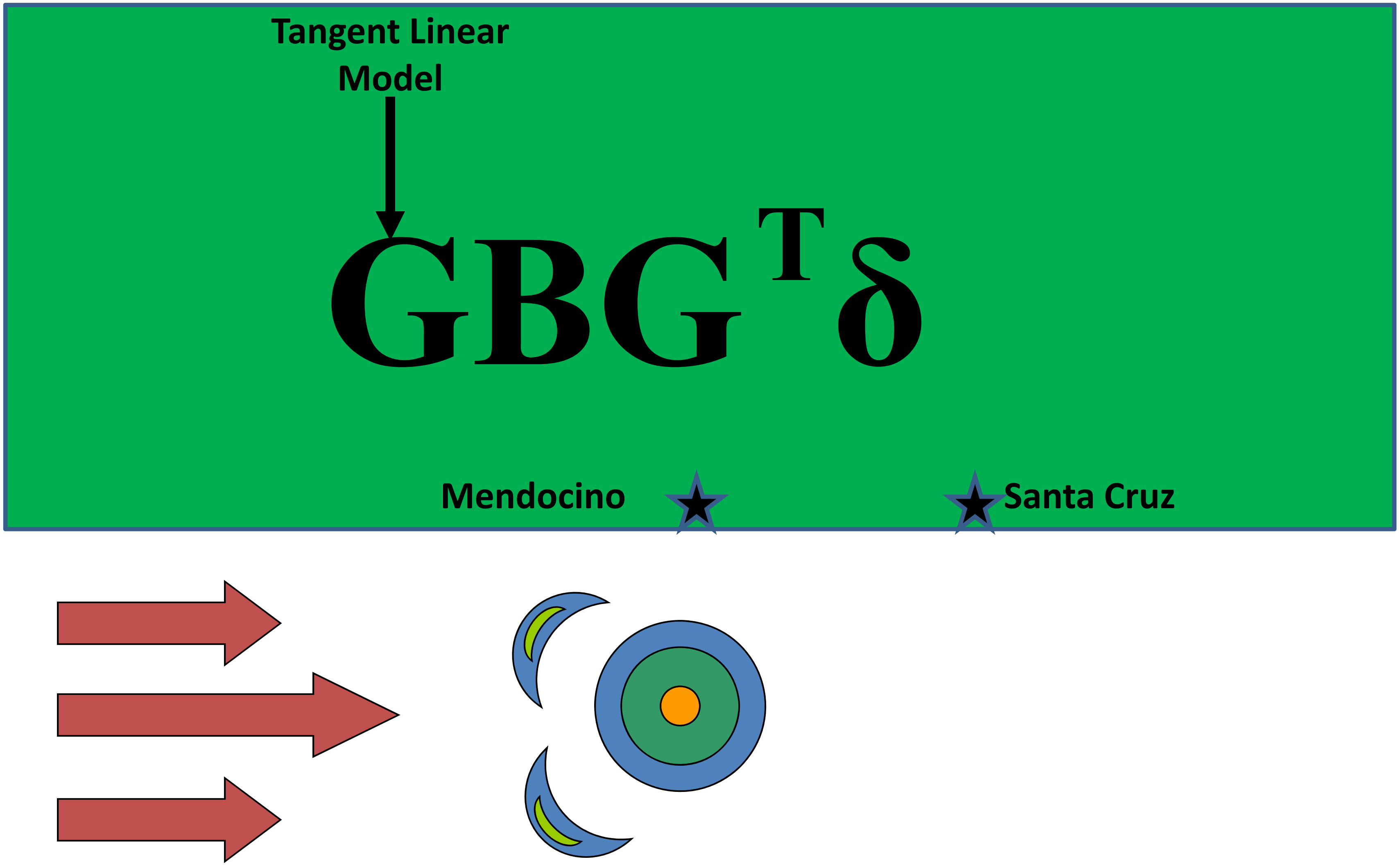
$$GBG^T\delta$$

Mendocino

Santa Cruz

California Current

Dynamic Interpolation



Tangent Linear Model

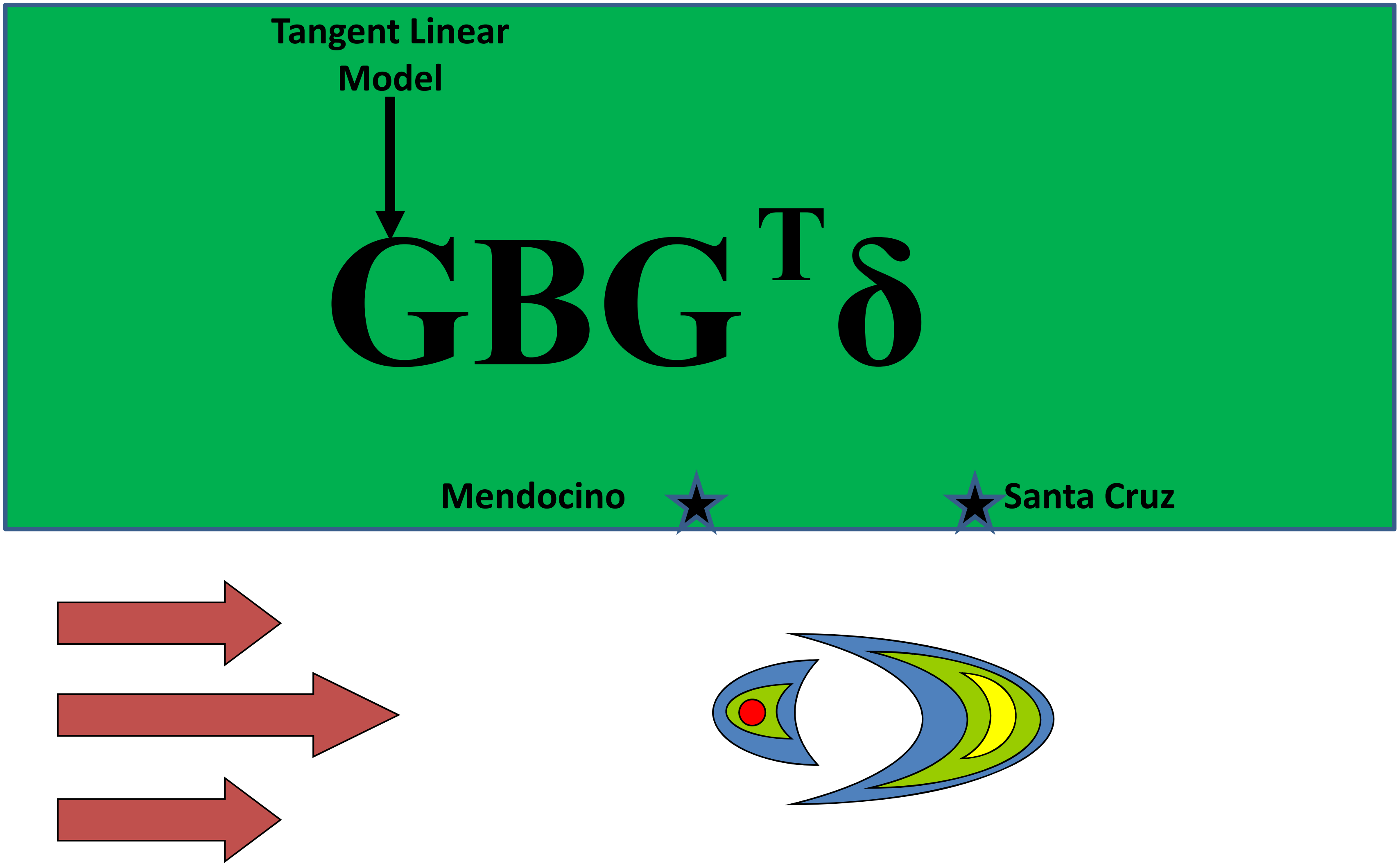
$$GBG^T\delta$$

Mendocino

Santa Cruz

California Current

Dynamic Interpolation



Tangent Linear Model

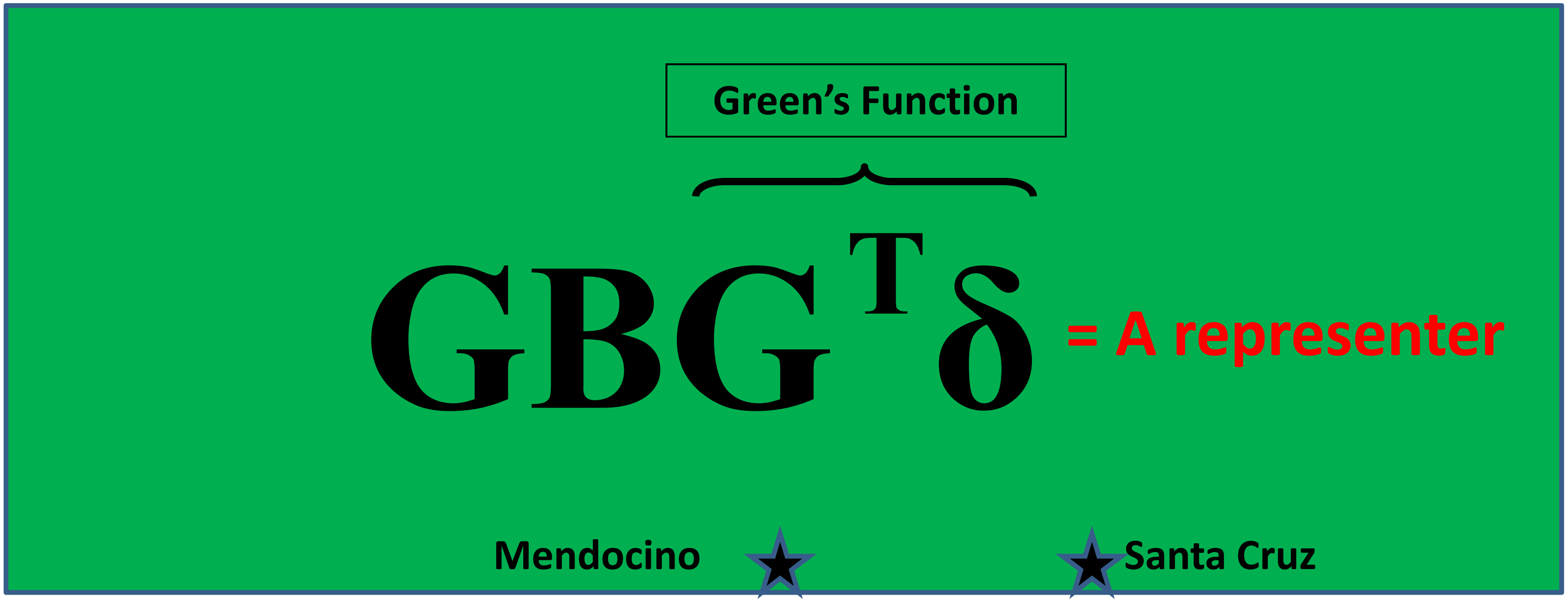
$$GBG^{T}\delta$$

Mendocino

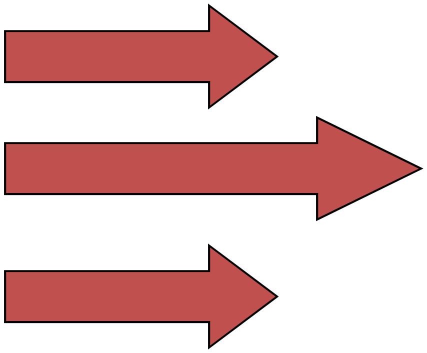
Santa Cruz

California Current

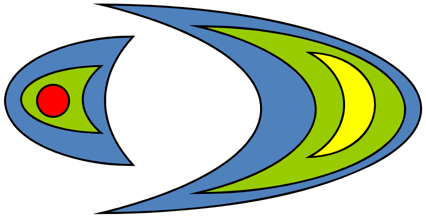
Dynamic Interpolation



$$\mathrm{GB}\overbrace{\mathrm{G}^{T}\delta}^{\text{Green's Function}}$$ = A representer



