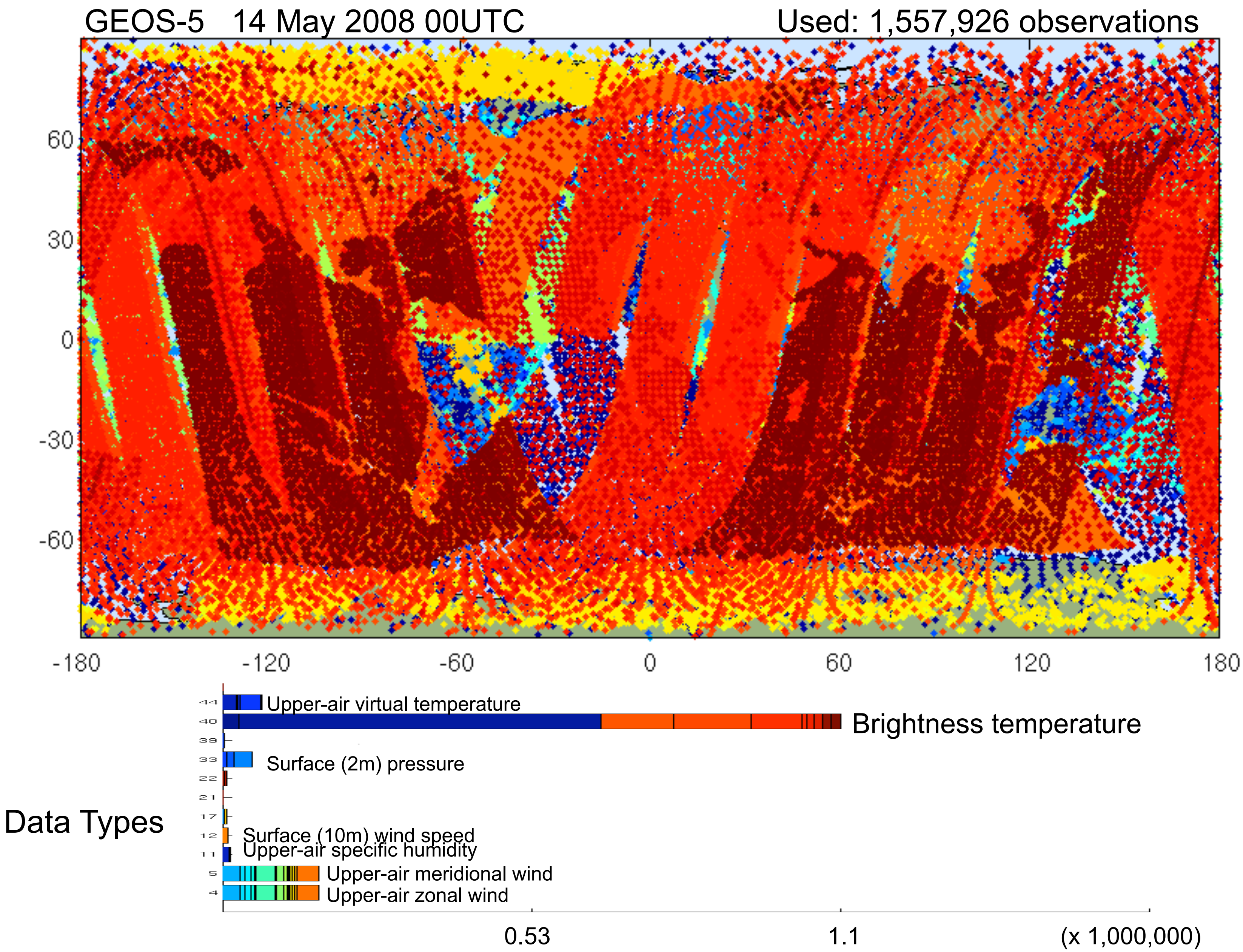
Estimating observation impact using the adjoint of a data assimilation system

Ron Gelaro
NASA Global Modeling and Assimilation Office

With thanks to R. Todling (GMAO), R. Langland (NRL) and
Y. Trémolet (ECMWF)

JCSDA Summer Colloquium on Data Assimilation
Stevenson WA, 14 June 2009

The observing system…a 6-hour snapshot



Background / Outline for this Talk

• Adjoint-based estimates of observation impact have become increasingly popular as an alternative/complement to traditional observing system experiments (OSEs)

✓ Used at several centers for experimentation or routine monitoring

✓ Inter-comparison project between centers in progress

• For linear analysis problems, observation impact is closely related to (is an extension of) observation sensitivity

…see Baker and Daley (2000)

• This talk touches on:

✓ Methodology, illustrative results for two systems

✓ Need for, implications of >1st order estimates of impact

✓ Extension to nonlinear analysis problems

✓ Comparison, complementarity with OSEs

The Data Assimilation System

- Consider a forecast model: $\mathbf{x}^f = \mathbf{m}(\mathbf{x}_0)$

 and atmospheric analysis: $\mathbf{x}_\mathrm{a} = \mathbf{x}_\mathrm{b} + \mathbf{K}[\mathbf{y} - H(\mathbf{x})]$

 where $\mathbf{x}_\mathrm{b}$ is a background state, $\mathbf{y}$ are observations, H is a (possibly nonlinear) observation operator and $\mathbf{K}$ determines the weight, or gain, given to each observation

 …the difference $\delta\mathbf{y} = \mathbf{y} - H(\mathbf{x})$ is the innovation vector, (~10^6)

- Assume, for now, that H is either <u>linear</u> or only a function of $\mathbf{x}_\mathrm{b}$, and define the analysis increment:

$$\delta\mathbf{x}_0 = \mathbf{x}_\mathrm{a} - \mathbf{x}_\mathrm{b} = \mathbf{K}\delta\mathbf{y} \qquad (1)$$

 Note that (1) may be viewed as a transformation between a perturbation $\delta\mathbf{x}_0$ in *state space* and a perturbation $\delta\mathbf{y}$ in *observation space*

Observation Sensitivity

Baker and Daley (2000) showed that the sensitivity of the analysis to observations could be computed using the adjoint of the DAS

$$\partial \mathbf{x}_{\mathrm{a}} \big/ \partial \mathbf{y} = \mathbf{K}^{\mathrm{T}}$$

- The sensitivity of a measure J with respect to the initial conditions (analysis) is then extended into observation space as

$$\frac{\partial J}{\partial \mathbf{y}} = \frac{\partial \mathbf{x}_{\mathrm{a}}}{\partial \mathbf{y}} \frac{\partial J}{\partial \mathbf{x}_{\mathrm{a}}} = \mathbf{K}^{\mathrm{T}} \frac{\partial J}{\partial \mathbf{x}_{\mathrm{a}}}$$

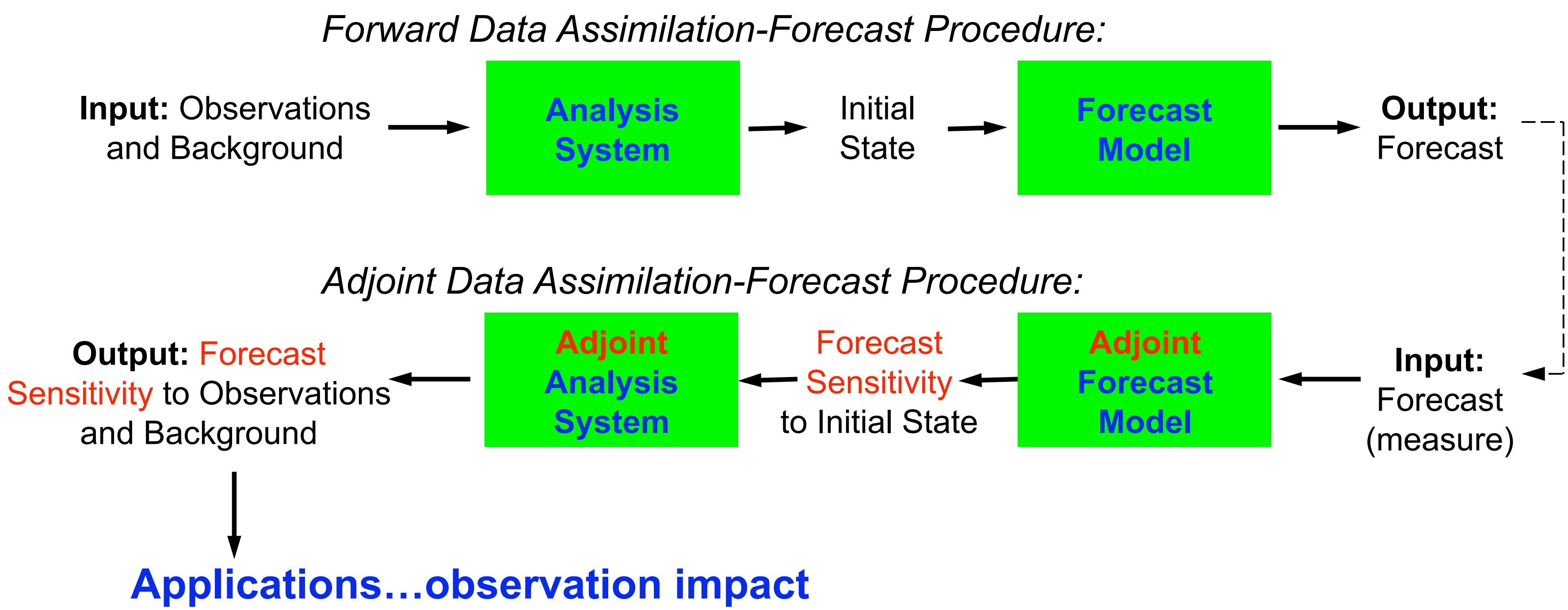
observation space state space

- If J is based on a model forecast, then the sensitivity of J with respect to the observations is

$$\frac{\partial J}{\partial \mathbf{y}} = \mathbf{K}^{\mathrm{T}} \mathbf{M}^{\mathrm{T}} \frac{\partial J}{\partial \mathbf{x}^{f}}$$

where $\mathbf{M}^{\mathrm{T}}$ is the adjoint of $\mathbf{m}$

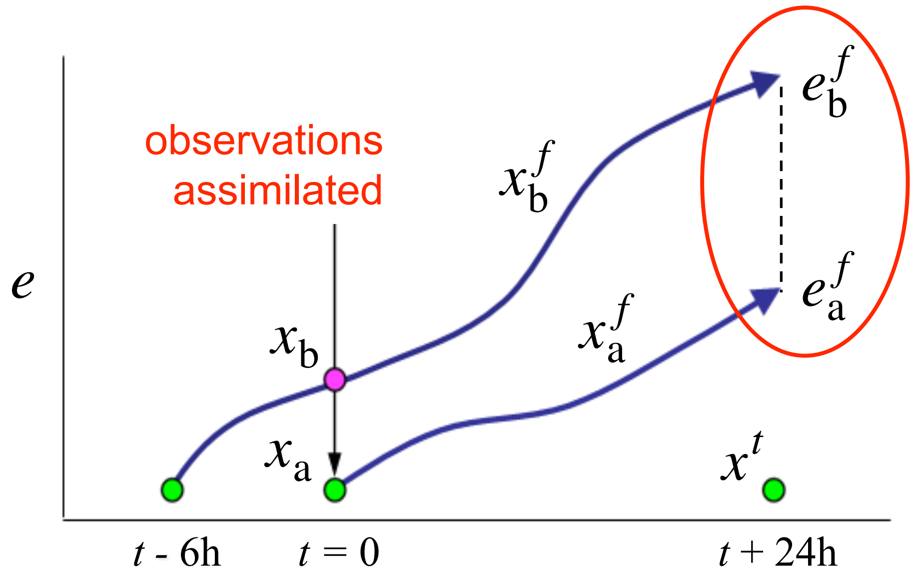
Adjoint Data Assimilation System



The computational cost of the adjoint system is roughly the same as that of the regular (forward) data assimilation system

