# Overview of the AIRS Science Team Algorithm.

by

**Dr. Christopher Barnet**

**NOAA/NESDIS/ORA**

**Mar 21, 2005**

**Complete Notes on remote sounding and the Atmospheric InfraRed Sounder (AIRS) Science Team retrieval methodology are on:**

**ftp ftp.orbit.nesdis.noaa.gov**

**cd pub/smcd/spb/cbarnet/reference**

| | |
|---|---|
| rs_notes.pdf | remote sounding notes & retrieval theory |
| phys640_s04.pdf | computational methods with chapters on linear and non-linear least squares |
| airsb_code.pdf | documentation on the code itself |

# Organizations That Have Contributed to AIRS

**AIRS-AMSU-HSB**

**Atmospheric Infrared Sounder**

**Advanced Microwave Sounding Unit**

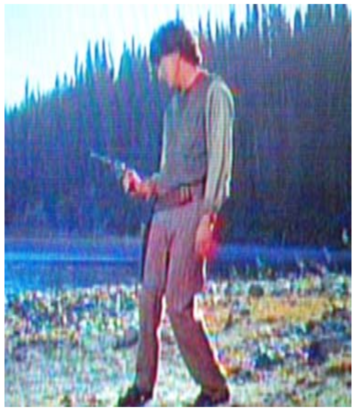
**Humidity Sounder for Brazil**

## Overview of This Talk

- **Introduction to regularization**
  - **How the theory of remote sensing has evolved.**
  - **What is the basis for the AIRS methodology.**
    - **Cloud clearing philosophy in an integral component**
    - **Minimization approach is optimized for hyper-spectral sounders.**
- **Details of the AIRS Science Team implementation.**
  - **Justification for use of channel sub-sets.**
  - **Justification for use of trapezoids.**
  - **Justification for use of finite differences.**
- **Some post-launch issues**
  - **Microwave side-lobe corrections**
  - **Tuning versus Error Term Experiments**
- **A few details on the code & development system**

## Need to Use “Snake-eyed Concentration


In the 1970 movie, Little Big Man, the character Jack Crabb learns from his sister than “it is possible to shoot a gun before you touch it; however, it takes lightnin’ reflexes and considerable snake-eyed concentration.”



**In this context:**

**It is possible to understand the methodology without understanding the details of the equations; however, it still takes lightnin’ reflexes and considerable snake-eyed concentration.**

## Notation, a.k.a., How I Learned to Embrace my OCD

- I adopt a notation of linear algebra that denotes the dimensions and row/column indices of the matrices. Yes, there is medication for this.
- For example, the kernel function, $\boldsymbol{K}$, is a rank-2 matrix, $\boldsymbol{K}(n, L)$ where $n$ is a channel index and $L$ is a parameter index (our pressure grid).
- I will write this matrix as $\boldsymbol{K}_{n,L}$. Having explicit indices makes transposes, $\boldsymbol{K}^T_{L,n}$, and inverses, $\boldsymbol{K}^{-1}_{L,n}$, more obvious.
- It is also useful when programming. Loop indices, $n, i, j, k, L, m$ etc. used in the FORTRAN code correspond to the matrix indices in the theory documents. Obviously, I am now off the medication!
- In this notation an implied summation rule is imposed

$$\begin{aligned} \boldsymbol{y}^T_n \cdot \boldsymbol{y}_n &= \boldsymbol{C}, \quad \text{a scalar} \\ \boldsymbol{y}_n \cdot \boldsymbol{y}^T_n &= \boldsymbol{C}_{n,n}, \quad \text{a rank 2 matrix} \end{aligned}$$

- I also use a superscript $i$ to denote items that are iterated. The concept of iteration evolves during the talk, so some items will begin without the iteration superscript and then will have it later.

## Least Squares (LSQ) Solution

- If we want to solve a linear system of equations of observations $\boldsymbol{y_n}$, where $\boldsymbol{n}$ is number of channels using $\boldsymbol{L}$ geophysical parameters, $\boldsymbol{X_L}$, we could write the relationship as

$$\boldsymbol{y_n = K_{n,L} \cdot X_L} \tag{1.1}$$

- We can begin by weighting the equation, if desired

$$\boldsymbol{W_{n,n} \cdot y_n = W_{n,n} \cdot K_{n,L} \cdot X_L} \tag{1.2}$$

- Then multiplying by, $\boldsymbol{K^T}$ we obtain

$$\boldsymbol{K^T_{L,n} \cdot W_{n,n} \cdot y_n = K^T_{L,n} \cdot W_{n,n} \cdot K_{n,L} \cdot X_L} \tag{1.3}$$

- And the solution is

$$\boldsymbol{X_L = \left[K^T_{L,n} \cdot W_{n,n} \cdot K_{n,L}\right]^{-1} \cdot K^T_{L,n} \cdot W_{n,n} \cdot y_n} \tag{1.4}$$

- Note that, $\boldsymbol{X_L = K^{-1}_{L,n} \cdot y_n}$ is the direct solution but for a non-square matrix $\boldsymbol{K^{-1} \equiv [K^T \cdot K]^{-1} \cdot K^T}$.

## Summary of Geophysical Products, $\boldsymbol{X_L}$

| | |
|---|---|
| $T(p)$ | vertical temperature profile |
| $q(p)$ | vertical water vapor profile ($\approx$ 8 g/kg @ surface) |
| $L(p)$ | vertical liquid water profile (f/ AMSU/HSB) |
| $O_3(p)$ | vertical ozone profile ($\approx$ 8 ppmv @ 6 mb) |
| $T_s$ | surface temperature |
| $\epsilon(\nu)$ | spectral surface emissivity, (*e.g.*, 0.95 @ 800 $\text{cm}^{-1}$) |
| $\rho_\odot(\nu)$ | spectral surface reflectivity of solar radiation |
| $P_{\text{cld}}$ | cloud top pressure for $\leq$ 2 cloud levels |
| $\alpha_{\text{cld,fov}}$ | cloud fraction for $\leq$ 2 cloud levels and 9 FOV's |
| $CO_2$ | total column carbon dioxide ($\approx$ 370 ppmv) |
| $CH_4(p)$ | methane profile ($\approx$ 1.65 ppmv) |
| $CO(p)$ | carbon monoxide profile ($\approx$ 0.11 ppmv) |

### Ancillary Information Needed for Retrieval

| | |
|---|---|
| $P_s$ | surface pressure (f/ AVN forecast) |
| $\theta$ | satellite zenith angle |
| $\theta_\odot$ | solar zenith angle |
| $\epsilon_{\text{cld},\nu} \equiv 1$ | spectral cloud emissivity for $\leq$ 2 cloud levels |

## The Cost Function

The idea of minimization of a cost function for the solution of linear equations dates back to Gauss (c. 1800). Most methods, include the AIRS Science Team method, minimize a cost function of the form:

$$\begin{aligned} J &= \left(f_n\left(X_L\right) - y_n\right)^T \cdot N_{n,n}^{-1} \cdot \left(f_n\left(X_L\right) - y_n\right) \\ &+ \left(X_L - X_L^0\right)^T \cdot H_{L,L} \cdot \left(X_L - X_L^0\right) \end{aligned} \quad (1.5)$$

We find the solution of $\frac{\partial J}{\partial X_L} = 0$. Since derivatives of the forward model are a function of the parameters, this problem is inherently non-linear. Therefore, we must iterate. For iteration=$i$ the solution is given by

$$K_{n,L}^i \equiv \frac{\partial f_n(X_L)}{\partial X_L}\Big|_{X_L^i} \quad (1.6)$$

$$\begin{aligned} X_L^{i+1} = X_L^0 \quad &+ \quad \left[K_{L,n}^{T^i} \cdot N_{n,n}^{-1} \cdot K_{n,L}^i + H_{L,L}\right]^{-1} \cdot K_{L,n}^{T^i} \cdot N_{n,n}^{-1} \cdot \\ &\left[y_n - f_n(X_L^i) + K_{n,L}^i \left(X_L^i - X_L^0\right)\right] \end{aligned} \quad (1.7)$$

However, this form can also be derived using linear algebra (similar to Eqn. 1.4) or maximum likelihood (Bayesian probability density functions).

## Reminder, What is the Forward Model

- **The non-scattering infrared forward model has components from the surface (mostly linear), the reflected solar component (mostly linear), a complicated, bet very small, down-welling term, and an atmospheric term of the form of a Fredholm integral equation of the $1^{st}$ kind:**

$$f_n(X_L^i) \simeq \int_\nu d\nu \cdot \Phi(\nu, \nu_0(n)) \cdot \int_p dp \cdot B_\nu(T^i(p)) \cdot \frac{\partial \exp\left(-\int_{z'=\infty}^{z(p)} \sum_i \kappa_i(\nu, X_L^i, p, \ldots) dz'\right)}{\partial p} \tag{1.8}$$

- **The optical depth for a given gas (denoted by a subscript $_i$), $\kappa_i$, is a complex interacting function of all the parameters, $X_L^i$, such as temperature $T^i(p)$, moisture, ozone, etc.**
- **$\Phi(\nu, \nu_0(n))$ is our instrument response function**
- **This equation can be highly non-linear for composition (*i.e.*, moisture, ozone, etc.) retrievals.**

# Damping and the “Background Term”

- **The matrix $\boldsymbol{H}_{L,L}$ is a form of regularization**
  - **Prevents inverse from being singular, *i.e.*, it stabilizes the inverse.**
  - **$\boldsymbol{X}^{i+1} - \boldsymbol{X}^{i}$ is smaller (*i.e.*,, damped) than the LSQ solution (see Eqn. 1.4), which I will denote as $\boldsymbol{X}^{i+1}(\boldsymbol{H} = 0)$ in these notes.**
- **The radiances are modified by the right hand term, $\boldsymbol{K}^{i}_{n,L}\left(\boldsymbol{X}^{i}_{L} - \boldsymbol{X}^{0}_{L}\right)$**
  - **When $\boldsymbol{H}$ is non-zero, the part that wasn’t believed on the $1^{st}$ iteration, $\boldsymbol{X}^{1}(\boldsymbol{H} = 0) - \boldsymbol{X}^{1}$, must be subtracted from the radiances.**
  - **If this term is neglected then the part that wasn’t believed would be introduced into subsequent iterations and the effect would be an un-damped LSQ retrieval, $\boldsymbol{X}^{i+1}(\boldsymbol{H} = 0)$.**
  - **This is why, a physical retrieval cannot use the results from a previous physical retrieval $\Rightarrow$ RULE #1 of Iteration**