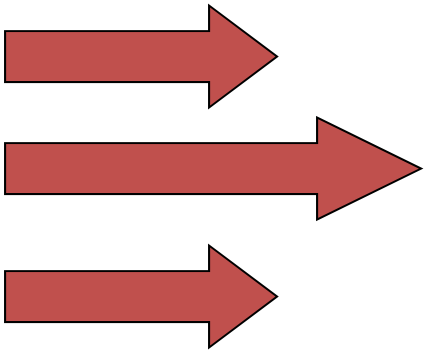
California Current



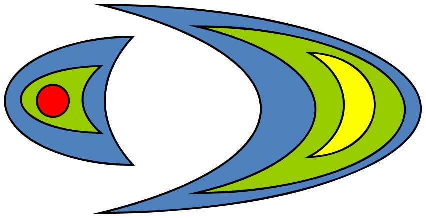
A covariance

Dynamic Interpolation

$GBG^T\delta$ = A representer

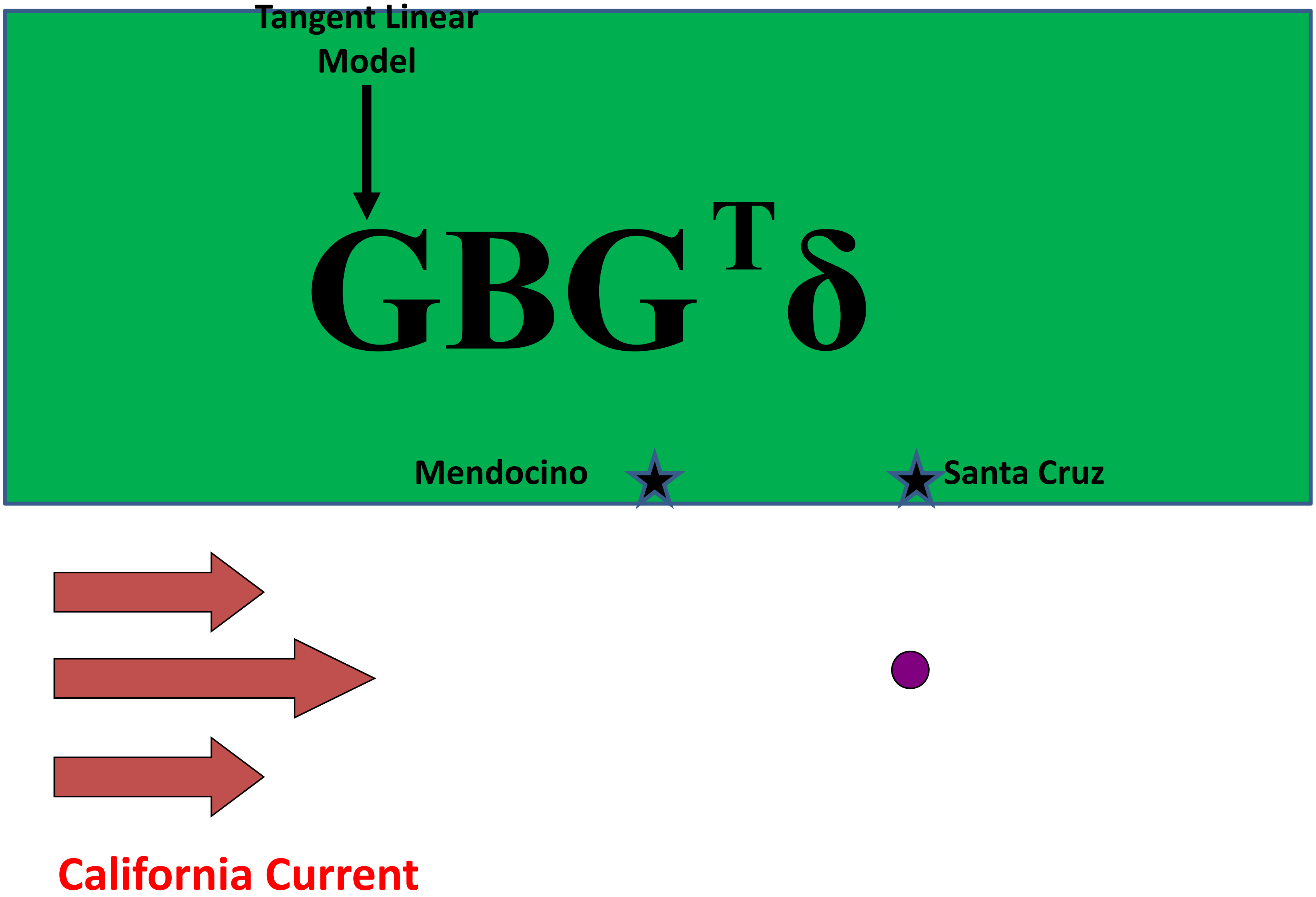


California Current



A covariance

Dynamic Interpolation



Tangent Linear Model

$$GBG^T\delta$$

Mendocino

Santa Cruz

California Current

ROMS CCS 30 Yr Analysis

CCMP+ERA
(1980-2010)
or COAMPS
(1999-2012)
forcing

$f_b(t), B_f$

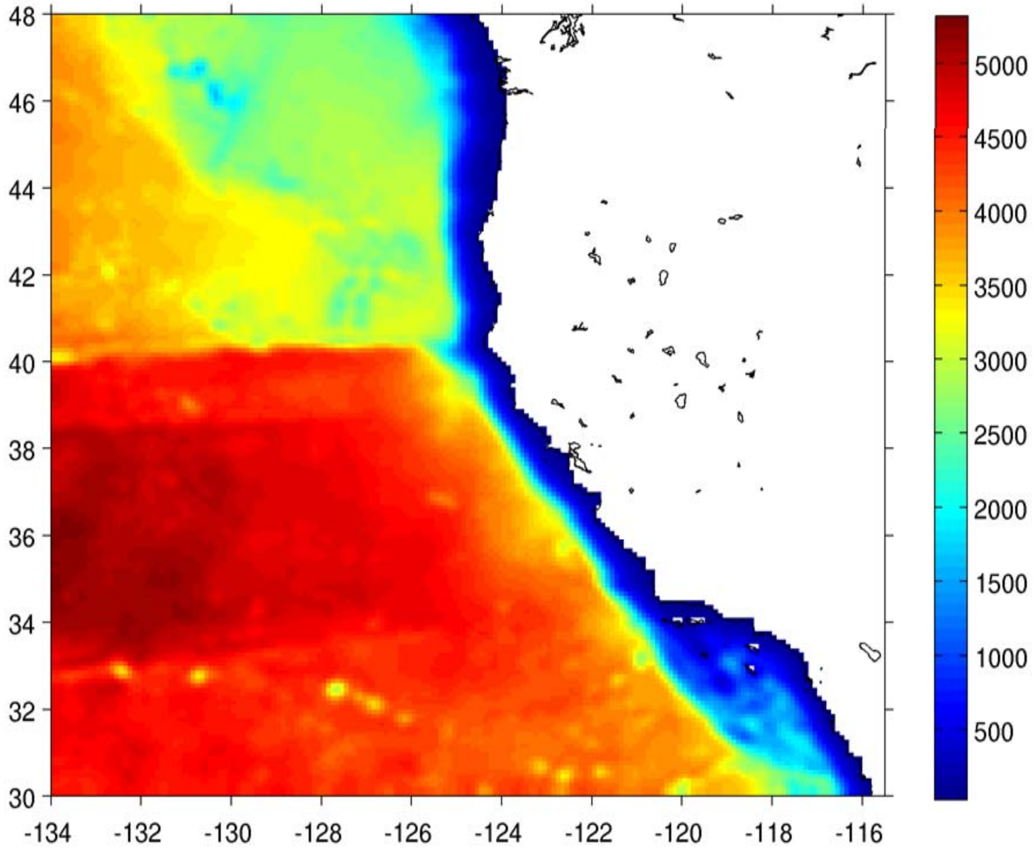
SODA open
boundary
conditions

$b_b(t), B_b$

$x_b(0), B_x$

↑

Previous
assimilation cycle
(8 day overlapping cycles)

