Adjoint Models as Analytical Tools

Dr. Ronald M. Errico

Goddard Earth Sciences and Technology Center (UMBC)
Global Modeling and Assimilation Office (NASA)

Outline

2. Sensitivity analysis: The basis for adjoint model applications
2. An example of adjoint-derived sensitivities
4. Development of adjoint model software
5. Nonlinear validation
6. Sensitivity to Observations
7. Problems with physics
8. Summary

Sensitivity Analysis:
The basis for adjoint model applications

Errico, R.M., 1997: What is an djoint model? *Bull. Am. Meteor. Soc.*, **78**, 2577-2591.

Adjoint Sensitivity Analysis for a Discrete Model

The Problem to Consider:

A possibly nonlinear model:

$$\mathbf{y} = \mathbf{m}(\mathbf{x}) \tag{1}$$

A differentiable scalar measure of model output fields:

$$J = J(\mathbf{y}) \tag{2}$$

The result of input perturbations

$$\Delta J = J(\mathbf{x} + \mathbf{x}') - J(\mathbf{x}) \tag{3}$$

A 1st–order Taylor series approximation to ΔJ

$$J' = \sum_i \frac{\partial J}{\partial x_i}\, x_i' \tag{4}$$

The goal is to efficiently determine $\frac{\partial J}{\partial x_i}$ for all i

Adjoint Sensitivity Analysis for a Discrete Model

The Tangent Linear Model (TLM)

Apply a 1st–order Taylor series to approximate the model output

$$y_i' = \sum_j \frac{\partial y_i}{\partial x_j} x_j' \tag{5}$$

$\partial y_i / \partial x_j$ is called the **Resolvant** matrix of the TLM or, less accurately, the **Jacobian** of the nonlinear model.

Approximate ΔJ by a 1st–order Taylor series about $\mathbf{y}'$

$$J' = \sum_i \frac{\partial J}{\partial y_i} y_i' \tag{6}$$

Adjoint Sensitivity Analysis for a Discrete Model

The Adjoint Model

(Adjoint of the TLM or adjoint of the nonlinear model)

Application of the "chain rule" yields

$$\frac{\partial J}{\partial x_i} = \sum_j \frac{\partial y_j}{\partial x_i} \frac{\partial J}{\partial y_j} \tag{9}$$

Contrast with the TLM

$$y'_i = \sum_j \frac{\partial y_i}{\partial x_j} x'_j \tag{10}$$

A. The variables are different in the two equations

B. The order of applications of the varaiables related to x and y differ

C. The indices i and j in the matrix operator are reversed

Adjoint Sensitivity Analysis for a Discrete Model

Additional Notes

1. Mathematically, the field $\partial J/\partial \mathbf{x}$ is said to reside in the dual space of $\mathbf{x}$
2. With the change of notation $\hat{\mathbf{x}} = \partial J/\partial \mathbf{x}$, $\mathbf{M} = \partial \mathbf{y}/\partial \mathbf{x}$, etc.,

$$J' = \hat{\mathbf{y}}^T \mathbf{y} = \hat{\mathbf{y}}^T (\mathbf{M}\mathbf{x}) = \left(\hat{\mathbf{y}}^T \mathbf{M}\right) \mathbf{x} = \left(\mathbf{M}^T \hat{\mathbf{y}}\right)^T \mathbf{x} = \hat{\mathbf{x}}^T \mathbf{x} \qquad (11)$$

3. The exact definition of the the adjoint depends on the quadratic expression used to define J'. If the simple Euclidean norm (or dot product) is used, then for a discrete model, the adjoint is simply a transpose. Such a simple norm may not be appropriate when the dual space fields are to be physical interpretated. (More on this later.)
4. The adjoint is not generally the inverse: in non–trivial atmospheric models, $\mathbf{M}^T \neq \mathbf{M}^{-1}$.
5. This is all 1st–year calculus and linear algebra. If examination of gradients is useful, then so are the adjoint models used to calculate them.

Examples of Adjoint-Derived Sensitivity

Errico, R.M., and T. Vukicevic, 1992: Sensitivity analysis using an adjoint of the PSU-NCAR mesoscale model. *Mon. Wea. Rev.,* **120**, 1644-1660.

Rabier, F., P. Courtier, and O. Talagrand, 1992: An application of adjoint models to sensitivity analysis. *Beitr. Phys. Atmos*., **65**, 177-192.

Langland, R.H., R.L. Elsberry, and R.M. Errico, 1995: Evaluation of physical processes in an idealized extratropical cyclone using adjoint sensitivity. *Quart. J. Roy. Meteor. Soc*., **121**, 1349-1386.

