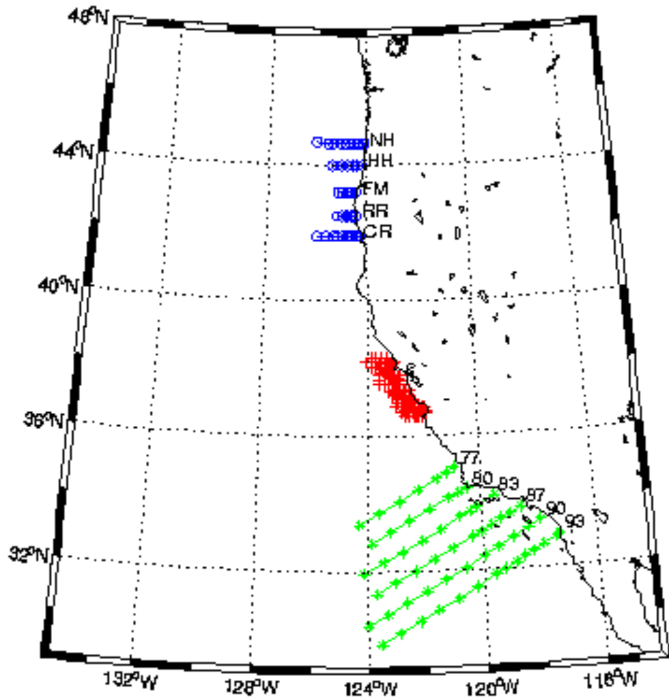
Lecture 5:
Observation Impact & Observation Sensitivity

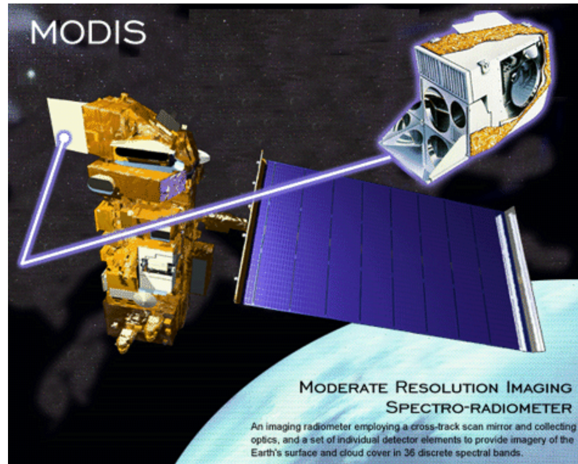
Outline

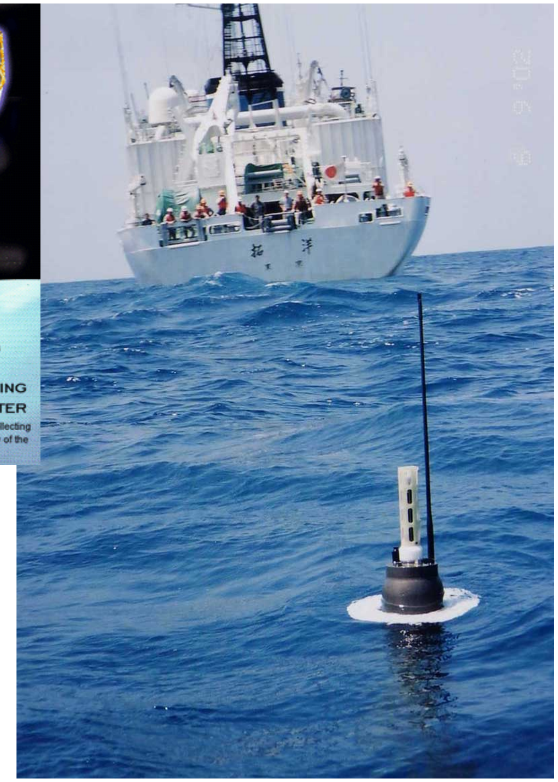
- Observation impacts:
  - (a) analysis cycle
  - (b) forecast cycle
- Adjoint 4D-Var: $(\text{4D-Var})^T$
- Observation sensitivity

Observation Impacts

(Useful references: Langland & Baker, 2004; Gelaro and Zhu, 2009; Tremolet, 2008)

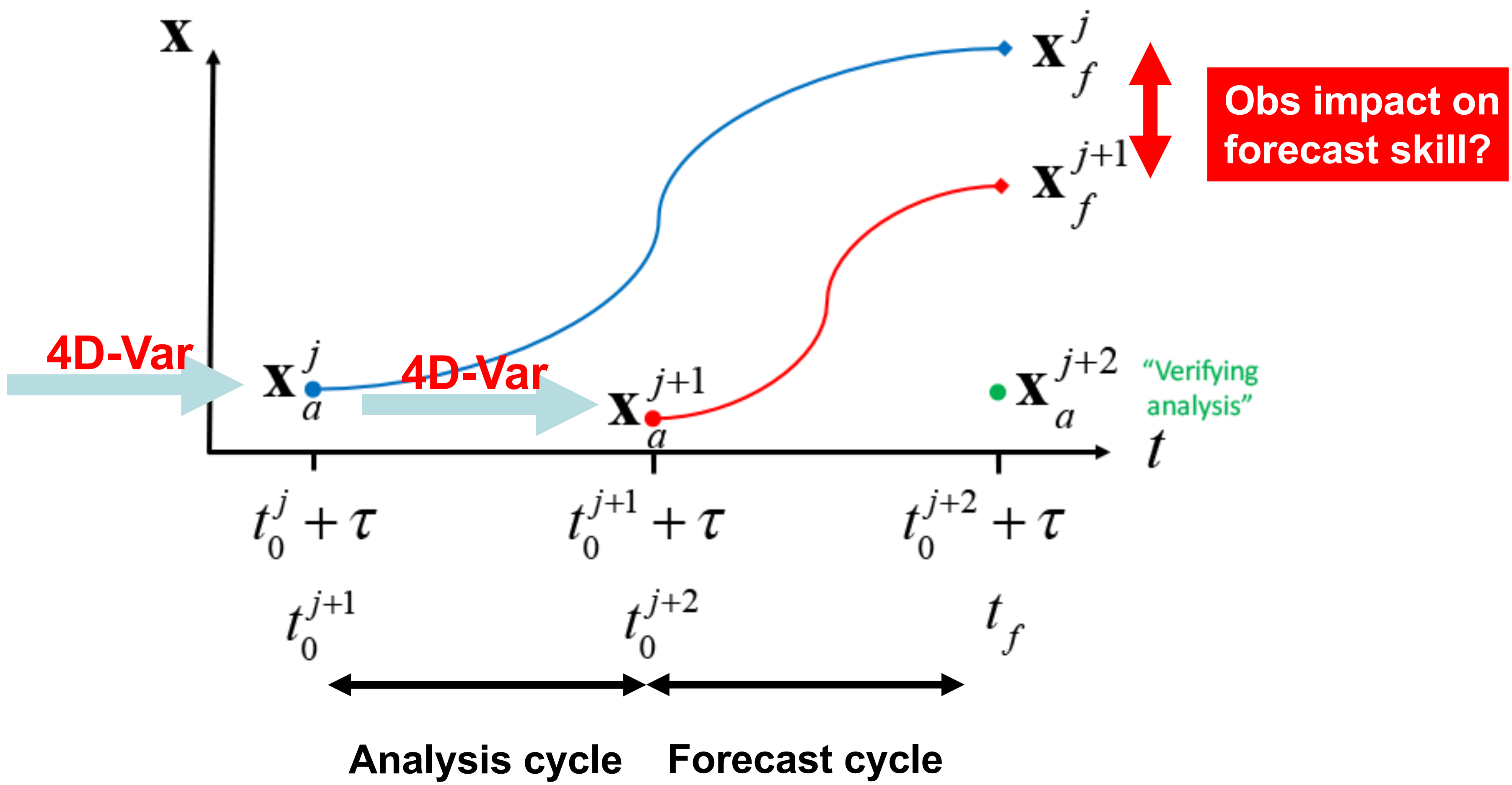






Given the plethora of different observation platforms, what impact does each have on the 4D-Var analysis?

A Typical Sequential Analysis-Forecast Procedure



Analysis Cycle Observation Impacts

Analysis Cycle Observation Impacts

Consider a scalar function of the ocean state vector:

$$I = I(\mathbf{x})$$

Prior	*Posterior*	*Increment*
$I_{\mathbf{b}} = I(\mathbf{x}_{\mathbf{b}})$	$I_{\mathbf{a}} = I(\mathbf{x}_{\mathbf{a}})$	$\Delta I = I_{\mathbf{a}} - I_{\mathbf{b}}$

$$\mathbf{x}_{\mathbf{a}}(\mathrm{t}) = \mathbf{x}_{\mathbf{b}}(\mathrm{t}) + \delta\mathbf{x}(\mathrm{t})$$

$$I_{\mathbf{a}} = I(\mathbf{x}_{\mathbf{b}} + \delta\mathbf{x}) \simeq I_{\mathbf{b}} + \delta\mathbf{x}^{\mathrm{T}} * (\partial I / \partial\mathbf{x})$$

$\Delta I \simeq \delta\mathbf{x}^{\mathrm{T}} * (\partial I / \partial\mathbf{x})$ **but** $\delta\mathbf{x}(t) = \mathbf{M}_{\mathbf{b}}(t, t_0)\tilde{\mathbf{K}}\mathbf{d}$

$$\boxed{\Delta I \simeq \mathbf{d}^{\mathrm{T}}\tilde{\mathbf{K}}^{\mathrm{T}}\mathbf{M}_{\mathbf{b}}^{\mathrm{T}} * (\partial I / \partial\mathbf{x})}$$

($\mathbf{M}_{\mathbf{b}}^{\mathrm{T}} * (\partial I / \partial\mathbf{x})$ **denotes a time convolution)**

Analysis Cycle Observation Impacts

Consider a scalar function of the ocean state vector:

$$I = I(\mathbf{x})$$

Prior	***Posterior***	***Increment***
$I_{\mathbf{b}} = I(\mathbf{x}_{\mathbf{b}})$	$I_{\mathbf{a}} = I(\mathbf{x}_{\mathbf{a}})$	$\Delta I = I_{\mathbf{a}} - I_{\mathbf{b}}$

$$\Delta I \simeq \mathbf{d}^{\mathrm{T}} \tilde{\mathbf{K}}^{\mathrm{T}} \mathbf{M}^{\mathrm{T}} * (\partial I / \partial \mathbf{x})$$

Innovations ($\mathbf{d}^{\mathrm{T}}$)