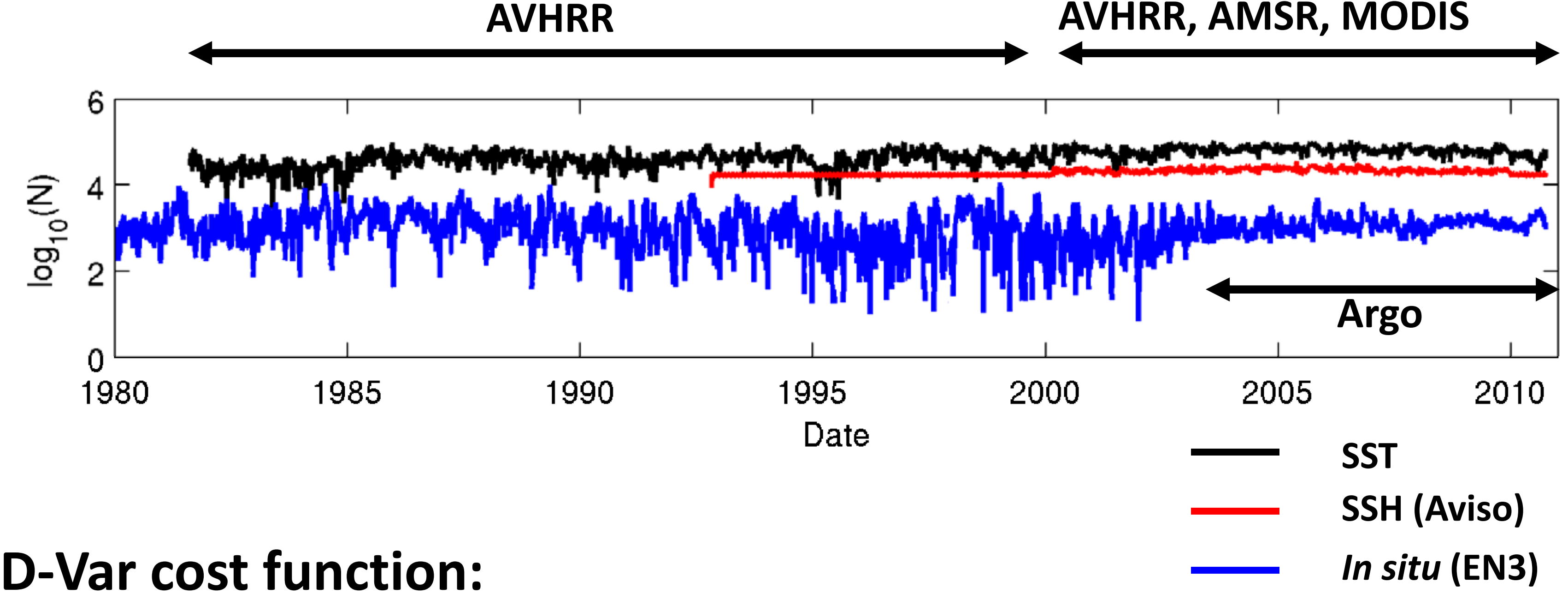


1/10° horizontal resolution, 42 levels

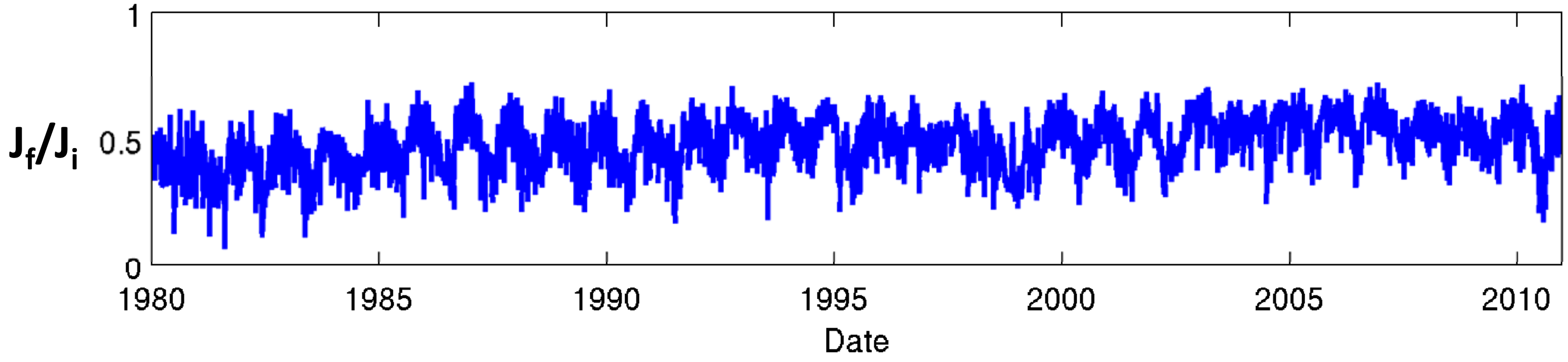
Veneziani et al (2009)
Broquet et al (2009)
Moore et al (2010)

Diagnostic Summary

Number of Obs:



4D-Var cost function:



Observation Impact vs Observation Sensitivity

$$\mathbf{x}_\mathrm{a} = \mathbf{x}_\mathrm{b} + \tilde{\mathbf{K}}(\mathbf{y} - G(\mathbf{x}_\mathrm{b}))$$

posterior=prior + gain×innovation

Observation impact

Scalar function: $I(\mathbf{x})$ **(e.g. transport)**

Change due to 4D-Var: $\Delta I = I(\mathbf{x}_\mathrm{a}) - I(\mathbf{x}_\mathrm{b})$

$$\Delta I = I(\mathbf{x}_\mathrm{b} + \tilde{\mathbf{K}}\mathbf{d}) - I(\mathbf{x}_\mathrm{b})$$

$$\simeq \mathbf{d}^T \tilde{\mathbf{K}}^T (\partial I / \partial \mathbf{x})\big|_{\mathbf{x}_\mathrm{b}}$$

$$= (\mathbf{y} - G(\mathbf{x}_\mathrm{b}))^T \tilde{\mathbf{K}}^T (\partial I / \partial \mathbf{x})\big|_{\mathbf{x}_\mathrm{b}}$$

Change in I can be uniquely attributed to each obs y_i.

Observation sensitivity

4D-Var as a function: $\mathbf{x}_\mathrm{a} = \mathbf{x}_\mathrm{b} + \mathcal{K}(\mathbf{d})$

Scalar function: $I(\mathbf{x})$ **(e.g. transport)**

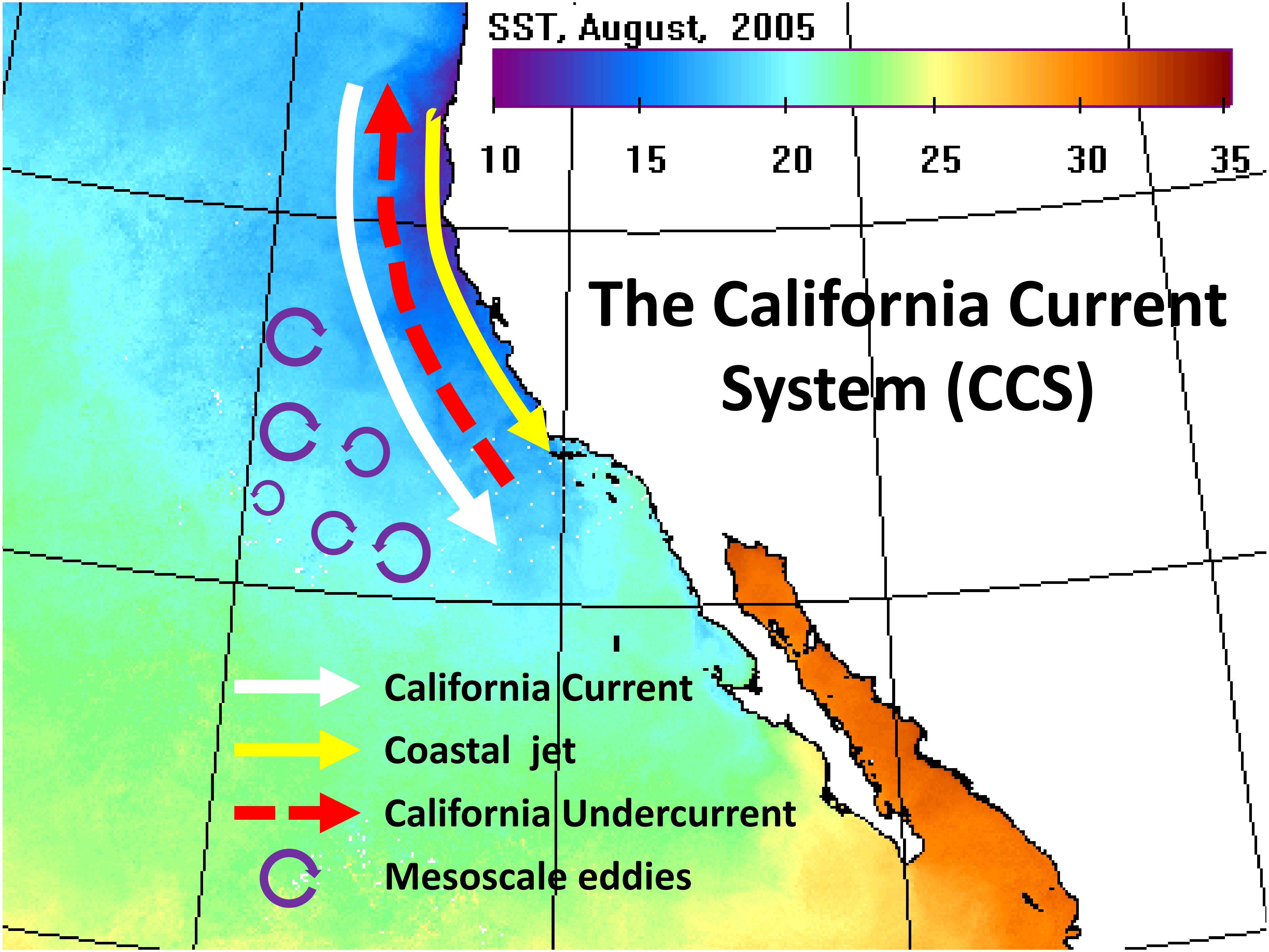
Change in I due to change δy in y:

$$\delta I \simeq \delta \mathbf{y}^T (\partial \mathcal{K} / \partial \mathbf{y})\big|_{\mathbf{x}_\mathrm{a}}^T (\partial I / \partial \mathbf{x})\big|_{\mathbf{x}_\mathrm{a}}$$