Adjoint Sensitivity Analysis for a Discrete Model

Example J

Consider J for northward moisture flux through a "window"

J for continuous fields

$$J = \frac{1}{M} \int q \ v \ dm \tag{13}$$

J for discretized model

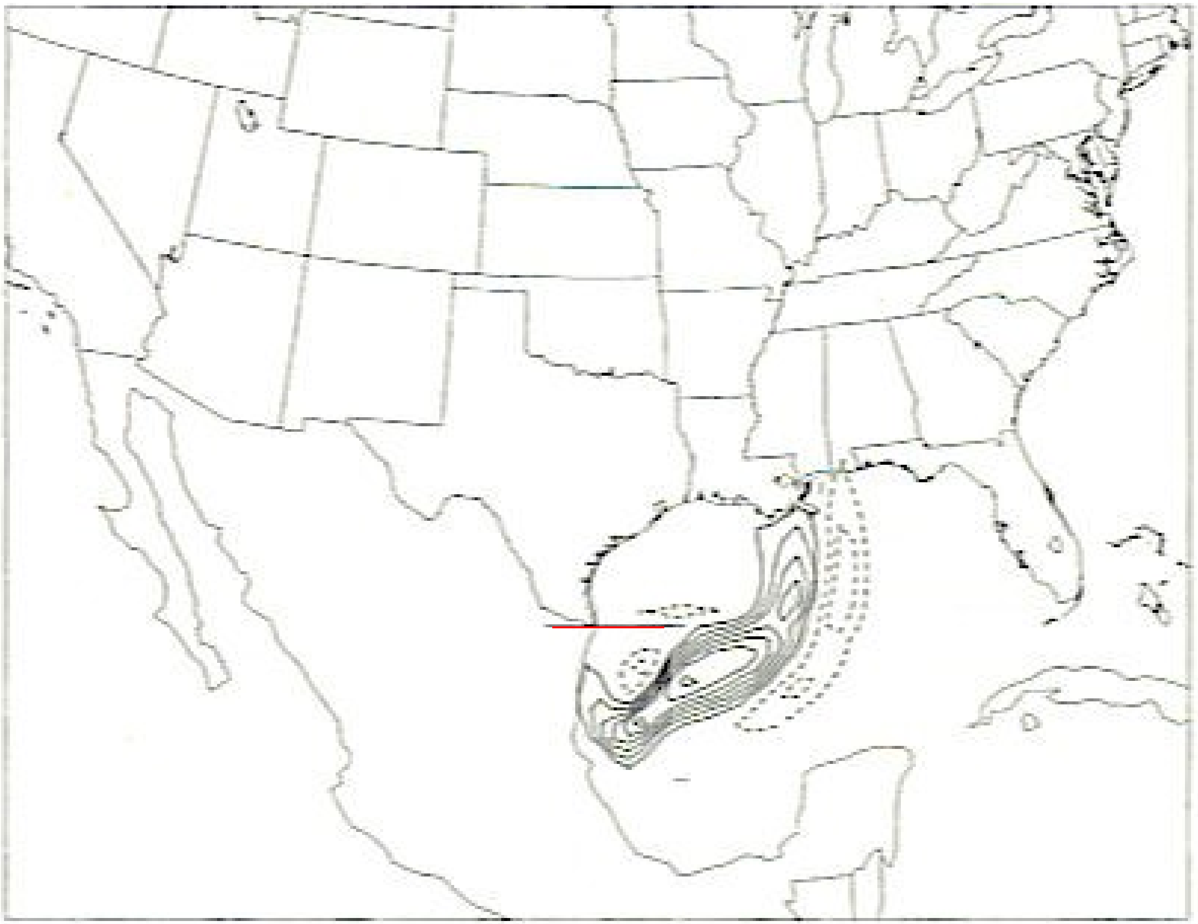
$$J = \sum w_{i,j,k} \ q_{i,j,k} \ v_{i,j,k} \tag{14}$$

$$\frac{\partial J}{\partial v_{i,j,k}} = w_{i,j,k} \ \tilde{q}_{i,j,k} \tag{15}$$

$$\frac{\partial J}{\partial q_{i,j,k}} = w_{i,j,k} \ \tilde{v}_{i,j,k} \tag{16}$$

From Lewis et al. 2001 Tellus

$$\frac{\partial \overline{qv}}{\partial T_s}(t = -48\ \mathrm{h})$$



Development of Adjoint Model Software from Line by Line Analysis of Computer Code

http://autodiff.com/tamc/
www.fastopt.com
http://imgi.uibk.ac.at/MEhrendorfer/work_7/present/

session3/giering.pdf

Development of Adjoint Model Software from Line by Line Analysis of Computer Code

Let the model be expressed in terms of a sequence of operators:

$$\mathbf{y} = \mathbf{m}(\mathbf{x}) = \mathbf{d}(\mathbf{c}(\mathbf{b}(\mathbf{a}(\mathbf{x}))))$$