Estimating the Impact of Observations on Forecasts

Langland and Baker (2004) showed that the adjoint of a data assimilation system could be used effectively to measure the impact of observations on forecast skill



- Consider forecasts from an analysis x_a and background state x_b , and energy-based measure of forecast error $e = (\mathbf{x}^f - \mathbf{x}^t)^\mathrm{T}\mathbf{C}(\mathbf{x}^f - \mathbf{x}^t)$ where x^t is a verification analysis state

- The difference $\delta e = e_a^f - e_b^f$ measures the combined impact of all obs assimilated at $t = 0$…

…it can be estimated as a sum of contributions from <u>individual</u> obs using information from the model and analysis adjoints

LB04 Observation Impact Estimate

$$\delta e \approx (\delta \mathbf{y})^{\mathrm{T}} \mathbf{K}^{\mathrm{T}} [\mathbf{M}_{\mathrm{b}}^{\mathrm{T}} \mathbf{C}(\mathbf{x}_{\mathrm{b}}^{f} - \mathbf{x}^{t}) + \mathbf{M}_{\mathrm{a}}^{\mathrm{T}} \mathbf{C}(\mathbf{x}_{\mathrm{a}}^{f} - \mathbf{x}^{t})]$$

analysis adjoint ($\mathbf{K}^{\mathrm{T}}$)   model adjoint ($\mathbf{M}_{\mathrm{b}}^{\mathrm{T}}$, $\mathbf{M}_{\mathrm{a}}^{\mathrm{T}}$)

- The impact of arbitrary subsets of observations can be estimated by summing only terms involving the desired elements of $\delta \mathbf{y}$

- The vector $\mathbf{K}^{\mathrm{T}}[\ldots]$ is computed only once and involves the **entire set** of observations

  ✓*The impacts of all observations are estimated* ***simultaneously***

  …contrasts with classical observing system (data denial) exps

- Application is subject to assumptions and simplifications in $\mathbf{M}^{\mathrm{T}}$

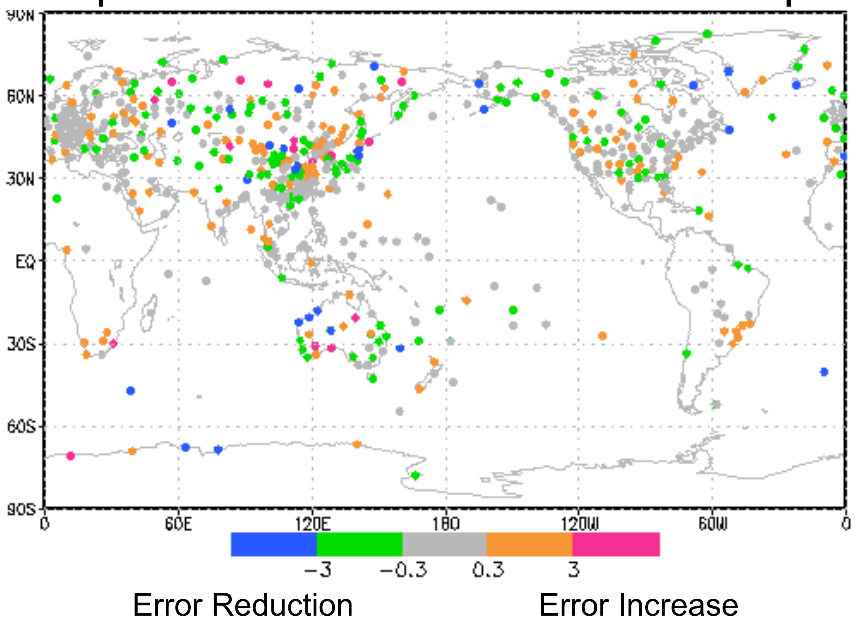
$\delta e < 0$ …the observation **improves** the forecast

$\delta e > 0$ …the observation **degrades** the forecast

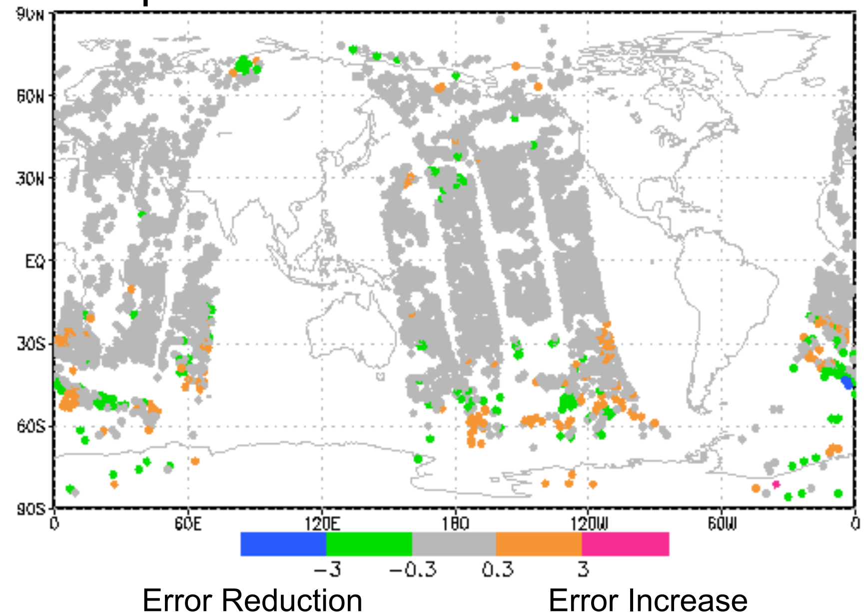
Observation Impact on GEOS-5 24h Forecast for a Single Case

(Global forecast error norm, 00UTC 10 July 2005)

Impact of 500mb **Radiosonde** Temps



Impact of **AIRS** Ch.221 Radiances



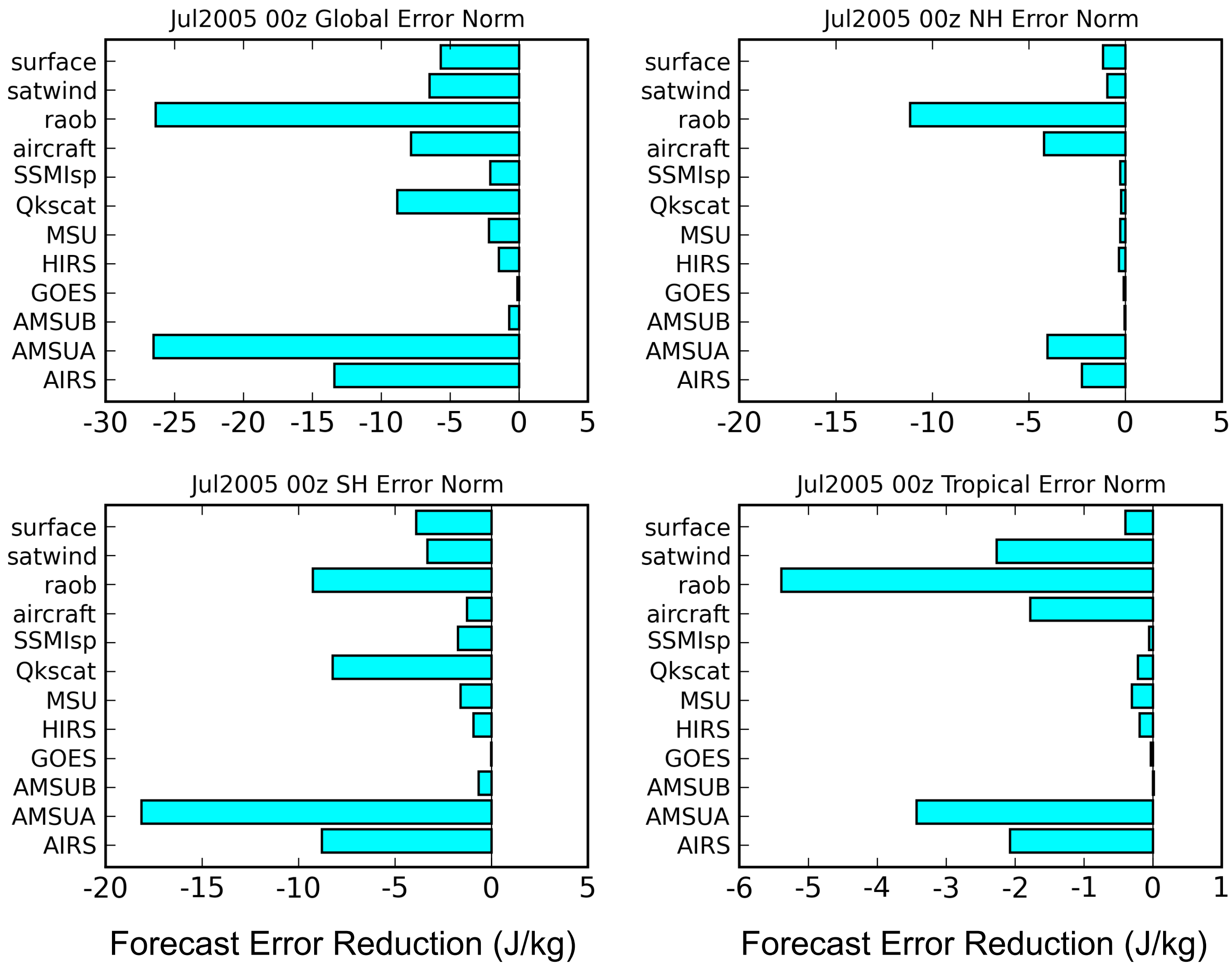
- Observations that **improved** the 24h forecast
- Observations that **degraded** the 24h forecast
- Observations that had small impact on 24h forecast

Note that there is a mixture of positive and negative impacts

…even for ‘good’ observations.