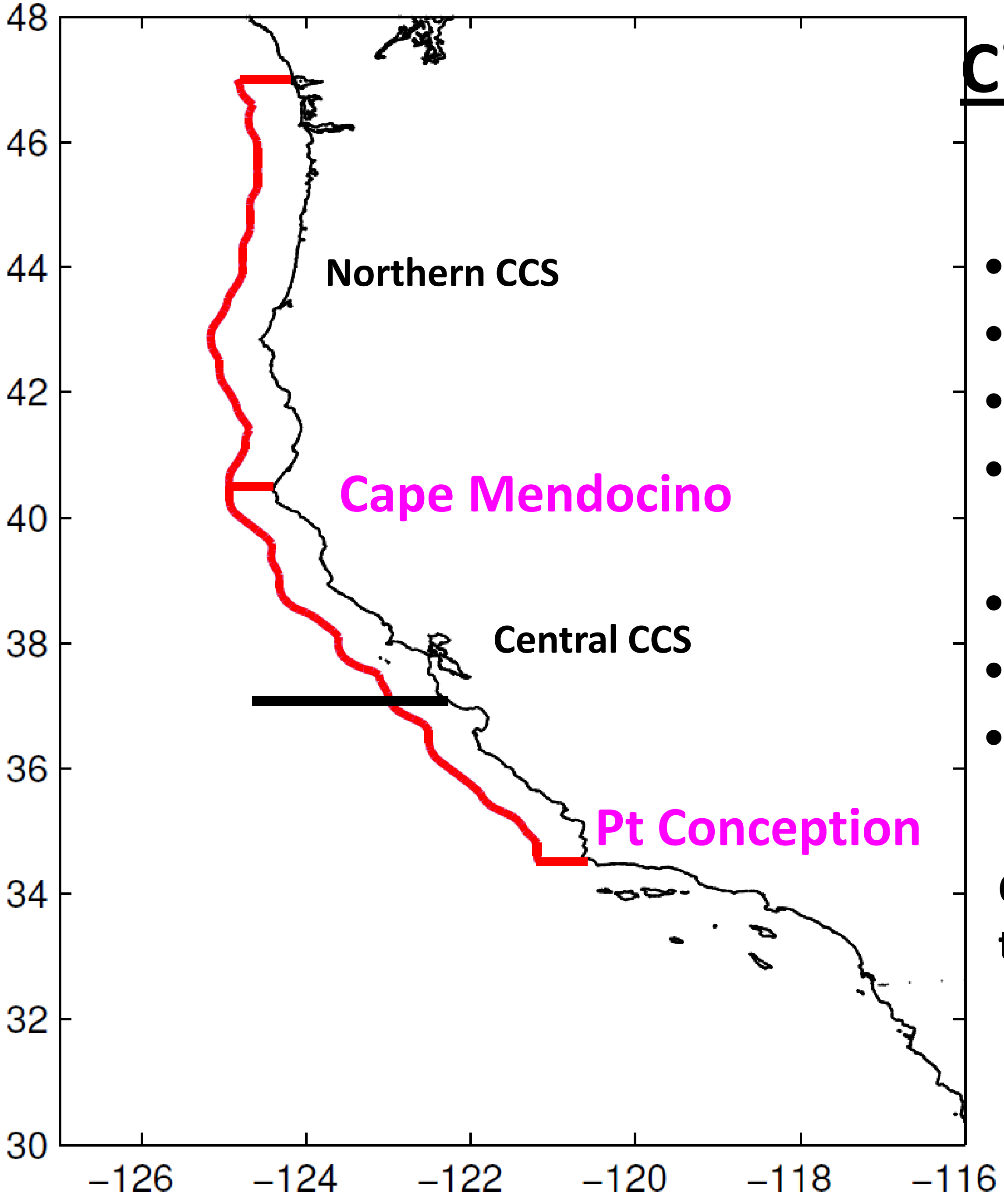
For exact arithmetic and complete convergence:

$$\tilde{\mathbf{K}} = (\partial \mathcal{K} / \partial \mathbf{y})\big|_{\mathbf{x}_\mathrm{a}} = \mathbf{K}$$

The California Current System (CCS)



Circulation Metrics, *I*(x)



- 37N transport
- CUC transport
- Upwelling transport
- σ=26 kg m^{-3} isopycnal depth

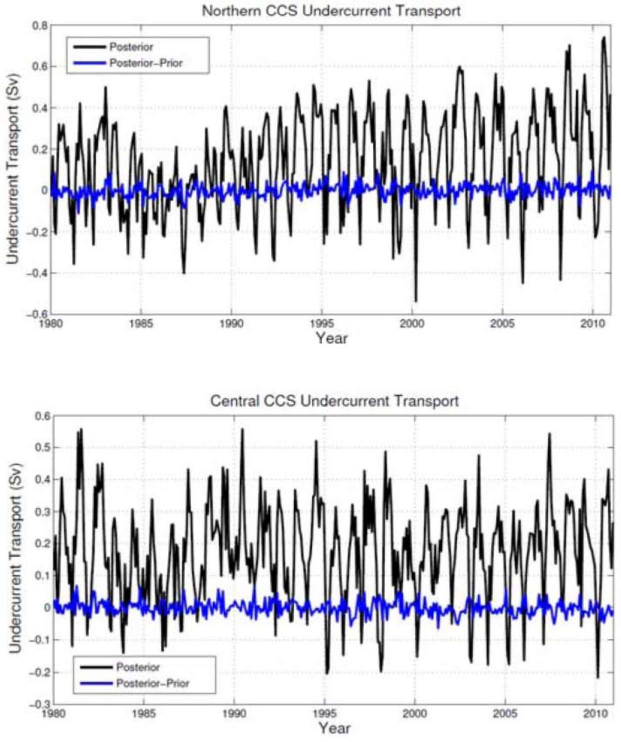
- Two regions
- 8 day averages
- Every 4D-Var cycle