Then, the TLM and Adjoint are decribed by sequences of linear operators

TLM:

$$\mathbf{y}' = \mathbf{DCBAx}'$$

Adjoint

$$\frac{\partial J}{\partial \mathbf{x}} = \mathbf{A}^T \mathbf{B}^T \mathbf{C}^T \mathbf{D}^T \frac{\partial J}{\partial \mathbf{y}}$$

Consider a 2 line example:

Z = m(X,W) = line 2(line 1 (X,W)), with a fixed parameter A

Parent NLM :

Y = X * (W**A)

Z = X * Y

TLM :

$(\partial Y/\partial X)_W$ $(\partial Y/\partial W)_X$

Ytlm = Xtlm * [W**A] + Wtlm *[A*X*(W**(A-1))]

$(\partial Z/\partial Y)_X$ $(\partial Z/\partial X)_Y$

Ztlm = Ytlm * [X] + Xtlm * [Y]

Adjoint :

Xadj = Xadj + Zadj * Y

Yadj = Yadj + Zadj * X

Xadj = Xadj + Yadj * (W**A)

Wadj = Wadj + Yadj * A*X *(W**(A-1))

Development of Adjoint Model From Line by Line Analysis of Computer Code

Parent NLM :

Y = X * (W**A)

Z = Y * X

TLM :

Ytlm = Xtlm * (W**A) + Wtlm *A* X *(W**(A-1))

Ztlm = Ytlm * X + Xtlm * Y

Xadj = Xadj + Zadj * Y

Yadj = Yadj + Zadj * X

Adjoint :

Xadj = Xadj + Yadj * (W**A)

Wadj = Wadj + Yadj * X *(W**(A-1))

Development of Adjoint Model From Line by Line Analysis of Computer Code

1. TLM and Adjoint models are straight-forward to derive from the parent NLM code.
3. The derivation can be tedious by hand, but automatic differentiation tools are available.
4. Code generated by AD tools may be incorrect or inefficient.
5. Intelligent approximations can be made to improve efficiency.
6. TLM and (especially) Adjoint codes are simple to test rigorously.
6. Some outstanding errors and problems in the NLM are typically revealed when a TLM and Adjoint are developed from it.
7. It is best to start from clean NLM code.
8. The TLM and Adjoint can be formally correct but useless!

Nonlinear Validation

Does the TLM or Adjoint model tell us anything about the behavior of meaningful perturbations in the nonlinear model that may be of interest?

Tangent Linear vs. Nonlinear Results

Is

$$y'_j \;=\; \sum_i \frac{\partial y_j}{\partial x_i}\, x'_i$$

a good approximation to

$$\Delta y_j \;=\; y_j(\mathbf{x}+\mathbf{x}') \;-\; y_j(\mathbf{x})$$

for meaningful $\mathbf{x}'$?

Errico, R. M., and K. D. Raeder, 1999: An examination of the accuracy of the linearization of a mesoscale model with moist physics. *Quart. J. Roy. Meteor. Soc.*, **125**, 169-195.

Tangent Linear vs. Nonlinear Results

In general, agreement between TLM and NLM results will depend on:

4. Amplitude of perturbations
5. Stability properties of the reference state
6. Structure of perturbations
7. Physics involved
8. Time period over which perturbation evolves
9. Measure of agreement

Adjoint vs. Nonlinear Results

The agreement between the Adjoint and NLM is exactly that of the TLM and NLM if the Adjoint is exact with respect to the TLM in the sense that

$$J' = \sum_j \frac{\partial J}{\partial y_j} y'_j = \sum_i \frac{\partial J}{\partial x_i} x'_i$$

Sensitivity to Observations
The use of an adjoint of a data analysis algorithm

Baker, N., 2000: *Observation adjoint sensitivity and the adaptive observation-targeting problem*. Ph. D. Thesis, Naval Post—Graduate School.

Gelaro, R., Y. Zhu, and R. Errico, 2007: Examination of various-order adjoint-based approximation of observation impact. *Meteorologische Zeitschrift*, in press.

Sensitivity to Observations

Using the Adjoint of a Data Assimilation System

Consider an analysis $\mathbf{x}_a$ determined from a background $\mathbf{x}_b$ and observations $\mathbf{y}_o$

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}\left[\mathbf{y}_o - H(\mathbf{x}_b)\right]$$

Consider a sensitivity field $\partial J/\partial \mathbf{x}_a$. Then a corresponding sensitivity to observations can be obtained by using the adjoint $\mathbf{K}^T$ of the data assimilation system

