Issues Regarding the Assimilation of Precipitation Observations

Dr. Ronald M. Errico

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

Outline

1. A general approach
2. Issues peculiar to precip. assimilation
3. Examples of some issues
4. Examples of neglected fundamentals
5. Summary and recommendations

A Bayesian Approach

Information from observations: $\rho_o(\mathbf{y}^o|\mathbf{y}^d)$

Information from models $\rho_m(\mathbf{y}^d|\mathbf{H}(\mathbf{x}))$

Information from prior $\rho_p(\mathbf{x}|\mathbf{x}^b)$

$$\rho_a(\mathbf{x}|\mathbf{x}^b,\mathbf{y}^o,\mathbf{H}) = \text{const} \times \rho_p(\mathbf{x}|\mathbf{x}^b) \int_Y \rho_o(\mathbf{y}^o|\mathbf{y}^d)\, \rho_m(\mathbf{y}^d|\mathbf{H}(\mathbf{x}))\, d\mathbf{y}^d$$

If Gaussian input statistics, then Bayesian result is:

$$\rho_a(\mathbf{x}|\mathbf{x}^b,\mathbf{y}^o,\mathbf{H}) = \text{const} \times \exp\left[-\frac{1}{2}J(\mathbf{x})\right]$$

where

$$\begin{aligned} J(\mathbf{x}) = {} & \left[\mathbf{x}-\mathbf{x}^b\right]^T \mathbf{B}^{-1} \left[\mathbf{x}-\mathbf{x}^b\right] \\ & + \left[\mathbf{H}(\mathbf{x})-\mathbf{y}^o)\right]^T (\mathbf{E}+\mathbf{F})^{-1} \left[\mathbf{H}(\mathbf{x})-\mathbf{y}^o\right] \end{aligned}$$

Implications of the Bayesian Approach

1. Unless the underlying distributions are simple, the problem is computationally intractable for large problems.
2. We see how the different information should be optimally combined.
3. We see what statistical knowledge is required as input.
4. We see that **E** and **F** may be equally important.
5. Results may depend on shapes of distributions, not only their means and variances.
6. We see that selection of a “best” analysis can be somewhat ambiguous.
7. Multi-modality of the PDF can occur, particularly due to model non-linearity.
8. While an explicit Bayesian approach may be impractical, the Bayesian implications of other techniques should be considered.

Examples of some issues peculiar to precipitation assimilation

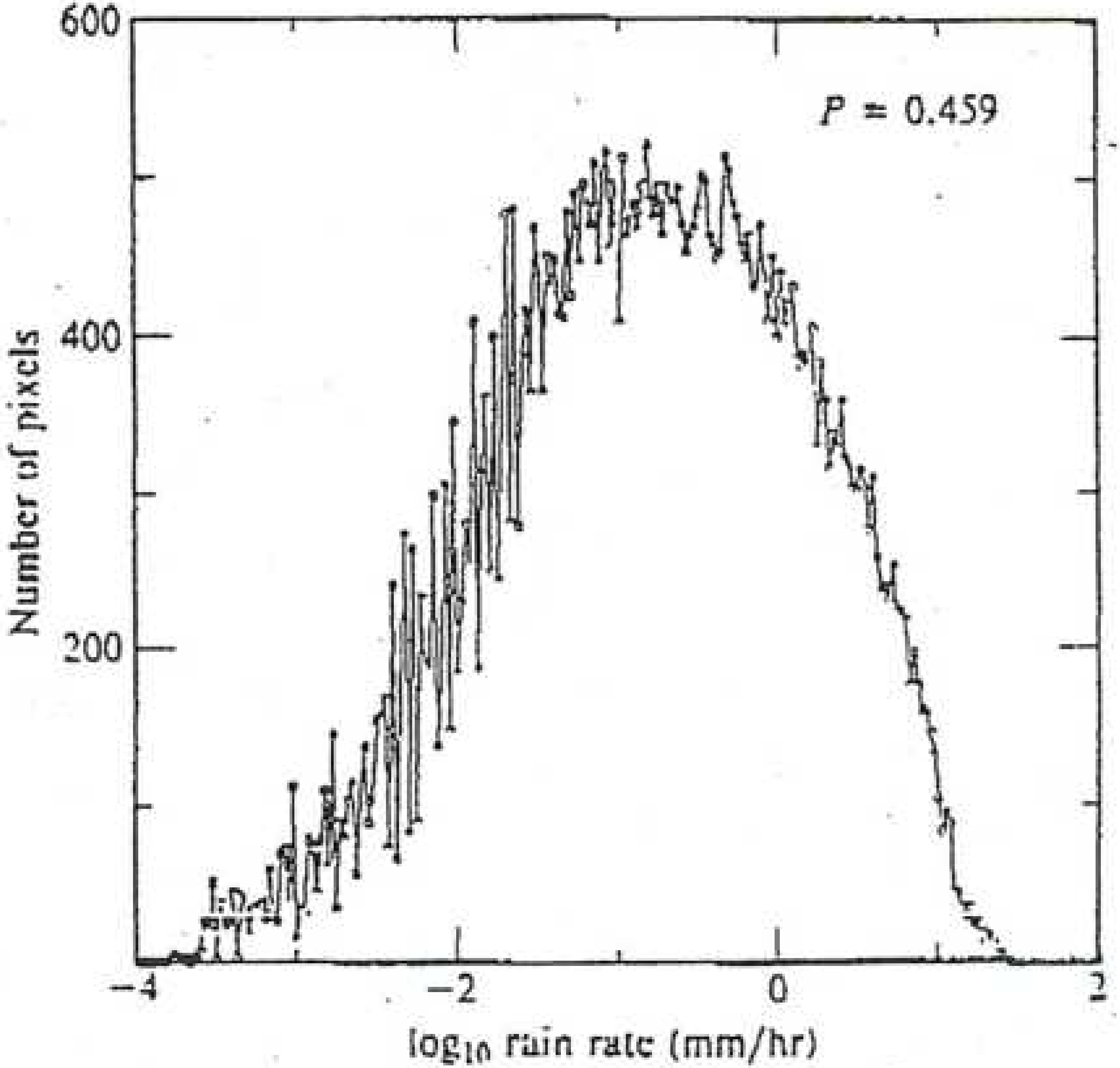


FIG. 2. A histogram of $\log_{10}$ of the rain rate obtained from a large number of 40 × 40 km² GATE pixels.