Adjoint of gain matrix ($\tilde{\mathbf{K}}^{\mathrm{T}}$)

Adjoint model ($\mathbf{M}^{\mathrm{T}}$)

Analysis Cycle Observation Impacts

Recall the dual form of the gain matrix:

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

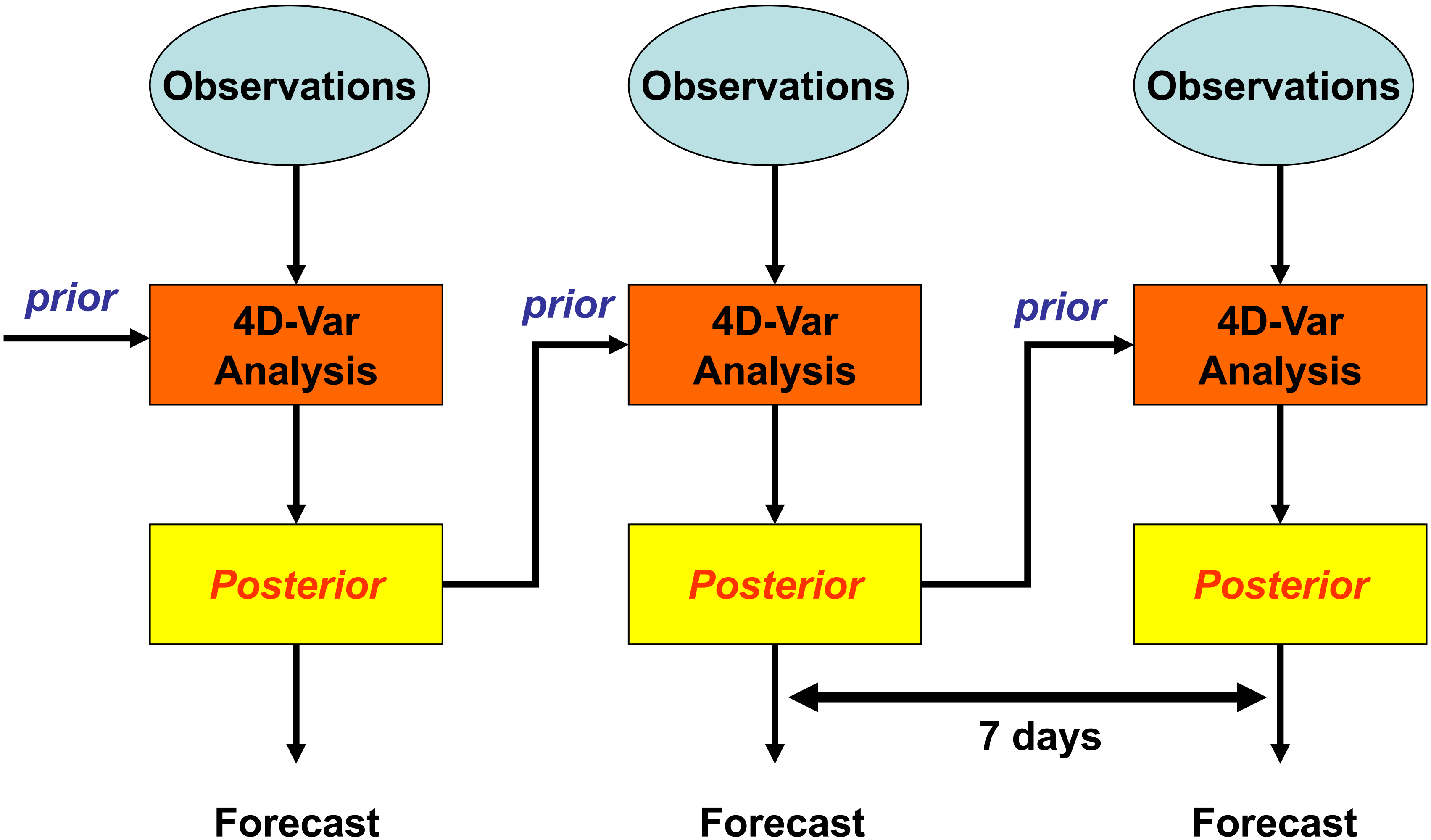
So:

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

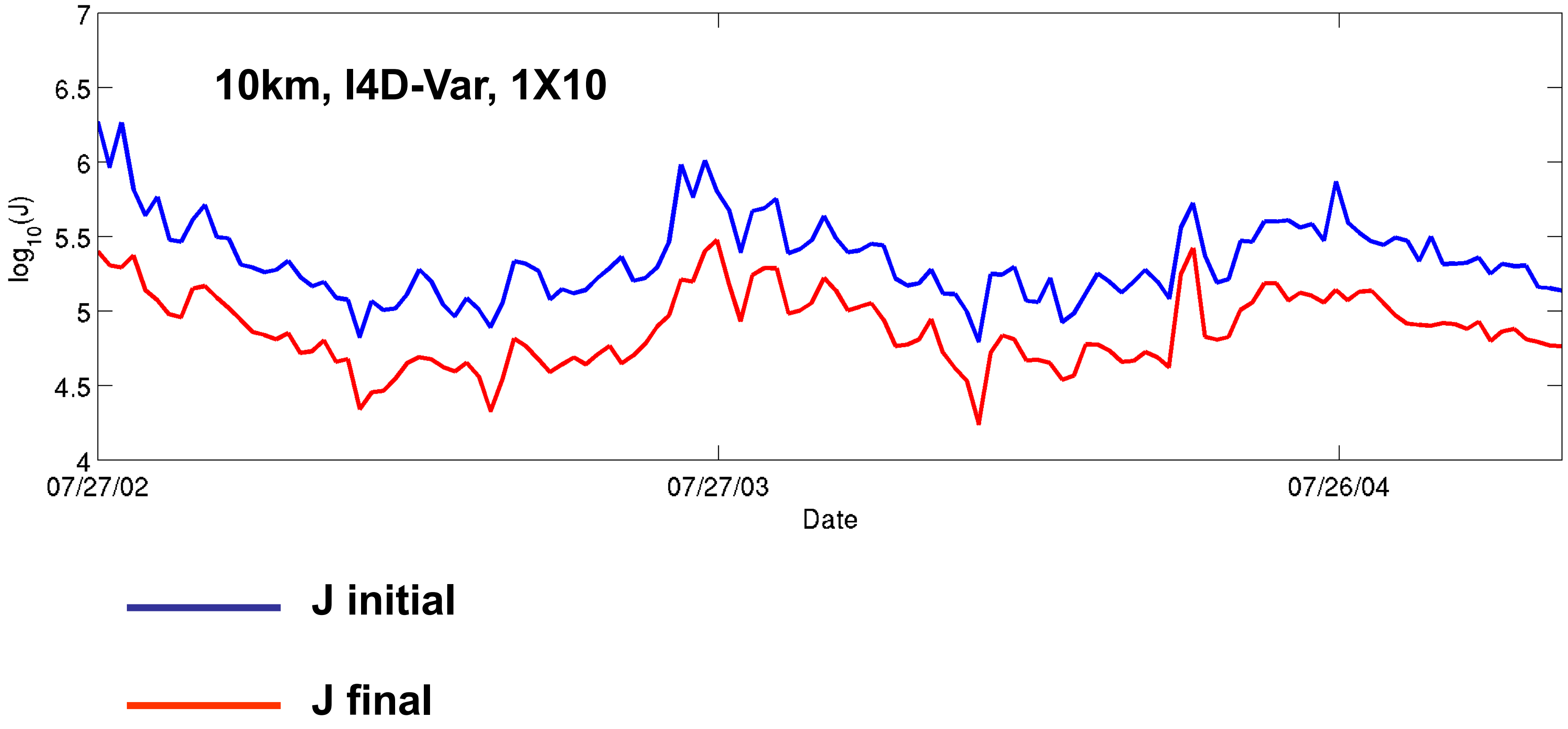
Therefore:

$$\Delta I \simeq \mathbf{d}^{\mathrm{T}}\mathbf{R}^{-1/2}\mathbf{V}_k\mathbf{T}_k^{-1}\mathbf{V}_k^{\mathrm{T}}\mathbf{R}^{-1/2}\mathbf{G}\mathbf{D}\mathbf{M}^{\mathrm{T}} * (\partial I/\partial \mathbf{x})$$