Impacts of Major Observing Systems in GEOS-5

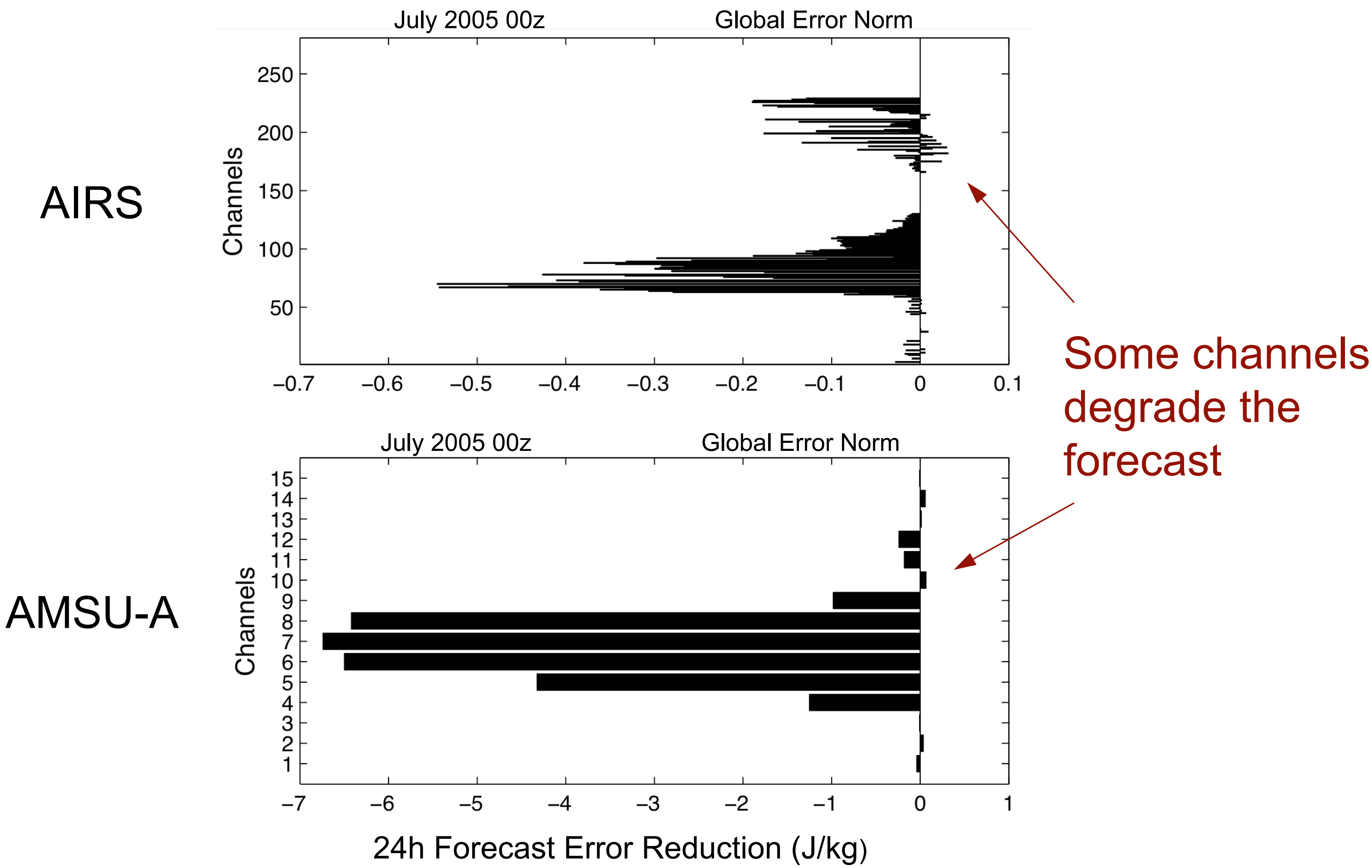
24h Error Norm (Globe, NH, SH, Tropics) – July 2005 00UTC Totals



Gelaro and Zhu (2009)

Assessing Impacts of Hyper-Spectral Instruments

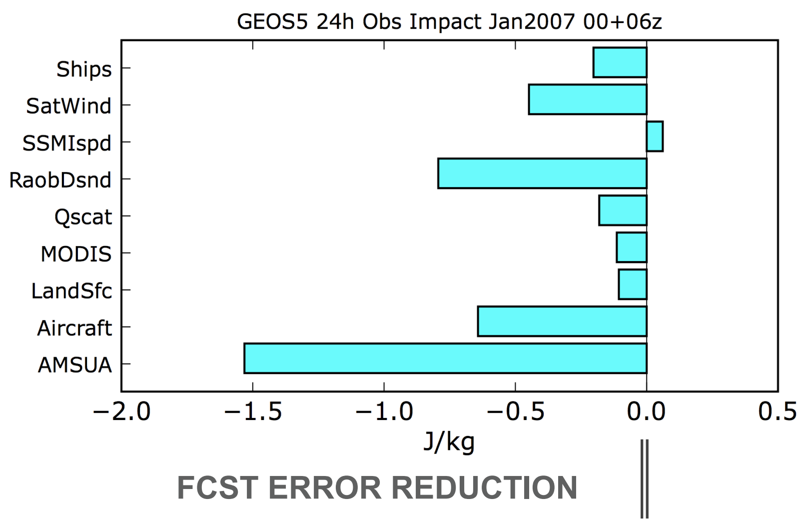
GEOS-5 Adjoint Data Assimilation System



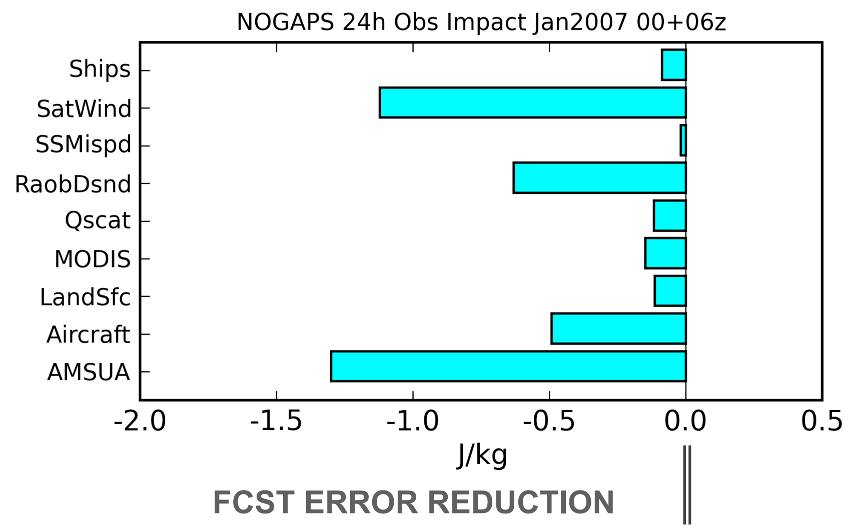
Comparison of Observation Impacts in Two Systems

24h Global Error Norm - Baseline Obs - Jan 2007 00+06 UTC

NASA GEOS-5



Navy NOGAPS



Overall impacts similar in NASA and Navy systems despite differences in algorithms, RT models, observation counts…

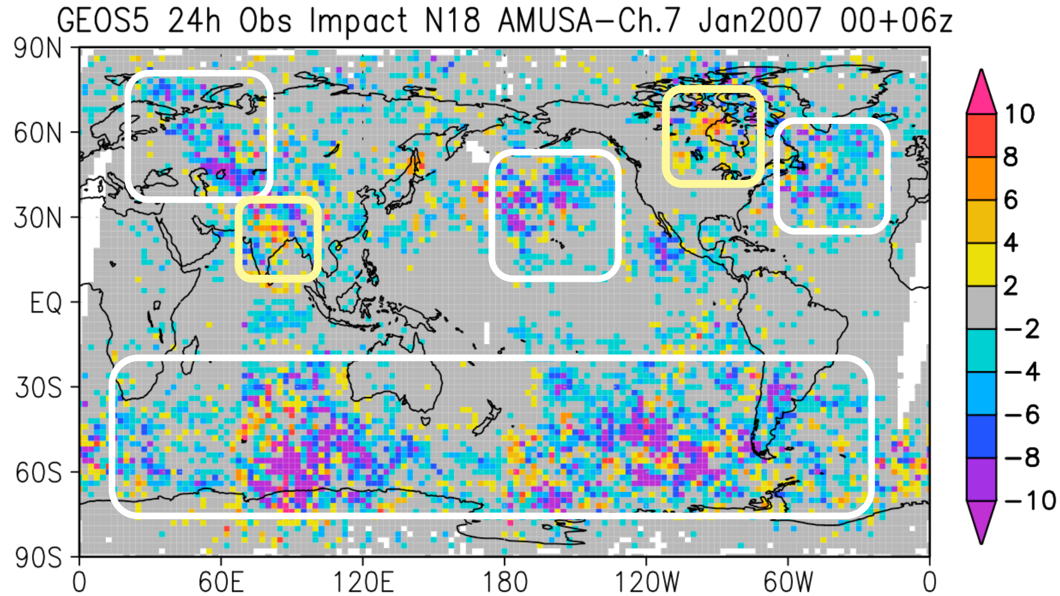
…notable differences in Satwinds, SSMI speeds

Impact of NOAA-18 AMSU-A Ch. 7 (binned by obs location)

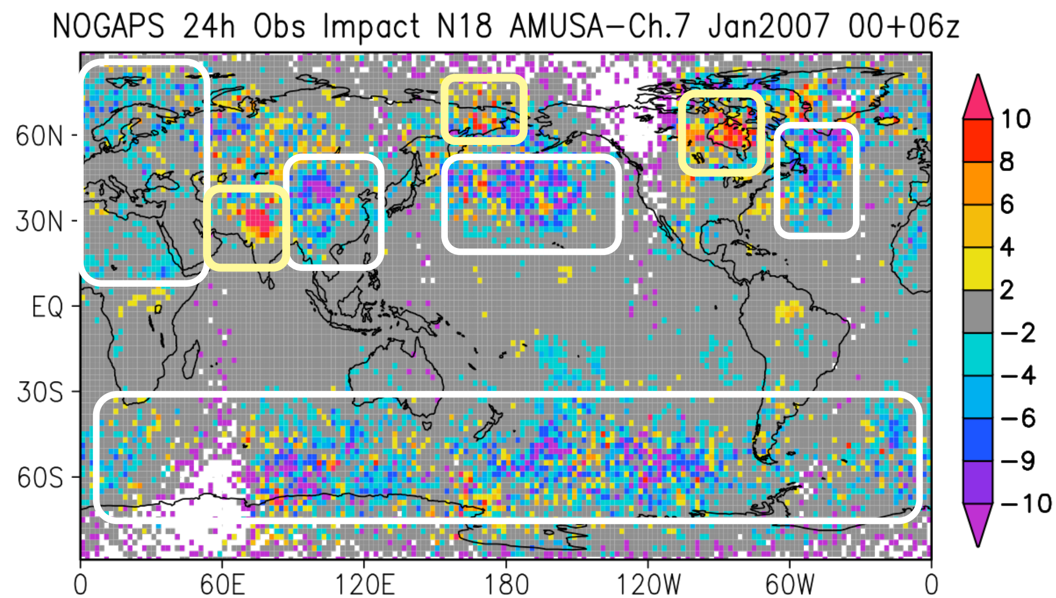
Observations that produce large forecast error reductions

Observations that produce forecast error increases in **both models**

Land or ice surface contamination of radiance data?