The distribution of precipitation

Kedem and Chiu 1986

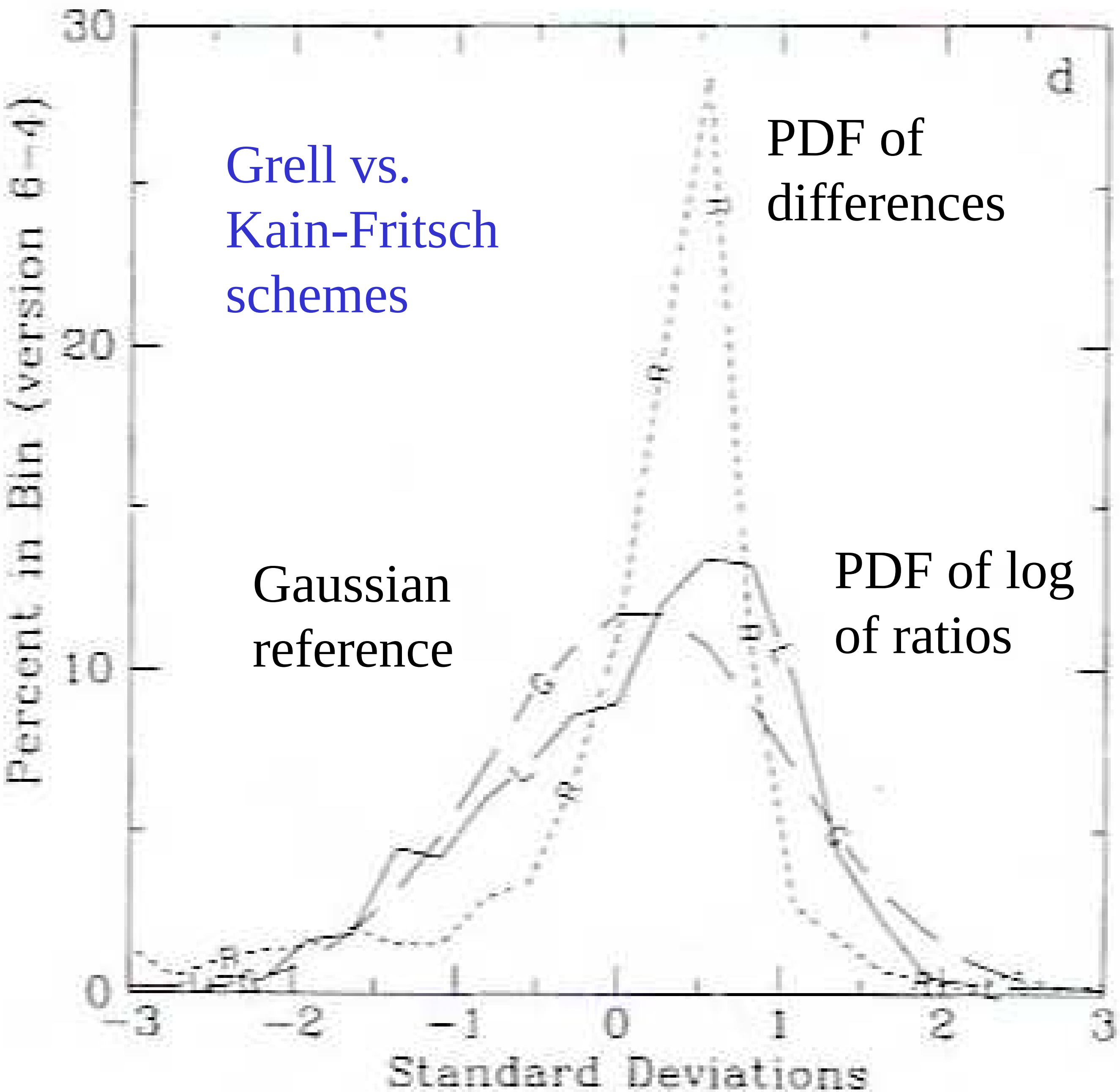
Kain - Fritsch

Example of Model Error:
Errico et al. *QJRMS* 2001

6-hour accumulated precip.
With 3 versions of MM5
Contour interval 1/3 cm

Betts - Miller

Grell



PDFs of
model
“errors”

Errico et al.
2001 *QJRMS*

A possible problem with use of logs in the cost function

Treating fits to precipitation observations P_i^o like other observations, we have as a contribution to the cost function

$$J_{\text{Precip}} = \sum_i \sigma_i^{-2}[P_i^o - H_i(\mathbf{x})]^2$$

An alternative treatment is

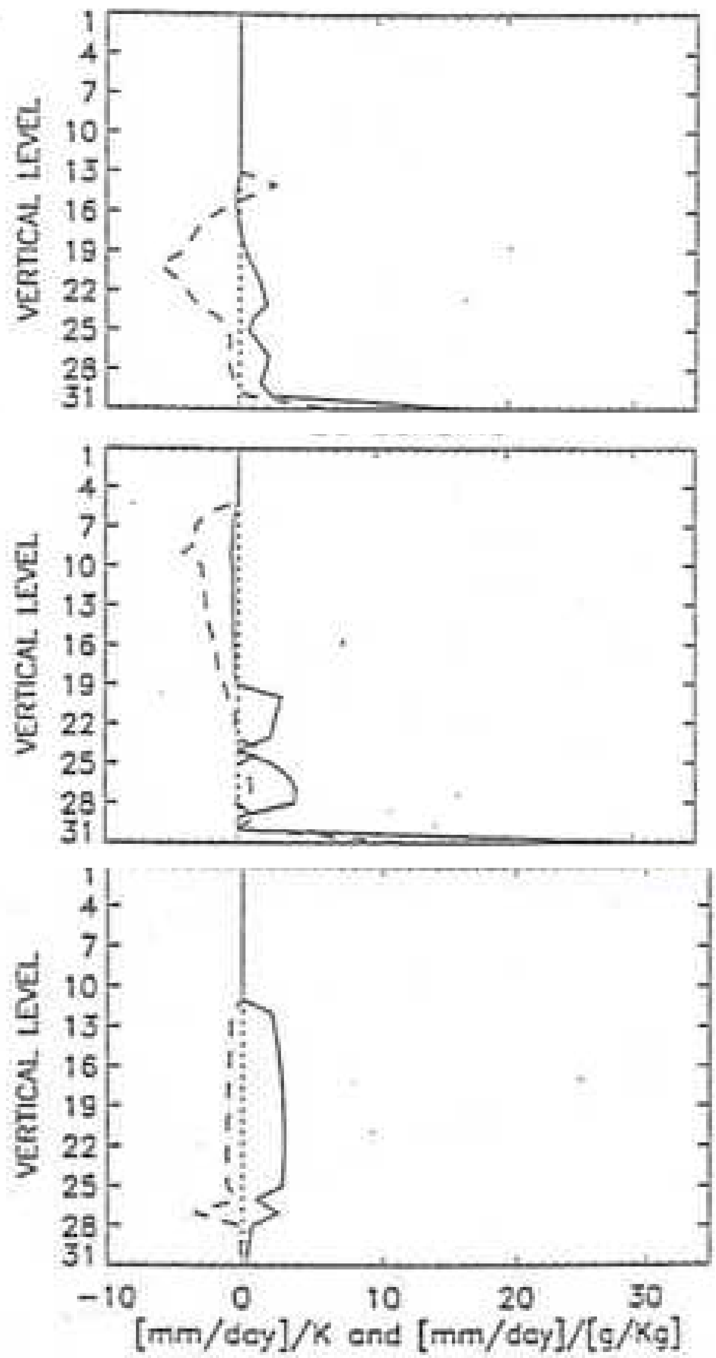
$$J_{\text{Precip}} = \sum_i w_i^{-2}[\ln P_i^o - \ln H_i(\mathbf{x})]^2$$