## Heritage of Regularization Approaches

| | year | authors | method of regularization |
|---|---|---|---|
| | 1943 | Levenberg | Steepest Descent & Newtonian Iteration |
| | 1963 | Marquart | Hessian operator, $\boldsymbol{H} = \boldsymbol{\Delta y} \cdot \frac{\partial^2 (y_n - f_n(X_L))^2}{\partial X_L^2}$ |
| | 1963 | Twomey, Tikhonov | $\boldsymbol{H} = \boldsymbol{\lambda} \cdot \boldsymbol{I}$, prevent singularities |
| $\approx$ | 1970 | Twomey | $\boldsymbol{H}$ minimizes vertical derivatives, *e.g.*, $\boldsymbol{\partial T / \partial z}$ |
| | 1970 | Backus, Gilbert | Compute Optimal Vertical Functions f/ sounding |
| | 1972 | Conrath | Trade-off: Vertical Resolution versus $\boldsymbol{\chi^2}$ error |
| $\approx$ | 1970 | Wark & Fleming | Use $\boldsymbol{H}$ = co-variance as a constraint |
| | 1976 | Rodgers | Use *a posteriori* PDF's as a constraint, $\boldsymbol{H} = \boldsymbol{S_a^{-1}}$ |
| | 1989 | Eyre | Formalization of forward model errors in $\boldsymbol{N_{n,n}^{-1}}$ |
| | 1992 | Hanel, Conrath | Optimal functions/vertical resolution by SVD |
| | 1992 | Hansen | L-curve, finding optimal $\boldsymbol{\lambda}$ via SVD |
| | 1996 | Phalippou | For $\boldsymbol{q(p)}$, use relative humidity as a constraint |
| | 1997 | Schimpf & Schreier | Use of SVD to determine $\boldsymbol{H}$ |
| | 1999 | Li | use residuals to derive $\boldsymbol{\lambda}$ |
| | 2000 | Peckham & Grippa | Use lapse rate as a constraint |

## Constraints on AIRS Science Team Algorithm

- **Retrieval methodology must be able to handle cloud cleared radiances (CCR's)**
  - **Random noise amplification, $\frac{1}{3} \leq A <\approx 3$**
  - **Large spectrally correlated component**
  - **Statistical *a-priori* difficult to implement**
- **Retrieval should have minimal sensitivity to first guess.**
  - **Maximize contribution from instrument radiances.**
  - **Maximize sensitivity to and understanding of climate signals**
  - **Trade-off: model background states $\Rightarrow$ to use or not to use.**
- **Retrieval should not artificially constrain problem.**
  - **Minimize sensitivity to incorrect statistics, *e.g.*, in frontal zones avoid statistical damping.**
  - **Trade-off: Stability versus Impact**

## Cloud Clearing Methodology (Chapter 7 of rs_notes.pdf)

**This cloud clearing methodology has a long heritage starting from the original papers (Smith, 1968, Chahine, 1974), Chahine, 1975, Chahine, 1977, Chahine** *et al.* **1977, McMillin and Dean 1982, Smith** *et al.* **1992, Susskind,** *et al.* **1998, Joiner and Rokker 2002, Susskind,** *et al.* **2003. The fundamental features of the AIRS cloud clearing algorithm are**

- Use the $\boldsymbol{J = 9}$ AIRS cloud scenes, $\boldsymbol{R_{n,j}}$, without any *a-priori* constraint, such as preferential grouping, to compute the extrapolation parameters, $\boldsymbol{\eta_j}$.

$$\boldsymbol{\hat{R}_n = < R_{n,j} >_j + (< R_{n,j} >_j - R_{n,j}) \cdot \eta_j} \tag{1.9}$$

- Determine the number of cloud *formations* and constrain the number of degrees of freedom for solution of $\boldsymbol{\eta_j}$ to the number of cloud formations.
- Compute both CCR's and error estimates for the CCR's, $\boldsymbol{\delta\hat{R}_n\delta\hat{R}_n}$, specifically taking into account the noise amplification induced by the linear extrapolation and the spectrally correlated component of the radiance error due to error covariance of the $\boldsymbol{\eta}$'s, $\boldsymbol{\delta\eta_j\delta\eta_j^T}$.
- Compare the clear state estimate with the AIRS retrieval products and reject cases that violate any of the assumptions of cloud clearing.

## A non-Traditional Look at the Cost Function, 1/3

$$\begin{aligned} \boldsymbol{J} \;=\; & \left(\boldsymbol{f_n}\left(\boldsymbol{X_L^i}\right) - \boldsymbol{y_n}\right)^{\boldsymbol{T}} \cdot \boldsymbol{N_{n,n}^{-1}} \cdot \left(\boldsymbol{f_n}\left(\boldsymbol{X_L^i}\right) - \boldsymbol{y_n}\right) \\ + \quad & \left(\boldsymbol{X_L^i} - \boldsymbol{X_L^0}\right)^{\boldsymbol{T}} \cdot \boldsymbol{H_{L,L}} \cdot \left(\boldsymbol{X_L^i} - \boldsymbol{X_L^0}\right) \end{aligned} \tag{1.10}$$

- **We can compute the error in cloud cleared observations, $\boldsymbol{\delta\hat{R}_n\delta\hat{R}_n}$ very well.**
- **We can estimate errors in the forward model. For parameters held constant, $\boldsymbol{X_b}$, the obs-calc error covariance is**

$$\boldsymbol{N_{n,n'}^{-1}} = \boldsymbol{K_{n,b}} \cdot \boldsymbol{\delta X_b \delta X_{b'}^T} \cdot \boldsymbol{K_{b'n'}^T} + \boldsymbol{\delta\hat{R}_n\delta\hat{R}_n} \tag{1.11}$$

- *a-priori* **information enters the system through statistical estimates of $\boldsymbol{\delta X_b \delta X_{b'}^T}$**
  - **In the sense of estimates of errors in $\boldsymbol{X_L^0}$**
  - **In the sense of null space errors, the minimum allowed value of $\boldsymbol{\delta X_b \delta X_{b'}^T}$**
- **But, we also compute the formal errors of the solution, $\left(\boldsymbol{\delta X_L \delta X_L^T}\right)^{\boldsymbol{s}}$ for each retrieval step = $\boldsymbol{s}$ (details in Section 21.9 of rs_notes.pdf).**

## A non-Traditional Look at the Cost Function, 2/3

- $\left(\delta X_L \delta X_L^T\right)^s$ **from step** $= s$ **becomes the** $\left(\delta X_b \delta X_b^T\right)^{s+1}$ **in step** $= s+1$, *e.g.*, **we solve for** $T(p)$ **and** $\delta T(p) \delta T(p)^T$ **and then use that error covariance when we solve for** $q(p)$, $O_3(p)$, **etc., in later steps.**
- **Therefore, an improvement in temperature errors, for example, can be used to improve our moisture retrieval (vice a versa).**
- **This formulation also brings spectral correlation (*i.e.*, *a priori* knowledge via the forward calculation) into the solution via** $K_{n,b}(X_L^i)$ **on a case-by-case basis.**
  - **Spectral correlation is a function of other state parameters. For example, temperature lapse rate changes sensitivity of all the composition derivatives.**
  - $N$ **equations of** $y_n$ **changed into** $N$ **new equations:** $N_{n,n'} \cdot y_{n'}$.
  - **This is a powerful concept that when used properly**
    - **allows separation of mixed signals,** *e.g.*, $T(p)$ **and** $CO_2$
    - **minimizes sensitivity to biases,** *e.g.*, **surface effects (**$T_{skin}$, $P_{surf}$, **emissivity) in** $T(p)$ **retrieval.**

# A non-Traditional Look at the Cost Function, 3/3

- We can compute $\boldsymbol{H}$ from information content of $(\boldsymbol{K}^T\boldsymbol{N}^{-1}\boldsymbol{K})$ by singular value decomposition (SVD, a.k.a. empirical orthogonal functions (EOF's) see Section 21.3, rs_notes.pdf)