Change due to 4D-Var: $$\Delta I = I(\mathbf{x}_a) - I(\mathbf{x}_b)$$

Circulation Indices & Target Areas

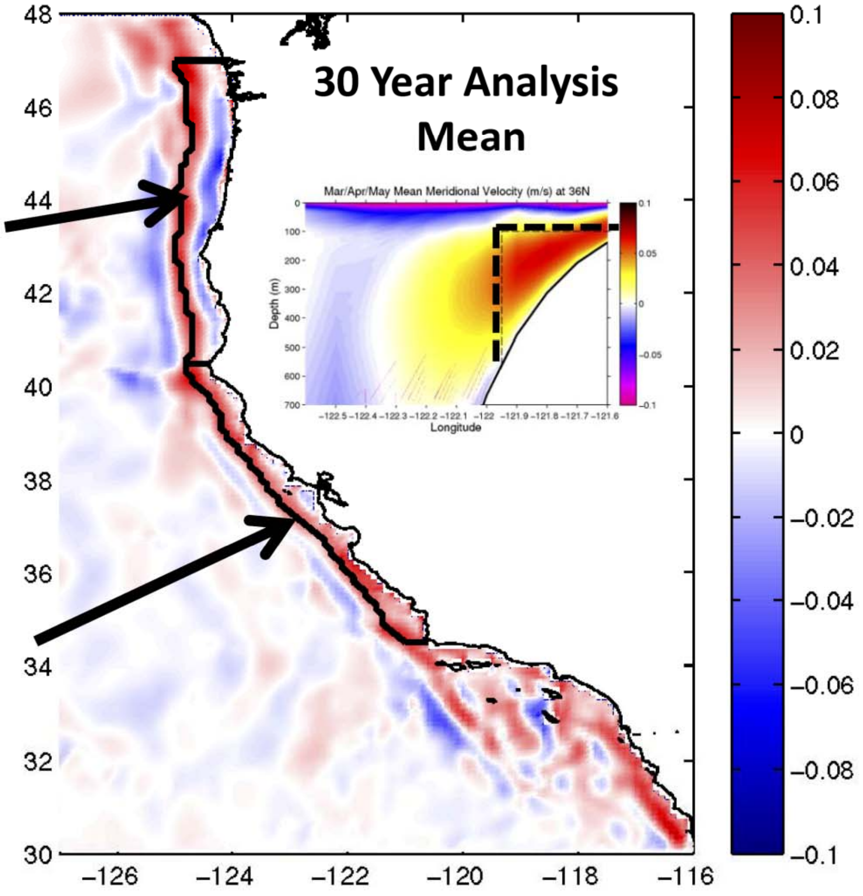
The California Undercurrent

California Undercurrent Transport



Analysis

Analysis - Background



30 Year Analysis Mean

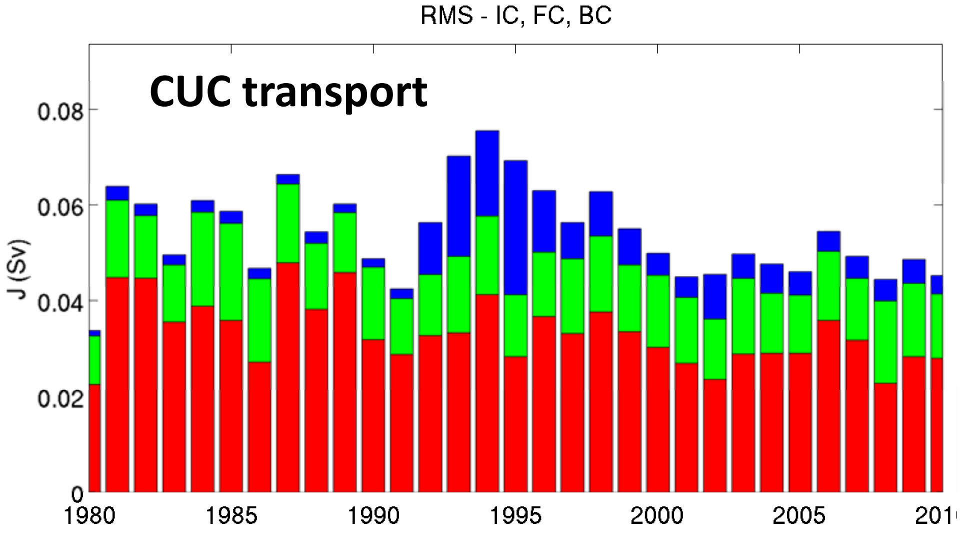
Alongshore v on s-level 16

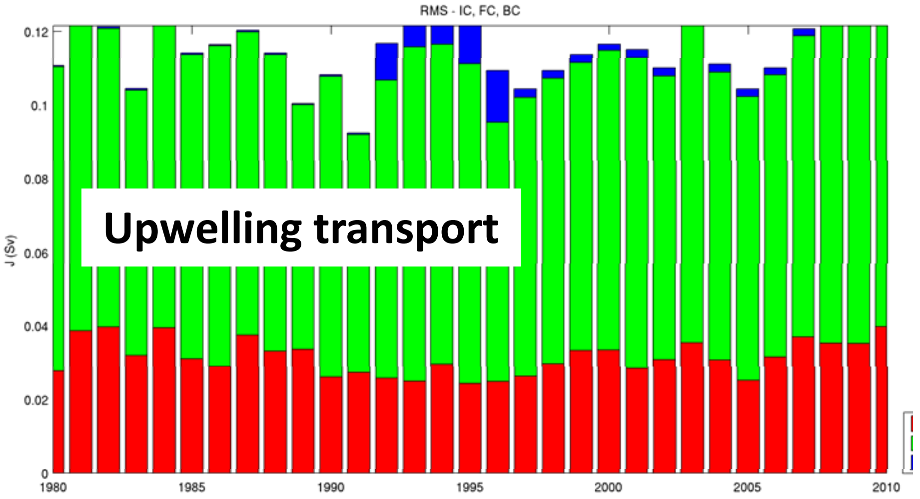
Control Vector Impacts

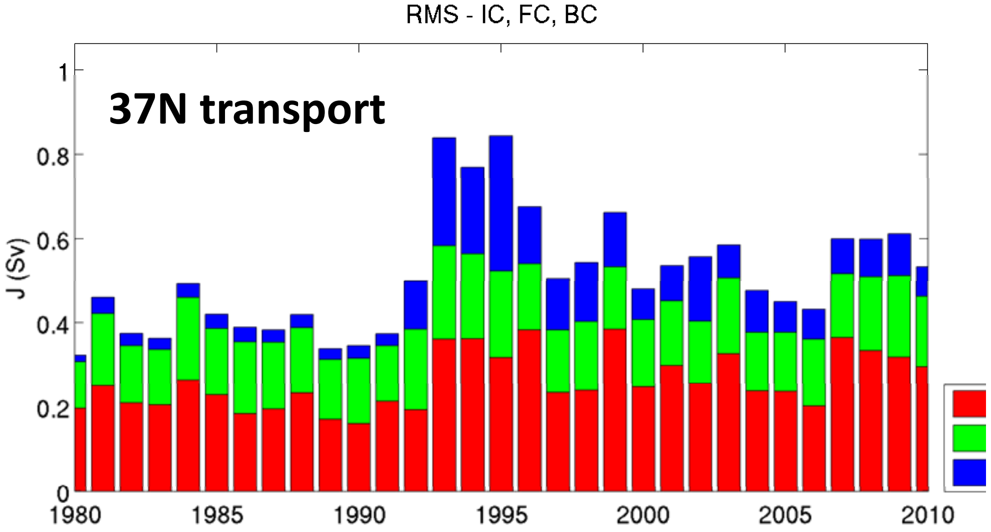
$$\Delta I = (\mathbf{y} - G(\mathbf{x_b}))^T \tilde{\mathbf{K}}^T \left. (\partial I / \partial \mathbf{x}) \right|_{\mathbf{x_b}}$$

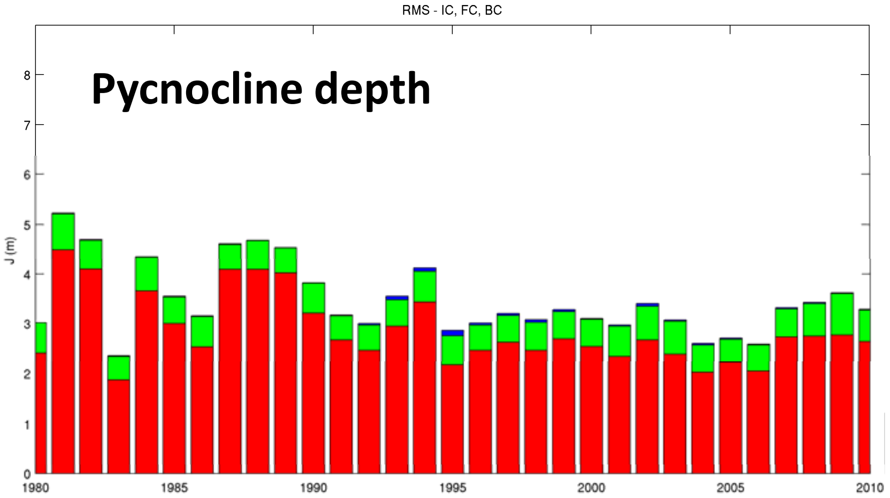
$$= \Delta I_{\mathbf{x}} + \Delta I_{\mathbf{f}} + \Delta I_{\mathbf{b}}$$

Control Vector Monitoring









🟥 **Initial conditions**

🟩 **Surface forcing**

🟦 **Open boundary conditions**

Observing Platform Impacts

$$\Delta I = (\mathbf{y} - G(\mathbf{x_b}))^T \tilde{\mathbf{K}}^T (\partial I / \partial \mathbf{x})\Big|_{\mathbf{x_b}}$$