$$= \sum_i w_i^{-2}\left[\ln \frac{P_i^o}{H_i(\mathbf{x})}\right]^2$$

Now, consider two observations, one with $P_i^o = 10$, $H_i(\mathbf{x}) = 5$, the other with values .01, .005.

Note that special treatment of 0 values are required.

Jacobians of Precipitation



RAS scheme

ECMWF scheme

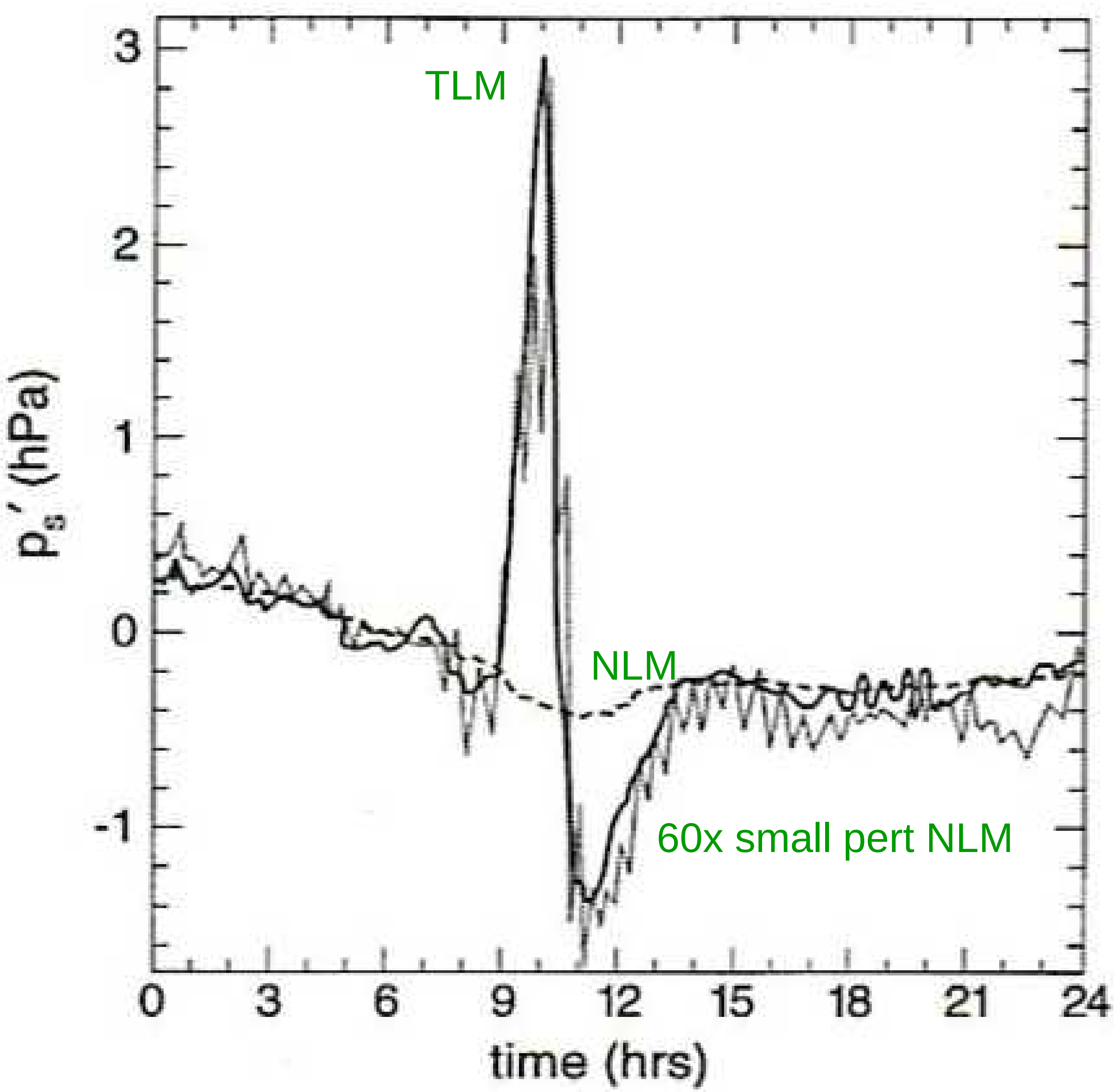
BM scheme

$$\frac{\partial R}{\partial T} \text{ dashed}$$

$$\frac{\partial R}{\partial q} \text{ solid}$$

Fillion and Mahfouf 1999 *MWR*

Tangent linear vs. nonlinear model solutions

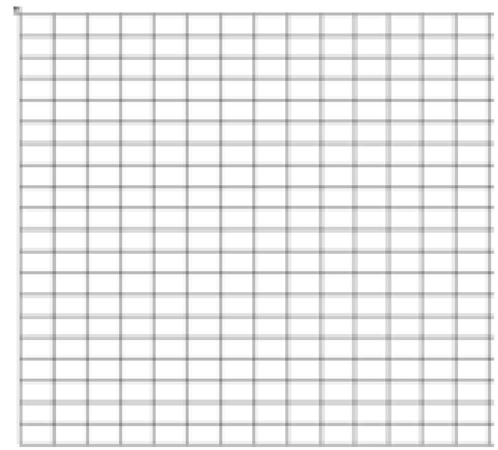


Errico and Raeder 1999 *QJRMS*

Statistically-Based Sub-Grid Parameterization

Model for small volume of mass Δm:

$$r_i = \begin{cases} a_i[q_i - q_s(T_i, p_i)] + \epsilon_i & \text{if } q_i > q_s(T_i, p_i) \\ \epsilon_i & \text{otherwise} \end{cases}$$