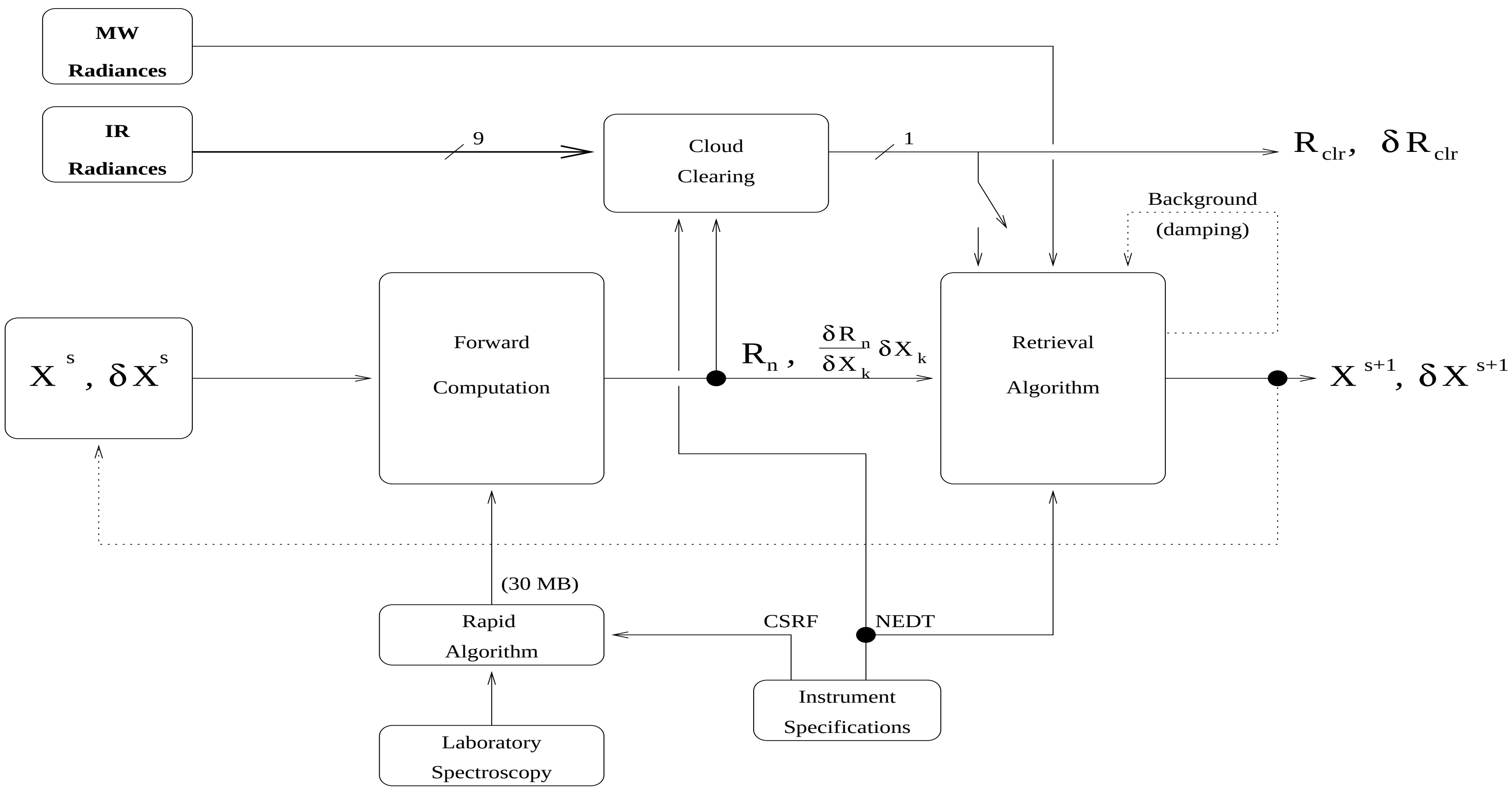
$$\boldsymbol{\Lambda}^i_{k,k} \equiv \boldsymbol{U}^{T^i}_{k,L}\left[\boldsymbol{K}^{T^i}_{L,n}\cdot\left(\boldsymbol{N}^s_{n,n}\right)^{-1}\cdot\boldsymbol{K}^i_{n,L}\right]\cdot\boldsymbol{U}^i_{L,k}\cdot \tag{1.12}$$

- $\Lambda^i_{k,k}$ is diagonal with elements equal to $\boldsymbol{\lambda}_k$
- When $\boldsymbol{\lambda}^i_k \gg 1$ the terms are well determined. $\boldsymbol{K}^i_{n,L}\cdot\boldsymbol{U}^i_{L,k}$ are new Jacobians with very high signal to noise: $\boldsymbol{H}=0$
- When $\boldsymbol{\lambda}^i_k \to 0$ the observations have no influence on the solution: $\boldsymbol{H}\to\infty$ and components of $\boldsymbol{X}^{i+1}_L \to \boldsymbol{X}^0_L$
- When $\boldsymbol{\lambda}^i_k$ is small and significant we add a $\boldsymbol{\Delta\lambda}^i_k$ (details in Section 21.4 of rs_notes.pdf) which is equivalent (see Section 21.8, rs_notes.pdf) to adding a case dependent $\boldsymbol{H}^i$ given by

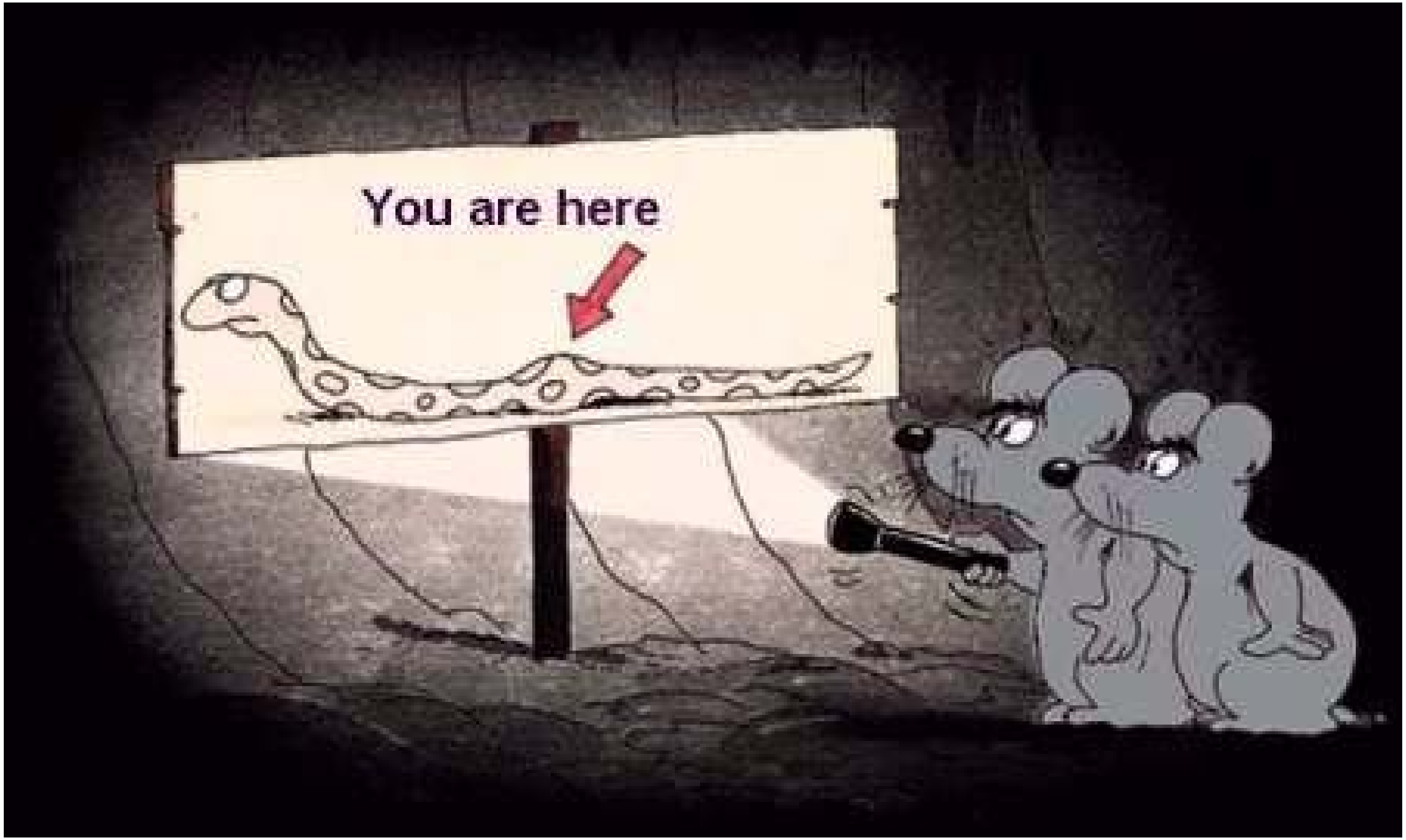
$$\boldsymbol{H}^{s,i}_{L,L'} = \boldsymbol{U}^i_{L,k}\cdot\boldsymbol{\Delta\Lambda}^i_{k,k}\cdot\boldsymbol{U}^{T^i}_{k,L} \tag{1.13}$$

- SVD determines the optimal fraction of the *a priori* information to use..
- We think this is more robust than using ensemble statistics of $\boldsymbol{I}$ cases to compute a static *a priori* covariance $\boldsymbol{S_a} \equiv \boldsymbol{\delta X}_{L,i}\boldsymbol{\delta X}^T_{i,L}$

# Overview of the AIRS Multi-spectral Physical Retrieval System




**The atmospheric state, $X_L^s$, and the error estimate of that state, $\delta X_L^s$, are used to minimize the residuals in observed minus computed radiances in each retrieval step=$s$.**




