Overview of the NOAA Operational Hyper-Spectral Infrared and Microwave Retrieval Algorithm

CrIS Atmospheric Chemistry Users Workshop

Thursday, Sep. 18, 2014

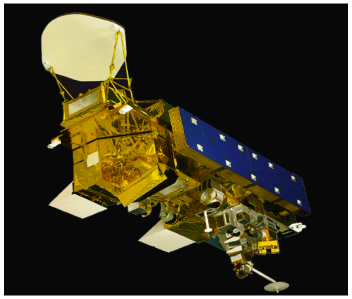
Chris Barnet

Discussion Points

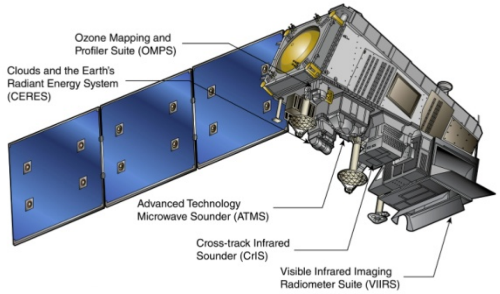
- Brief Introduction to the NOAA Unique CrIS/ATMS Processing System (NUCAPS) algorithm.
  - 1DVAR vs. Sequential
  - SVD vs. OE approaches (*i.e.*, geophysical *a-priori*)
- Separability of state parameters
  - CO2 and temperature separability
  - O3 tropopause relative first guess
- New Product ideas
  - Tracer-tracer correlation – indices for STE, etc.

Original Concept: Exploit existing *operational* assets to provide long-term trace-gas products

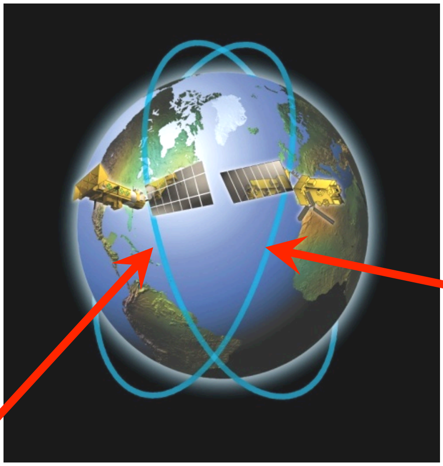
NASA/Aqua
1:30 pm orbit (May 4, 2002)

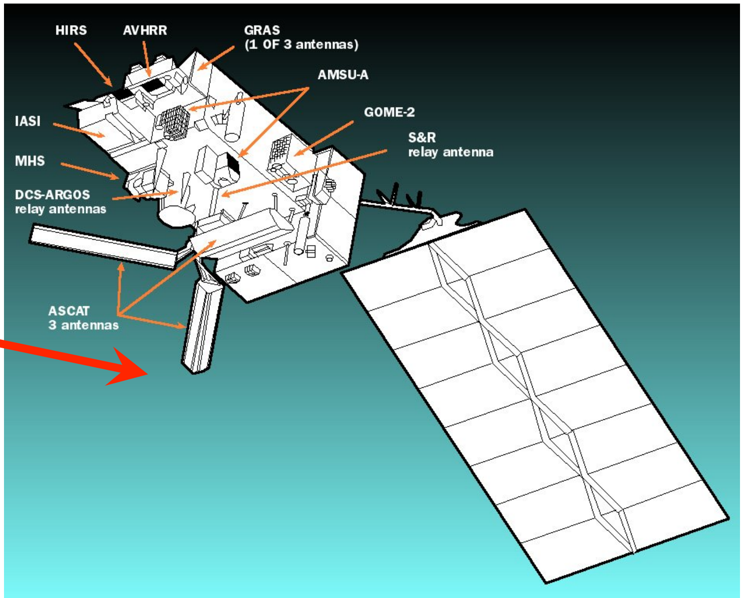


Suomi-NPP & JPSS
1:30 pm orbit
(Oct. 28, 2011, 2017, 2021)



See Barnet and Susskind 1999 Tech. Proc. Int'l TOVS Study Conf. v.10 p.22-33.