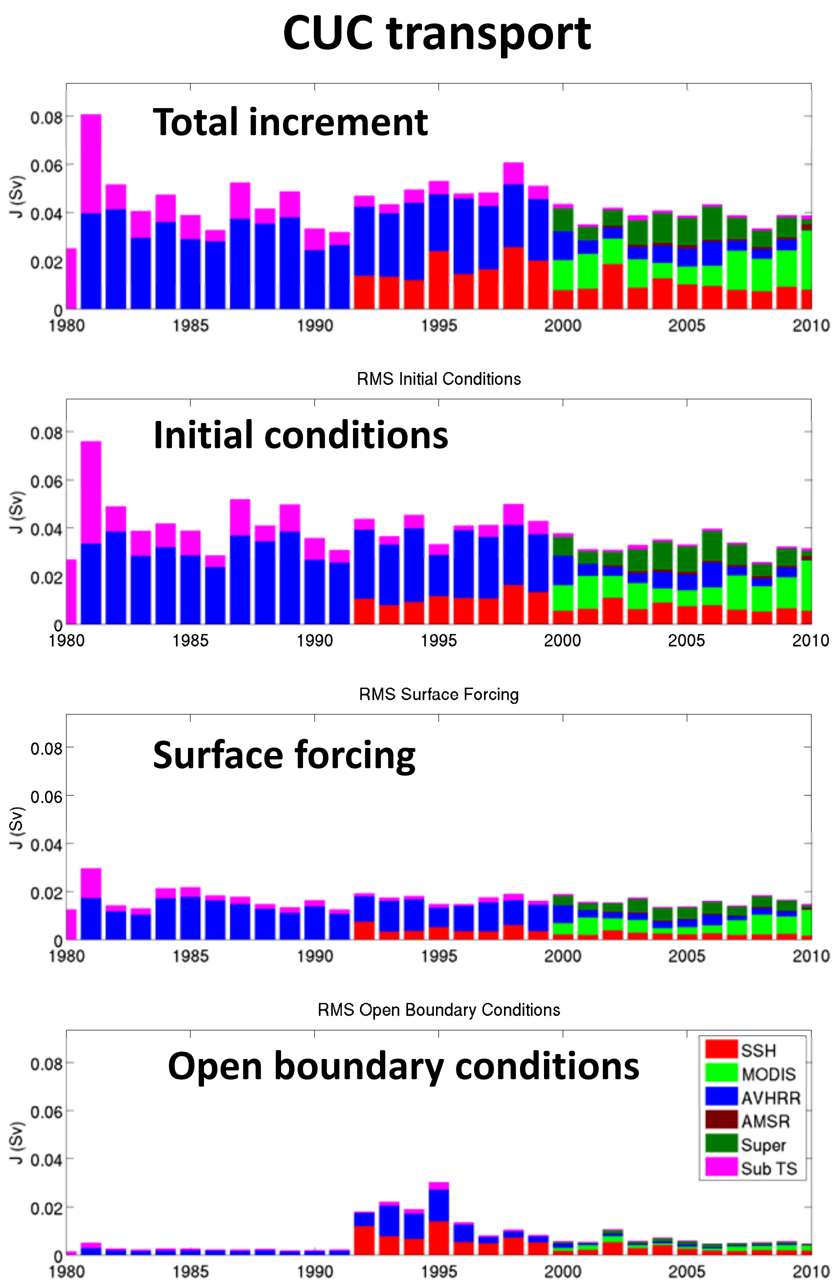
$$= \Delta I_{SST} + \Delta I_{SSH} + \Delta I_{T,S}$$

$$\Delta I = (\mathbf{y} - G(\mathbf{x_b}))^T \tilde{\mathbf{K}}^T (\partial I / \partial \mathbf{x})\Big|_{\mathbf{x_b}}$$

