

# **Analysis of Systematic Errors in Climate Products**

**John Bates**

**NOAA's National Climatic Data Center**

**Thanks: John Eyre, Lei Shi, Dongsoo Kim**



# Outline

- Introduction – Best practices for generation of Climate Data Records (CDRs)
- Analytic assessment - Systematic errors in retrievals using the Eyre method
- Application – Retrieval of mean tropospheric temperatures using MSU and false trend identification
- Conclusions

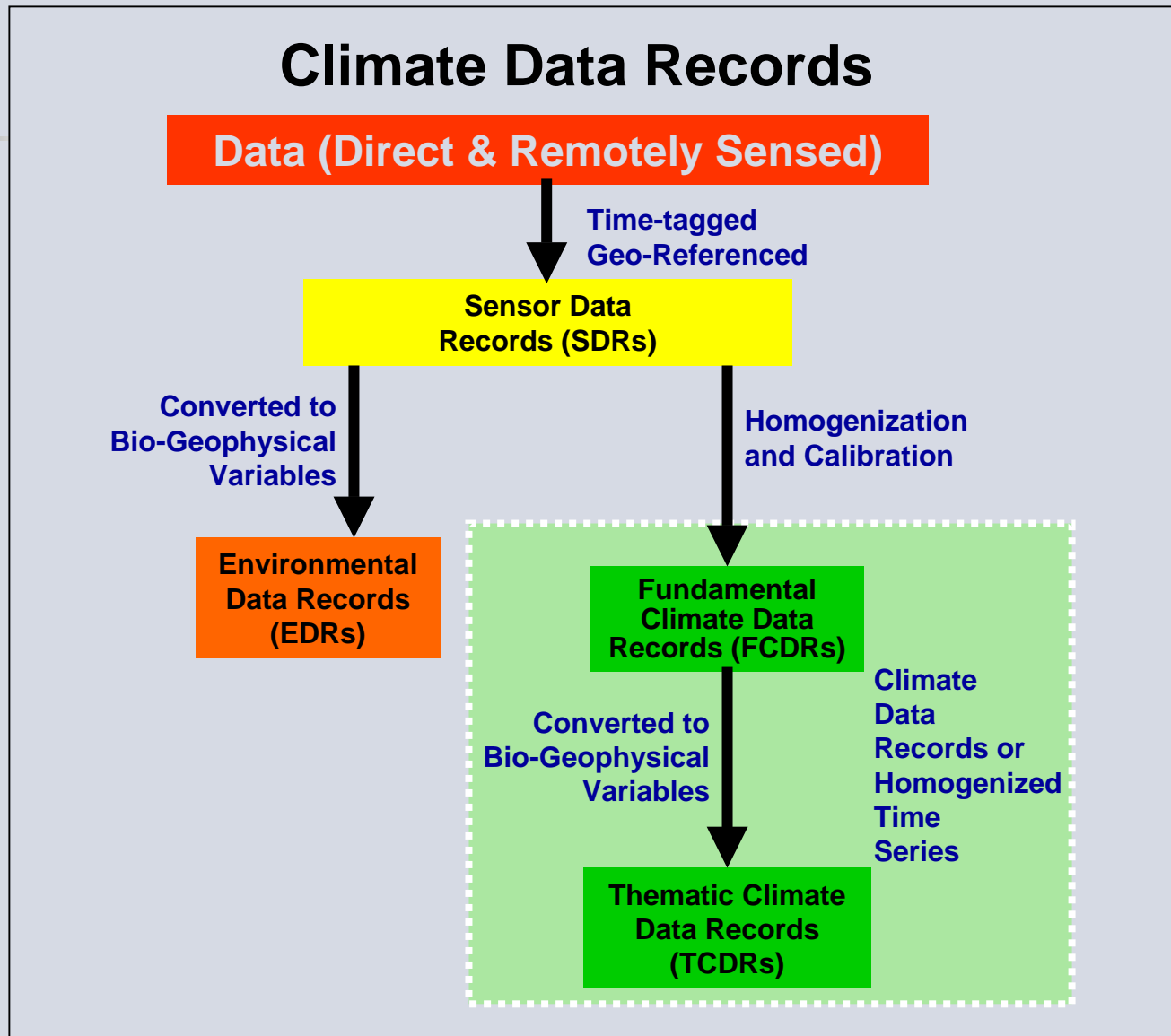


# Defining Climate Data Records (CDRs)

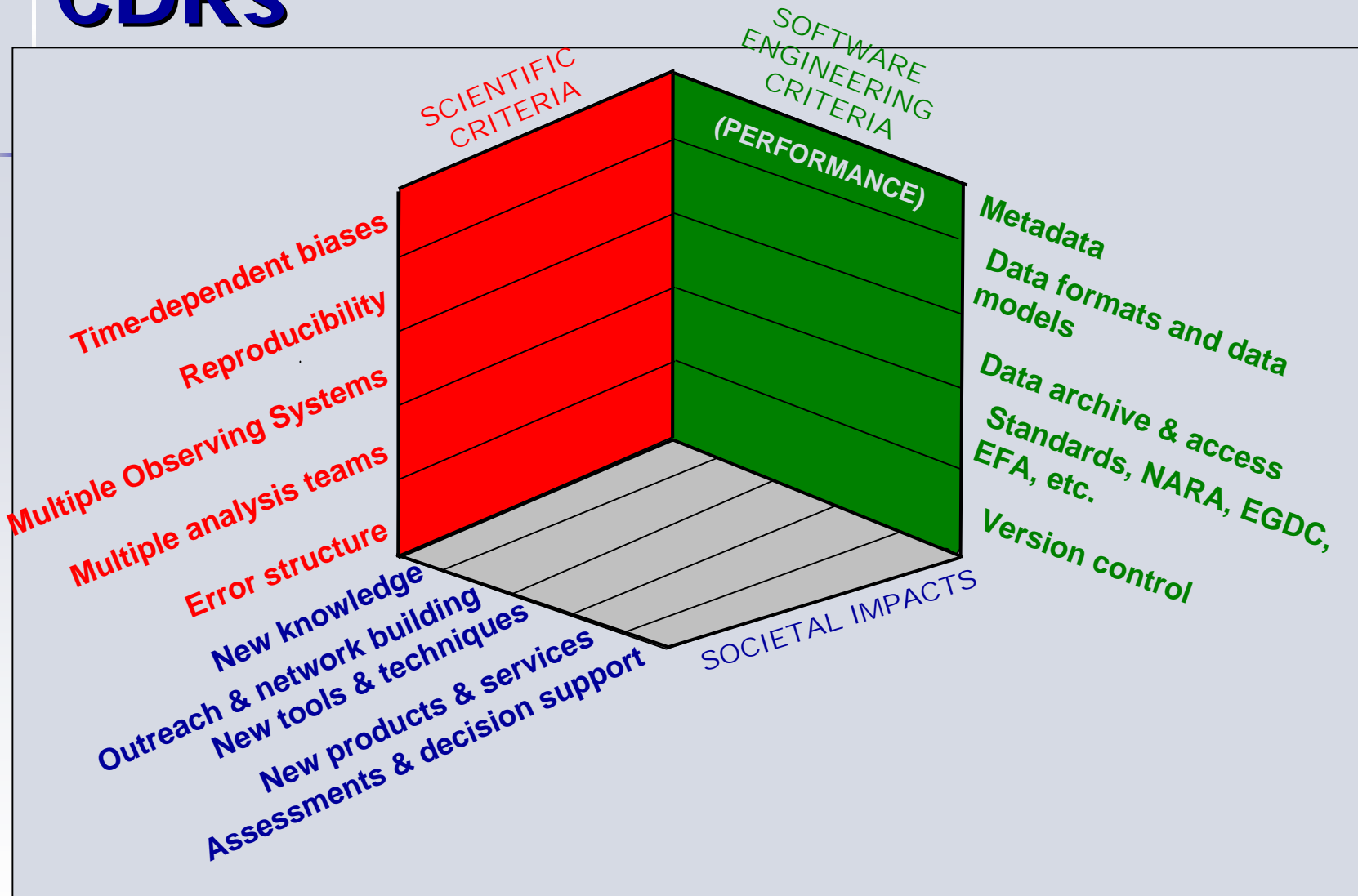
- A CDR is a time series of sufficient length, consistency, and continuity to determine climate variability and change.
- Fundamental CDRs (FCDRs) are sensor data (e.g., calibrated radiances, brightness temperatures, radar backscatter) that have been improved and quality-controlled over time, together with the ancillary data used to calibrate them.
- Thematic CDRs (TCDRs) are geophysical variables derived from the FCDRs, specific to various disciplines, and often generated by blending satellite observations, in-situ data, and model output.