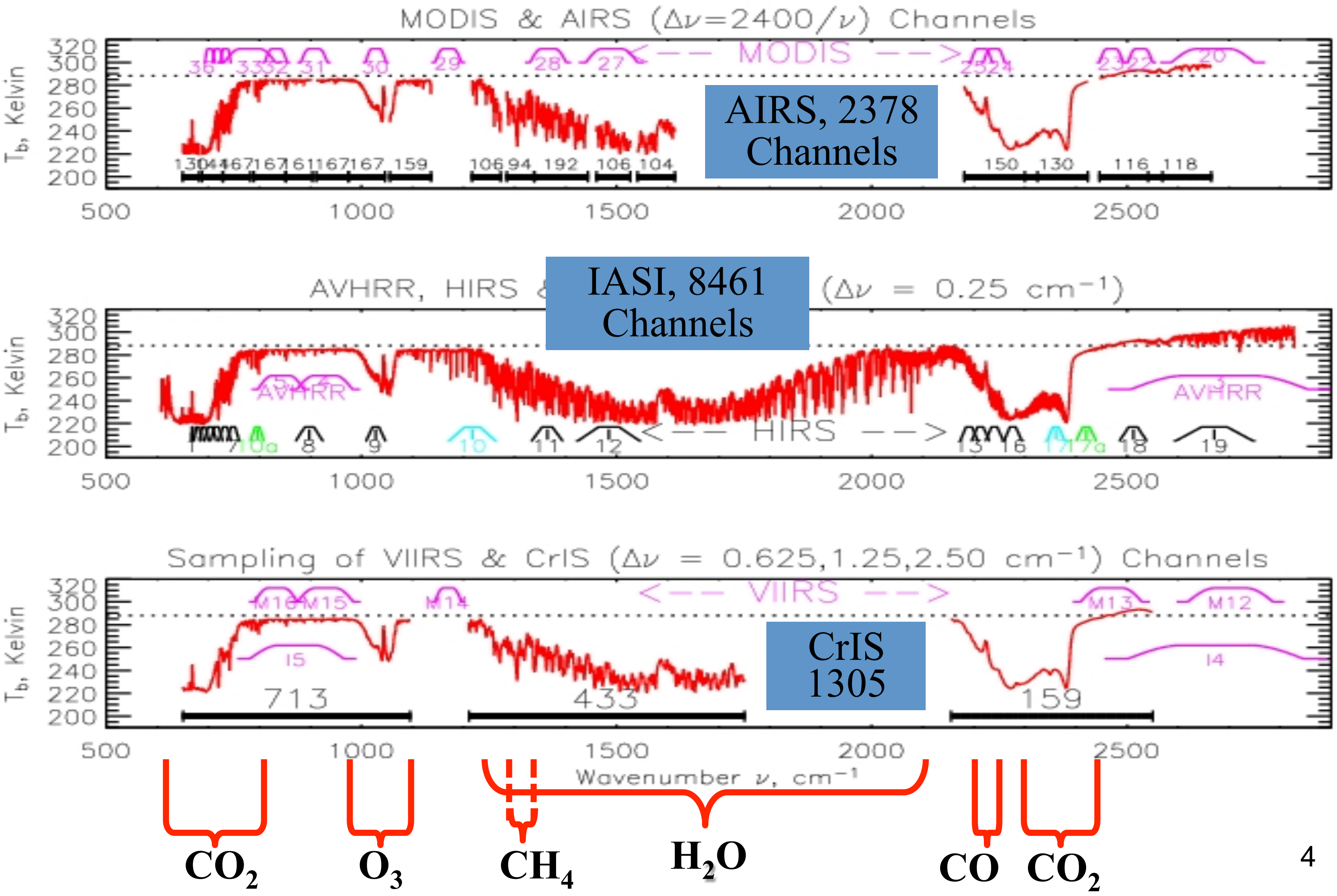




EUMETSAT/METOP-A+B
9:30 am orbit (Oct. 19, 2006, Sep. 17, 2012, 2017)

20+ years of hyperspectral sounders are already funded for weather applications

Spectral Coverage of Thermal Sounders & Imagers (Example Aqua, Metop, Suomi-NPP)