**It feels like this, eh?**

## Philosophy of AIRS *Team* Algorithm

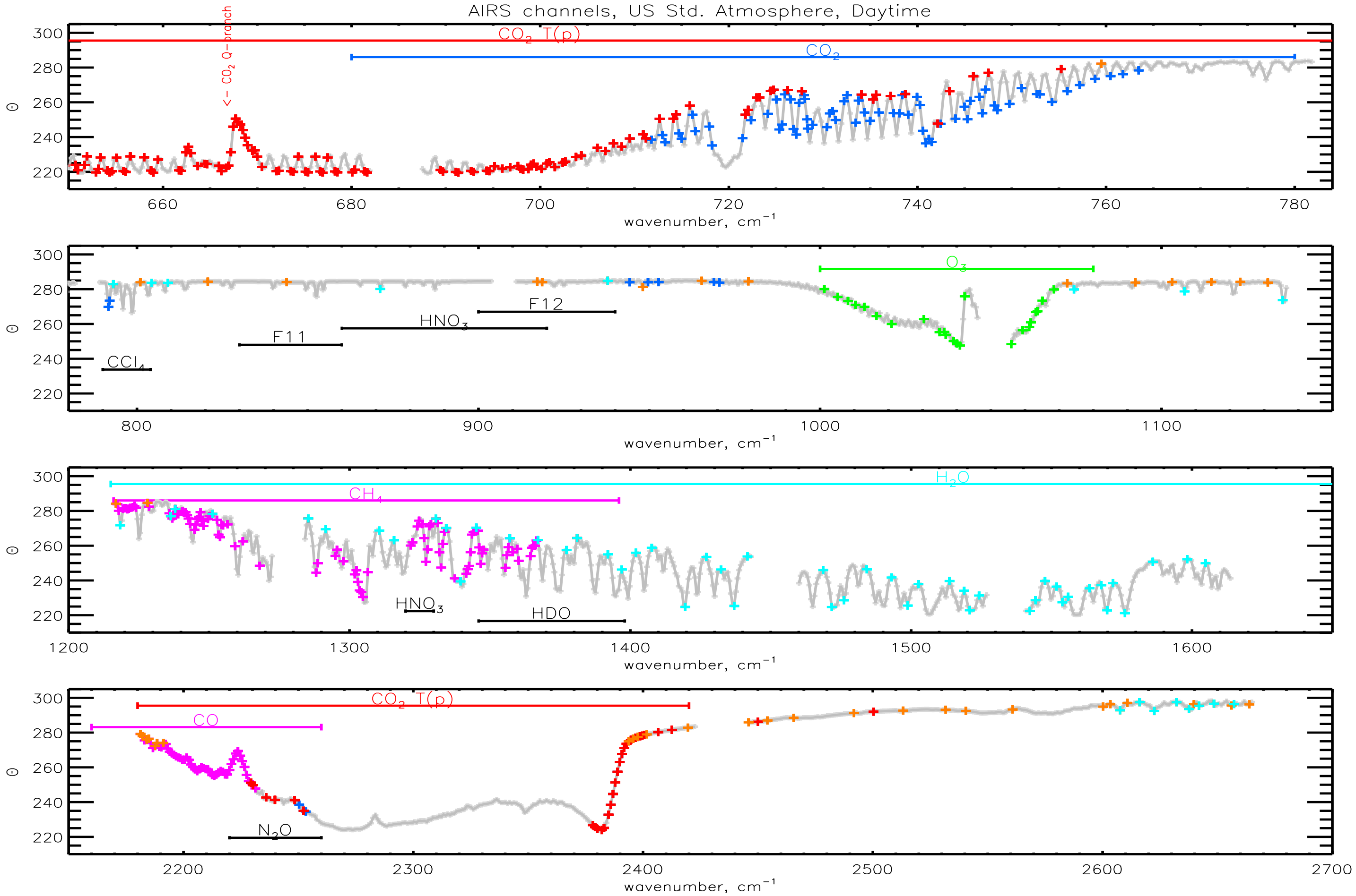
- **Utilize microwave-only product (MIT maximum likelihood algorithm $\Rightarrow$ $T(p)$, $q(p)$, $T_{skin}$) to estimate the infrared clear radiance estimate for initial cloud clearing.**
- **Utilize eigenvector regression (NOAA/NESDIS) to provide the first guess state for the physical algorithm. This solution contains the fine vertical structure information based on $\approx$ 1600 AIRS channels..**
- **Utilize a physical retrieval (NOAA/GSFC) to improve the state.**

1. Microwave and infrared observations are used in each step.
2. Use microwave observations and products to reject cases with poor cloud clearing.
   - Reject if Obs-Calc of coupled retrieval is too large
   - Perform microwave-only retrieval using coupled retrieval as first guess. Reject if $\Delta T(p)$ is too large.
3. All observations are used at their native angles of observation and the forward model is computed at the correct angles.
4. For cloud clearing we do perform a local angle correction of 1.1° with each 3 x 3 set of FOV's

# Philosophy of AIRS Physical Algorithm

1. **Embed an information content analysis into each step to determine the optimal damping (regularization) for each case.**
   - **Cloud cleared radiances are both case and iteration dependent.**
   - **Propagate a formal geophysical error estimate through each step.**
   - **Compute an estimate of the *a-priori* covariance at each step.**
2. **Take advantage of parameters that are separable (*i.e.*, pay attention to spectroscopy and radiative transfer features in our spectrum)**
   - **For example, we can solve for $T_{surf}$ holding all other variables (*e.g.*, water) constant, since $T_{surf}$ is quite linear. Steps are denoted as a superscript $s$ in my notes.**
   - **BUT, if a step is repeated (*e.g.*, when an error estimate has been improved) NEVER use the products from the previous step.**
3. **Select channels that are "spectrally pure", that is**
   - **Have a high sensitivity to what is being solved for.**
   - **Have a low sensitivity to the parameters held constant (*i.e.*, keep our error estimates small when separating variables)**
4. **Use optimal number of parameters for each retrieval step are determined in simulation (Backus-Gilbert optimization) to speed up processing.**

# Channels used in the AIRS retrieval algorithm



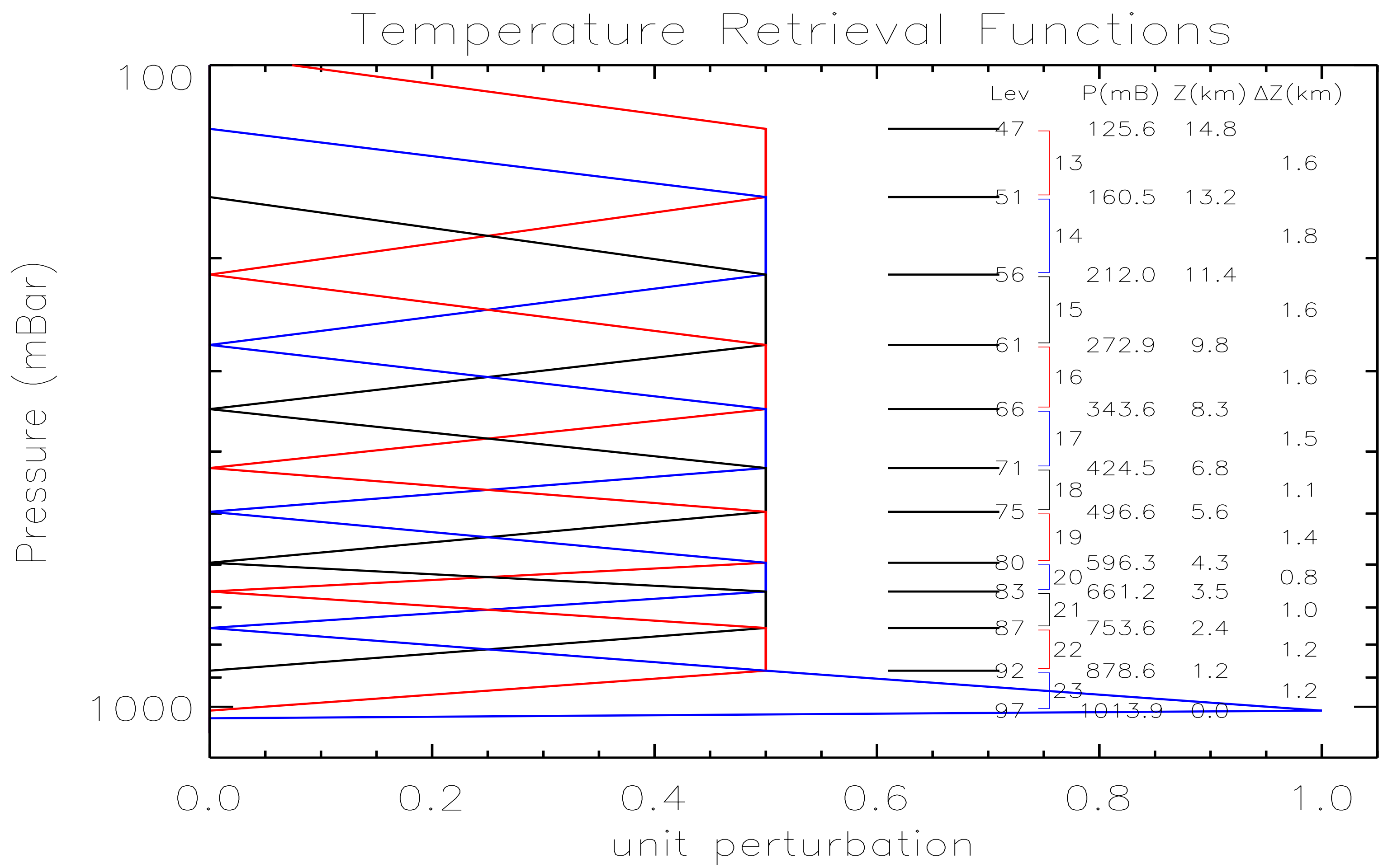
## Justification for Vertical and Spectral Functions

- A fine vertical grid is required for accurate computation of the absorption coefficients, $\kappa_i(\nu, p(z), X^{s,i}, \theta)$ and radiances.
- There are about 85 independent pieces of information in AIRS (2378 chl's), IASI (8461 chl's), and CrIS (1305 chl's)
- Solving for 100 vertical levels of T(p), q(p), $O_3$(p), CO(p), $CH_4$(p), and $CO_2$(p) wastes time and can destabilize the retrieval.
- If we allowed 2378 emissivities, there would be nothing left to solve for.
- Functions, $F^s_{L,j}$, and associated parameters, $A^{s,i}_j$, are chosen in a trade-off between resolution and stability for each retrieval step. Analogous to Backus & Gilbert (1970) trade-off. Also discussed in Hanel, (1992).
- Major issue for code execution time and improves stability.