$$= \Delta I_{\mathbf{x}} + \Delta I_{\mathbf{f}} + \Delta I_{\mathbf{b}}$$

SST	SST	SST
SSH	SSH	SSH
T,S	T,S	T,S

Observation Impacts


CUC transport

🟥 **SSH**

🟩 **MODIS SST**

🟦 **AVHRR SST**

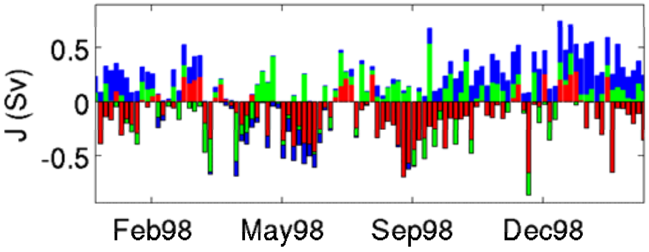
🟫 **AMSR SST**

🟩 **SST super obs**

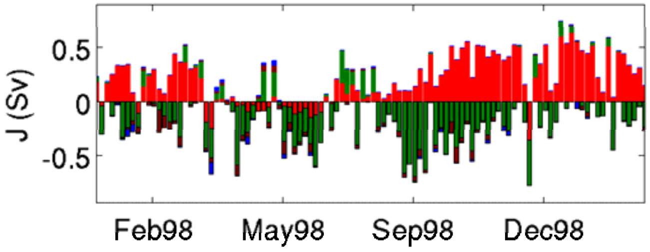
🟪 **Hydrographic obs**

37N transport

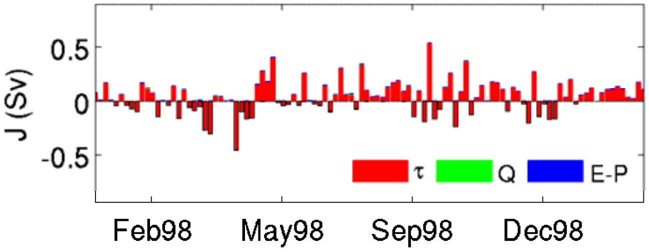
Control Vector Impacts



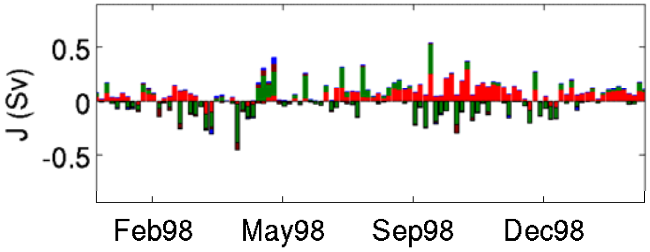
Total Transport Impacts



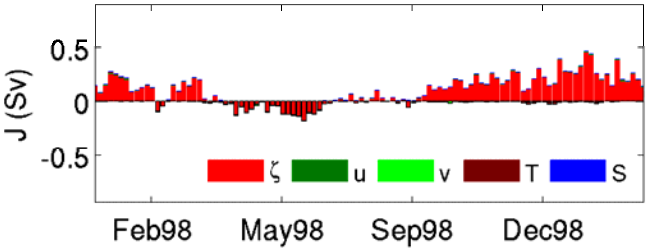
Surface Forcing Impacts



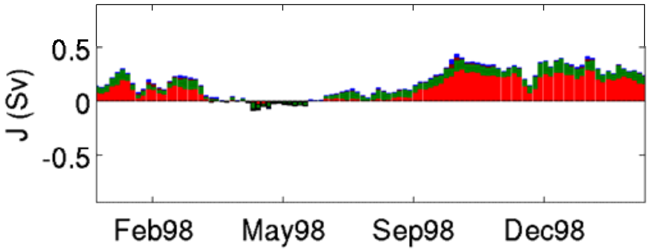