# Defining CDRs (cont.)



# Performance Measures for CDRs



# Analytic assessment - Systematic errors in retrievals using the Eyre method

Following the derivation by Rogers and further work by Eyre, we view the direct inversion solution to the radiative transfer equation as fundamentally an ill-posed problem and, instead, seek to optimally solve an alternative problem. That is, we have a measurement system, in this case remote sensing observations from space, that provides information (with error) and prior information (with error) and we seek to find the optimal combination of the two:

$$\hat{x} - x_0 = W \bullet (y_m - y_c \{x_0\}) \quad (1)$$

where,

$\hat{x}$  is the vector of retrieved parameters

$x_0$  is the first-guess value of the vector

$y_m$  is the vector of measurements

$y_c \{x_0\}$  is the corresponding prior information vector

$W$  is a matrix operator



# Analytic assessment (cont.)

The matrix operator  $W$  can be obtained through a minimum variance solution via:

$$W = (K \cdot C)^T \cdot (K \cdot C \cdot K^T + E)^{-1} \quad (2)$$

where,

$C$  is the error covariance of the prior information,  $x_0$

$E$  is the error covariance of the measurements,  $y_m$

$K$  contains the partial derivatives of the measurements with respect to the profile evaluated at  $x_0$

superscripts T and  $-1$  denote matrix transpose and inverse



# Analytic assessment (cont.)

The linear approximation to the forward radiative transfer problem is:

$$y_m - y_c \{x_0\} = K \bullet (x_T - x_0) + \mathcal{E}_m \quad (3)$$

where,

$x_T$  is the vector of true geophysical parameters

$\mathcal{E}_m$  is the vector of measurement errors, assumed to be random, Gaussian, unbiased and includes unbiased errors in the forward radiative transfer model

Combining the forward and inverse radiative transfer equations and rewriting it in terms of the retrieval, first-guess and measurement errors yields,

$$\hat{x} - x_T = (I - R) \bullet (x_0 - x_T) + W \bullet \mathcal{E}_m \quad (4)$$

where I is the identity matrix and  $R = W \bullet K$

By taking large averages, as is done in climate studies,  $\mathcal{E}_m \rightarrow 0$  so the mean retrieval errors are determined by the R matrix and the mean errors in the first guess.





# Application to MSU Temperature Trends

- The MSU channels include contributions from both the troposphere and stratosphere
- Some researchers have suggested using linear regression of multiple channels to retrieve layer mean temperature and trends
- The systematic error involved in applying a priori information (regression coefficients generated from mean radiosonde profiles) has not been rigorously evaluated for these methods



# Application to MSU (cont)

We apply,

$$\hat{x} - x_T = (I - R) \bullet (x_0 - x_T) + W \bullet \varepsilon_n$$

To evaluate the following regression equations of Fu and Johnson (2005)

$$T_{TLT} = a_{23} T_2 + (1 - a_{23}) T_3, \quad (1a)$$

$$T_{TT} = a_{24} T_2 + (1 - a_{24}) T_4, \quad (1b)$$

Where,  $T_{TLT}$  is the mean lower tropical tropospheric temperature (scf-250hPa),