Consider average over large volume V

$$\bar{r} = \frac{1}{I\Delta m} \sum_{i=1}^{I} r_i \Delta m$$

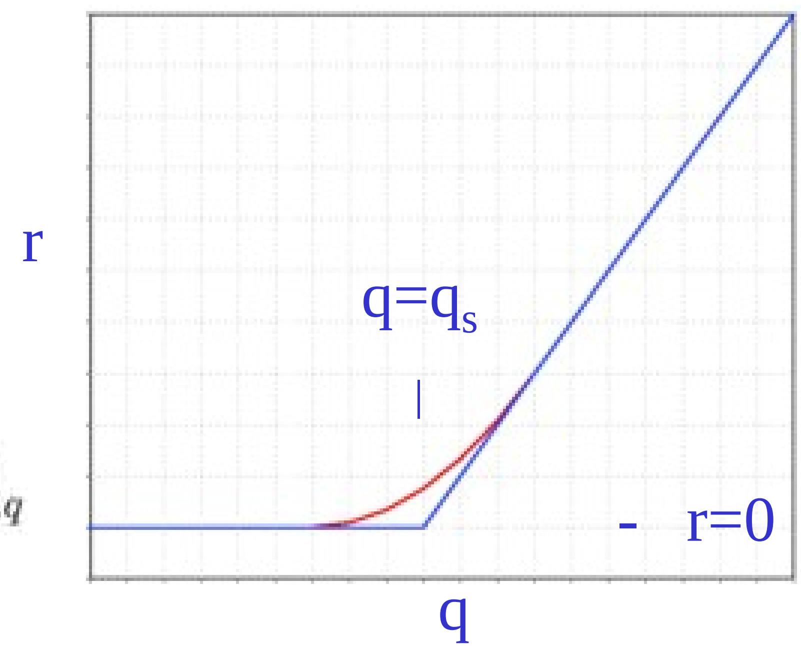
In general, $\bar{r} \neq r_m(\bar{q}, \bar{T}, \bar{p})$

Consider a uniform distribution of $q - q_s$ within V:

$$-\Delta q \leq q_i - q_s(T_i, p_i) \leq \Delta q$$

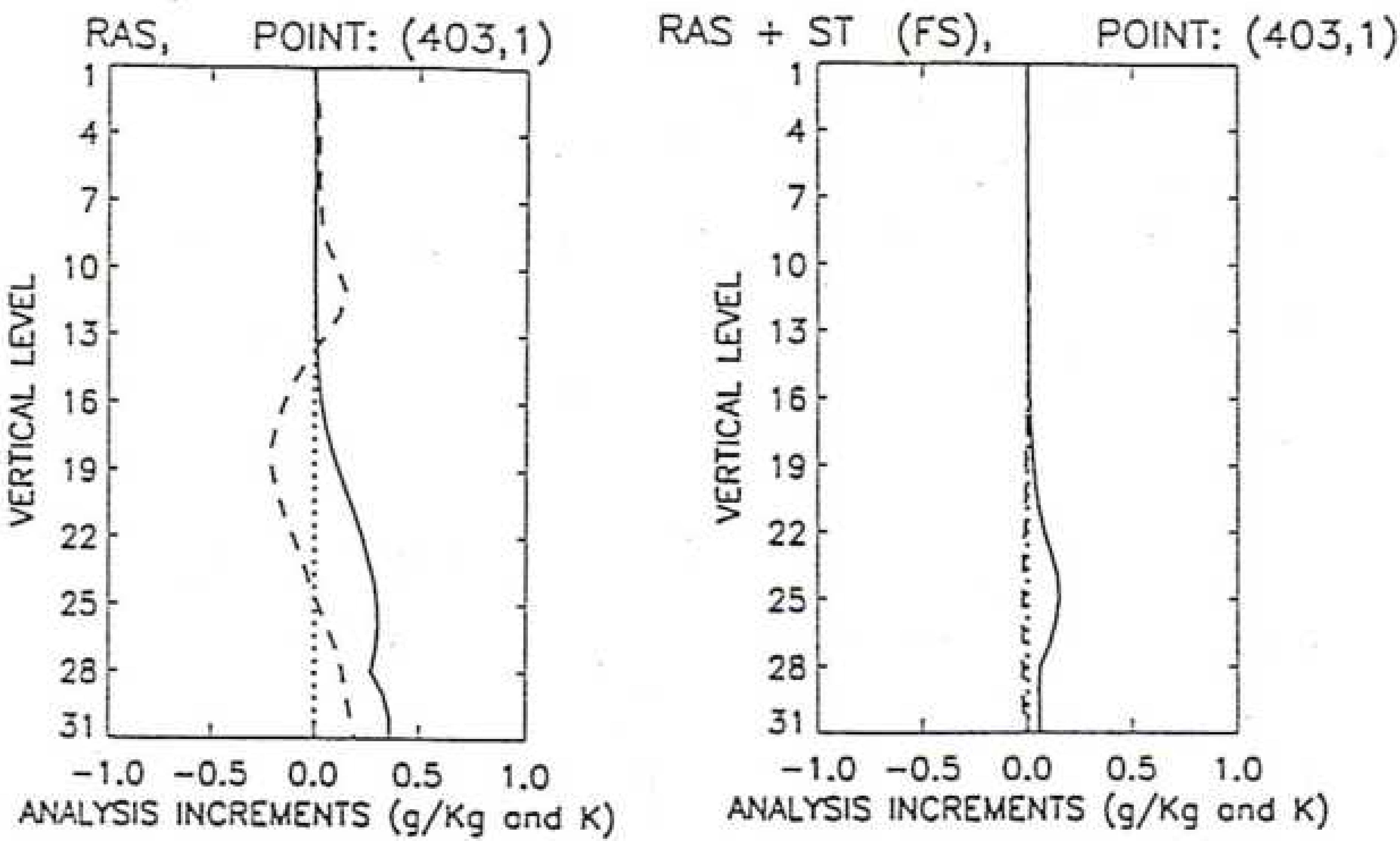
A new model:

$$r_{m*} = \begin{cases} a(\bar{q} - \overline{q_s}) & \text{if } \bar{q} - \overline{q_s} > \Delta q \\ 0 & \text{if } \bar{q} - \overline{q_s} < -\Delta q \\ \frac{a}{4\Delta q}(\bar{q} - \overline{q_s} + \Delta q)^2 & \text{otherwise} \end{cases}$$