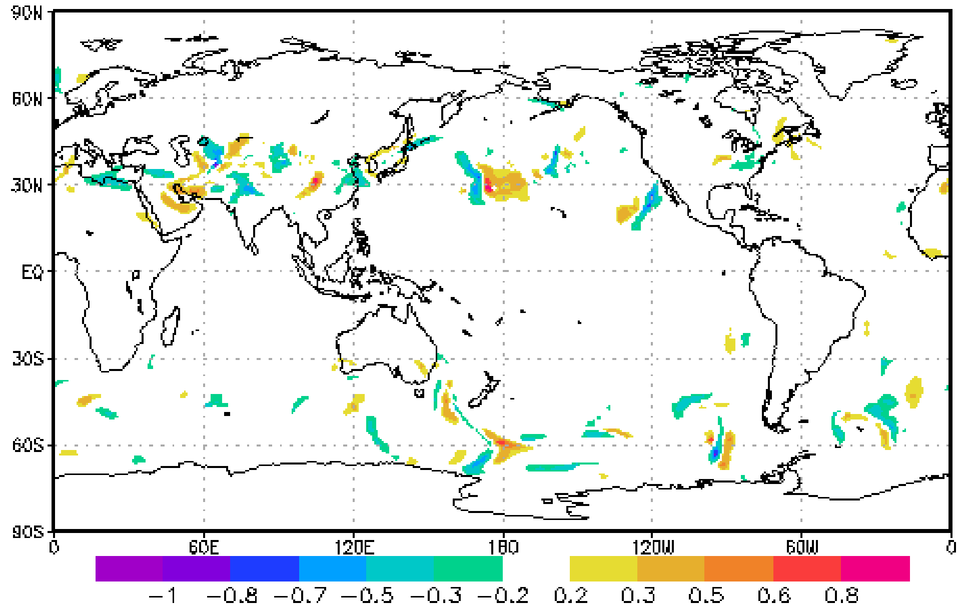
$$\left(\frac{\partial \mathbf{x}_a}{\partial \mathbf{y}_o}\right)^T = \mathbf{K}^T$$

With the result

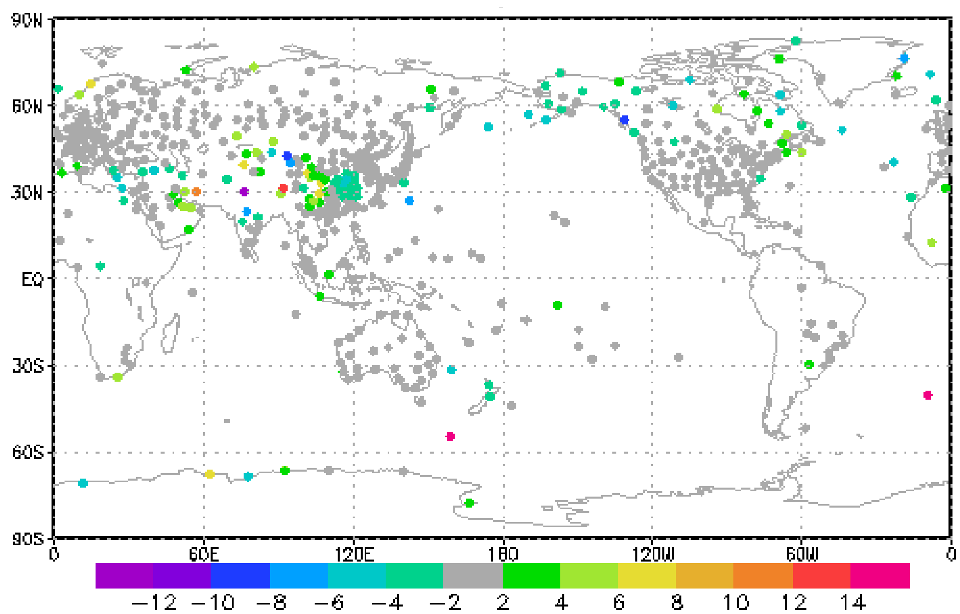
$$\frac{\partial J}{\partial \mathbf{y}_o} = \mathbf{K}^T \frac{\partial J}{\partial \mathbf{x}_a}$$