AIRS, 2378 Channels

IASI, 8461 Channels

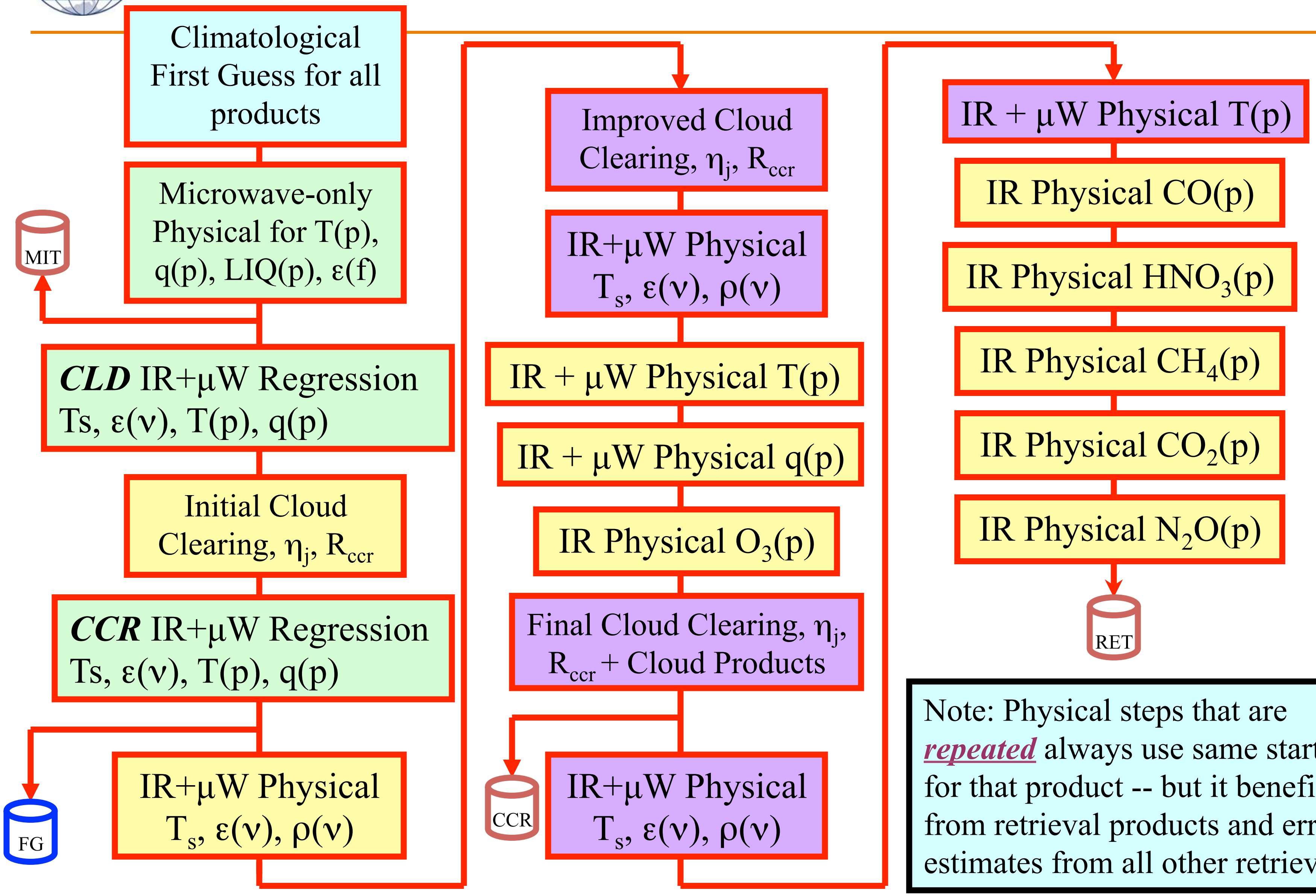
CrIS 1305

CO_2 O_3 CH_4 H_2O CO CO_2

Constraints and Assumptions for the AIRS Science Team (AST) Algorithm

- Must be able to process, end-to-end (using ≤ 10 250 MHz CPU's in 2002)
  - NUCAPS does ~1 retrieval per 0.12 seconds on modern CPUs
  - AIRS, IASI, and CrIS all acquire 1 FOR in ~0.27 seconds
- Only static data files can be used
  - One exception: model surface pressure.
  - Cannot use output from model or other instrument data.
  - Maximize information coming from AIRS radiances.
- Cloud clearing will be used to "correct" for cloud contamination in the radiances.
  - Amplification of Noise, A, is a function of scene $0.33 \le A < \approx 5$
  - Spectral Correlation of Noise is a function of scene
  - IR retrievals must be available for all Earth conditions within the assumptions/limitations of cloud clearing.