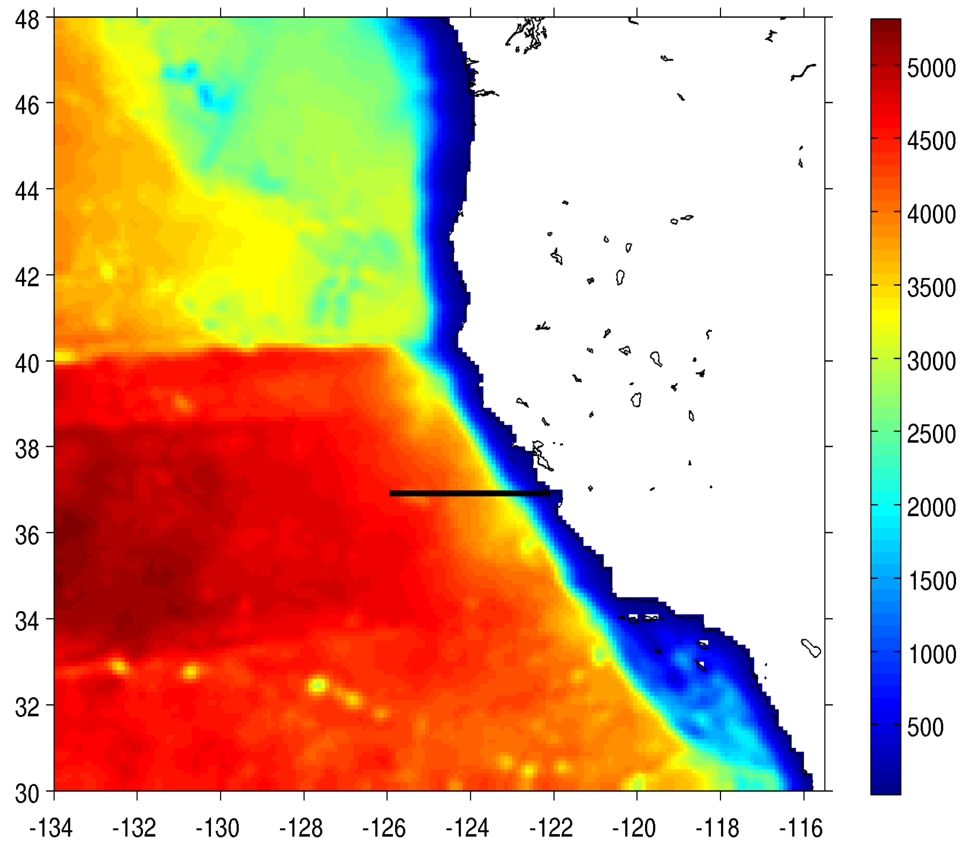
Sequential 4D-Var with 10km CCS ROMS



Sequential 4D-Var CCS ROMS



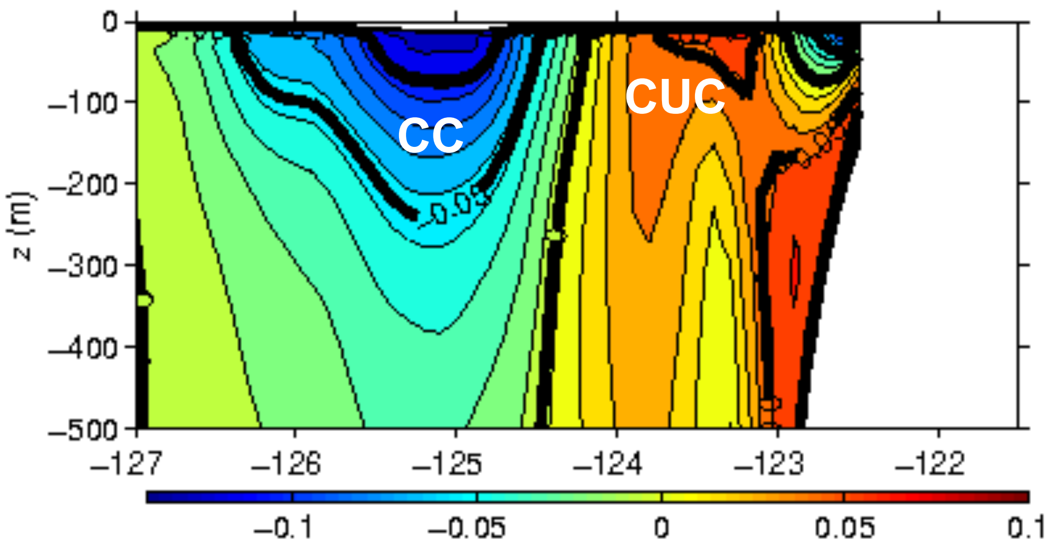
Example: 37N Transport



10km, CCS ROMS

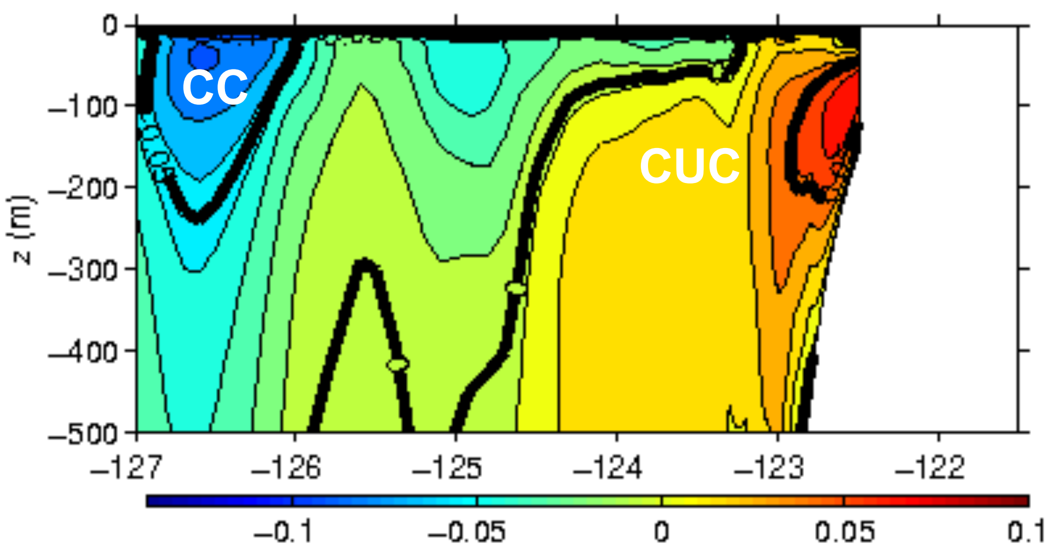
Example: 37N Transport

No assim



JAS time mean alongshore Flow (10km, 42 lev)

Primal Strong



CC = California Current

CUC = California Under Current

37N Transport Observation Impacts

The time average 37N transport can be written as:

$$I_{37N} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{h}^{\mathrm{T}}\mathbf{x}_i$$

where: $\mathbf{x}_i \equiv \mathbf{x}(i\Delta t) = \mathbf{x}(t)$

↑ **Model timestep**

therefore:

$$\Delta I_{37N} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{h}^{\mathrm{T}}\left((\mathbf{x_a})_i - (\mathbf{x_b})_i\right)$$

$\mathbf{M}_\mathbf{b}^{\mathrm{T}} * (\partial I / \partial \mathbf{x})$

$$\simeq \frac{1}{N}\sum_{i=1}^{N} \mathbf{h}^{\mathrm{T}}(\mathbf{M_b})_i \tilde{\mathbf{K}}\mathbf{d} = \mathbf{d}^{\mathrm{T}}\tilde{\mathbf{K}}^{\mathrm{T}} \boxed{\sum_{i=1}^{N} \frac{1}{N}(\mathbf{M_b})_i^{\mathrm{T}}\mathbf{h}}$$

where: $(\mathbf{M_b})_i \equiv \mathbf{M}(t_0 + i\Delta t, t_0) = \mathbf{M}(t, t_0)$

37N Transport Observation Impacts

37N time averaged transport increment:

$$\Delta I_{37N} \simeq \frac{1}{N} \mathbf{d}^{\mathrm{T}} \tilde{\mathbf{K}}^{\mathrm{T}} \underbrace{\sum_{i=1}^{N} (\mathbf{M_b})_i^{\mathrm{T}} \mathbf{h}}_{\text{ADROMS forced by h}}$$

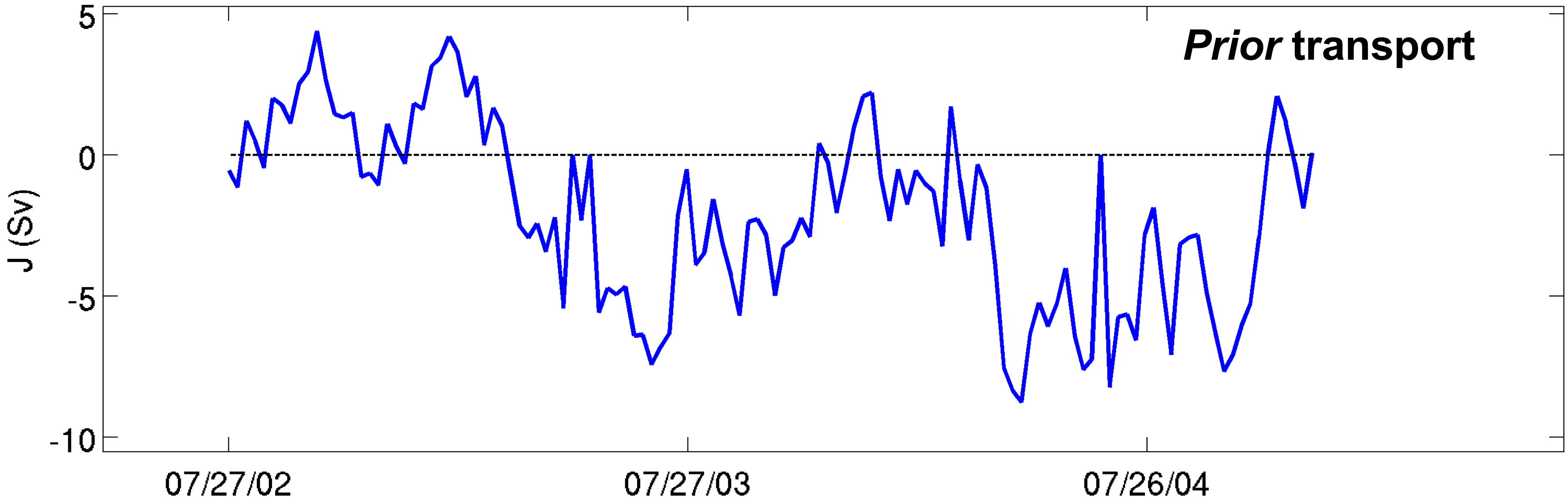
ADROMS forced by h

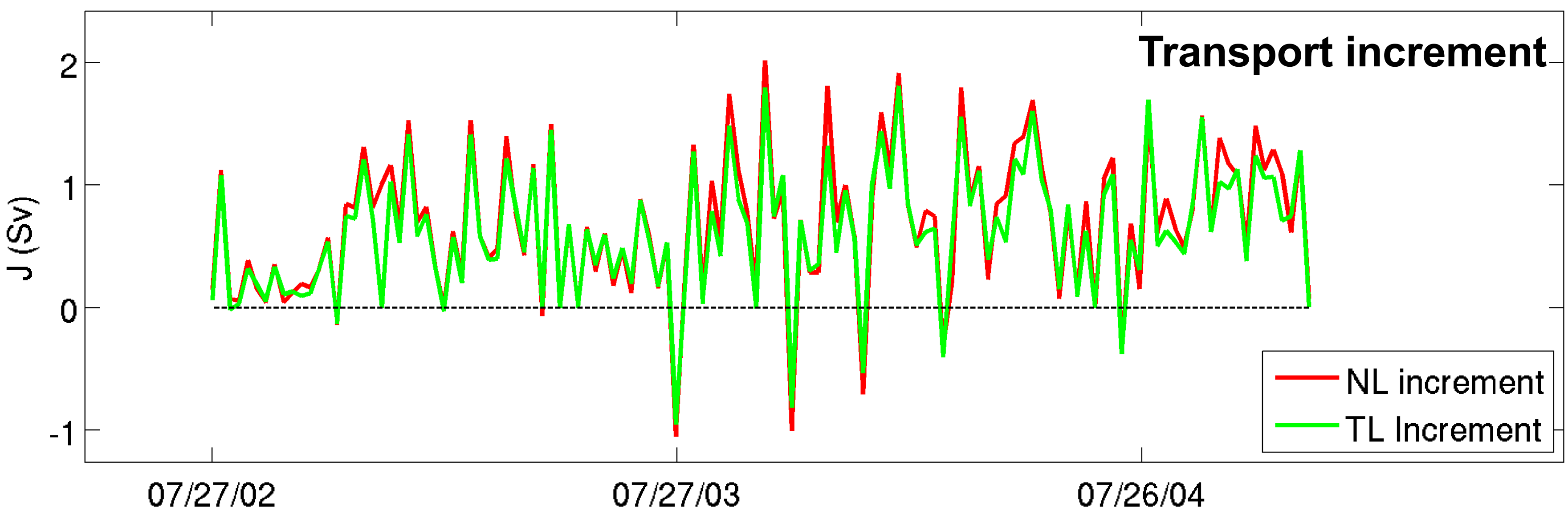
$$\tilde{\mathbf{K}}_k^{\mathrm{T}} = \mathbf{R}^{-1/2} \underbrace{\mathbf{V}_k \mathbf{T}_k^{-1} \mathbf{V}_k^{\mathrm{T}}}_{\text{Dual space Lanczos vectors}} \mathbf{R}^{-1/2} \mathbf{G} \mathbf{D}$$