## Formulation of Vertical and Spectral Functions

- **Temperature functions are additive vertical trapezoids.**
- **Composition functions are multiplicative vertical trapezoids.**
  - Radiance kernel is $\propto \exp(\boldsymbol{\kappa}(\boldsymbol{X}_L^{s,i}))$,
  - $\boldsymbol{\kappa}(\boldsymbol{X}_L^{s,i})$, is the optical depth $\propto \boldsymbol{X}_L^{s,i}$.
  - Therefore, composition variables are more linear in $\ln\left(\boldsymbol{X}_L^{s,i}\right)$
  - $\partial \ln(\boldsymbol{X}_L^{s,i}) \propto \frac{\partial X_L^{s,i}}{X_L^{s,i}}$ which is a % change in $\boldsymbol{X}_L^{s,i}$.
- **Emissivity functions are additive spectral triangles.**
- **A scaling parameter $\hat{\boldsymbol{A}}_j^s$ is used to create dimensionless parameters and adjust scale between different functional groups (*e.g.*, when mixing T(p), q(p), and emissivity in one retrieval).**
- **The Jacobian, $\boldsymbol{K}_{n,L}^{s,i}$, becomes a set of new derivatives, $\boldsymbol{S}_{n,j}^{s,i}$, in which groups of parameters in $\boldsymbol{L}$ space are grouped together in $\boldsymbol{J}$ space.**
- **Sub-sets (*e.g.*, temperature) of vertical and spectral functions must sum to unity: $\sum_j \left(\boldsymbol{F}_{L,j}^s\right) = 1$ for a group of functions.**

# Example of functions, $F_{L,j}$, for $T(p(L))$ retrieval



# Kernel Function replaced by Sensitivity Matrix

- **For additive functions the S-matrix is given by**

$$\boldsymbol{S}^{s,i}_{n,j} \equiv \frac{\partial \boldsymbol{f}_n\left(\boldsymbol{X}^{s,i}_L + \boldsymbol{F}^s_{L,j} \cdot \boldsymbol{A}^s_j\right)}{\partial \boldsymbol{A}^s_j} \cdot \boldsymbol{\Delta}\hat{\boldsymbol{A}}^s_j \tag{2.1}$$

$$\simeq \boldsymbol{f}_n\left(\boldsymbol{X}^{s,i}_L + \boldsymbol{F}^s_{L,j} \cdot \hat{\boldsymbol{A}}^s_j\right) - \boldsymbol{f}_n\left(\boldsymbol{X}^{s,i}_L\right) \tag{2.2}$$

- **For multiplicative functions the S-matrix is given by**

$$\boldsymbol{S}^{s,i}_{n,j} \equiv \frac{\partial \boldsymbol{f}_n\left(\boldsymbol{X}^{s,i}_L \cdot \left(1 + \boldsymbol{F}^s_{L,j} \cdot \boldsymbol{A}^s_j\right)\right)}{\partial \boldsymbol{A}^s_j} \cdot \boldsymbol{\Delta}\hat{\boldsymbol{A}}^s_j \tag{2.3}$$

$$\simeq \boldsymbol{f}_n\left(\boldsymbol{X}^{s,i}_L \cdot \left(1 + \boldsymbol{F}^s_{L,j} \cdot \hat{\boldsymbol{A}}^s_j\right)\right) - \boldsymbol{f}_n\left(\boldsymbol{X}^{s,i}_L\right) \tag{2.4}$$

- **Analytic derivatives on the RT grid do not help our algorithm, $\boldsymbol{\delta}$ function perturbations are sub-optimal (Backus+Gilbert).**
- **Single sided finite difference is currently used, we will explore the benefit of double-sided and dynamically scaled derivatives someday. This is not our biggest error source!!!**

## So . . . we really solve for dimensionless parameters

**The minimization equation for both additive and multiplicative forms is given by**

$$
\begin{aligned}
\boldsymbol{\Delta A_j^{s,i+1}} &\equiv \boldsymbol{A_j^{s,i+1} - A_j^{s,0}} \\
&= \left[\left(\boldsymbol{S_{j,n}^{T^{s,i}}}\right) \cdot \left(\boldsymbol{N_{n,n}^{s}}\right)^{-1} \cdot \boldsymbol{S_{n,j}^{s,i}} + \boldsymbol{H_{j,j}^{s,i}}\right]^{-1} \cdot \boldsymbol{S_{j,n}^{T^{s,i}}} \cdot \left(\boldsymbol{N_{n,n}^{s}}\right)^{-1} \cdot \\
&\qquad \left[\boldsymbol{y_n - f_n(X_L^{s,i}) + S_{n,j}^{s,i}}\left(\boldsymbol{A_j^{s,i} - A_j^{s,0}}\right)\right] \qquad (2.5)
\end{aligned}
$$

**And to return to parameter $\boldsymbol{L}$ space is done by combining the components and dividing by the scaling factors, $\boldsymbol{\hat{A}_j^s}$.**

$$\boldsymbol{\Delta X_L^{s,i+1}} = \sum_j \boldsymbol{F_{L,j}^s} \cdot \left(\boldsymbol{\Delta A_j^{s,i+1} \cdot I_{j,j} \cdot \Delta \hat{A}_j^{-1^s}}\right) \qquad \textbf{Additive} \qquad (2.6)$$

$$\boldsymbol{\Delta X_L^{s,i+1}} = \prod_j \left(\boldsymbol{1 + F_{L,j}^s}\right) \cdot \left(\boldsymbol{\Delta A_j^{s,i+1} \cdot I_{j,j} \cdot \Delta \hat{A}_j^{-1^s}}\right) \quad \textbf{Multiplicative} \qquad (2.7)$$

## Finally, we regularize

- We perform the information content analysis on our new functions

$$\boldsymbol{\Lambda}^{s,i}_{k,k} \equiv \boldsymbol{U}^{T^{s,i}}_{k,j} \cdot \left[\boldsymbol{S}^{T^{s,i}}_{j,n} \cdot \left(\boldsymbol{N}^{s}_{n,n}\right)^{-1} \cdot \boldsymbol{S}^{s,i}_{n,j}\right] \cdot \boldsymbol{U}^{s,i}_{j,k} \tag{2.8}$$

- This transformation is equivalent to new Jacobians, $\boldsymbol{S}^{s,i}_{n,j} \cdot \boldsymbol{U}^{s,i}_{j,k}$
- We determine $\boldsymbol{\Delta\lambda}$ as follows (see Section 21.8, rs_notes.pdf)

$$\begin{aligned}
\boldsymbol{\Delta\lambda}^{s,i}_{k} &= 0 && \text{for } \boldsymbol{\lambda}^{s,i}_{k} \geq \boldsymbol{\lambda}^{s}_{c} && (2.9)\\
&= \sqrt{\boldsymbol{\lambda}^{s,i}_{k}} \cdot \left(\sqrt{\boldsymbol{\lambda}^{s}_{c}} - \boldsymbol{\lambda}^{s,i}_{k}\right) && \text{for } \boldsymbol{\lambda}^{s,i}_{k} < \boldsymbol{\lambda}^{s}_{c} && (2.10)\\
&= \infty && \text{for } \boldsymbol{\lambda}^{s,i}_{k} \leq (0.05)^2 \cdot \boldsymbol{\lambda}^{s}_{c} && (2.11)
\end{aligned}$$

## And Solve for Our Parameters

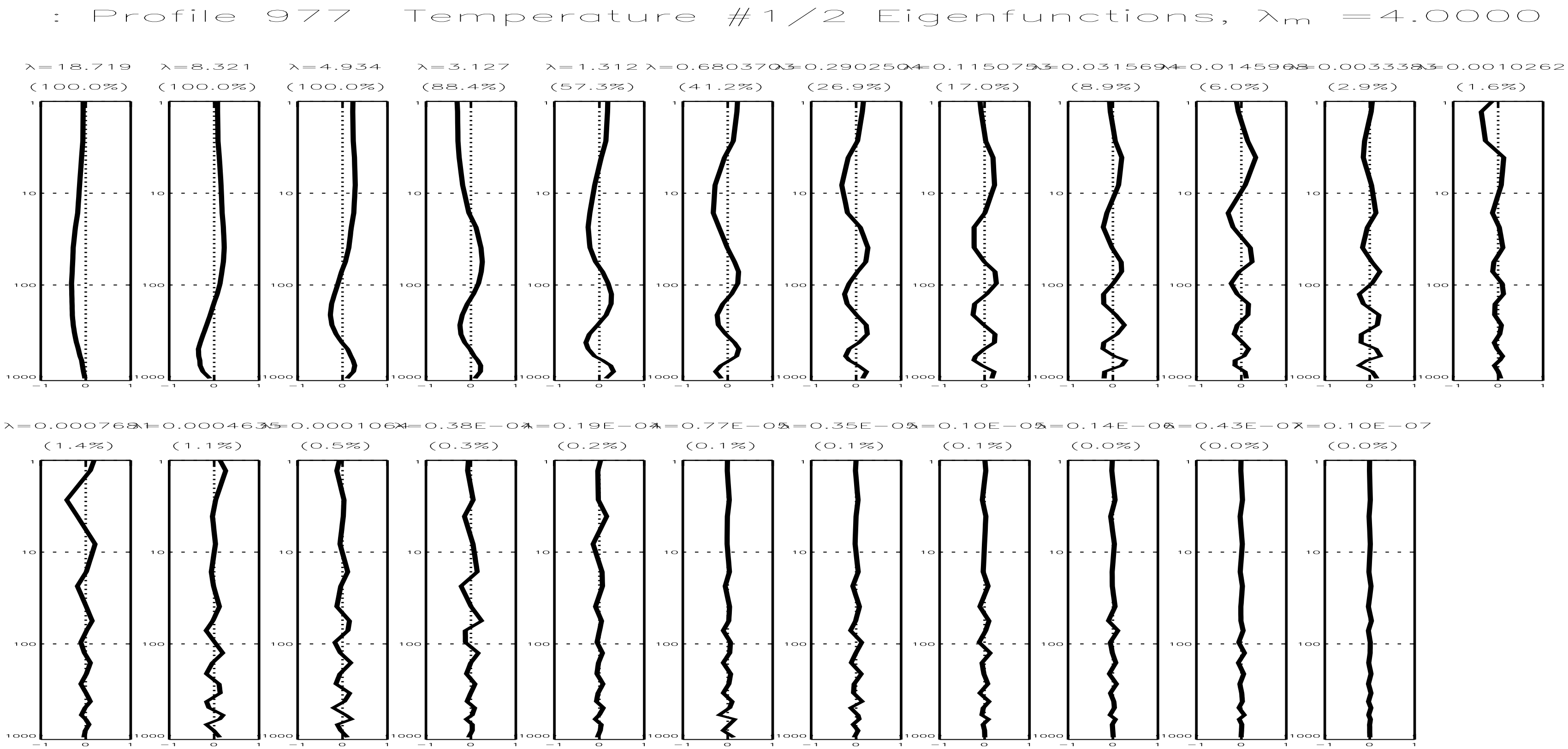
- **We solve for the new parameters**

$$\boldsymbol{\Delta A_j^{s,i+1}} = \boldsymbol{U_{j,k}^{s,i} \cdot \Delta B_k^{s,i+1}} \tag{2.12}$$

$$\boldsymbol{\Delta B_k^{s,i+1}} = \left(\frac{\mathbf{1}}{\boldsymbol{\lambda_k^{s,i} + \Delta\lambda_k^{s,i}}}\right) \cdot \boldsymbol{U_{k,j}^{T^{s,i}} \cdot S_{j,n}^{T^{s,i}}} \cdot \left(\boldsymbol{N_{n,n}^{s}}\right)^{-1} \cdot \left[\boldsymbol{y_n - f_n(X_L^{s,i}) + S_{n,j}^{s,i}}\left(\boldsymbol{A_j^{s,i} - A_j^{s,0}}\right)\right] \tag{2.13}$$