τ Impacts



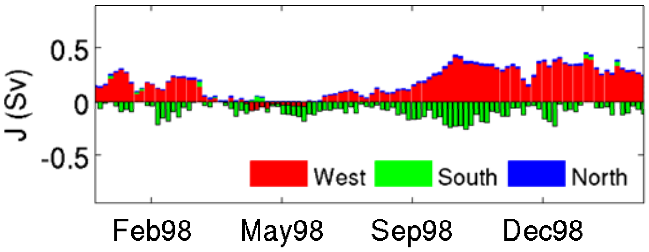
Open Boundary Impacts



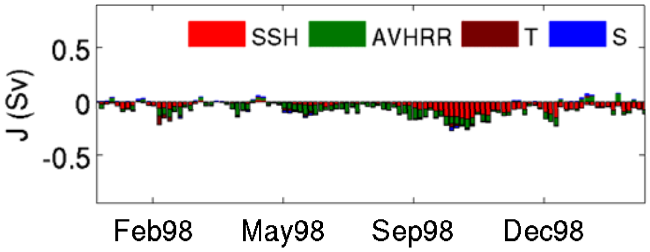
ζ West Impacts



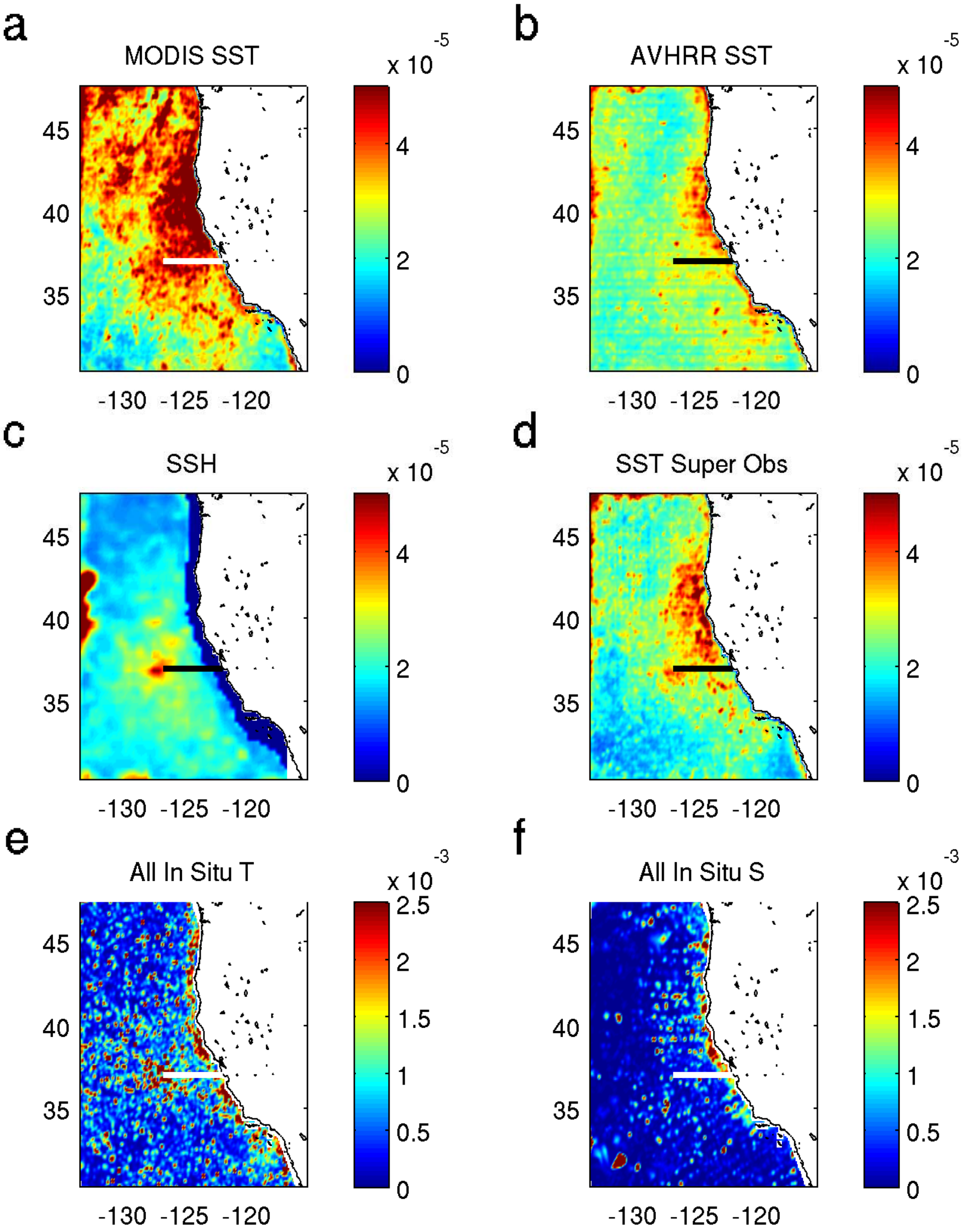
Open Boundary Impacts



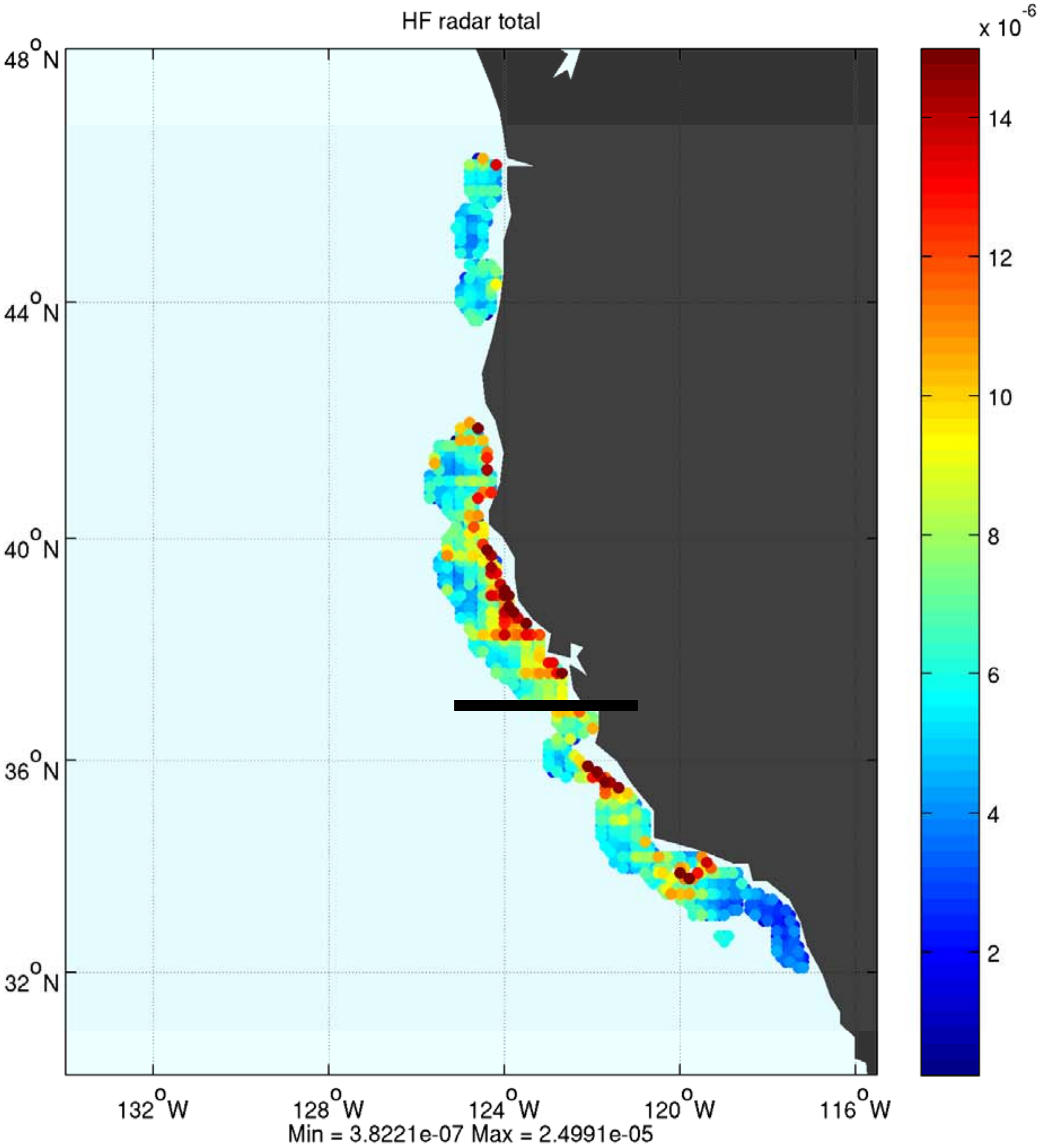
ζ South Impacts



37N transport



Impact of HF Radar on 37N transport



Information Horizons

For 8 day assimilation cycles:

- Advection: ~70 km (u ~ 0.1 m/s)
- 1st baroclinic mode waves: ~1700 km (c~2.5 m/s)
- Coastal waveguides: ~1700 km
- Barotropic waves – whole domain
- SSH pressure gradient – gyre scale
- Covariance regularization: ~300 km