Dual space
Lanczos vectors

TLROMS sampled at
observation points

37N Transport





Control Vector Impacts

37N time averaged transport increment:

$$\Delta I_{37N} \simeq \frac{1}{N}\mathbf{d}^{\mathrm{T}}\tilde{\mathbf{K}}^{\mathrm{T}}\sum_{i=1}^{N}(\mathbf{M_b})_i^{\mathrm{T}}\mathbf{h}$$

$$= \mathbf{d}^{\mathrm{T}}\mathbf{g} = \mathbf{d}^{\mathrm{T}}\left(\mathbf{g_x} + \mathbf{g_f} + \mathbf{g_b}\right)$$

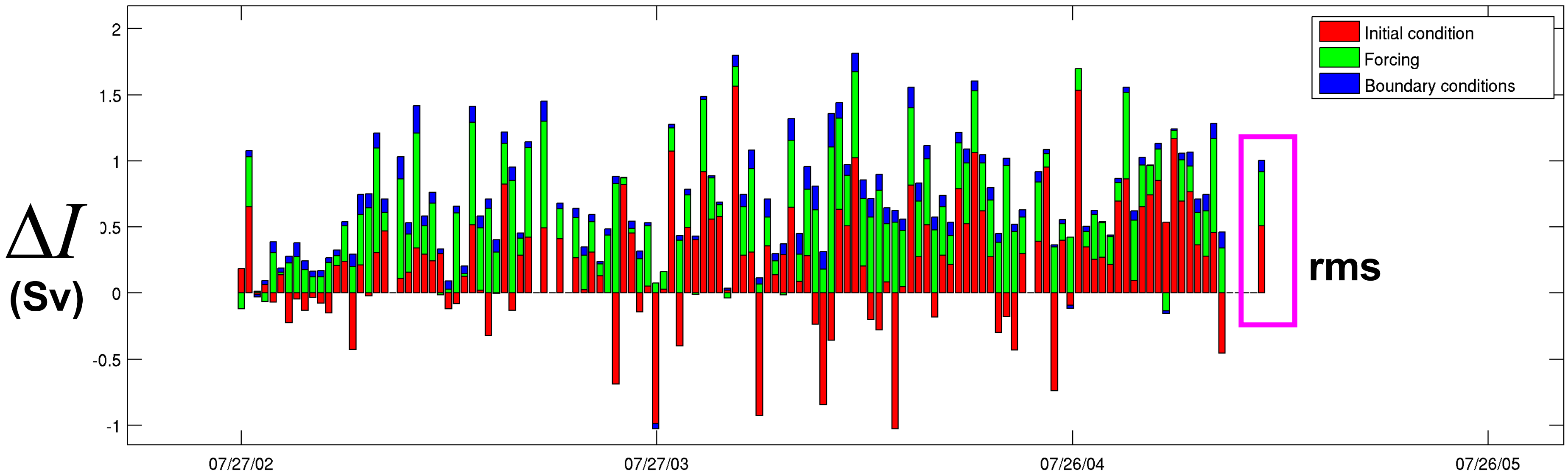
where: $$\mathbf{g} \simeq \frac{1}{N}\tilde{\mathbf{K}}\sum_{i=1}^{N}(\mathbf{M_b})_i^{\mathrm{T}}\mathbf{h}$$

$\mathbf{g_x}$ **- contribution from initial condition increments**

$\mathbf{g_f}$ **- contribution from surface forcing increments**

$\mathbf{g_b}$ **- contribution from open boundary increments**

37N Transport Control Vector Impacts



Observation Impacts

37N time averaged transport increment:

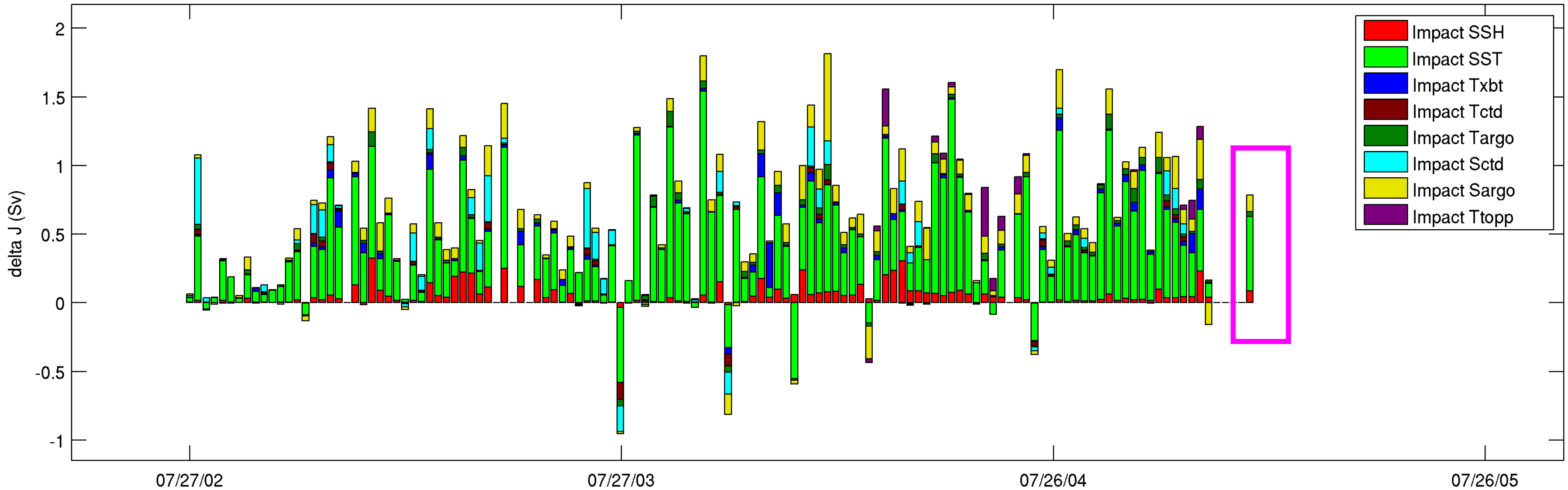
$$\Delta I_{37N} \simeq \frac{1}{N} \mathbf{d}^{\mathrm{T}} \tilde{\mathbf{K}}^{\mathrm{T}} \sum_{i=1}^{N} (\mathbf{M_b})_i^{\mathrm{T}} \mathbf{h}$$

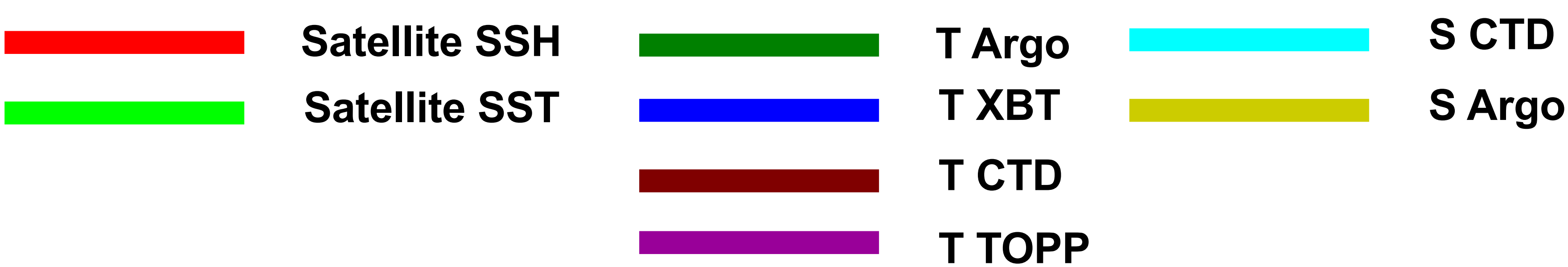
$$= \mathbf{d}^{\mathrm{T}} \mathbf{g} = \sum_{i=1}^{N_{obs}} d_i g_i$$

$$= \sum_{i=1}^{N_{obs}} \underbrace{\left(y_i - H_i \left(\mathbf{x_b}(t) \right) \right)}_{} g_i$$