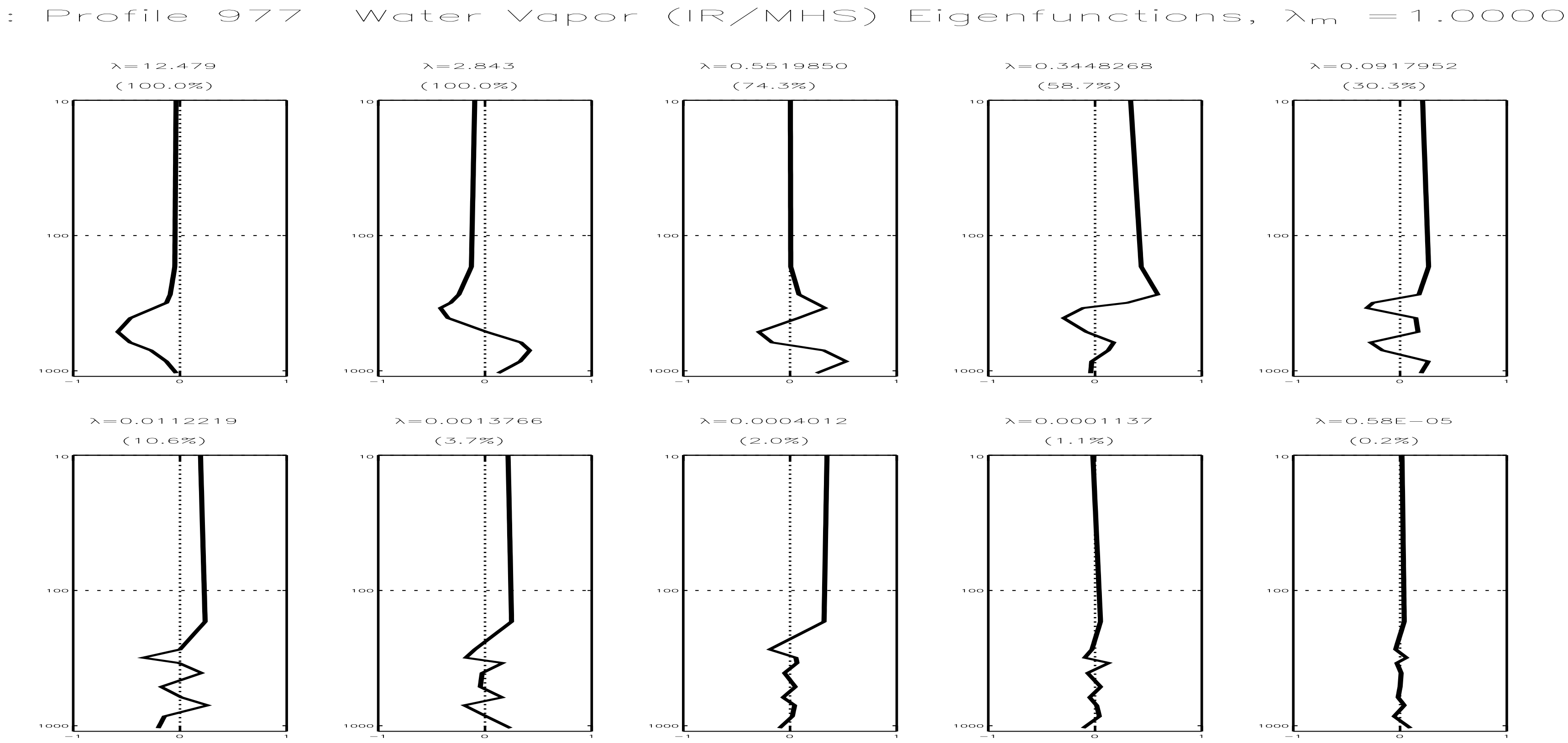
- **Note that the change in the parameters, $\boldsymbol{\Delta A_j^{s,i+1}}$ associated with original functions, $\boldsymbol{F_{L,j}^{s}}$ is equivalent to a transformed parameter change, $\boldsymbol{\Delta B_k^{s,i+1}}$, associated with a new function $\boldsymbol{G_{L,k}^{s,i} = F_{L,j}^{s} \cdot U_{j,k}^{s,i}}$.**
- **It is illustrative to visualize the transformed functions.**

# Example of SVD *optimal* $\boldsymbol{T(p)}$ Functions



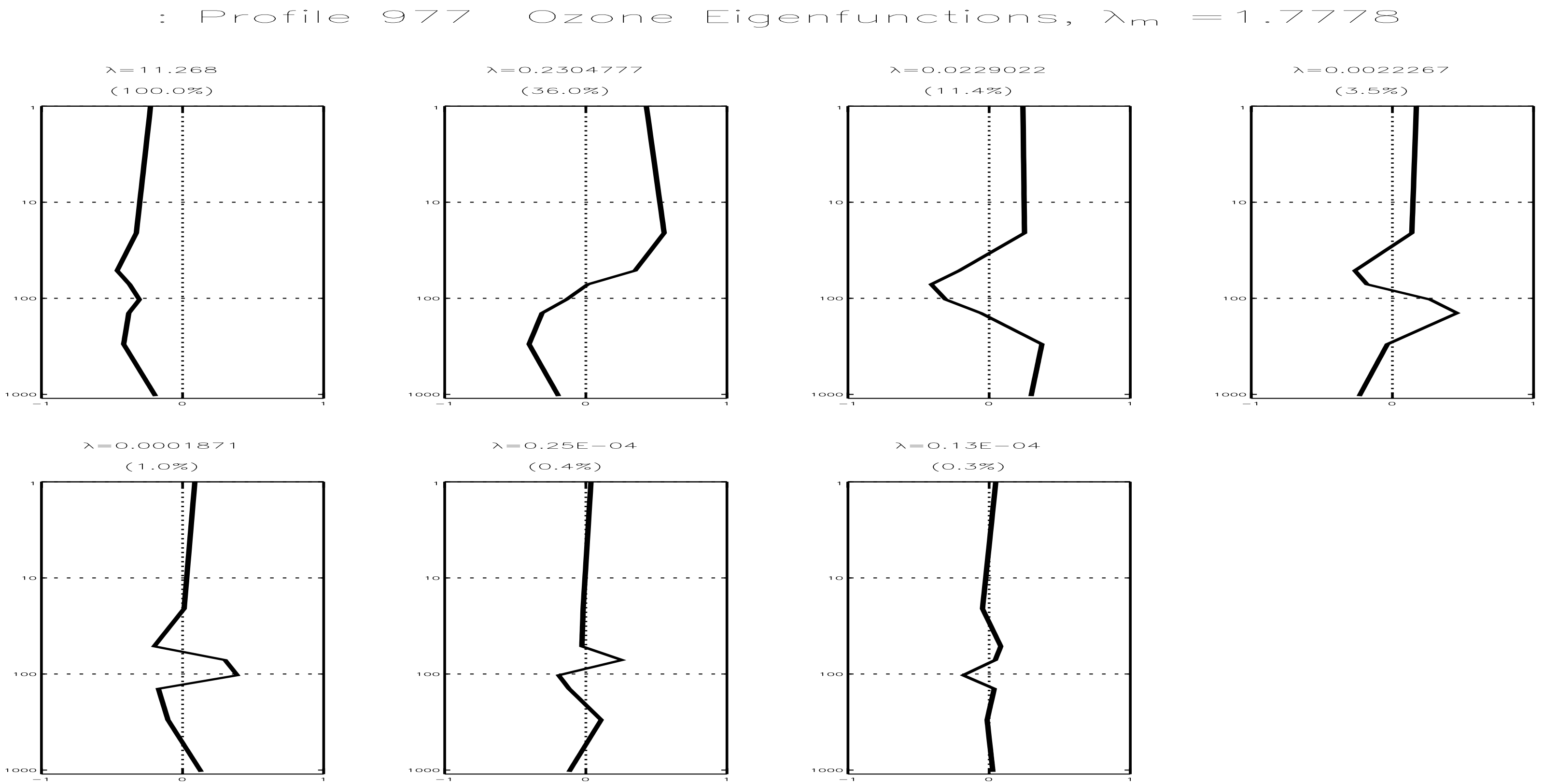

*For temperature functions, the new set of vertical functions,* $\boldsymbol{F^{s}_{L,j} \cdot U^{s,i}_{j,k}}$ *are shown for the AIRS temperature retrieval information content analysis. In this case, ten functions of the 23 functions are determined to better than 5%.*

# Example of SVD *optimal* $\boldsymbol{q(p)}$ Functions




*For water functions, the new set of vertical functions,* $\boldsymbol{F^{s}_{L,j} \cdot U^{s,i}_{j,k}}$ *are shown for the AIRS temperature retrieval information content analysis. In this case, six functions of the ten functions are determined to better than 5%.*

# Example of SVD *optimal* $O_3(p)$ Functions




*For ozone functions, the new set of vertical functions,* $\boldsymbol{F^s_{L,j}} \cdot \boldsymbol{U^{s,i}_{j,k}}$ *are shown for the AIRS temperature retrieval information content analysis. In this case, three functions of the seven functions are determined to better than 5%.*

## Propagation of Formal Errors

- The uncertainty in $\Delta B_k^{s,i+1}\left(\Delta\lambda_k^{s,i}=0\right)$ are uncorrelated and equal to $\left(\lambda_k^{s,i}\right)^{-\frac{1}{2}}$.
- The fraction of $\Delta B_k^{s,i}$ solved for is equal to $\phi_k^{s,i}=\lambda_k^{s,i}/\left(\lambda_k^{s,i}+\Delta\lambda_k^{s,i}\right)$
- The propagated error in $B$ space is the RSS of the transformed first guess error, $\delta B^{s,0}$, and the error from the radiances.