NASA GEOS-5



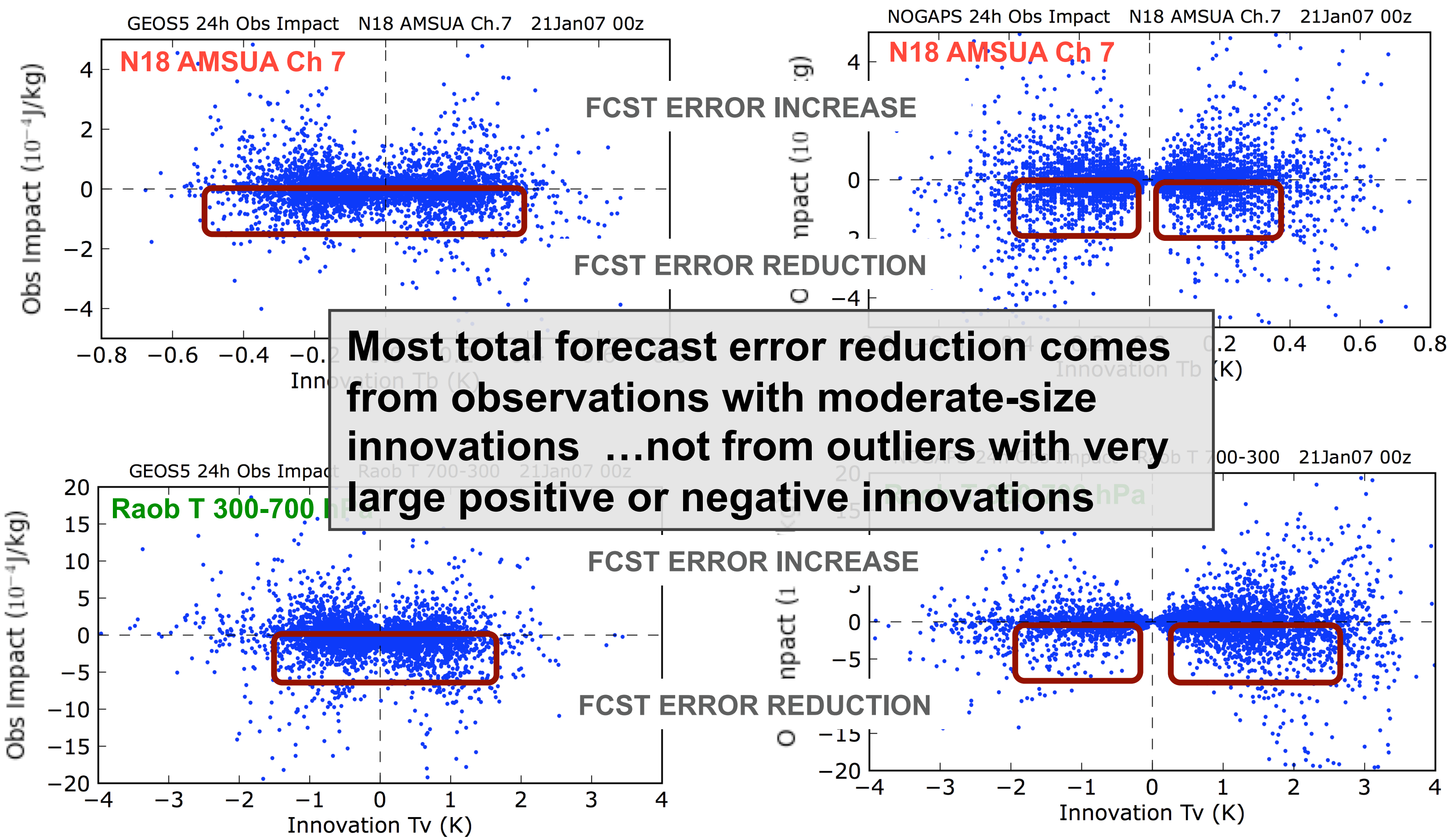
Navy NOGAPS

24h Global Error Norm
Baseline Obs
Jan 2007 00+06 UTC

Scatter of Observation Impact vs Innovation

24h Global Error Norm - Baseline Obs - 21Jan 2007 00UTC

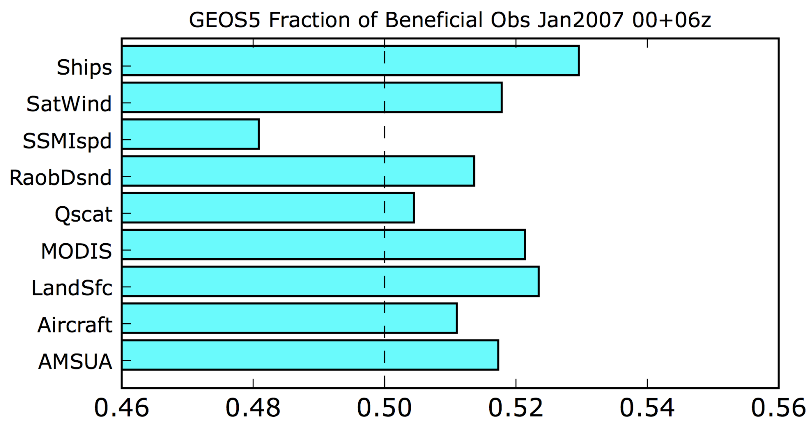
NASA GEOS-5 | **Navy NOGAPS**



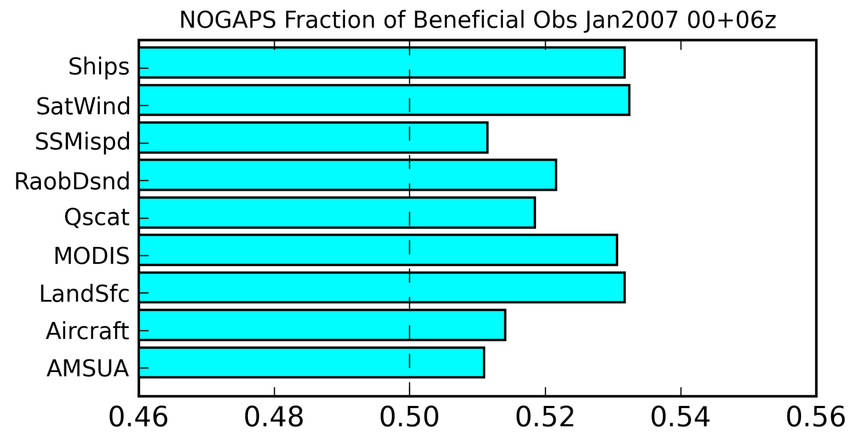
Fraction of observations that reduce forecast error

24h Global Error Norm - Baseline Obs - Jan 2007 00+06 UTC

NASA GEOS-5



Navy NOGAPS



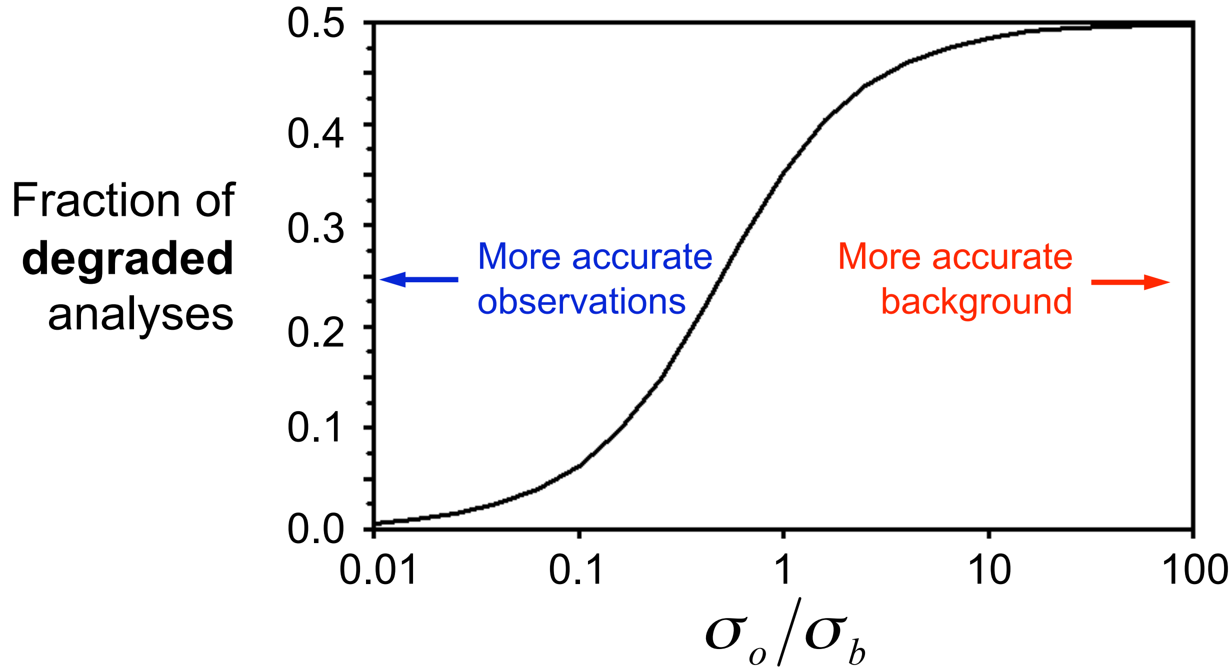
Only a small majority (50-55%) of observations are beneficial, while the rest degrade the forecast!

How can 'good observations' have a negative impact?

The fact that data assimilation relies on ***statistics*** *of background and observation errors implies a distribution of beneficial and non-beneficial impacts…*

…the fact that we don't know these statistics accurately increases the likelihood of there being non-beneficial impacts

- Single-ob, scalar analysis: $x^a = x^b + k\left(y - x^b\right)$ where $k = \sigma_b^2 / \left(\sigma_b^2 + \sigma_o^2\right)$
- Expected impact is positive: $E\left(\varepsilon_a^2 - \varepsilon_b^2\right) = -k\sigma_b^2 < 0$
- **But sometimes, the impact is negative:**



Calculation by Mike Fisher, ECMWF

Further interpretation of the impact measure

Errico (2007) placed the LB04 measure in the context of various-order Taylor series approximations of δe in terms of $\delta \mathbf{y}$

- Forecast error measure:

$$e = (\mathbf{x}_0^f - \mathbf{x}_\mathrm{v})^\mathrm{T} \mathbf{C} (\mathbf{x}_0^f - \mathbf{x}_\mathrm{v})$$

- Taylor expansion of change in e due to change in $\mathbf{x}_0$:

$$\delta e = \delta \mathbf{x}_0 \left(\frac{\partial e}{\partial \mathbf{x}_0} + \frac{1}{2} \frac{\partial^2 e}{\partial \mathbf{x}_0^2} \delta \mathbf{x}_0 + \frac{1}{6} \frac{\partial^3 e}{\partial \mathbf{x}_0^3} \delta \mathbf{x}_0^2 + \ldots \right) = (\delta \mathbf{x}_0)^\mathrm{T} \mathbf{g}$$

- Analysis equation allows transformation to **observation-space**:

$$\delta \mathbf{x}_0 = \mathbf{x}_\mathrm{a} - \mathbf{x}_\mathrm{b} = \mathbf{K} \delta \mathbf{y}$$

- n^{th} order approximation of δe in **observation space**:

$$\delta e_n = (\delta \mathbf{x}_0)^\mathrm{T} \mathbf{g}_n = (\delta \mathbf{y})^\mathrm{T} \tilde{\mathbf{g}}_n$$

Orders of approximation of δe

1st order:

$$\delta e_1 = \delta \mathbf{y}^{\mathrm{T}} \underbrace{2\mathbf{K}^{\mathrm{T}}\mathbf{M}_{\mathrm{b}}^{\mathrm{T}}\mathbf{C}(\mathbf{x}_{\mathrm{b}}^{f} - \mathbf{x}^{t})}_{\tilde{\mathbf{g}}_1}$$

2nd order:

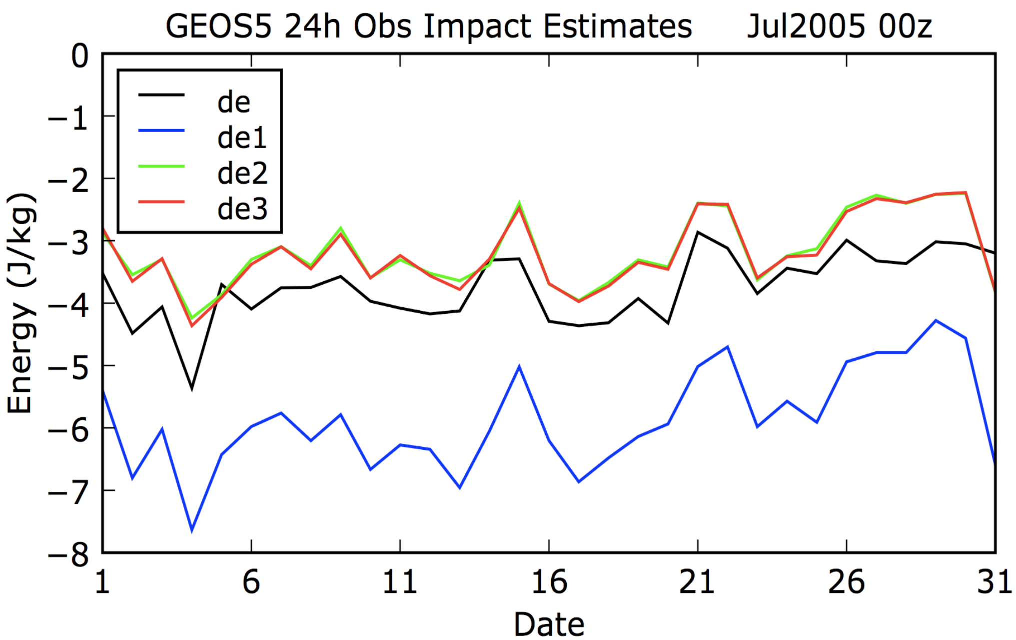
$$\delta e_2 = \delta \mathbf{y}^{\mathrm{T}} \underbrace{\mathbf{K}^{\mathrm{T}}[\mathbf{M}_{\mathrm{b}}^{\mathrm{T}}\mathbf{C}(\mathbf{x}_{\mathrm{a}}^{f} - \mathbf{x}^{t}) + \mathbf{M}_{\mathrm{a}}^{\mathrm{T}}\mathbf{C}(\mathbf{x}_{\mathrm{b}}^{f} - \mathbf{x}^{t})]}_{\tilde{\mathbf{g}}_2}$$

3rd order: **(LB04)**

$$\delta e_3 = \delta \mathbf{y}^{\mathrm{T}} \underbrace{\mathbf{K}^{\mathrm{T}}[\mathbf{M}_{\mathrm{b}}^{\mathrm{T}}\mathbf{C}(\mathbf{x}_{\mathrm{b}}^{f} - \mathbf{x}^{t}) + \mathbf{M}_{\mathrm{a}}^{\mathrm{T}}\mathbf{C}(\mathbf{x}_{\mathrm{a}}^{f} - \mathbf{x}^{t})]}_{\tilde{\mathbf{g}}_3}$$ *+ a higher order term*

Note that $\tilde{\mathbf{g}}_1$ is a gradient and independent of δy, but $\tilde{\mathbf{g}}_2$ and $\tilde{\mathbf{g}}_3$ are weights that depend on all δy through $\mathbf{x}_{\mathrm{a}}$

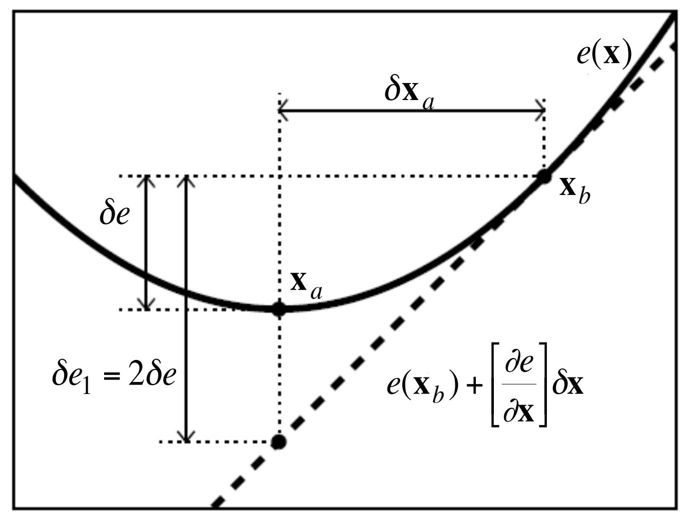
First- vs. Higher-Order Approximations of δe



Gelaro et al. (2007)

- Higher-than-first-order approximation of impact required due to quadratic nature of e

Trémolet (2007)



- If $\mathbf{x}_a$ is near the minimum of e, then the first order approximation will be twice the correct value*

*$\delta e \approx \frac{1}{2}\delta e_1$ *is a tempting approximation, but dangerous if the forecast is poor*

A Caveat with Higher-Order Approximations

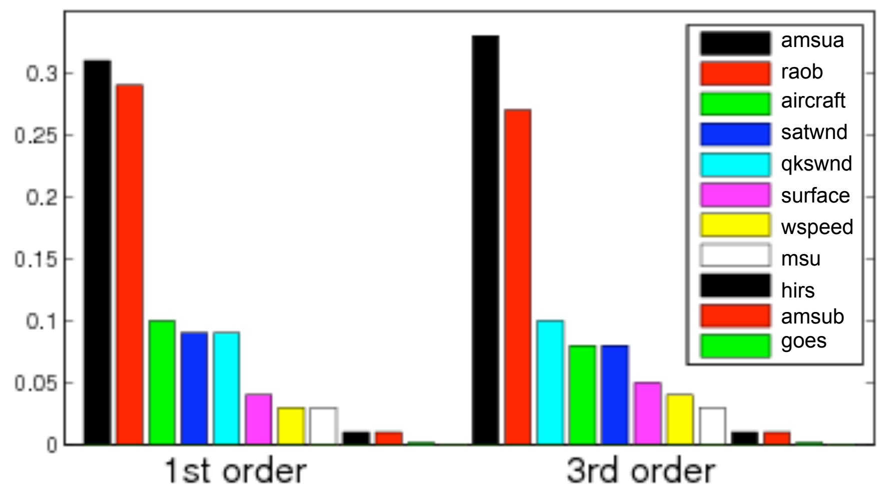
Terms beyond first-order in the approximation δe_3 have the form:

$$\delta e_3 - \delta e_1 \approx (\delta \mathbf{y})^T \mathbf{K}^T \mathbf{M}^T \mathbf{C} \mathbf{M} \mathbf{K} (\delta \mathbf{y})$$

Errico (2007) noted that the nonlinear dependence of these terms on δy means partial sums of δe_3 involve **cross terms** with other observations, which may render results **ambiguous**

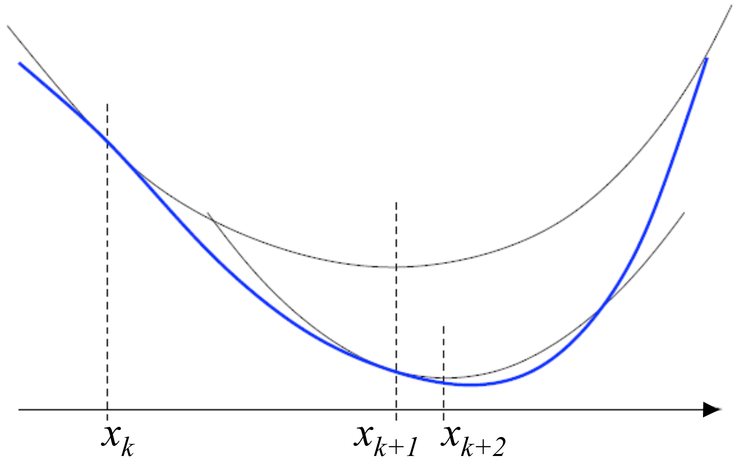
Gelaro et al. (2007) found this effect to be small when measuring impacts of the major observing systems …smaller subsets?

Ranked fractional contributions to 24-h forecast error reduction



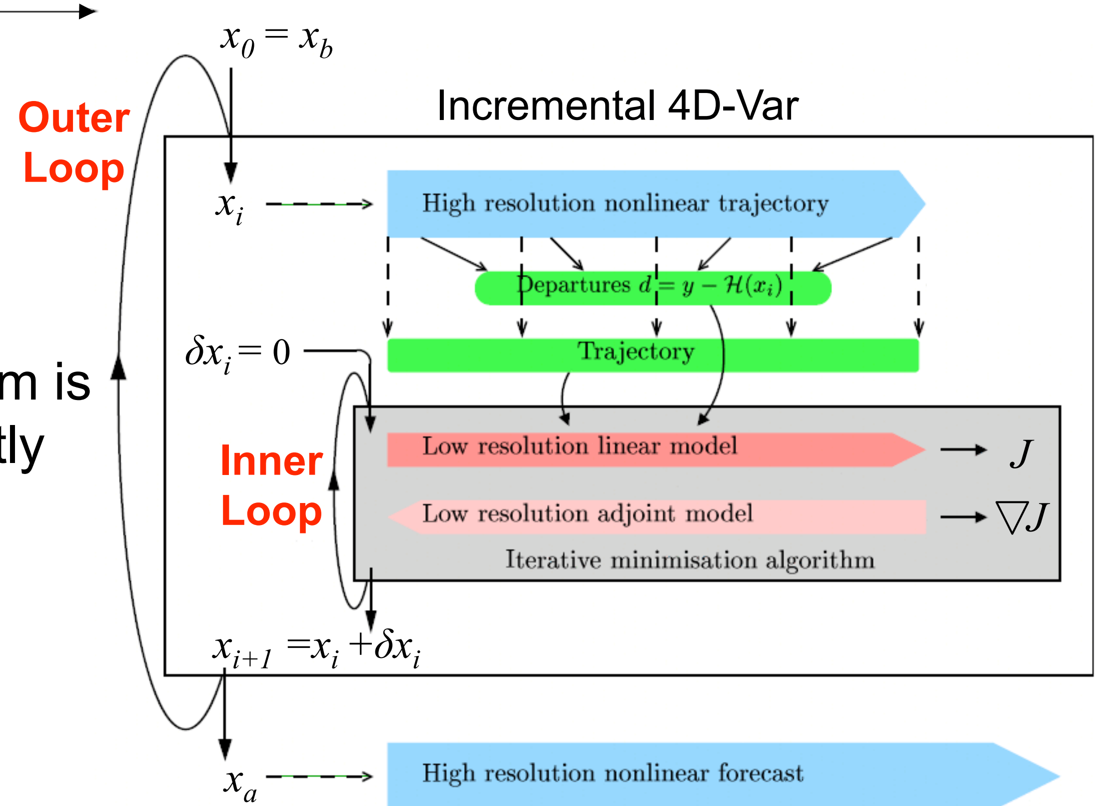
Little change in relative impacts of observing systems suggests cross-term affects are small

Nonlinear Analysis Problems



• In general, the analysis cost function is nonlinear and difficult to minimize

• In an incremental analysis system, one complex problem is replaced by a series of slightly simpler ones (**outer loops**)



Graphics courtesy of Y. Trémolet

Observation Impact in Incremental Variational Data Assim.