J = mean squared 24 hour forecast error using E-norm

Sensitivity to analyzed potential temperature at 500 hPa



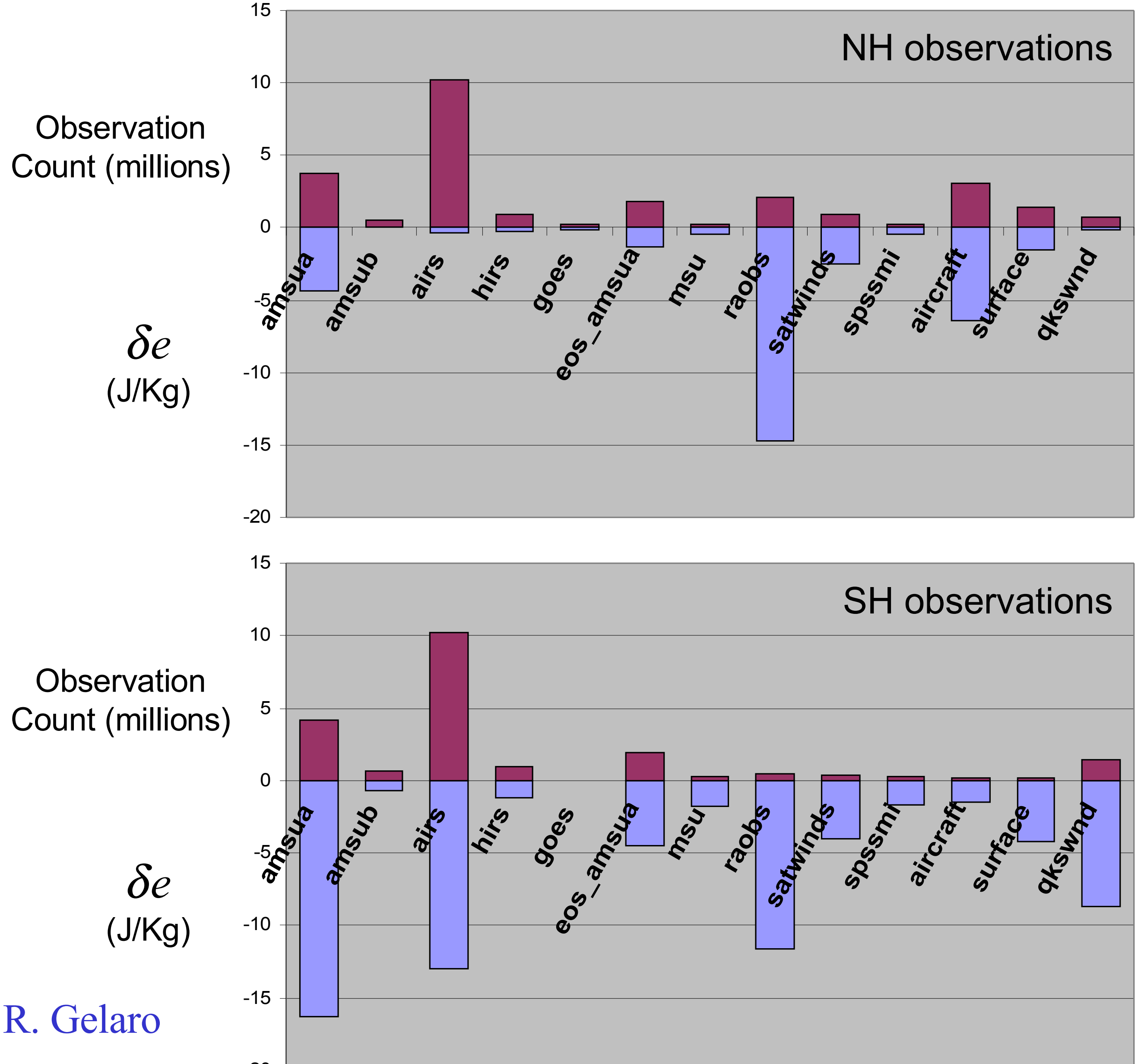
Sensitivity to raob temperatures at 500 hPa



From Gelaro and Zhu 2006

Impacts of various observing systems Totals

GEOS-5 July 2005



$\delta e < 0$

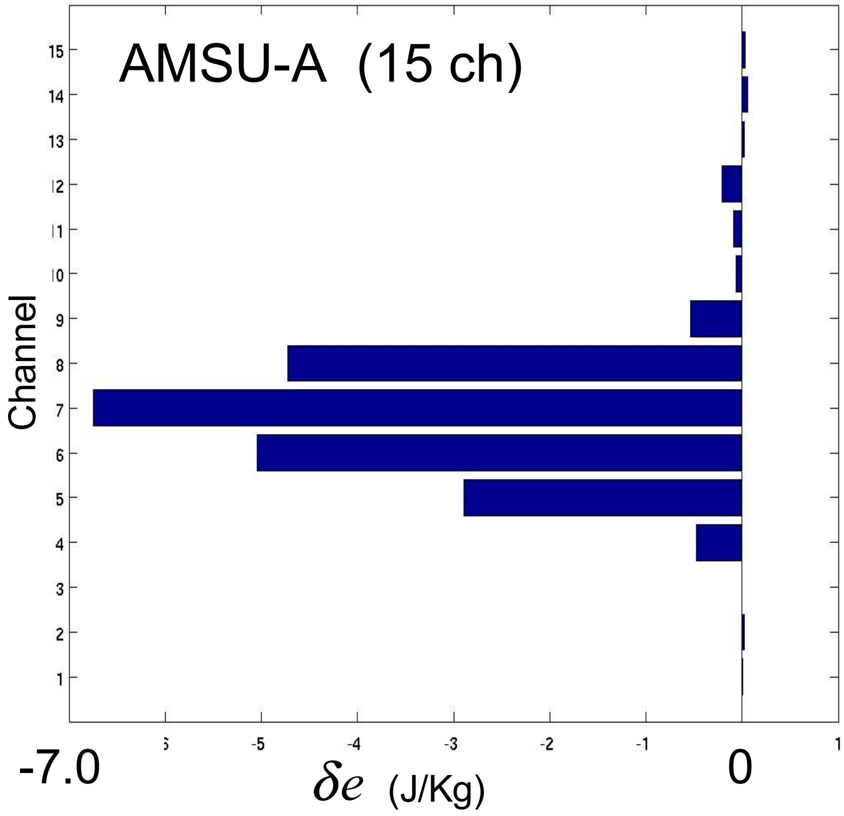
…all observing systems provide total monthly benefit