Dependence on precipitation type

Fillion and Mahfouf 1999 *MWR*



Adjoint-derived, optimal perturbations

Errico, Raeder and Fillion, 2003 *Tellus*

Consider $J = J(\mathbf{x})$

Determine initial perturbation $\mathbf{x}'$ that maximizes:

$$J' = \left(\frac{\partial J}{\partial \mathbf{x}}\right)^{\mathrm{T}} \mathbf{x}'$$

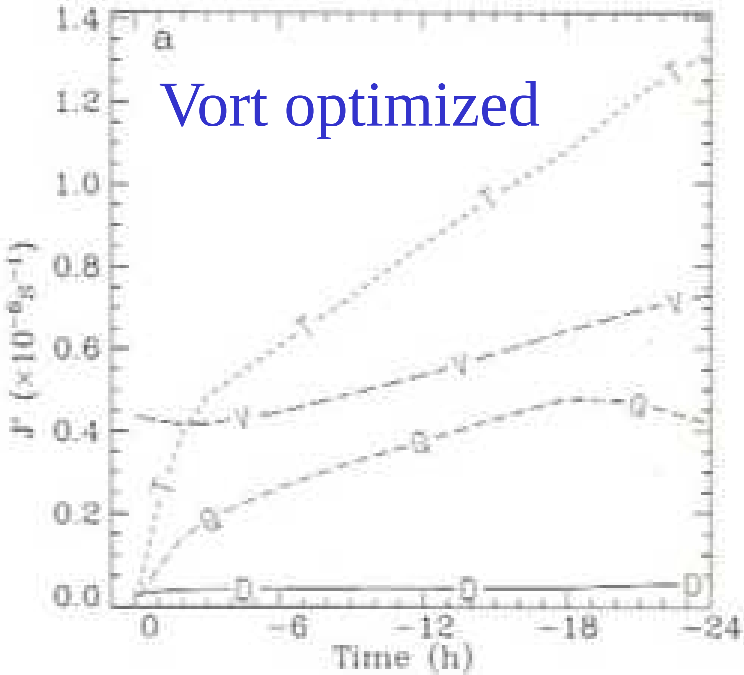
Given initial constraint:

$$C = \frac{1}{2}\mathbf{x}'^{\mathrm{T}}\mathbf{B}^{-1}\mathbf{x}'$$

Solution:

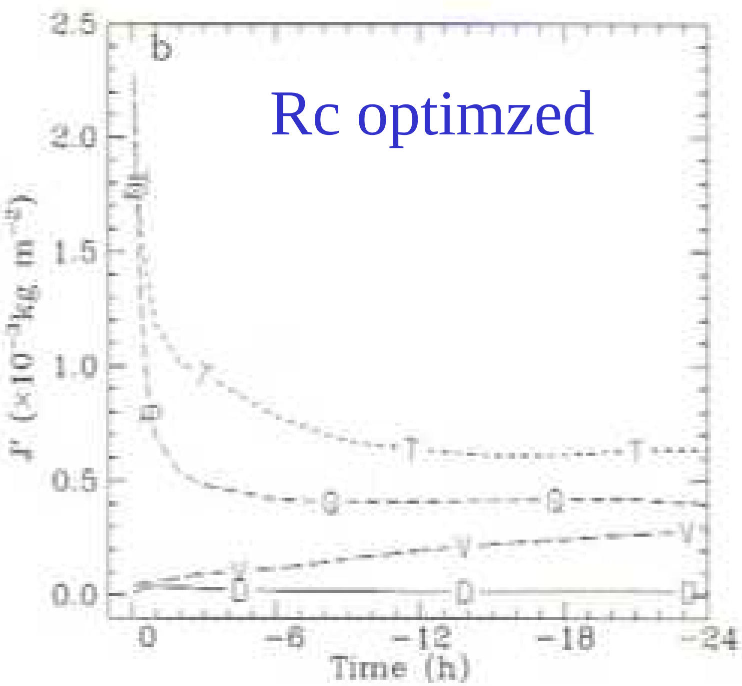
$$\mathbf{x}' = \lambda^{-1}\mathbf{B}\frac{\partial J}{\partial \mathbf{x}}$$

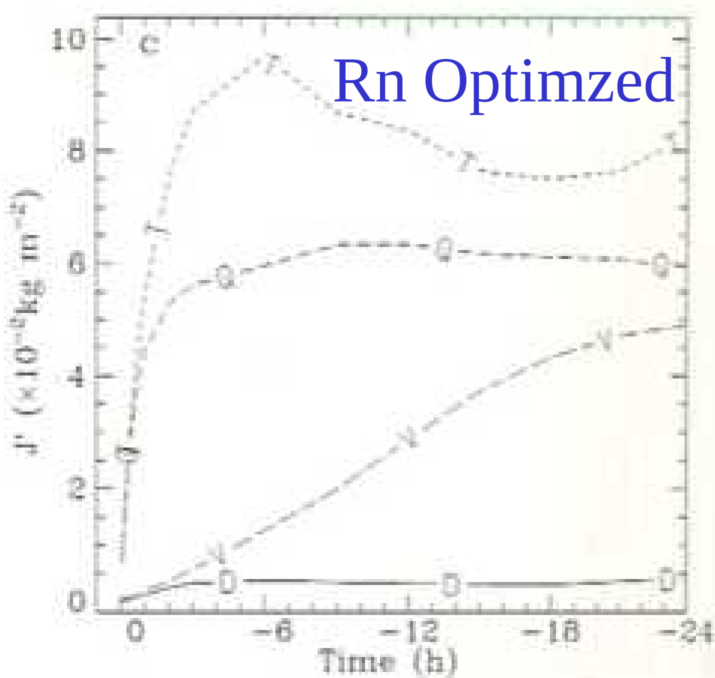
$$\mathrm{Max}(J') = \sqrt{2C\left(\frac{\partial J}{\partial \mathbf{x}}\right)^{\mathrm{T}} \mathbf{B}\frac{\partial J}{\partial \mathbf{x}}}$$