Trémolet (2008) examined observation impact in a variational data assimilation system, accounting for $j = 1,\dots,m$ outer loops

- Increment is not: $\mathbf{x}_\mathrm{a} - \mathbf{x}_\mathrm{b} = \mathbf{K}\,\delta\mathbf{y}$

  It is, after loop j : $\mathbf{x}_j - \mathbf{x}_\mathrm{b} = \mathbf{K}_j \mathbf{d}_j + \mathbf{K}_j \mathbf{H}_j (\mathbf{x}_{j-1} - \mathbf{x}_\mathrm{b})$

  or $\mathbf{x}_\mathrm{a} - \mathbf{x}_\mathrm{b} = \sum_{j=1}^{m} \mathbf{L}_j \mathbf{K}_j \mathbf{d}_j$

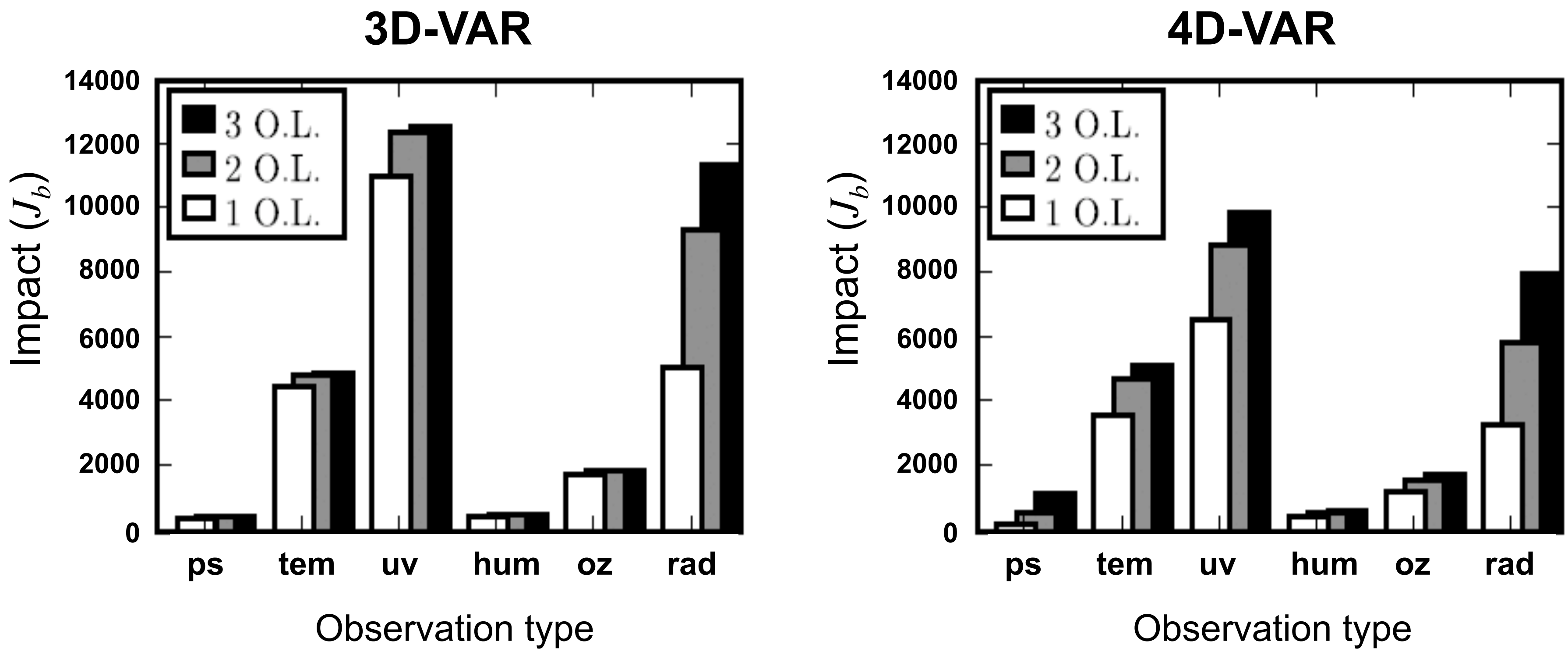
  where $\mathbf{d}_j = \mathbf{y} - H(\mathbf{x}_{j-1})$, $\mathbf{L}_j = \mathbf{K}_m \mathbf{H}_m \dots \mathbf{K}_{j+1} \mathbf{H}_{j+1}$ and $\mathbf{L}_m = \mathbf{I}$

- Then observation impact is: $I = \sum_{j=1}^{m} \left\langle \mathbf{K}_j^\mathrm{T} \mathbf{L}_j^\mathrm{T} \mathbf{g}, \mathbf{d}_j \right\rangle$

  where $\mathbf{g}$ is a gradient or weight in model space

- For example, with m=2 outer loops: $I = \left\langle \mathbf{K}_1^\mathrm{T} \mathbf{H}_2^\mathrm{T} \mathbf{K}_2^\mathrm{T} \mathbf{g}, \mathbf{d}_1 \right\rangle + \left\langle \mathbf{K}_2^\mathrm{T} \mathbf{g}, \mathbf{d}_2 \right\rangle$

Observation Impact with Outer Loops



Impact per observation type on the analysis increment with 1, 2, and 3 outer loop iterations

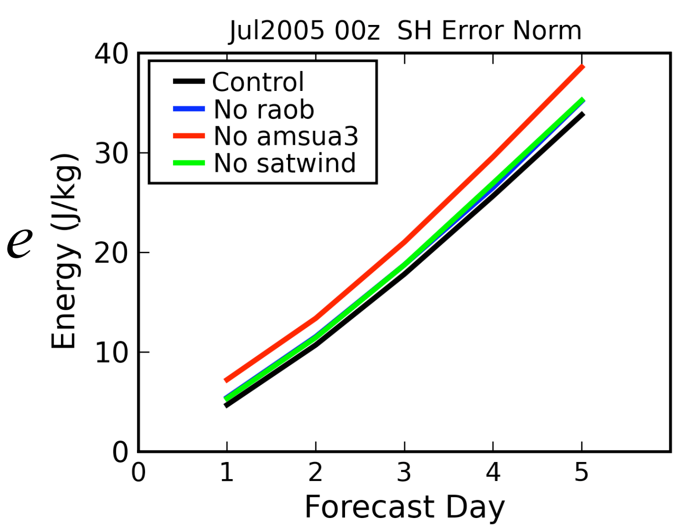
- Outer loop (nonlinear) effects are larger in 4D-Var
- Overall observation impact is smaller in 4D-Var

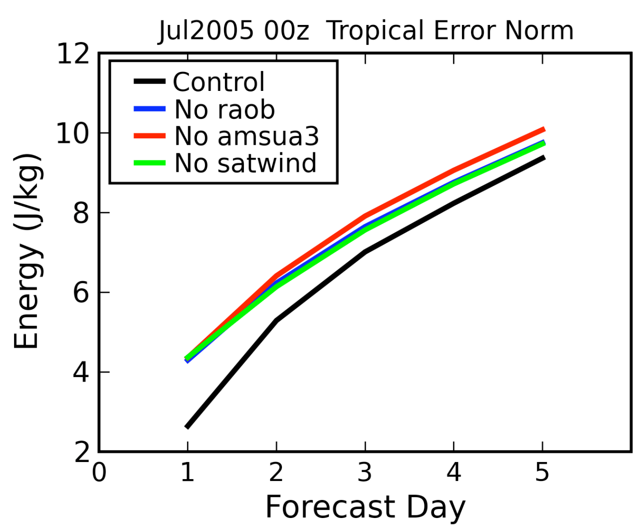
Trémolet (2008)

Observing System Experiments (OSEs)

- Subsets of observations are **added or removed** from the assimilation system and forecasts are compared against a control system that includes all observations

- Because of expense, usually involve a relatively small number of **separate experiments**, each considering a relatively large subset of observations





Gelaro and Zhu (2009)

Comparison and Interpretation of ADJ and OSE Results

…a few things to keep in mind…

ADJ: measures the impacts of observations in the context of all other observations present in the assimilation system

OSE: removal of observations changes or degrades the system…$\mathbf{K}$ differs for each experiment

❖

ADJ: measures the impact of observations in each analysis cycle separately and against the control background

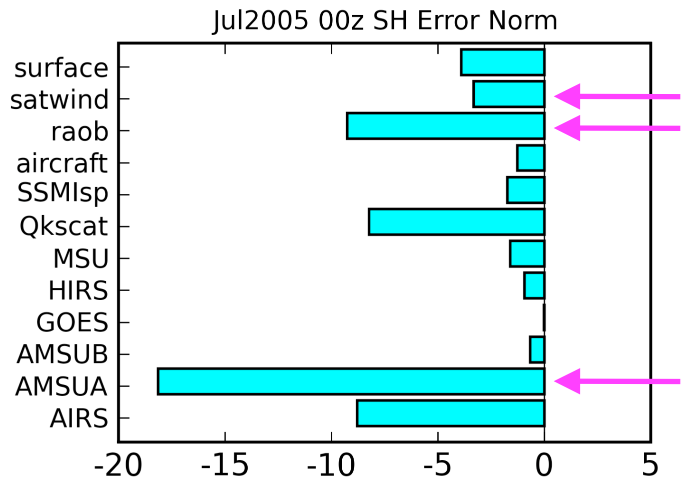
OSE: measures the impact of removing information from both the background and analysis in a cumulative manner

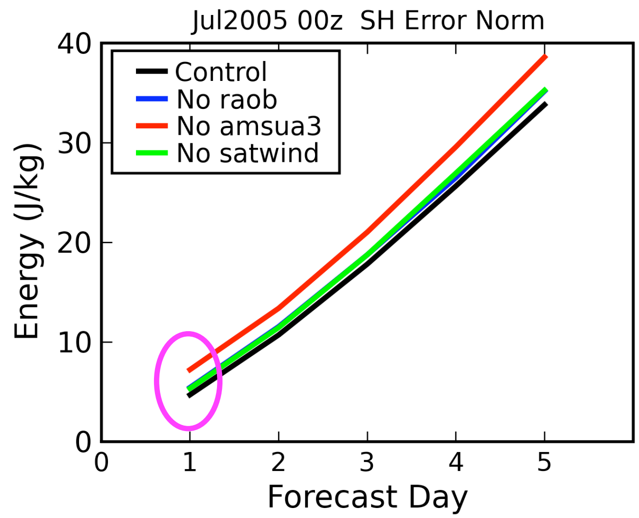
❖

ADJ: measures the response of a single forecast metric to all perturbations of the observing system

OSE: measures the effect of a single perturbation on all forecast metrics

Quantitative Comparison of ADJ and OSE Results





- Strictly speaking, quantitative comparison is limited to the forecast range and metric for which the ADJ results are valid on the one hand (e.g. 24h SH e-norm) and to the selected observing systems removed in the OSEs on the other hand

- Even then, comparisons between the ADJ and OSE results are complicated by the fact that values/changes in e measured in the OSE context are not directly comparable to values of δe measured in the ADJ context

GEOS-5 Observing System Experiments (OSEs)

• The following observation sets were individually excluded from the data assimilation system during July 2005 and January 2006:

- all AMSU-A radiances from 1 satellite (N-16): ***no amsua1***
- all AMSU-A radiances from 2 satellites (N-15,16): ***no amsua2***
- all AMSU-A radiances from 3 satellites (N-15,16, Aqua): ***no amsua3***
- all AIRS radiances: ***no airs***
- all rawinsonde observations: ***no raob***
- all satellite winds (AMVs): ***no satwind***
- all aircraft observations: ***no aircraft***
- all scatterometer winds from QuikSCAT: ***no qkscat***

• Control analysis used for verification, impact measured using the **same energy metric as in the ADJ experiments** for the globe, NH, SH and tropics

Quantitative Comparison of ADJ and OSE Results

OSE: $$e = (\mathbf{x}_0^f - \mathbf{x}^t)^{\mathrm{T}} \mathbf{C} (\mathbf{x}_0^f - \mathbf{x}^t)$$