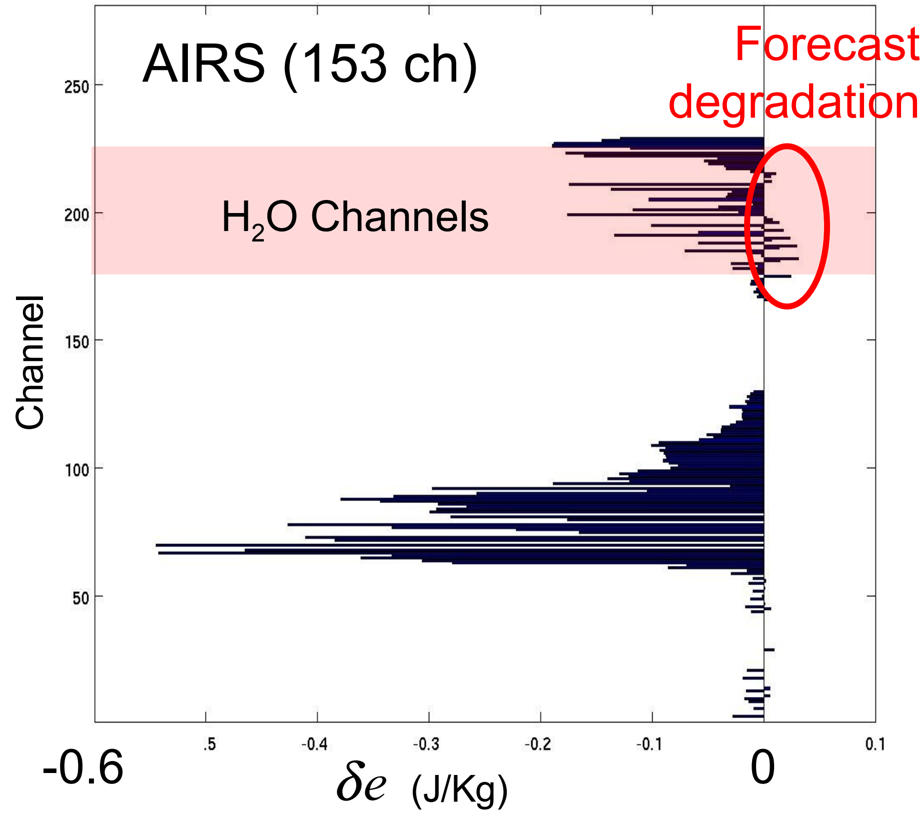
$\delta e < 0$

From R. Gelaro

Diagnosing impact of hyper-spectral observing systems

GEOS-5 July 2005 00z Totals





…some AIRS water vapor channels currently degrade the 24h forecast in GEOS-5…

From R. Gelaro

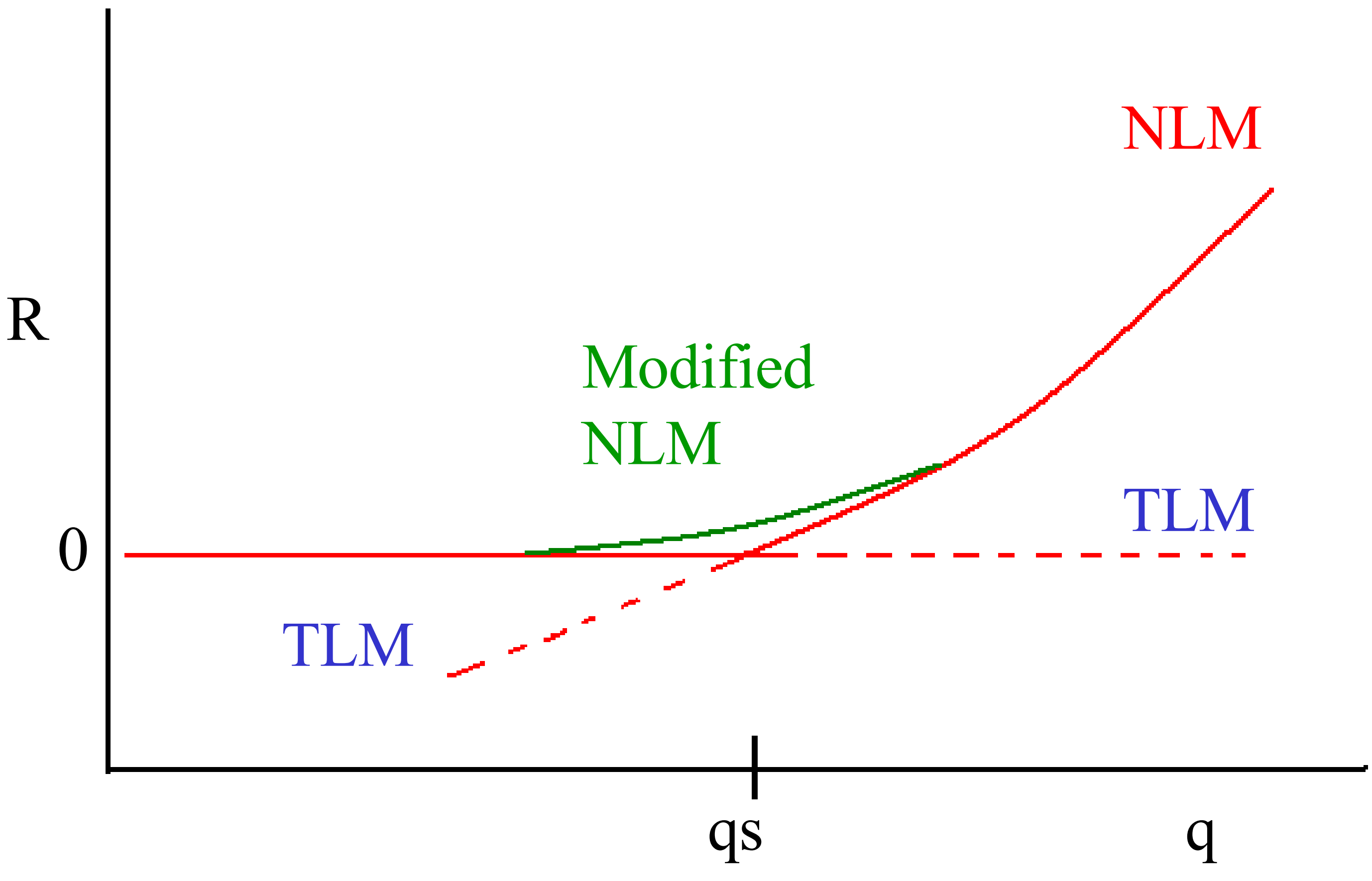
Problems with Physics

1. The model may be non-differentiable.
2. Unrealistic discontinuities should be smoothed after reconsideration of the physics being parameterized.
3. Perhaps worse than discontinuities are numerical instabilities that can be created from physics linearization.
4. It is possible to test the suitability of physics components for adjoint development before constructing the adjoint.
5. Development of an adjoint provides a fresh and complementary look at parameterization schemes.

Errico, R. M., and K. D. Raeder, 1999: An examination of the accuracy of the linearization of a mesoscale model with moist physics. *Quart. J. Roy. Meteor. Soc.*, **125**, 169-195.

Problems with Physics

Consider Parameterization of Stratiform Precipitation



Summary

1. Adjoint models estimate the sensitivities (gradient) of a single measure of output (J) with respect to all input values simultaneously and efficiently.