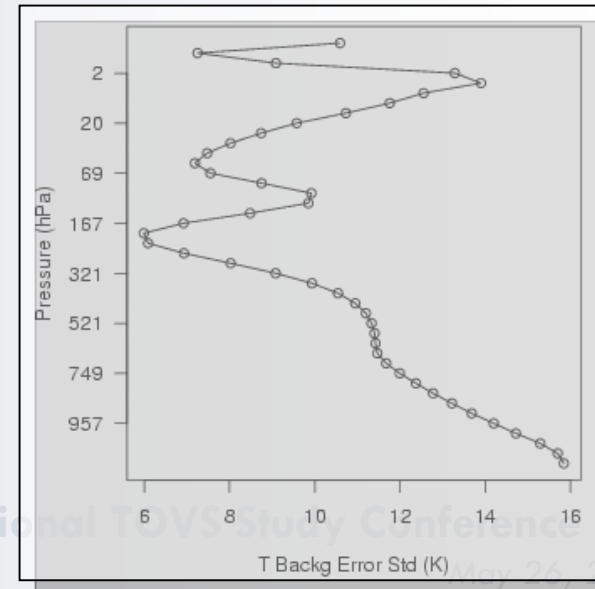
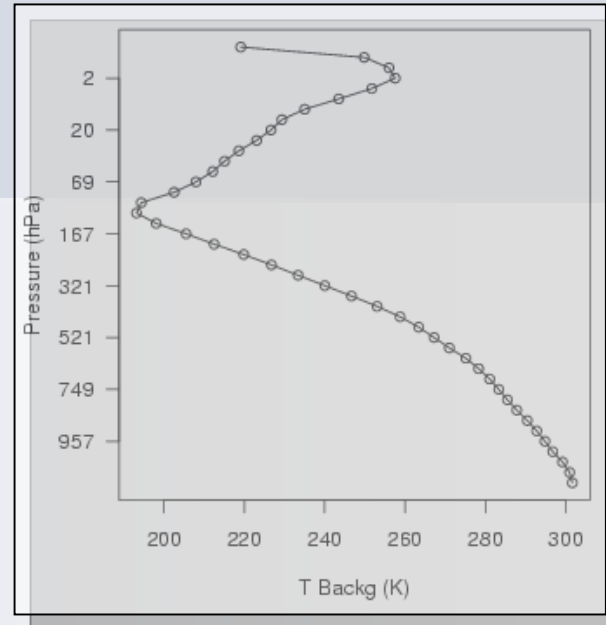
$T_{TT}$  is the mean tropical tropospheric temperature (scf-100hPa)

And  $T_2$ ,  $T_3$ , and  $T_4$  are MSU brightness temperatures for channels 2, 3, and 4



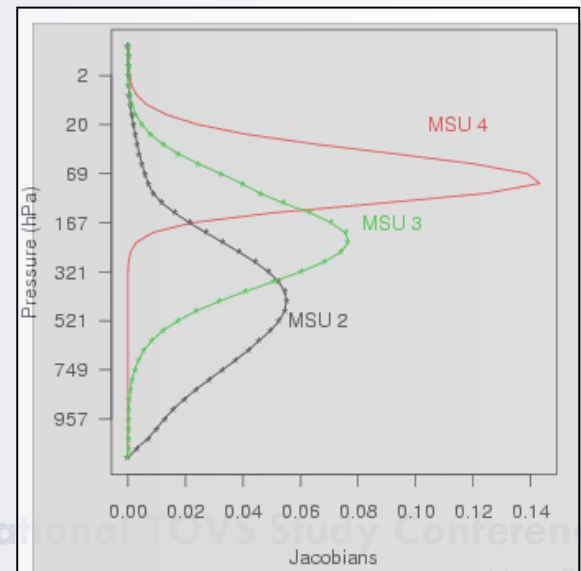
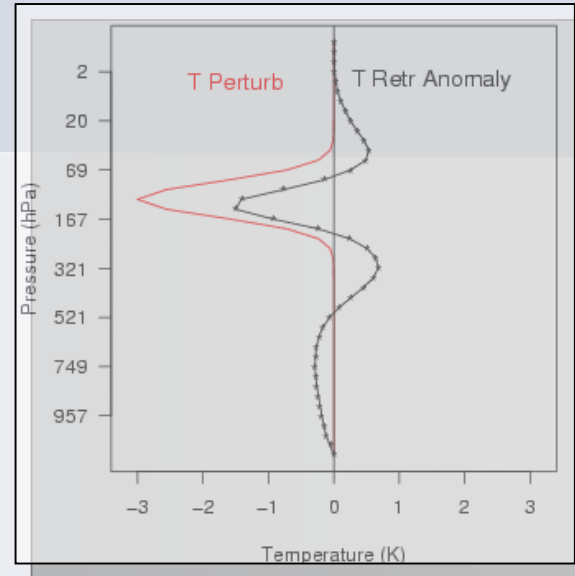
# Solving Equation 4

- Reference temperature profile of tropics used in simulations,  $x_7$  of Eq. 4.
- Temperature error standard deviations used in the computation of retrieval, that is diagonal component of matrix  $C$ .