$$\delta B_k^{s,i+1}=\sqrt{\left(\left(1-\phi_k^{s,i}\right)\cdot\delta B_k^{0,s}\right)^2+\left(\phi_k^{s,i}\cdot\frac{1}{\sqrt{\lambda_k^{s,i}}}\right)^2} \tag{2.14}$$

- In $\Delta A$ space the errors are correlated and can be computed from

$$(\delta A_j\delta A_j)^{s,i+1} = U_{j,k}^{s,i}\cdot\left(\delta B_k^{s,i}\right)^2 U_{k,j}^{T^{s,i}} \tag{2.15}$$

$$\delta A^{s,i+1}(j) \simeq \sqrt{\sum_k U^{s,i}(j,k)^2\cdot\left(\delta B^{s,i+1}(k)\right)^2} \tag{2.16}$$

# Propagation of Formal Errors

- **The errors in the geophysical products are computed in the root-sum-squared (RSS) sense from parameter errors:**

$$\left(\boldsymbol{\delta X_L \delta X_L^T}\right)^{s,i+1} = \boldsymbol{F^s_{L,j}} \cdot \left(\boldsymbol{\delta A_j \delta A_j}\right)^{s,i+1} \cdot \boldsymbol{F^{T^s}_{j,L}} \tag{2.17}$$

$$\boldsymbol{\delta X^{s,i+1}(L)} \simeq \sqrt{\sum_j \boldsymbol{\delta} \left(\boldsymbol{A^{s,i}(j)}\right)^2 \cdot \boldsymbol{F^s(j,L)^2}} \tag{2.18}$$

- **These error estimates can be used to compute $\left(\boldsymbol{N^s_{n,n}}\right)^{-1}$ and, therefore, propagated into the next retrieval step; however, in practice we only keep the diagonal components of these errors.**

- **It is a bit more complicated than this, due to handling of null space errors, ⇒ but I figure we have ALL had enough equations.**

# Some Post-Launch Issues

**Now we have the theory, we are ready to work with real data.**

## Issues with AMSU's Estimate of CLEAR State

- Microwave side-lobe corrections (SLC's) for the Aqua platform are more complex than the POES platforms and have NOT been applied to date.
- A large microwave tuning (empirical side-lobe correction) has been employed to mitigate SLC issues.
- A poor AMSU first guess has a negative impact on cloud clearing and, therefore, all AIRS products.
- To understand the impact to AIRS products, we are using a model analysis to increase our information content.
  1. To assess the impact of AMSU SLC issues on the AIRS products.
  2. To assess the need for tuning and/or RTA improvements.
- We are building the ability to bring in MODIS clear pixels co-located to AIRS FOV's to improve our QA & information content.

## TUNING and MODEL ERROR TERMS

**For discussion, assume a retrieval equation looks like**

$$\Delta X_j^{s,i+1} = \left[ S_{j,n'}^{T^{s,i}} \cdot \left( N_{n',n}^s \right)^{-1} \cdot S_{n,j}^{s,i} + H_{j,j}^{s,i} \right]^{-1} \cdot S_{j,n'}^{T^{s,i}} \cdot \left( N_{n',n}^s \right) -1 \cdot \left[ y_n - f_n(X_L^{s,i}) + \Psi_n^{s,i} + T(n) \right] \quad (2.19)$$

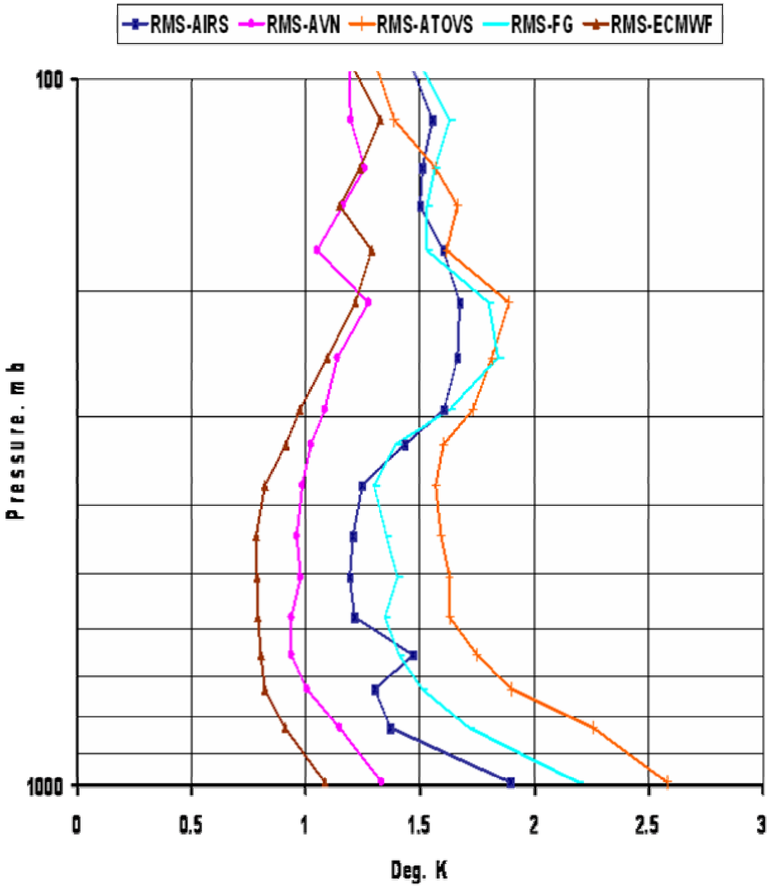
- $S_{n,j}^{s,i}$ **is the sensitivity of channel** $n$ **to parameter** $A_j^s$
- $\Psi_n^{s,i}$ **is the background term derived from** *a-priori* **contribution.**
- $T(n)$ **is radiance tuning, if applied.**

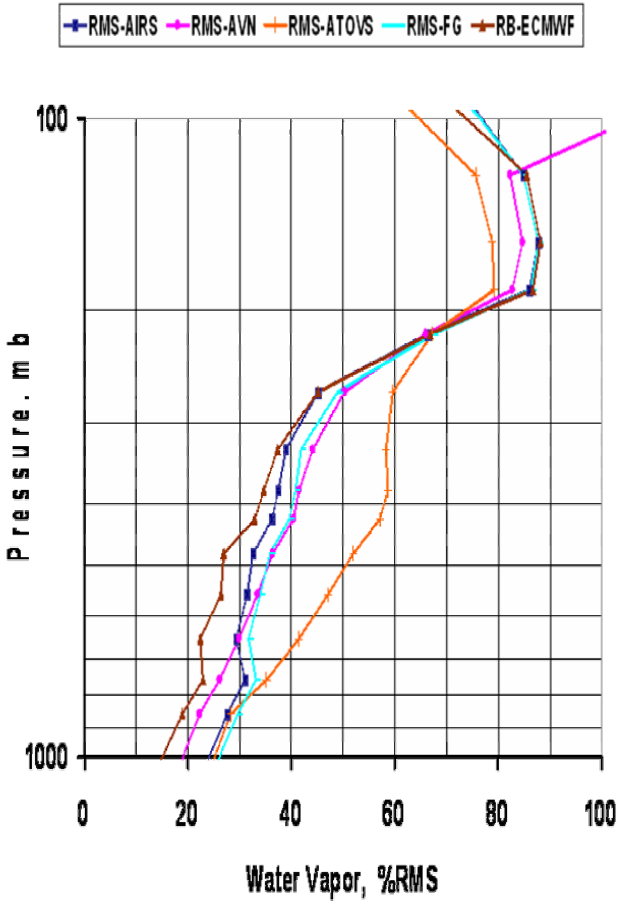
**With real data we have other error sources, such as rapid transmittance algorithm (RTA) and spectroscopy errors that we can write as** $E_{n',n}$**.**

$$N_{n',n}^s = N_{n',n}^s + E_{n',n} \quad (2.20)$$

# Example of RMS Statistics versus RAOB's

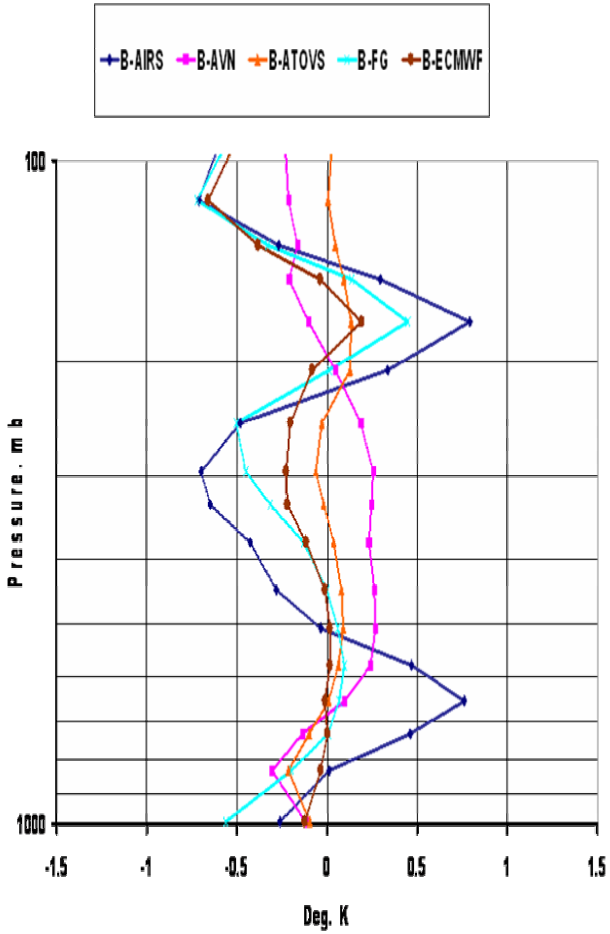
## 32,000 co-located cases from Sep. 2002 to Sep. 2004