ADJ: $$\delta e = (\delta \mathbf{y})^{\mathrm{T}} \mathbf{K}^{\mathrm{T}} [\mathbf{M}_{\mathrm{b}}^{\mathrm{T}} \mathbf{C} (\mathbf{x}_{\mathrm{b}}^f - \mathbf{x}^t) + \mathbf{M}_{\mathrm{a}}^{\mathrm{T}} \mathbf{C} (\mathbf{x}_{\mathrm{a}}^f - \mathbf{x}^t)]$$

Gelaro and Zhu (2009) defined a fractional impact F_j of observing system j for each approach:

$$F_j(\mathrm{ADJ}) = \delta e_j / \delta e$$

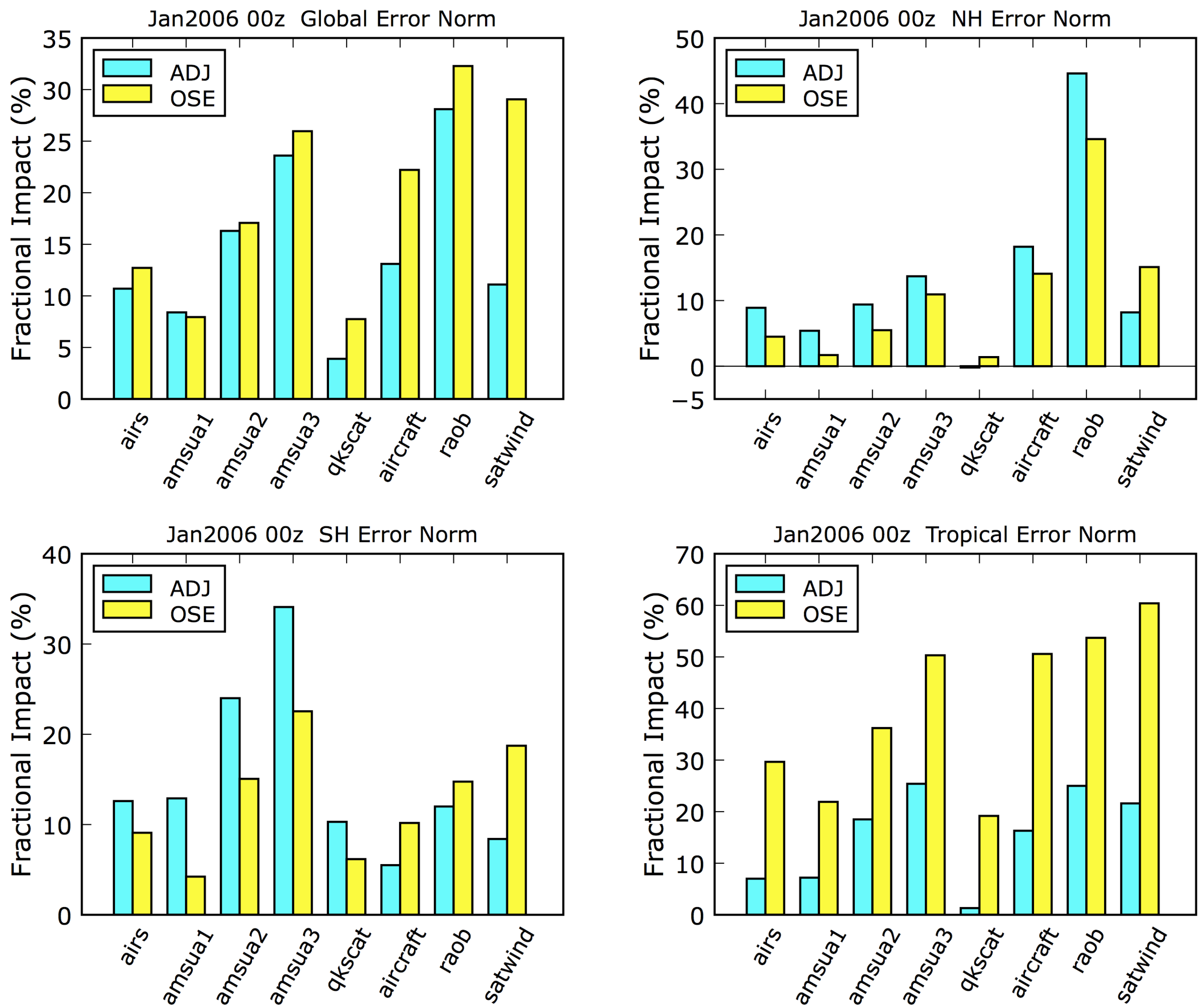
- Measures the % **decrease** in error due to the **presence** of obs system j with respect to the background forecast
- $\sum_j F_j(\mathrm{ADJ}) = 1$

$$F_j(\mathrm{OSE}) = (e_{\mathrm{no}\,j} - e_{\mathrm{ctl}}) / e_{\mathrm{ctl}}$$

- Measures the % **increase** in error due to the **removal** of obs system j with respect to the control forecast
- $\sum_j F_j(\mathrm{OSE}) \neq 1$

% Contributions to 24hr Forecast Error Reduction

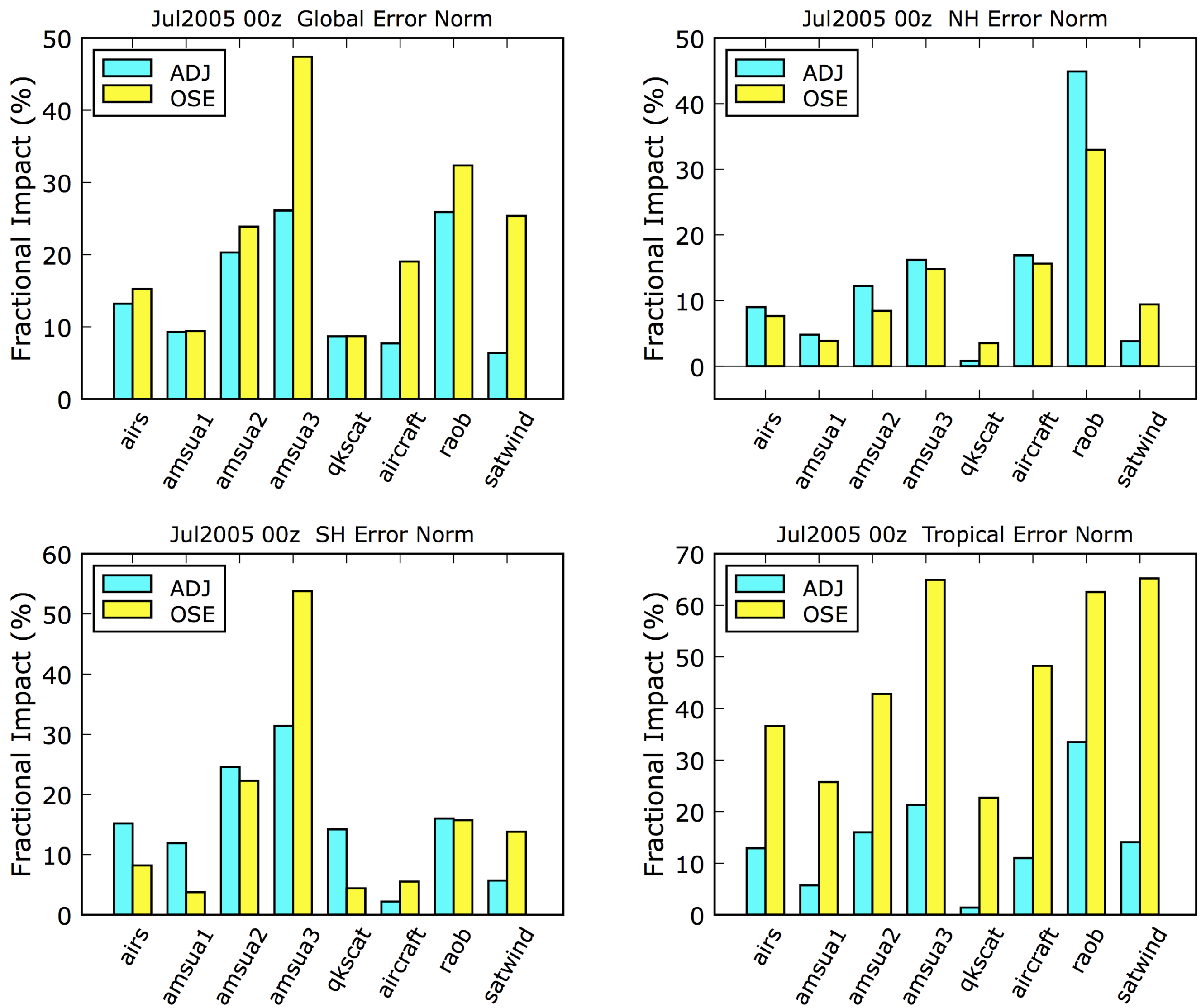
GEOS-5 January 2006



Gelaro and Zhu (2009)

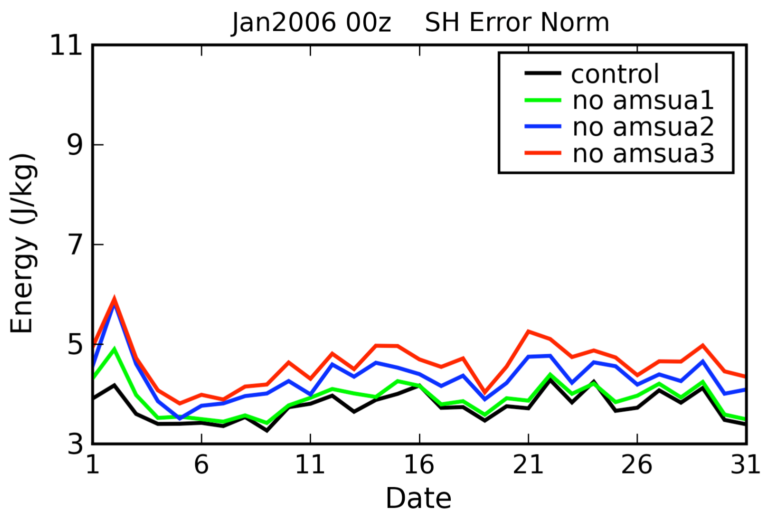
% Contributions to 24hr Forecast Error Reduction

GEOS-5 July 2005



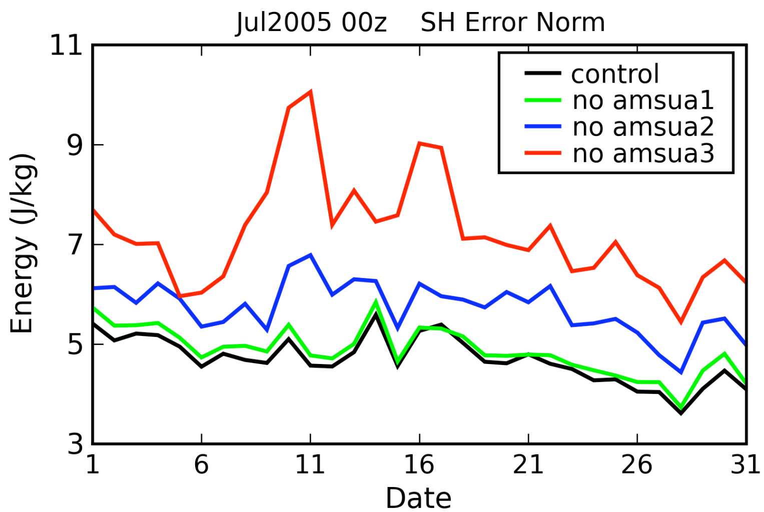
Gelaro and Zhu (2009)