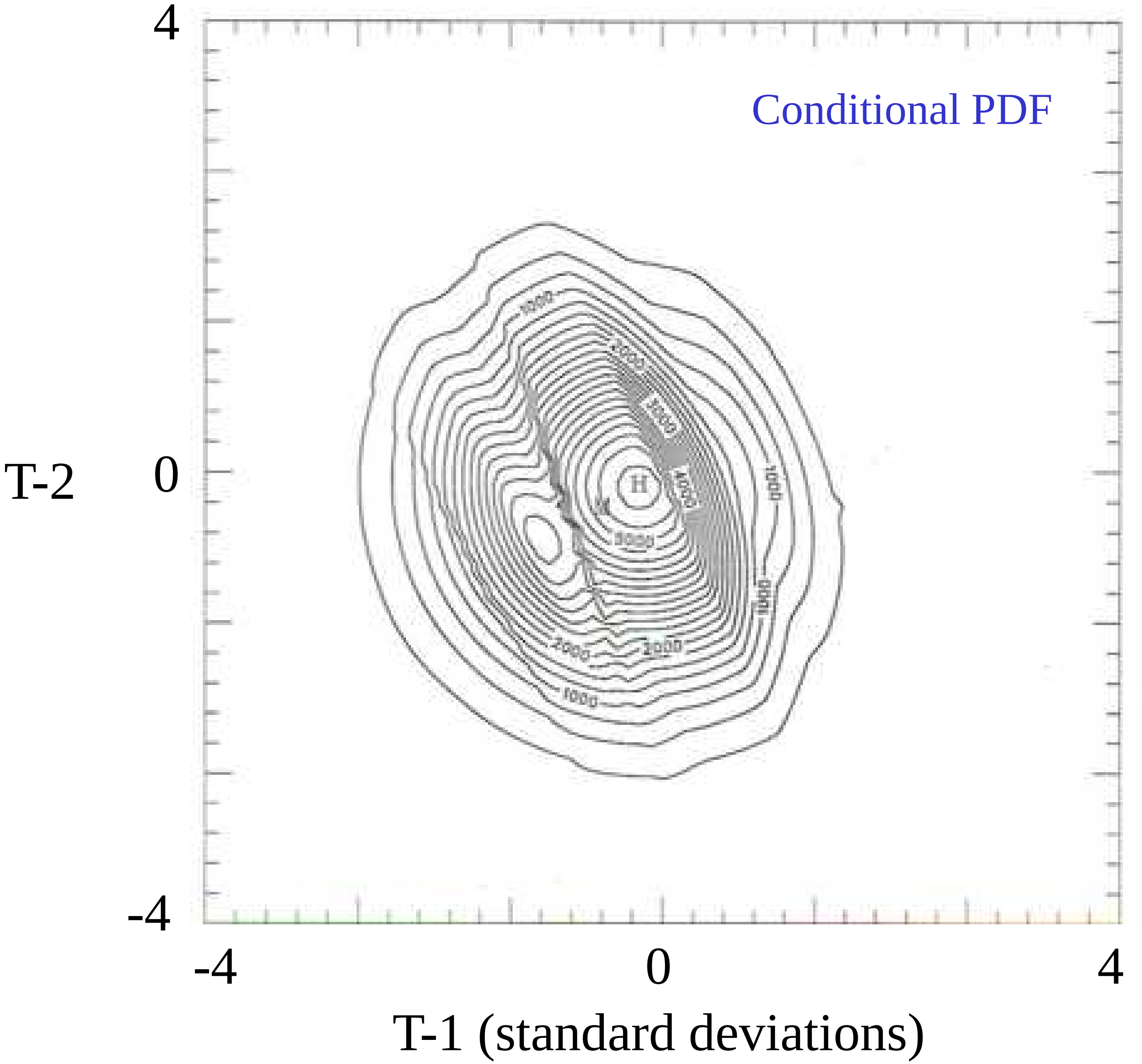
Errico et al.
2003 *Tellus*

Impacts for adjoint-derived optimal perturbations for forecasts starting indicated hours in the past.





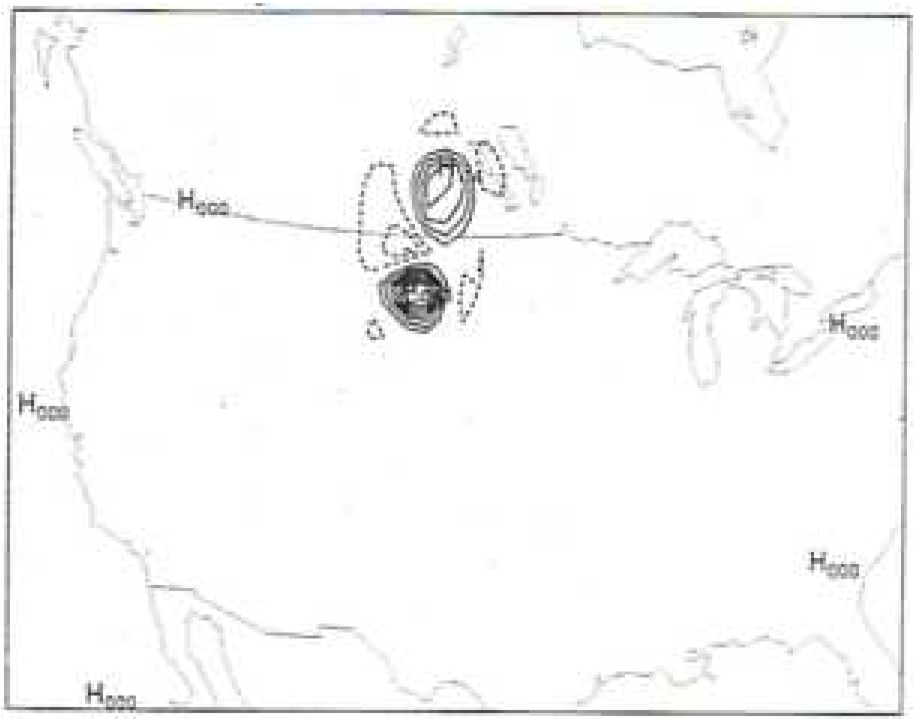
Posterior (analysis) PDF of 1DVAR of Convection