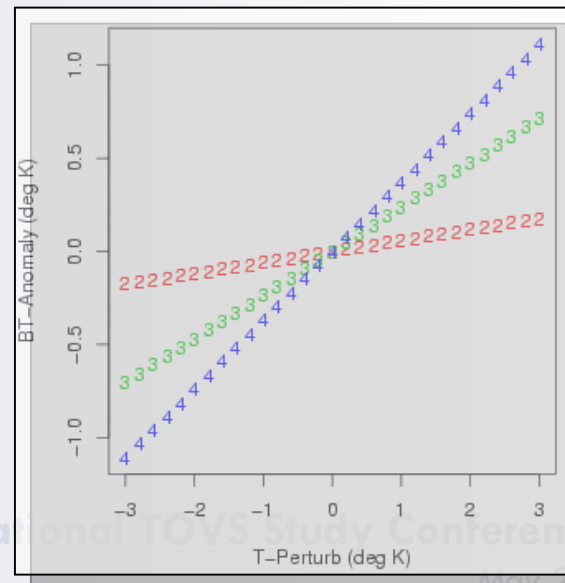
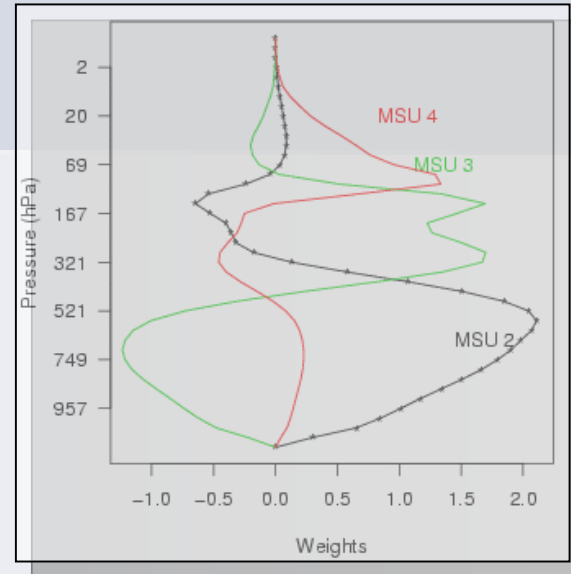
# Solving Equation 4 (cont)

- Perturbed temperature, that is,  $\mathbf{x}_o - \mathbf{x}_T$  of Eq. 4, and increment of retrieved temperature profiles, that is  $\mathbf{x}^{\wedge} - \mathbf{x}_T$  of Eq. 4. This example is at the 3 K cooling at the 100 hPa, and the retrieval anomaly is the result of applying  $(I - R)$  matrix in Eq. (4)
- $K$  matrices (Jacobian), whose elements are partial derivatives of radiances with respect to state vectors.



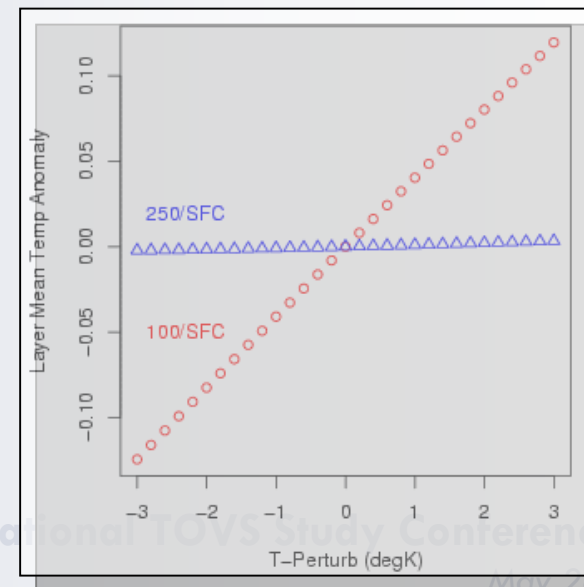
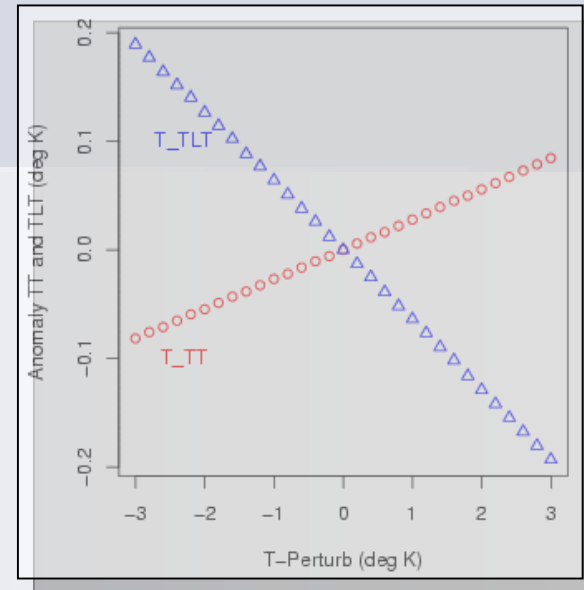
# Solving Equation 4 (cont)