4. Adjoint models have revealed new aspects of dynamics, requiring a paradigm shift.

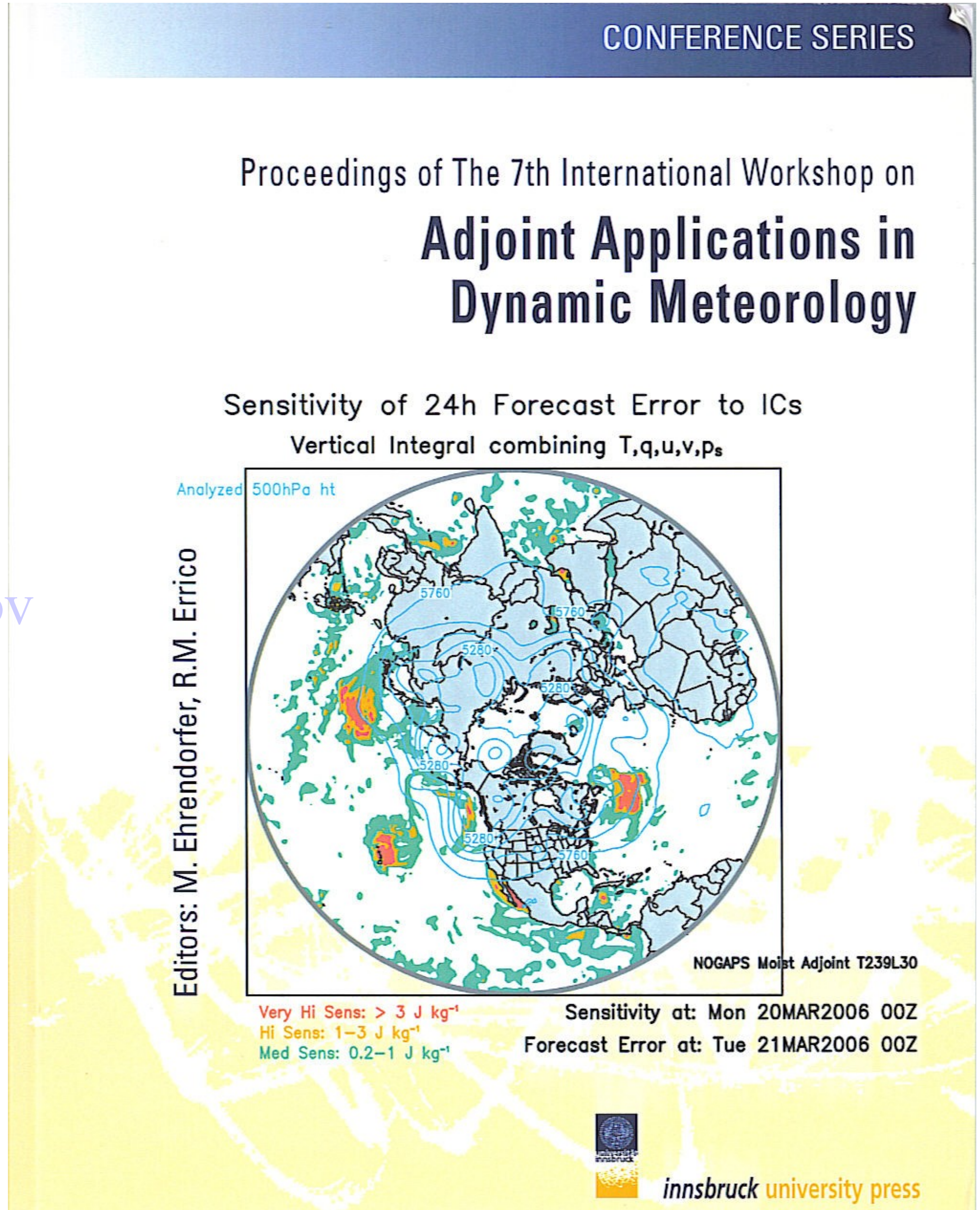
7. The development of an adjoint model from a nonlinear model is generally straightforward, but the results may be unsatisfactory, and approximations or model modifications may be required.

11. Adjoint models are powerful tools that remain underutilized, awaiting application to many problems inherently concerned with sensitivity.

Adjoint Workshop

The 8th will be in fall 2008 or spring 2009.

Contact Dr. R. Errico
rerrico@gmao.gsfc.nasa.gov
to be put on mailing list



Sensitivity to Observations

Estimating the Error Reduction Due to Observations

Consider an analysis increment $\delta\mathbf{x} = \mathbf{x}_a - \mathbf{x}_b = \mathbf{K}\delta\mathbf{y}$

Determined from background $\mathbf{x}_b$ and innovation $\delta\mathbf{y} = \mathbf{y}_o - H(\mathbf{x}_b)$

Consider 2 forecasts, $\mathbf{x}_a^f = \mathbf{m}(\mathbf{x}_a)$ and $\mathbf{x}_b^f = \mathbf{m}(\mathbf{x}_b)$ valid at the same time and a forecast verification data set $\mathbf{x}_v$

Consider a measure of forecast error $e = (\mathbf{x}^f - \mathbf{x}_v)^T\mathbf{E}(\mathbf{x}^f - \mathbf{x}_v)$

Consider the 3rd–order Taylor series expression for $e(\mathbf{x}_a) - e(\mathbf{x}_b)$:

$$\begin{aligned}
\delta_3 e &= (\delta\mathbf{x})^T\left[\mathbf{M}_b^T\left(\mathbf{x}_b^f - \mathbf{x}_v\right) + \mathbf{M}_a^T\left(\mathbf{x}_a^f - \mathbf{x}_v\right)\right] \\
&= (\delta\mathbf{y})^T\mathbf{K}^T\left[\mathbf{M}_b^T\left(\mathbf{x}_b^f - \mathbf{x}_v\right) + \mathbf{M}_a^T\left(\mathbf{x}_a^f - \mathbf{x}_v\right)\right]
\end{aligned}$$

Using $(\delta\mathbf{x})^T\mathbf{g} = (\mathbf{K}\delta\mathbf{y})^T\mathbf{g} = (\delta\mathbf{y})^T\mathbf{K}^T\mathbf{g} = (\delta\mathbf{y})^T\tilde{\mathbf{g}};\ \ \tilde{\mathbf{g}} = \mathbf{K}^T\mathbf{g}$

This is of the form $\delta_3 e = \sum_i \delta y_i\ \tilde{g}_i$

Langland and Baker 2004; Errico 2007 *Tellus*; Gelaro, Zhu and Errico 2007 *Meteorol. Z.*