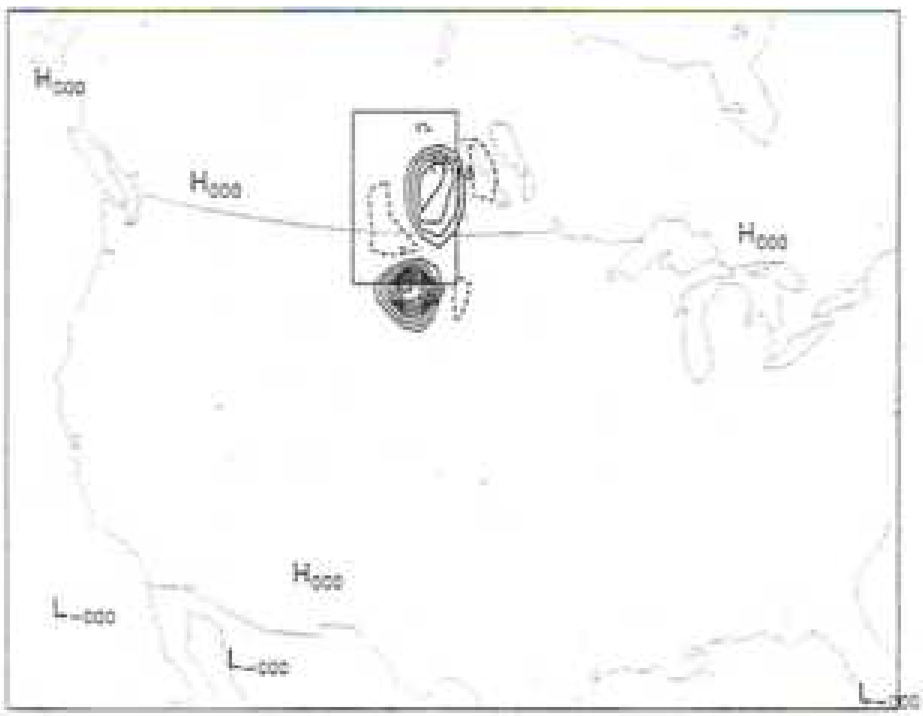
$$\partial(\text{Precipitation})/\partial T(500\text{hPa})$$

6-hour forecast

Full Field



Gravitational-mode field



Contour interval 0.0025 mm/K

Issues Regarding the Assimilation of Cloud and Precipitation Data

RONALD M. ERRICO

Goddard Earth Sciences and Technology Center, University of Maryland, Baltimore County, Baltimore, and Global Modeling and Assimilation Office, National Aeronautics and Space Administration, Greenbelt, Maryland

PETER BAUER

European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom

JEAN-FRANÇOIS MAHFOUF

Environment Canada, Dorval, Quebec, Canada

(Manuscript received 25 January 2006, in final form 26 June 2006)

ABSTRACT

The assimilation of observations indicative of quantitative cloud and precipitation characteristics is desirable for improving weather forecasts. For many fundamental reasons, it is a more difficult problem than the assimilation of conventional or clear-sky satellite radiance data. These reasons include concerns regarding nonlinearity of the required observation operators (forward models), nonnormality and large variances of representativeness, retrieval, or observation–operator errors, validation using new measures, dynamic and thermodynamic balances, and possibly limited predictability. Some operational weather prediction systems already assimilate precipitation observations, but much more research and development remains. The apparently critical, fundamental, and peculiar nature of many issues regarding cloud and precipitation assimilation implies that their more careful examination will be required for accelerating progress.

Review of state-of-the-art at ECMWF

QUARTERLY JOURNAL OF THE ROYAL METEOROLOGICAL SOCIETY
Q. J. R. Meteorol. Soc. **134**: 1513–1525 (2008)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/qj.304

RMetS
Royal Meteorological Society

Lessons learnt from the operational 1D+4D-Var assimilation of rain- and cloud-affected SSM/I observations at ECMWF

Alan J. Geer*, Peter Bauer and Philippe Lopez
European Centre for Medium-Range Weather Forecasts, Reading, UK

ABSTRACT: Rain- and cloud-affected Special Sensor Microwave/Imager (SSM/I) observations are assimilated operationally at the European Centre for Medium-Range Weather Forecasts (ECMWF). The four-dimensional variational analysis (4D-Var) assimilates total column water vapour (TCWV) derived from one-dimensional variational retrievals (1D-Var). From the SSM/I radiances, 1D-Var retrieves surface wind and the vertical profiles of temperature, humidity, cloud and precipitation. The main shortcoming of the '1D+4D-Var' technique is that, of all this information, only TCWV gets into the 4D-Var analysis. More information could be used: the rainwater path agrees well, in an instantaneous comparison, with observations from the precipitation radar on the Tropical Rainfall Measuring Mission. There are other issues, however: the simplified moist physics operators used in 1D-Var produce roughly twice the observed amount of rain, but the problem is masked by a sampling bias, which comes from applying 1D+4D-Var when the observations are cloudy or rainy, but not when the first guess is rainy or cloudy and the observations are clear. The shortcomings of 1D+4D-Var will be addressed by moving to a direct 4D-Var assimilation which includes all SSM/I observations, whether clear, cloudy or rainy, in the same stream.

KEY WORDS precipitation; numerical weather prediction; microwave radiance; satellite

Received 19 March 2008; Revised 9 July 2008; Accepted 11 July 2008

1. Introduction

In terms of both modelling and observations, it is perhaps in the area of clouds and precipitation that numerical weather prediction (NWP) is least well developed.

which derives the clear signal in areas of scattered cloud, and that of Pavelin *et al.* (2008), which infers information above cloud tops. Cloud clearing is important in the infrared (IR) where as few as 5% of observations can be

The apparent neglect of many fundamentals

Few statistical considerations

- background estimates ignored
- background error correlations ignored
- observations considered too accurate (and Gaussian)
- representativeness (forward model) error ignored