Contribution of each observation to ΔI

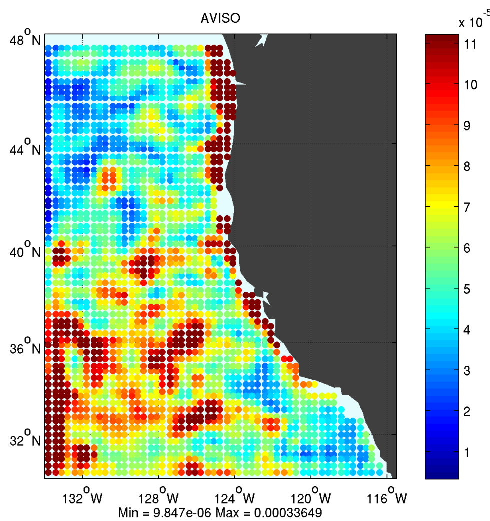
37N Transport Observation Impacts



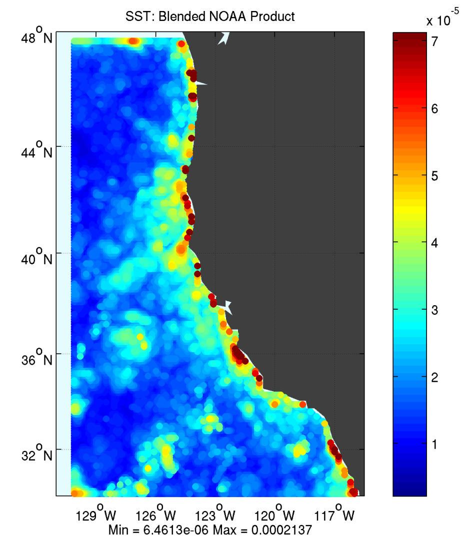


37N Transport Observation Impacts

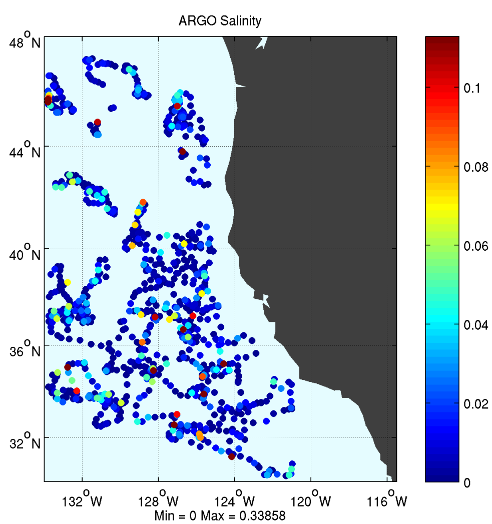
SSH



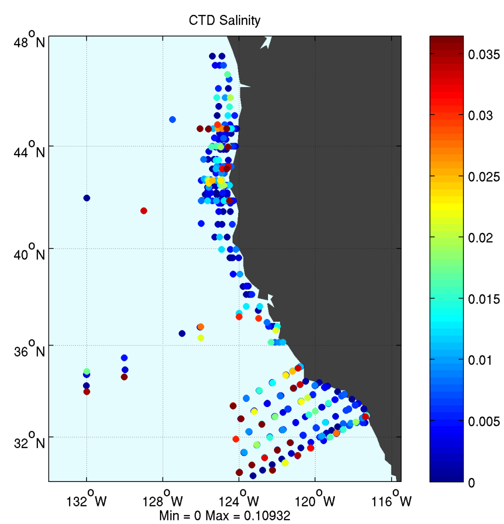
SST



Argo S



CTD S



Two Spaces: Obs Impact

$\mathbf{K}$ maps from observation (dual) space
to model (primal) space

$\mathbf{K}^T$ maps from model (primal) space
to observation (dual) space

Identifies the part of model space that controls 37N transport
<u>and</u> that is activated by the observations

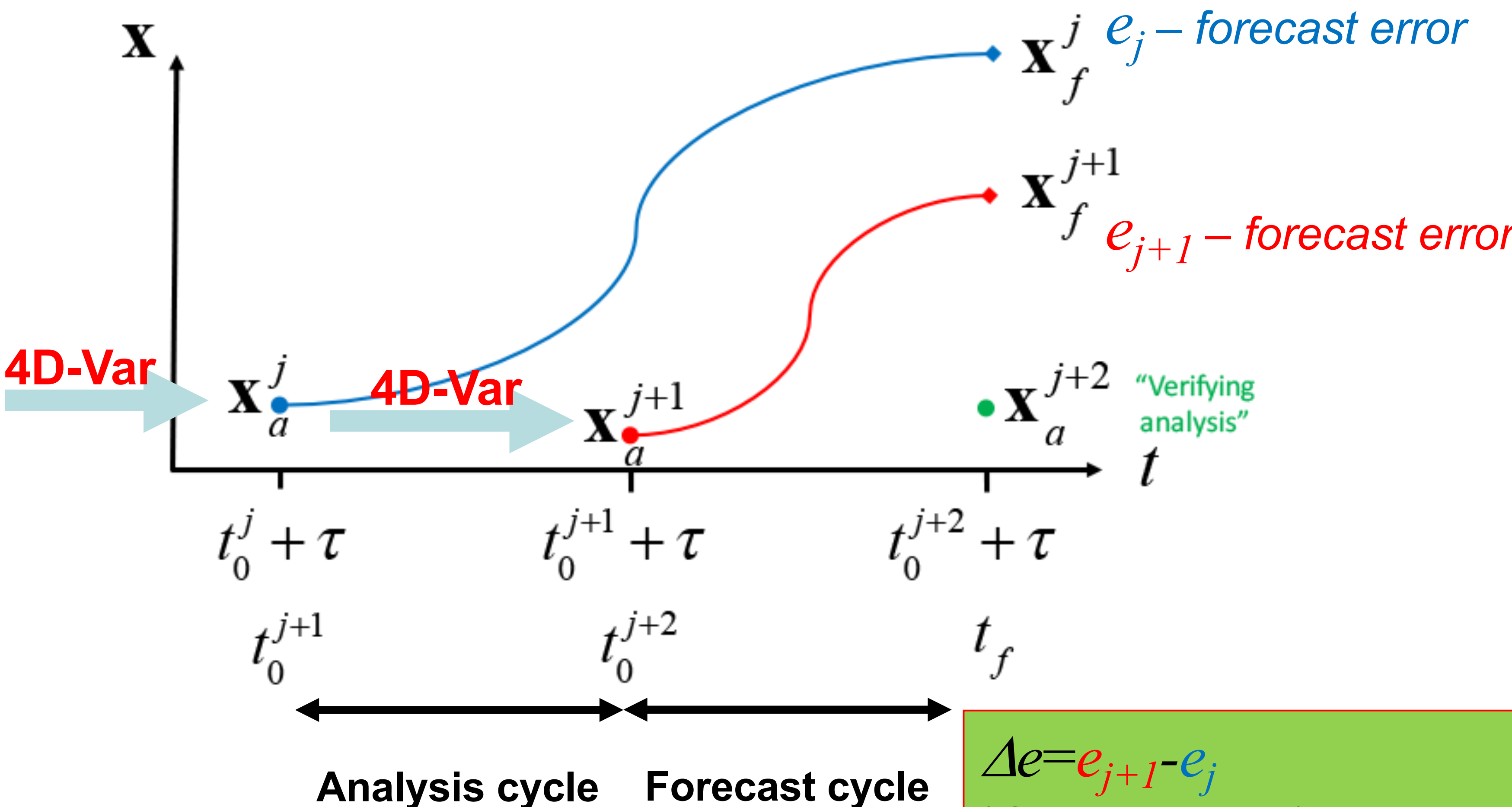
Analysis Cycle Observation Impacts: ROMS Implementation

- Primal (I4D-Var) and dual (4D-PSAS) forms available. Use 4D-Var cpp options plus:
  - define IS4DVAR_SENSITIVITY
    undef RECOMPUTE_4DVAR
    Drivers/obs_sen_is4dvar.h

  - define W4DPSAS_SENSITIVITY
    define OBS_IMPACT
    define OBS_IMPACT_SPLIT
    undef RECOMPUTE_4DVAR
    define SKIP_NLM
    Drivers/obs_sen_w4dpsas.h

Forecast Cycle Observation Impacts

A Typical Sequential Analysis-Forecast Procedure



$\Delta e = e_{j+1} - e_j$

if $\Delta e < 0$, DA improves fcst

if $\Delta e > 0$, DA degrades fcst

Forecast Cycle Observation Impacts

Choose a forecast error metric:

$$e = (\boldsymbol{x}_f - \boldsymbol{x}_t)^T \boldsymbol{C} (\boldsymbol{x}_f - \boldsymbol{x}_t)$$

One possibility of to use the verifying analysis as the truth.

To 3[rd]-order:

$$\Delta e_3 = \boldsymbol{d}^T \boldsymbol{K}^T \boldsymbol{M}_b^T \left[\boldsymbol{M}_j^T \boldsymbol{C} \left(\boldsymbol{x}_f^j - \boldsymbol{x}_a^{j+2} \right) + \boldsymbol{M}_{j+1}^T \boldsymbol{C} \left(\boldsymbol{x}_f^{j+1} - \boldsymbol{x}_a^{j+2} \right) \right]$$

$\boldsymbol{x}_a^{j+2}$ = verifying analysis

$\boldsymbol{M}_b^T$ = the adjoint model run backwards over the 4D-Var analysis cycle $[t_0^{j+1}, t_0^{j+1} + \tau]$ and linearized about 4D-Var background $\boldsymbol{x}_b$

$\boldsymbol{M}_{j+1}^T$ = the adjoint model linearized about the forecast solution $\boldsymbol{x}_f^{j+1}$

Forecast Cycle Observation Impacts

Another possibility of to use independent obs as the truth:

$$e = \left(\boldsymbol{y}_f - \boldsymbol{y}\right)^T \boldsymbol{C}\left(\boldsymbol{y}_f - \boldsymbol{y}\right)$$

To 3rd-order:

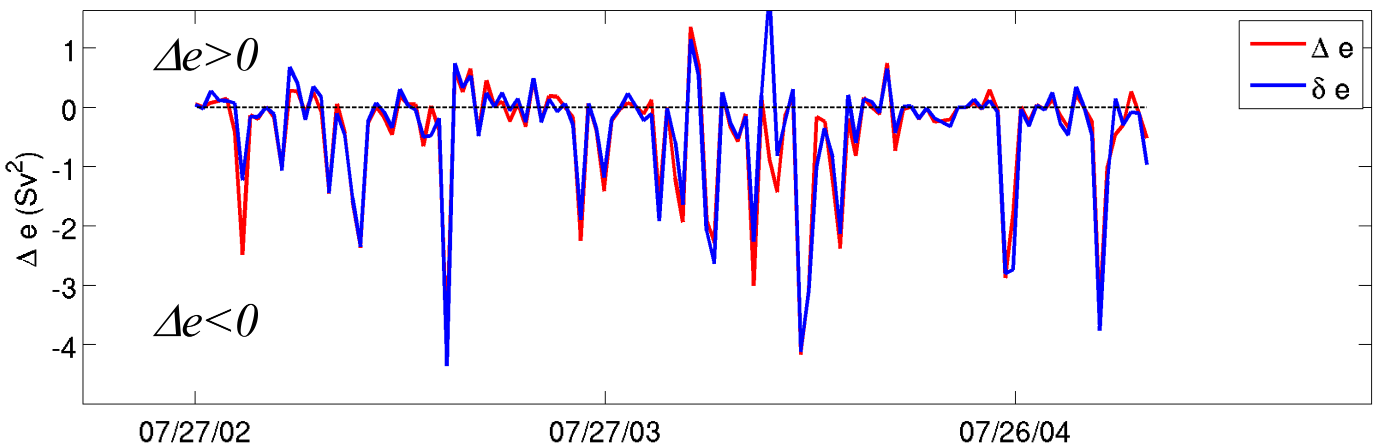
$$\Delta e_3 = \boldsymbol{d}^T \boldsymbol{K}^T \boldsymbol{M}_b^T \left[\boldsymbol{G}_j^T \boldsymbol{C}\left(\boldsymbol{y}_f^j - \boldsymbol{y}^{j+2}\right) + \boldsymbol{G}_{j+1}^T \boldsymbol{C}\left(\boldsymbol{y}_f^{j+1} - \boldsymbol{y}^{j+2}\right)\right]$$