OSE Time Series of SH 24-hr Forecast Error Norm



January 2006

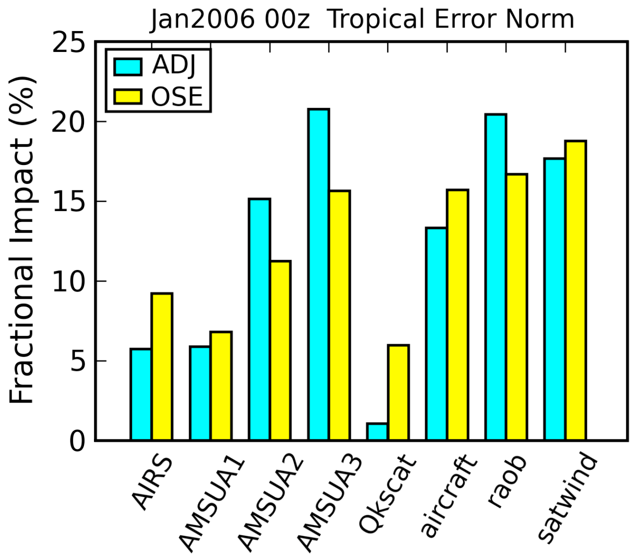
Skill collapses when all AMSUA removed during SH winter…OSE and ADJ results become difficult to compare



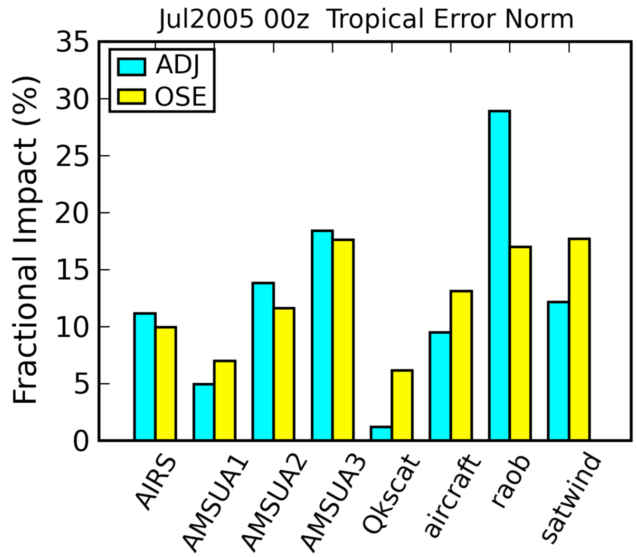
July 2005

Normalized % Contributions to 24hr Forecast Error Reduction

January 2006



July 2005



…ADJ and OSE responses differ in magnitude in the tropics, but assign similar relative ‘value’ to the various observing systems

Digging deeper into the ‘mix of observations’

• Both OSEs and ADJ measure the net effect of observations on the forecast

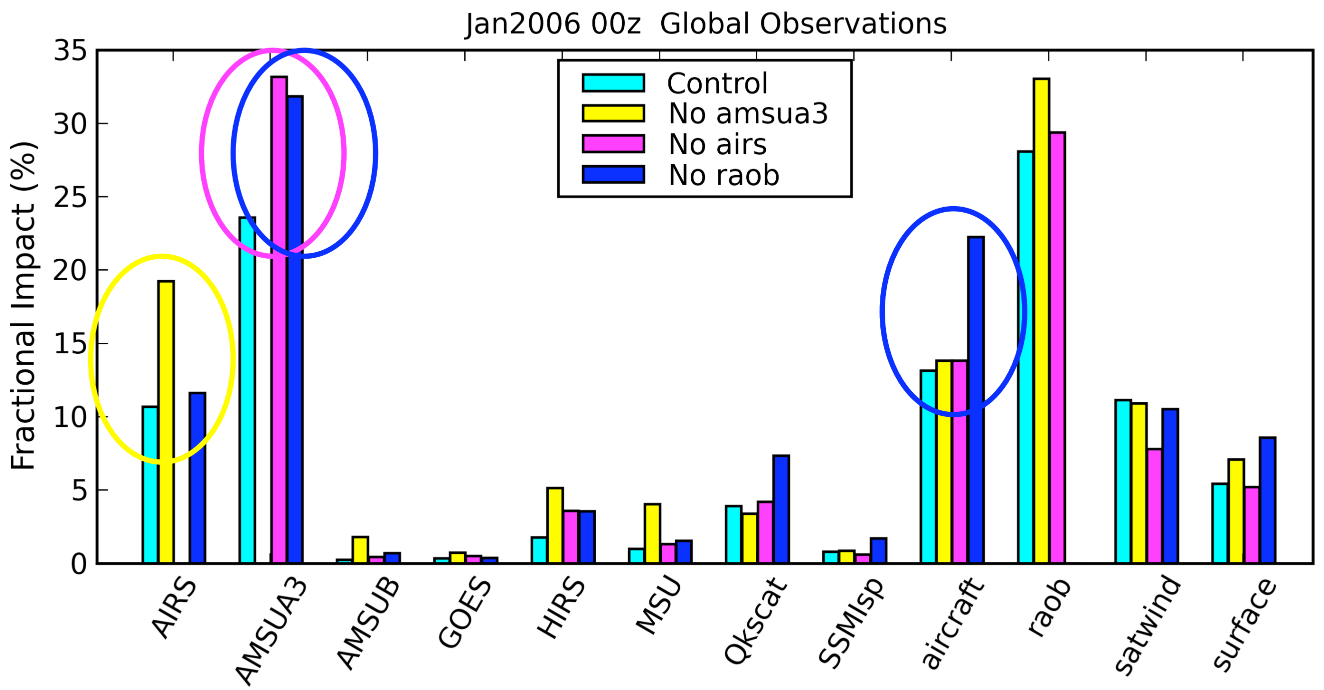
• We are also interested in dependencies and redundancies between observing systems as observations are added or removed …inform current data selection, future data needs

• Such information is implicitly available in an OSE in terms of the responses of the remaining observing systems when a given set of observations is removed

• These responses can be measured through the **combined use of OSEs and ADJs**, by applying the ADJ to the perturbed (vs. only the control) members of an OSE

Combined Use of ADJ and OSEs

ADJ applied to perturbed OSE members to examine how changing the mix of observations influences their impacts

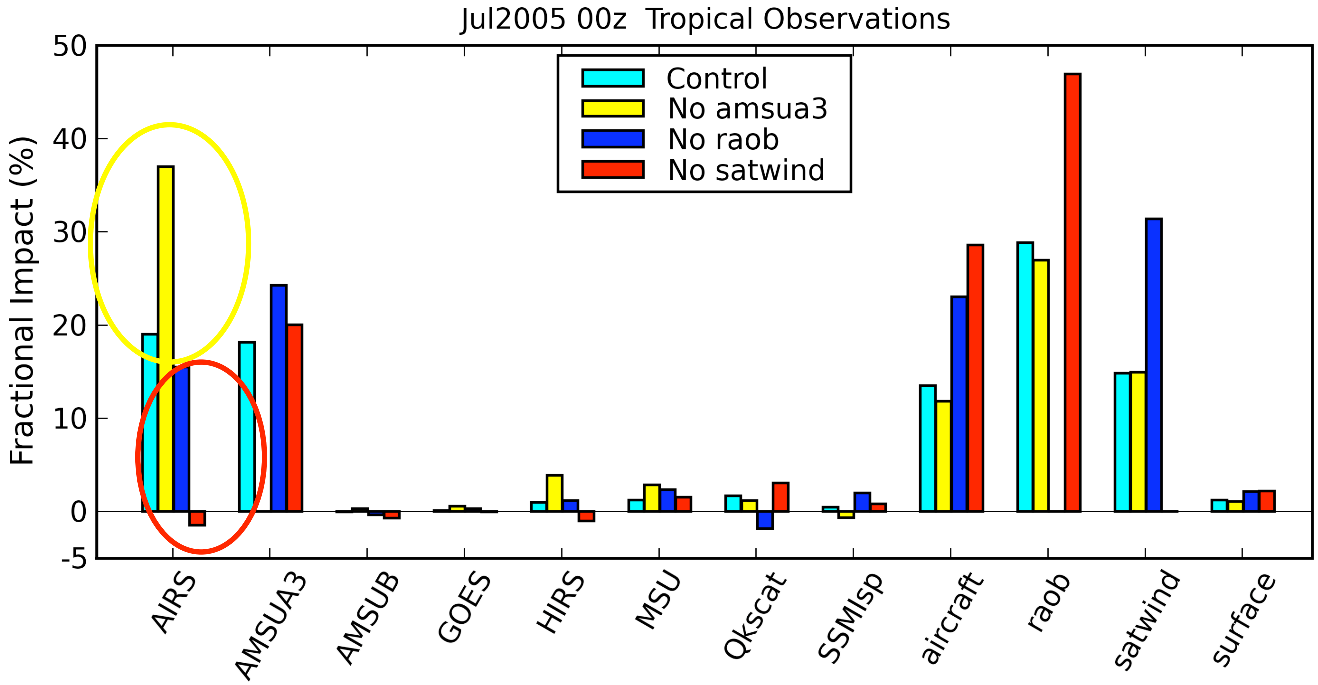


Gelaro and Zhu (2009)

- Removal of AMSUA results in large increase in AIRS (and other) impacts
- Removal of AIRS results in significant increase in AMSUA impact
- Removal of raobs results in significant increase in AMSUA, aircraft and other impacts (but not AIRS)

Combined Use of ADJ and OSEs

ADJ applied to perturbed OSE members to examine how changing the mix of observations influences their impacts



Gelaro and Zhu (2009)

- Removal of AMSUA results in large increase in AIRS impact in tropics
- **Removal of wind observations** results in significant **decrease** in AIRS impact in tropics (in fact, AIRS degrades forecast without satwinds!)

Conclusions on the Complementarity of ADJ and OSE

- Despite fundamental differences in how impact is measured, ADJ and OSE methods provide comparable estimates of the overall ‘value’ of most observing systems

- Differences in OSE and ADJ results should be expected and do not necessarily point to shortcomings in either:
  - ✓ different treatment of background information
  - ✓ removal of whole observing systems that contribute disproportionately to analysis quality (AMSU-A)

- Information gleaned from OSEs and ADJs should be viewed as complementary; ADJ extends, not replaces, OSEs:
  - ✓ applicable forecast range, metrics differ
  - ✓ ADJ well suited for routine monitoring

- The combined use of ADJs and OSEs illuminates the complex, complementary nature of how observations are used by the assimilation system

References Cited in this Talk

Baker, N. and Daley, R. 2000. Observation and background adjoint sensitivity in the adaptive observation-targeting problem. *Quart. J. R. Meteorol. Soc.* **126,** 1431–1454*.*

Errico, R. M. 2007. Interpretation of an adjoint-derived observational impact measure. *Tellus* **59A,** 273–276*.*

Gelaro, R. and Zhu Y. 2009. Examination of observation impacts derived from observing system experiments (OSEs) and adjoint models. *Tellus* **61A,** 179–193*.*

Gelaro, R., Zhu,Y. and Errico, R.M. 2007. Examination of various-order adjoint-based approximations of observation impact. *Meteorologische Zeitschrift* **16,** 685–692*.*

Langland, R. H. and Baker, N. 2004. Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus* **56A,** 189–201*.*

Trémolet, Y. 2008. Computation of observation sensitivity and observation impact in incremental variational data assimilation. *Tellus* ***60A,*** 964–978.

Trémolet, Y. 2007. First-order and higher-order approximations of observation impact. *Meteorologische Zeitschrift* **16,** 693–694.