- $W$  matrices corresponding to perturbed profile in Fig.3. The  $W$  matrix is function of background matrix  $C$ , Jacobian  $K$ , and measurement error  $E$  as seen in Eq. (2).
- Anomaly of three MSU simulated radiances (brightness temperatures) from the base state with respect to temperature perturbations. Numbers stand for MSU channels. Channels 3 and 4 have higher sensitivities than channel 2 as perturbation is made in the stratosphere.



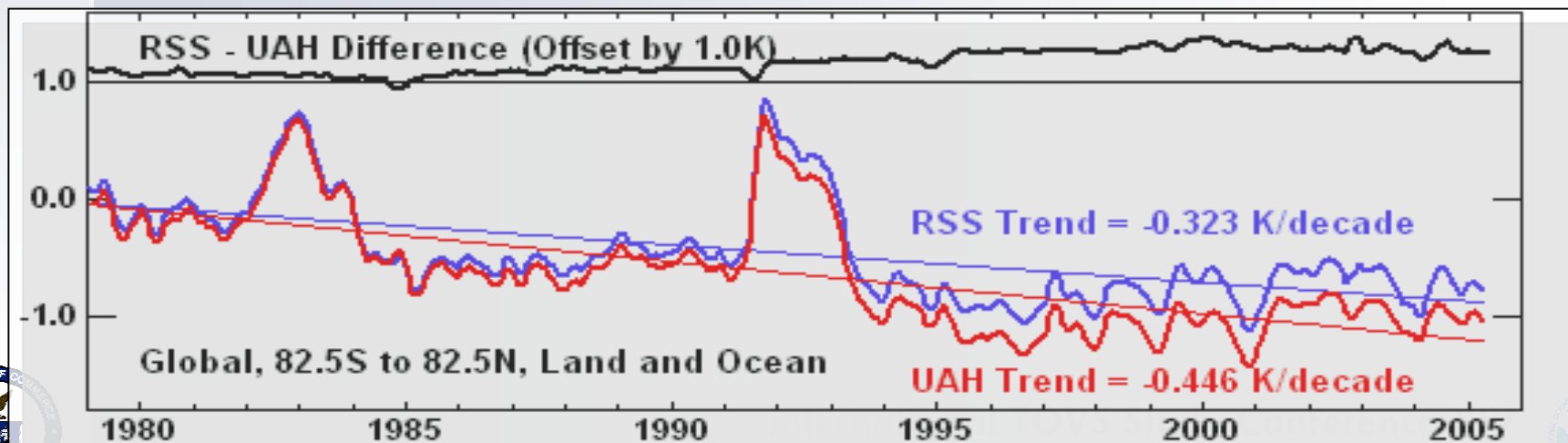
# Solving Equation 4 (cont)

- Empirically retrieved layer temperatures according to Equations 1a and 1b of Fu and Johanson.
- Layer mean temperature anomalies for layer of 100/SFC and 250/SFC with respect to stratospheric perturbations. The layer of 250 hPa – SFC is an equivalent parameter to compare with Fu and Johanson's parameter  $T\_TLT$ .
- Note that  $T\_TLT$  is opposite in sign and much greater value than the imposed physical anomaly (i.e., slope of  $-0.3825\text{K}$  vs. actual value of  $0.0055\text{K}$ )

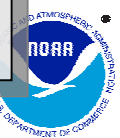


# Discussion

- Fu and Johanson claim the  $T_{TLT}$  trend from 1987-2003 is about 0.19K
- If we take the trend in MSU4 and view it as a systematic error we get,  $-0.3825 * -0.5168 = 0.1976$
- Thus, this trend is false and simply results from the imposition of first guess information upon the decreasing trend in the stratosphere of MSU4



May 26, 2005



# Conclusion

- Climate data records require careful attention to minimizing all sources of systematic bias
- Systematic errors are to be expected in the retrieval of geophysical information from satellite observations
- All trend analysis should critically evaluate this source of systematic error
- Failure to account for this systematic error can lead to false trend identification