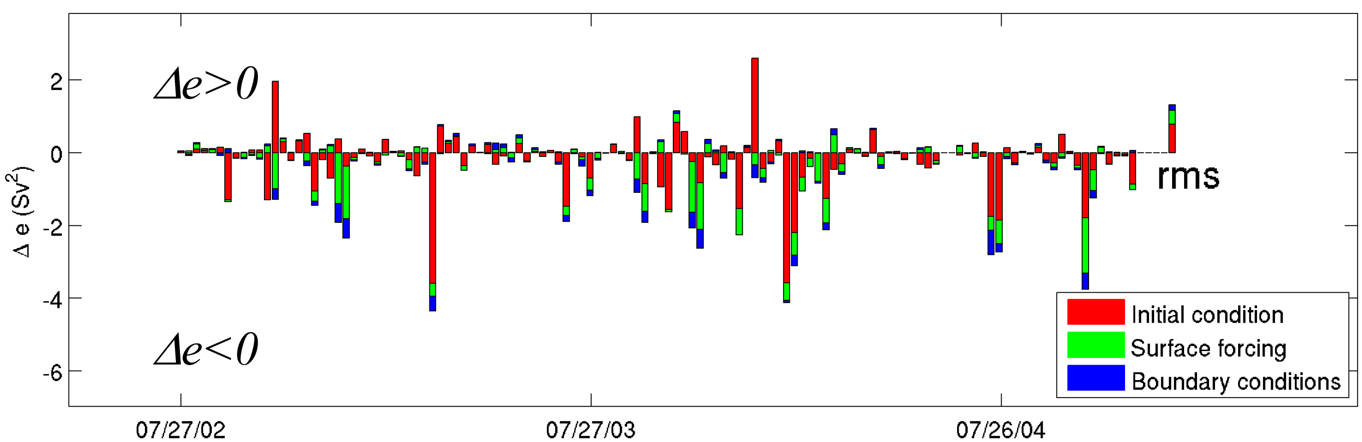
where $\boldsymbol{G}_j^T$ and $\boldsymbol{G}_{j+1}^T$ denote the adjoint model forced at the observation points and linearized about $\boldsymbol{x}_f^j$ and $\boldsymbol{x}_f^{j+1}$ respectively.

To use this option: define OBS_SPACE

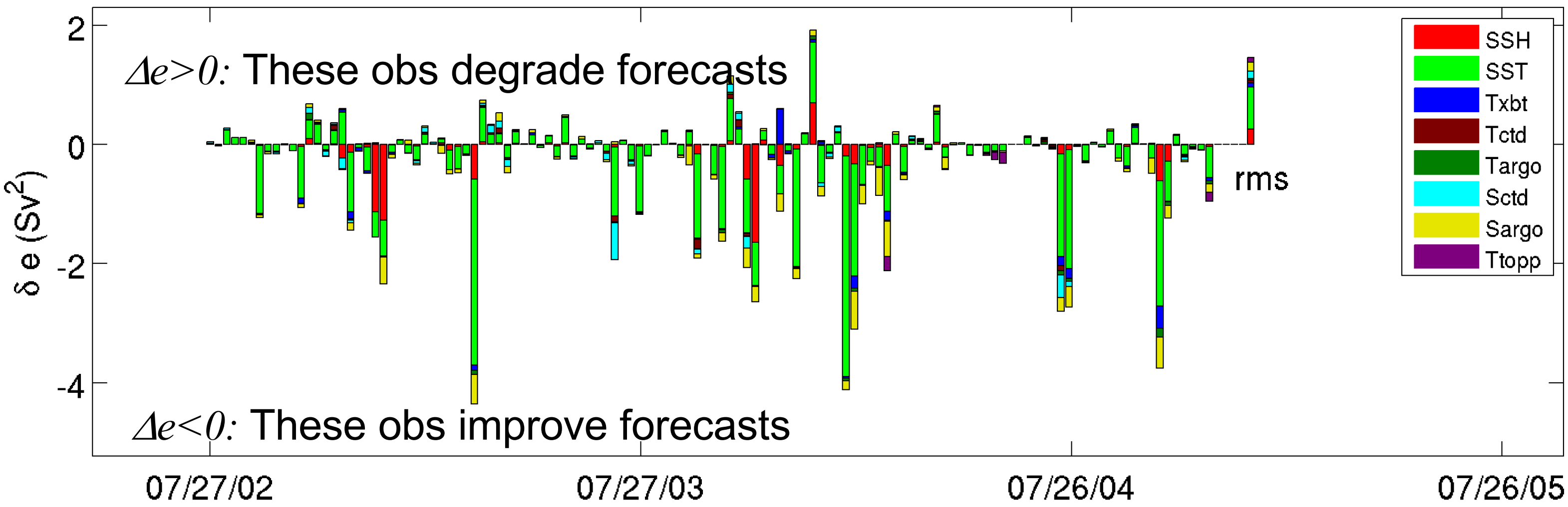
Forecast Cycle Observation Impacts

e **= mean squared error in 7-day forecast error in 37N transport**





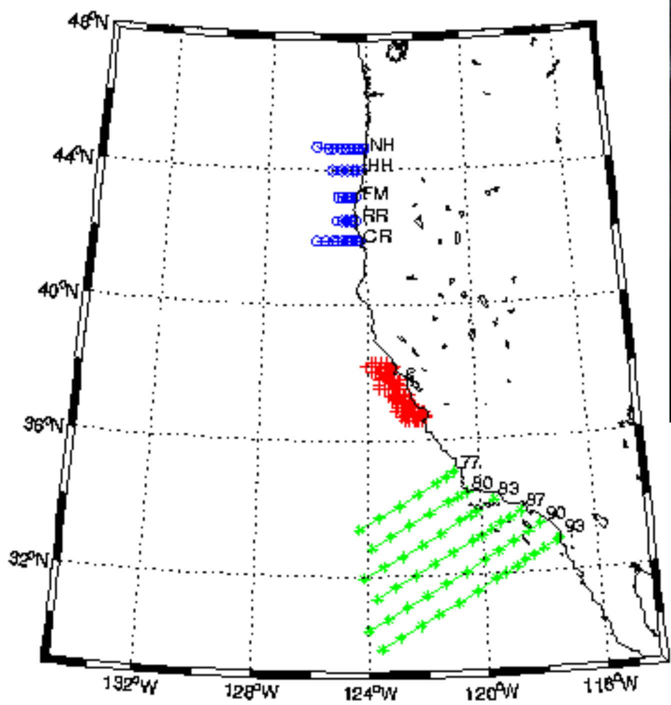
Forecast Cycle Observation Impacts

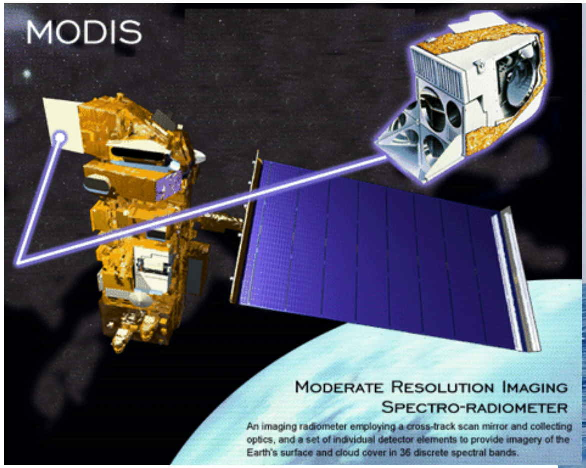


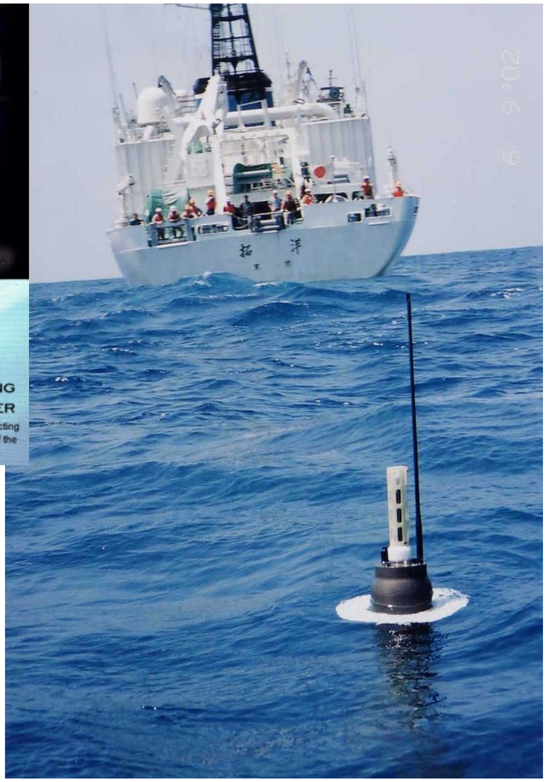
Forecast Cycle Observation Impacts: ROMS Implementation

- Only dual (4D-PSAS) form available:
  - define W4DPSAS_FCT_SENSITIVITY
    define OBS_IMPACT
    define OBS_IMPACT_SPLIT
    undef RECOMPUTE_4DVAR
    define SKIP_NLM
    define OBS_SPACE
      Drivers/obs_sen_w4dpsas_forecast.h

Adjoint 4D-Var & Observation Sensitivity







How will the circulation analysis change if some of the observations or the observation array change?