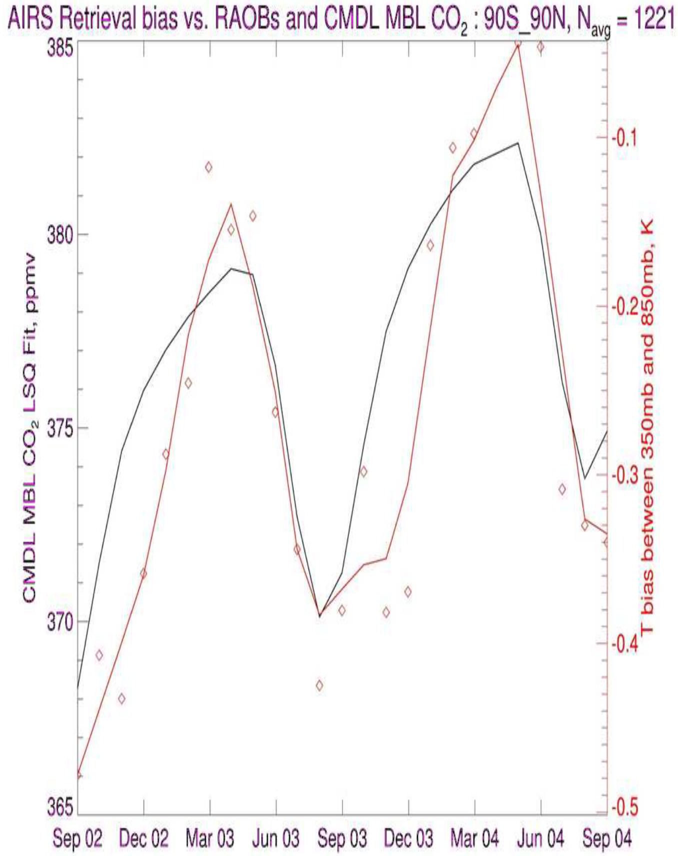



**These figures provided by Murty Divakarla.**

# Example of T(p) BIAS Statistics versus RAOB's
## 32,000 co-located cases from Sep. 2002 to Sep. 2004



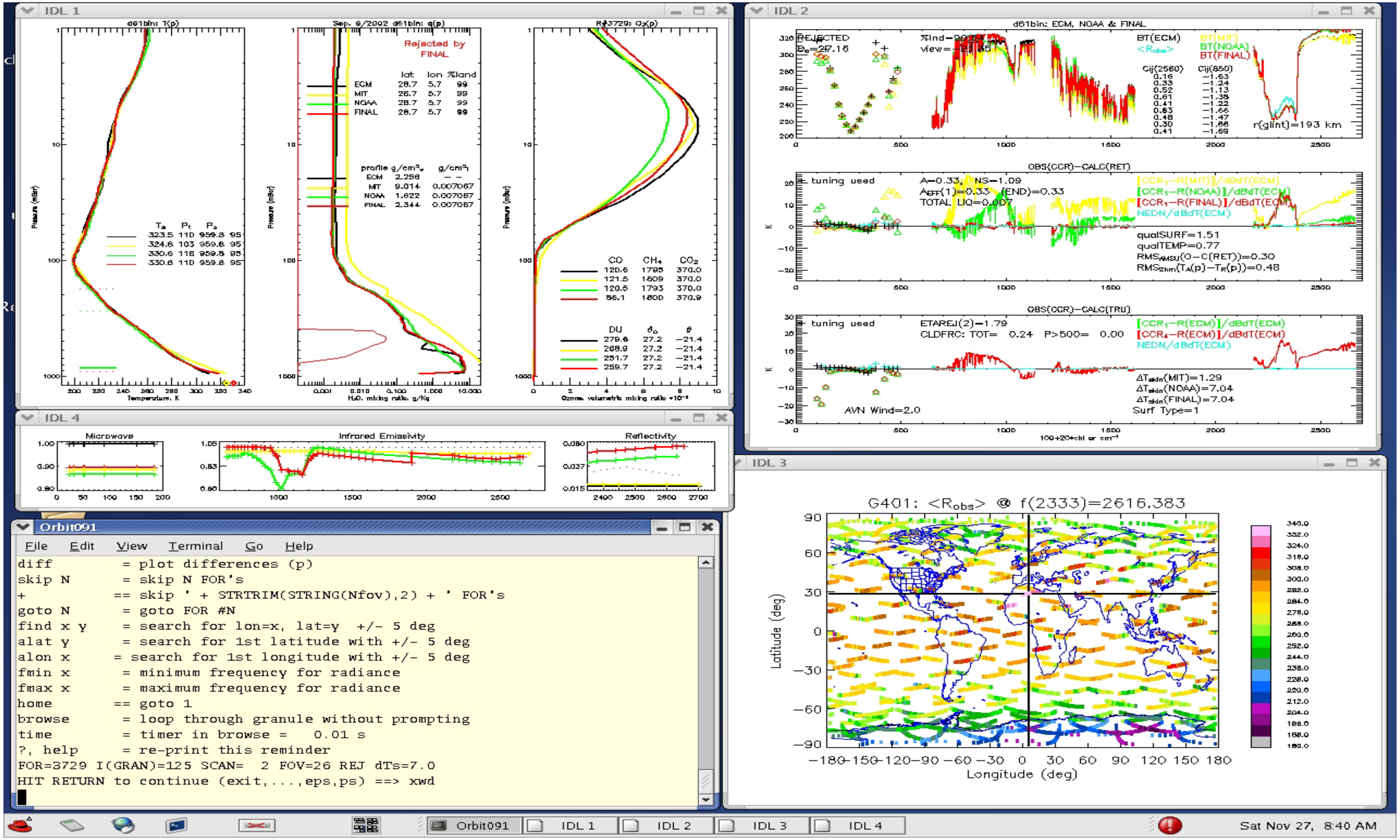




**These figures provided by Murty Divakarla & Eric Maddy.**

# Code Development

- **There is one code for AIRS, IASI, and CrIS.**
- **The retrieval system code is a total of about 106,000 lines of FORTRAN code.**
- **We have simulation capability (another 24,000 lines of code).**
  - **Can perform instrument trade studies.**
  - **Can study retrieval theory in simulation.**
  - **Simulation can be built from models and/or from AIRS retrieval products.**
- **Scientific evaluation is built into the code.**
  - **Each retrieval case can be compared to a reference state (AVN forecast, ECMWF forecast, RAOB, or truth (in simulation)) at every single iteration and step.**
  - **Diagnostics exist in radiance space and geophysical space.**
- **There are over 100,000 lines of IDL code for display and analysis of diagnostic output.**

# Example of Diagnostic Capabilities



## References for Today's Talk


Chahine, M.T. 1977. Remote sounding of cloudy atmospheres. II. Multiple cloud formations. J. Atmos. Sci. v.34 p.744-757.

Chahine, M.T., H.H. Aumann and F.W. Taylor 1977. Remote sounding of cloudy atmospheres. III. Experimental verifications. J. Atmos. Sci. v.34 p.758-765.

Chahine, M.T. 1975. An analytic transformation for remote sensing of clear-column atmospheric temperature profiles. J. Atmos. Sci. v.32 p.1946-1952.

Chahine, M.T. 1974. Remote sounding of cloudy atmospheres. I. The single cloud layer. J. Atmos. Sci. v.31 p.233-243.

Conrath, B.J. 1972. Vertical resolution of temperature profiles obtained from remote sensing radiation measurements. J. Atmos. Sci. v.29 p.1262- 1271.

Fleming, H.E., M.D. Goldberg and D.S. Crosby 1988. Operation implementation of the minimum variance simultaneous retrieval method. in 3rd Conf. Sat. Meteor. & Oceanography p.16-23.

Hanel, R.A., B.J. Conrath, D.E. Jennings and R.E. Samuelson 1992. Exploration of the solar system by infrared remote sensing. Cambridge Univ. Press 458 pgs.

Hansen, P.C. 1992. Analysis of discrete ill-posed problems by means of the L-curve. SIAM Review v.34 p.561-580.

Joiner, J. and L, Rokker 2000. Variational cloud-clearing with TOVS data. Quart. J. Roy. Meteor. Soc. v.126 p.725-748.

McMillin, L.M. and C. Dean 1982. Evaluation of a new operational technique for producing clear radiances. J. Appl. Meteor. v.21 p.1005- 1014.

References for Non-linear Least Squares (continued)


Peckham, G.E. and M. Grippa 2000. Improved retrieval of tropospheric temperatures from remote measurement of thermal radiation using the adiabatic lapse rate as a constraint. Quart. J. Roy. Meteor. Soc. v.126 p.749-760.

Phalippou, L. 1996. Variational retrieval of humidity profile, wind speed and cloud liquid water path with the SSM/I: potential for numerical weather prediction. *Q.J.R.Meteorol. Soc.* **122**, p.327-355.

Schimpf, B. and F. Schreier 1997. Robust and efficient inversion of vertical sounding atmospheric high-resolution spectra by means of regularization. J. Geophys. Res. v.102 p.16037-16055.

Smith, W.L., X.L. Ma, S.A. Ackerman, H.E. Revercomb and R.O. Knuteson 1992. Remote sensing cloud properties from high spectral resolution infrared observations. J. Atmos. Sci. v.50 p.1708-1720.

Smith, W.L. 1968. An improved method for calculating tropospheric temperature and moisture from satellite radiometer measurements. Monthly Weather Review v.96 p.387-396.

Susskind, J., C.D. Barnet and J.M. Blaisdell 2003. Retrieval of atmospheric and surface parameters from AIRS, AMSU, HSB data in the presence of clouds. IEEE Trans. Geosci. Remote Sens. v.41 p.390-409.

Susskind, J., C.D. Barnet and J. Blaisdell 1998. Determination of atmospheric and surface parameters from simulated AIRS/AMSU sounding data: Retrieval methodology and cloud clearing methodology. Adv. Space Res. v.21 p.369-384.

Twomey, Sean 1996. Introduction to the mathematics of inversion in remote sensing and indirect measurements. Dover Publ. Inc., Mineola NY. 243 pgs.