Few balance considerations

- univariate error statistics
- unbalanced reference states

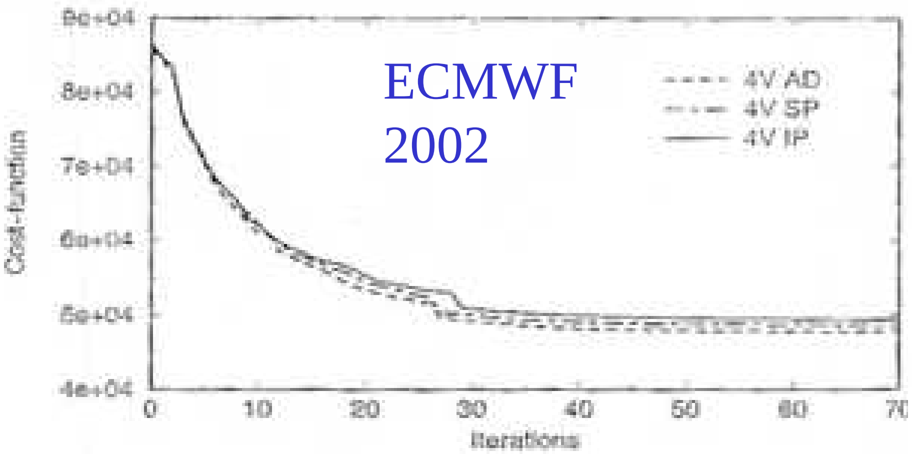
Limited evaluation

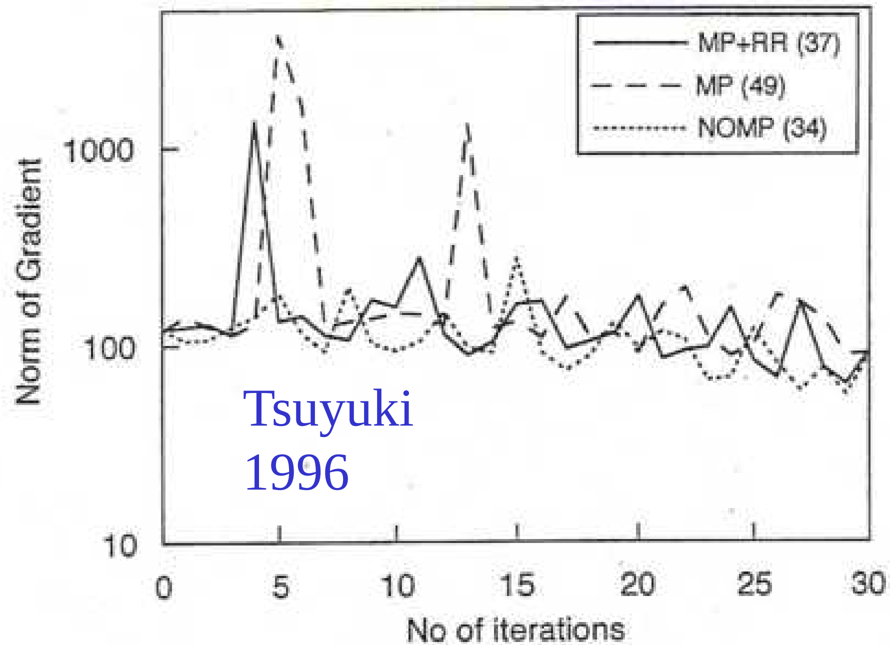
- limited cases
- limited measures

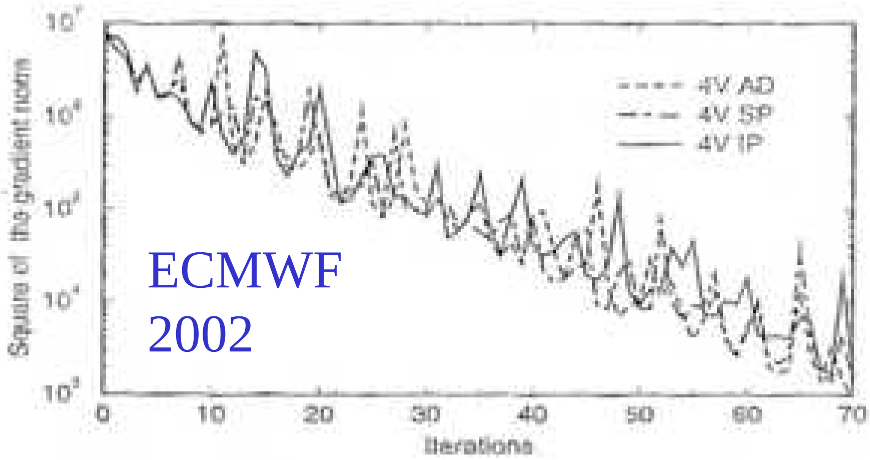
Some strange results

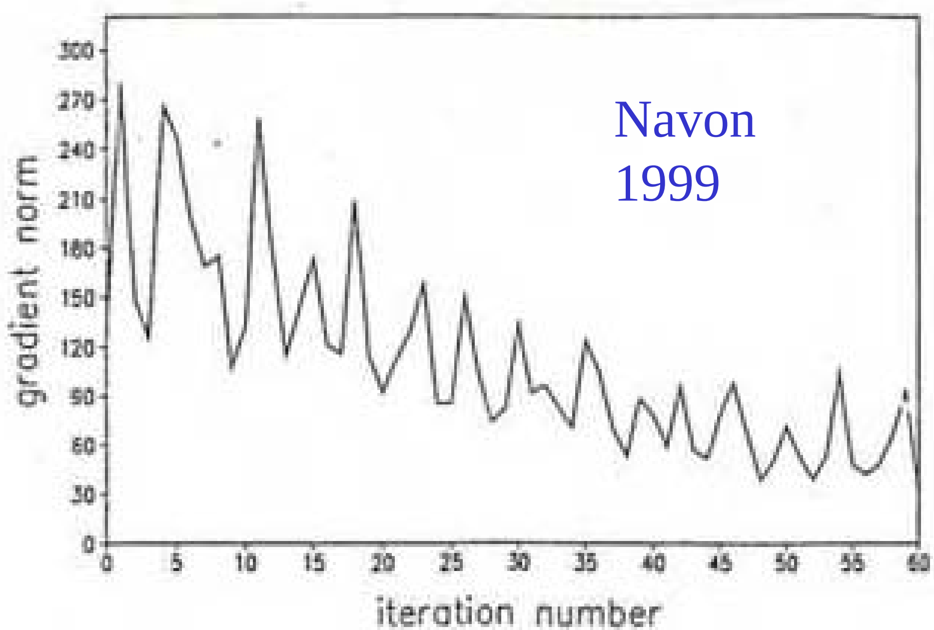
- ultra rapid convergence rates
- mis-characterization of sizes of terms
- little decrease of norm of *J*-gradient

Convergence of 4DVAR









Summary

1. I am confused!
2. How can so many apparently fundamental aspects of the problem be neglected, yet such good results be reported?
3. Only in rare cases is enough information provided to help explain question 2.
4. Since only an improvement over some baseline is required, it is not necessary that “correct” procedures are used, just “useful” ones.
5. Successes may reveal more about the baseline results than about the correctness of a new assimilation procedure.
6. With both observation and forward model errors likely very large, what should be a realistic expectation of the usefulness of precipitation information and how can this be realized?

Recommendations

1. Be skeptical.
2. Ask lots of questions.
3. Consider Bayesian implications.
4. Determine reasonable error estimates.
5. Estimate what issues are generally important.
6. Explain results.
7. Encourage research at research institutions.
8. Entrain some interested experts.