Adjoint 4D-Var & Observation Sensitivity

The analysis increments are a nonlinear function of the innovation vector d:

4D-Var

$$\mathbf{z_a} = \mathbf{z_b} + \boxed{K(\mathbf{d})}$$

where:

$$\mathbf{d} = \mathbf{y} - H\left(\mathbf{z_b}(t)\right)$$

Consider variations in the observation vector δy:

$$\delta\mathbf{d} = \delta\mathbf{y}; \quad \mathbf{z_a} + \delta\mathbf{z_a} = \mathbf{z_b} + K(\mathbf{d} + \delta\mathbf{d})$$

$$\approx \mathbf{z_b} + K(\mathbf{d}) + \left(\partial K / \partial \mathbf{y}\right)\delta\mathbf{y}$$

$$\delta\mathbf{z_a} \approx \boxed{\frac{\partial K}{\partial \mathbf{y}}}\delta\mathbf{y}$$

Tangent linearization of 4D-Var

Adjoint 4D-Var & Observation Sensitivity

Consider a scalar function of the *posterior* control vector $\mathbf{z}_a$:

$$I_{\mathbf{a}} = I(\mathbf{z}_{\mathbf{a}}) = I(\mathbf{z}_{\mathbf{b}} + K(\mathbf{d}))$$

A change $\delta \mathrm{y}$ in the observations yields a change in ΔI_{a} :

$$I_{\mathbf{a}} + \Delta I_{\mathbf{a}} = I(\mathbf{z}_{\mathbf{b}} + K(\mathbf{d} + \delta \mathbf{y}))$$

$$\approx I\left(\mathbf{z}_{\mathbf{b}} + K(\mathbf{d}) + (\partial K/\partial \mathbf{y})\,\delta \mathbf{y}\right)$$

$$\approx I(\mathbf{z}_{\mathbf{a}}) + \left((\partial K/\partial \mathbf{y})\,\delta \mathbf{y}\right)^{\mathrm{T}} \left(\partial I/\partial \mathbf{z}\right)$$

Therefore:

$$\Delta I_{\mathbf{a}} \approx \delta \mathbf{y}^{\mathrm{T}} \left(\partial K/\partial \mathbf{y}\right)^{\mathrm{T}} \left(\partial I/\partial \mathbf{z}\right)$$

Adjoint 4D-Var & Observation Sensitivity

$$\Delta I_{\mathbf{a}} \approx \delta \mathbf{y}^{\mathrm{T}} \boxed{\left(\partial K / \partial \mathbf{y}\right)^{\mathrm{T}}} \left(\partial I / \partial \mathbf{z}\right)$$

Adjoint of 4D-Var

Observation System Experiments (OSEs)

Suppose that during a particular assimilation cycle the satellite altimeter goes offline.

How would this have impacted the analysis?

We could run 4D-Var again with SSH obs removed.

Or let $\delta y_i = -d_i$ **for all SSH obs.**

The change in the analysis is: $\delta \mathbf{z_a} \approx (\partial K / \partial \mathbf{y}) \delta \mathbf{y}$

The change in $\Delta I_{\mathbf{a}}$ **is:** $\Delta I_{\mathbf{a}} \approx \delta \mathbf{y}^{\mathrm{T}} (\partial K / \partial \mathbf{y})^{\mathrm{T}} (\partial I / \partial \mathbf{z})$

Observation System Experiments (OSEs)

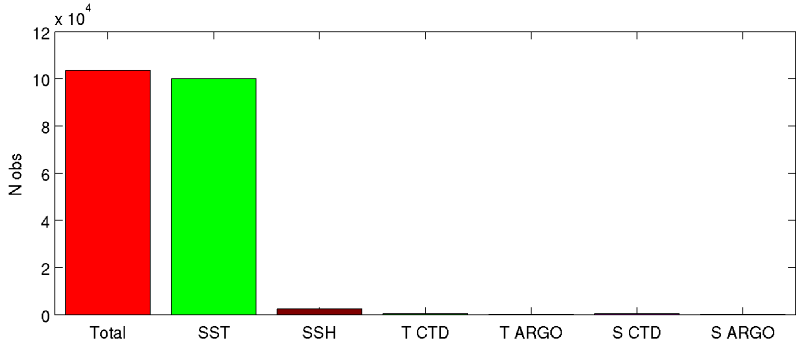
The cost of $(4D\text{-}Var)^T$ = cost of 4D-Var

But ONLY one run of $(4D\text{-}Var)^T$ is needed for ALL OSEs.

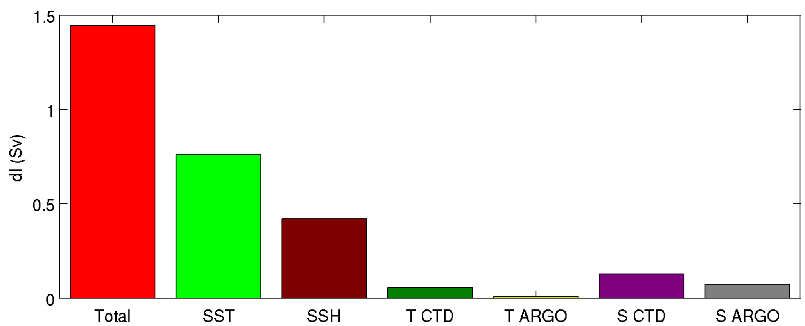
Example: 37N transport

(10km, CCS ROMS)

N_{obs}

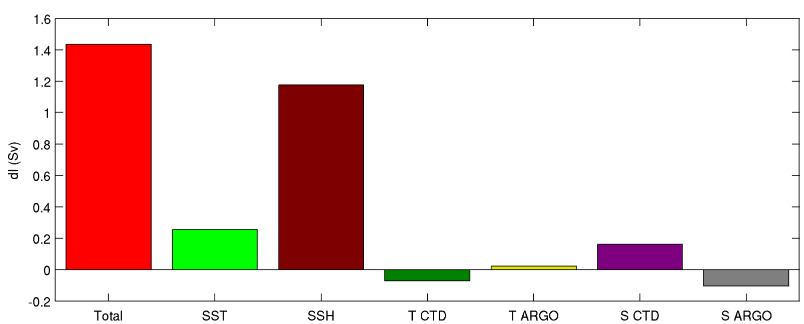


Obs Impact



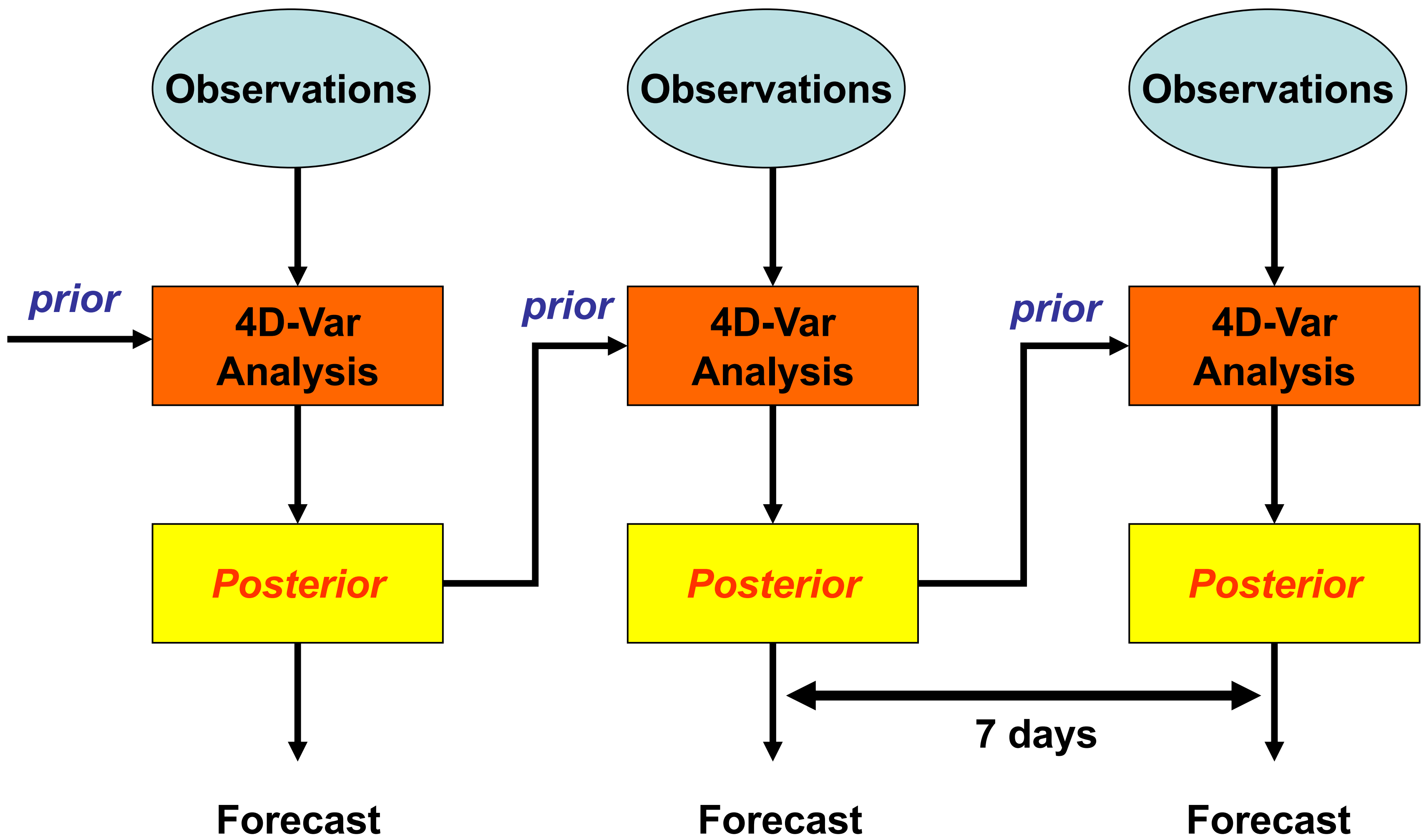
$$\Delta I = \mathbf{d}^{\mathrm{T}} \tilde{\mathbf{K}}^{\mathrm{T}} \sum_{i=1}^{N} \frac{1}{N} (\mathbf{M_b})_i^{\mathrm{T}} \mathbf{h}_i$$

Obs Sens

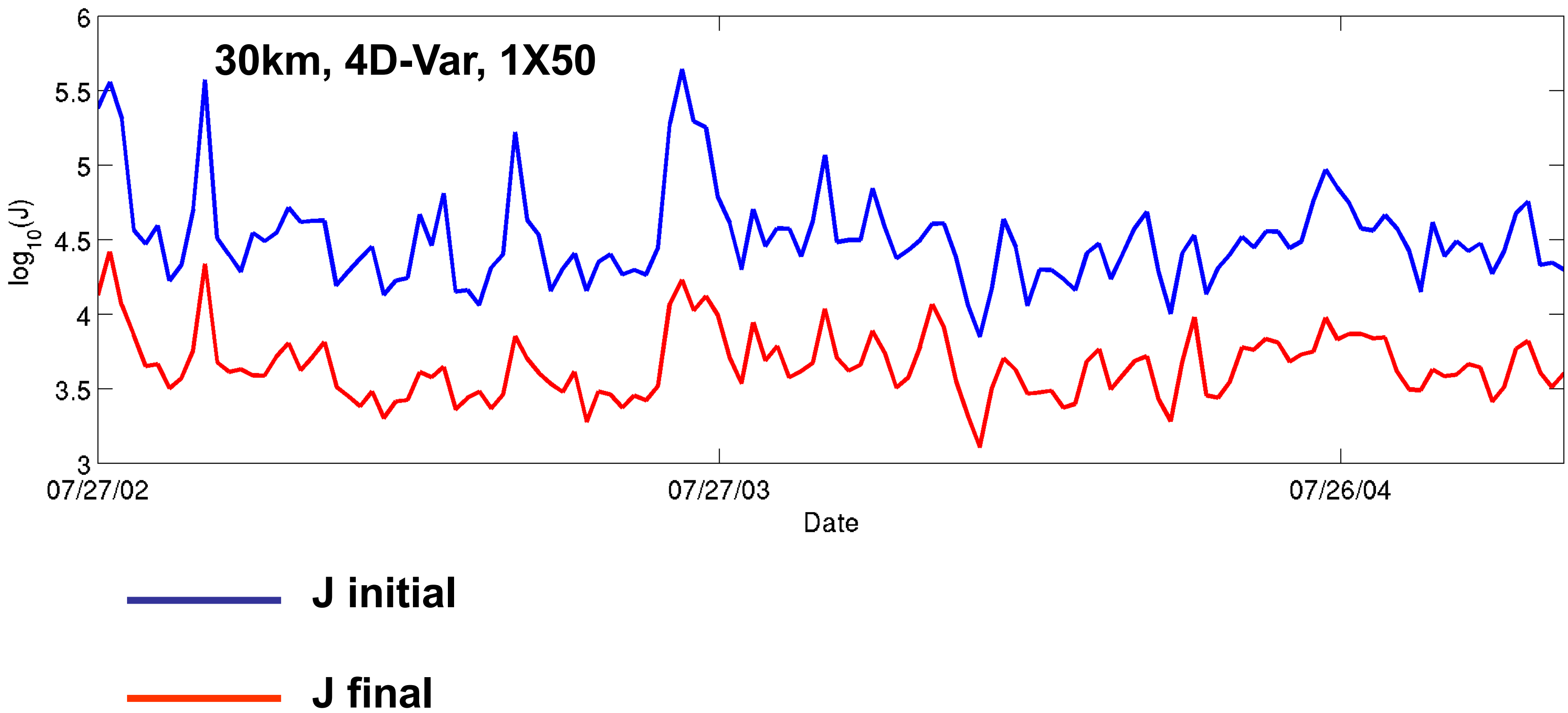


$$\Delta I = \mathbf{d}^{\mathrm{T}} \left(\frac{\partial K}{\partial \mathbf{y}} \right)^{\mathrm{T}} \sum_{i=1}^{N} \frac{1}{N} (\mathbf{M_b})_i^{\mathrm{T}} \mathbf{h}_i$$

Sequential 4D-Var with 30km CCS ROMS

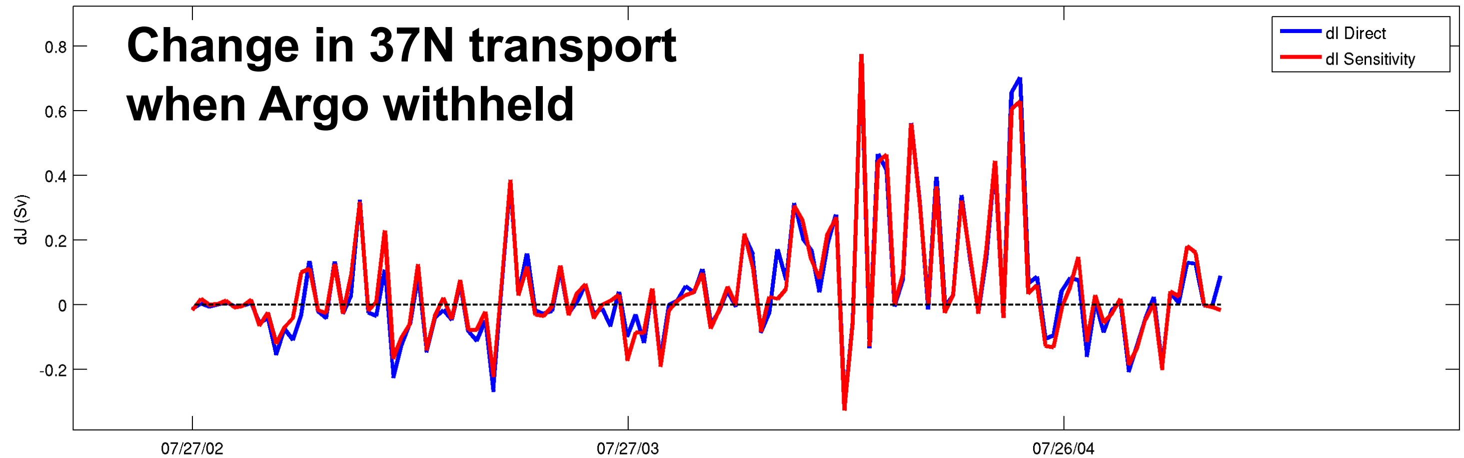


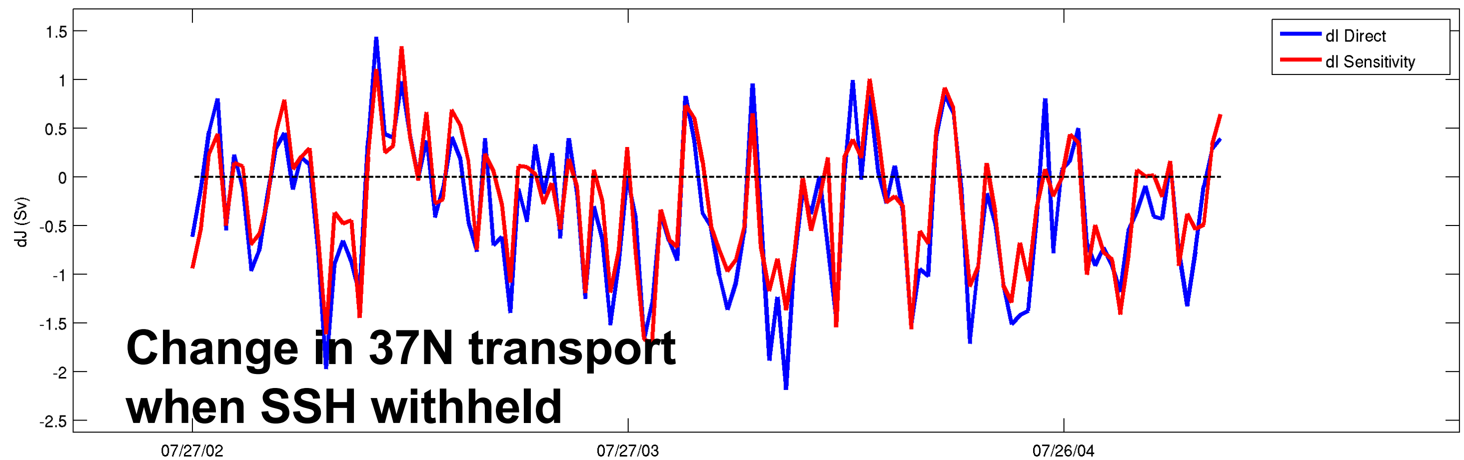
Sequential 4D-Var CCS ROMS



Observing System Experiments (OSEs)

(30km, CCS ROMS)





Two Spaces: Obs Sensitivity

$\partial K / \partial \mathbf{y}$ maps from observation (dual) space to model (primal) space

$\left(\partial K / \partial \mathbf{y} \right)^{\mathrm{T}}$ maps from model (primal) space to observation (dual) space

Identifies the part of model space that controls 37N transport
<u>and</u> that is activated by the observations during 4D-Var

Observation Sensitivity: ROMS Implementation

• Dual (4D-PSAS) form available. Use same 4D-Var cpp options plus:

Analysis cycle-

- define W4DPSAS_SENSITIVITY
  (undef RECOMPUTE_4DVAR)
  Drivers/obs_sen_w4dpsas.h

Forecast cycle-

- define W4DPSAS_FSCT_SENSITIVITY
  (undef RECOMPUTE_4DVAR)
  Drivers/obs_sen_w4dpsas_forecast.h

Summary

- Observation impact is based on $\tilde{\mathbf{K}}^{\mathrm{T}}$ and yields the actual contribution of each obs to the circulation increments.
- Observation sensitivity is based on $(4\text{D-Var})^{\mathrm{T}}$ and yields the change in circulation due to changes in obs (or array)
  - useful for efficient generation of OSEs.
- Both obs impact and obs sensitivity can be applied in during analysis and forecast cycle.
- $(4\text{D-Var})^{\mathrm{T}}$ is a *very* powerful operator.

References

- Langland, R.H. and N.L. Baker, 2004: Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus*, **56A**, 189-201.
- Gelaro, R. and Y. Zhu, 2009: Examination of observation impacts derives from Observing System Experiments (OSEs) and adjoint models. *Tellus*, **61A**, 179-193.
- Moore, A.M., H.G. Arango, G. Broquet, B.S. Powell, J. Zavala-Garay and A.T. Weaver, 2011a: The Regional Ocean Modeling System (ROMS) 4-dimensional data assimilation systems: Part I – System overview. *Ocean Modelling*, 91, 34-49.
- Moore, A.M., H.G. Arango, G. Broquet, C.. Edwards, M. Veneziani, B.S. Powell, D. Foley, J. Doyle, D. Costa and P. Robinson, 2011b: The Regional Ocean Modeling System (ROMS) 4-dimensional data assimilation systems: Part II – Performance and application to the California Current System. *Ocean Modelling*, 91, 50-73.

References

- Moore, A.M., H.G. Arango, G. Broquet, C.. Edwards, M. Veneziani, B.S. Powell, D. Foley, J. Doyle, D. Costa and P. Robinson, 2011c: The Regional Ocean Modeling System (ROMS) 4-dimensional data assimilation systems: Part III – Observation impact and observation sensitivity the California Current System. *Ocean Modelling*, 91, 74-94.
- Trémolet, Y., 2008: Computation of observation sensitivity and observation impact in incremental variational data assimilation. *Tellus*, **60A**, 964-978.