Flow Diagram of NUCAPS Retrieval Steps



Climatological First Guess for all products

Microwave-only Physical for T(p), q(p), LIQ(p), ε(f)

MIT

CLD IR+μW Regression Ts, ε(ν), T(p), q(p)

Initial Cloud Clearing, η_j, R_{ccr}

CCR IR+μW Regression Ts, ε(ν), T(p), q(p)

FG

IR+μW Physical T_s, ε(ν), ρ(ν)

Improved Cloud Clearing, η_j, R_{ccr}

IR+μW Physical T_s, ε(ν), ρ(ν)

IR + μW Physical T(p)

IR + μW Physical q(p)

IR Physical O_3(p)

Final Cloud Clearing, η_j, R_{ccr} + Cloud Products

CCR

IR+μW Physical T_s, ε(ν), ρ(ν)

IR + μW Physical T(p)

IR Physical CO(p)

IR Physical HNO_3(p)

IR Physical CH_4(p)

IR Physical CO_2(p)

IR Physical N_2O(p)

RET

Note: Physical steps that are ***repeated*** always use same startup for that product -- but it benefits from retrieval products and error estimates from all other retrievals.

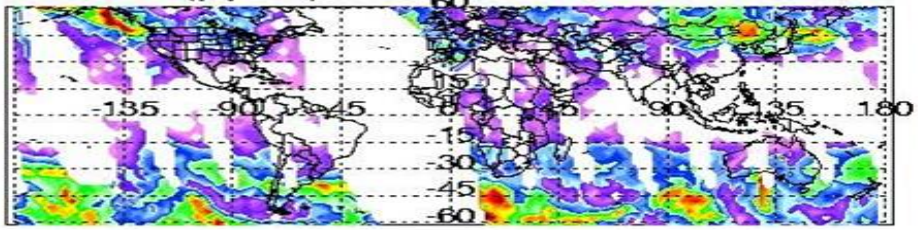
Summary of products from AIRS, IASI and NUCAPS Algorithm

gas	Range (cm^{-1})	Precision	d.o.f.	Interfering Gases	Sensitivity
T	650-800 2375-2395	1.5K/km	6-10	H2O,O3,N2O emissivity	surface to ~1 mb
H_2O	1200-1600	15%	4-6	CH4, HNO3	surf to 300 mb
Cloud P, T, fraction	700-900	25 mbar, 1.5K, 5%	≈2	CO2, H2O	surface to tropopause
O_3	1025-1050	10%	1+	H2O,emissivity	Lower strat.
CO	2080-2200	15%	≈ 1	H2O,N2O	Mid-trop
CH_4	1250-1370	1.5%	≈ 1	H2O,HNO3,N2O	Mid-trop
CO_2	680-795 2375-2395	0.5%	≈ 1	H2O,O3 T(p)	Mid-trop
Volcanic SO_2	1340-1380	50% ??	< 1	H2O,HNO3	flag
HNO_3	860-920 1320-1330	50% ??	< 1	emissivity H2O,CH4,N2O	Upper trop
N_2O	1250-1315 2180-2250	5% ??	< 1	H2O H2O,CO	Mid-trop

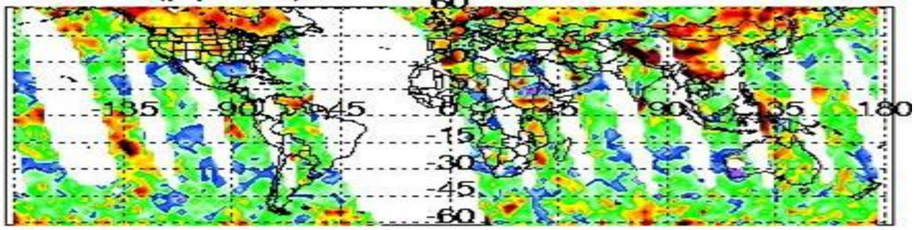
Example of AIRS Trace Gas Products

(Ascending Orbit, 1:30pm, Single Day)

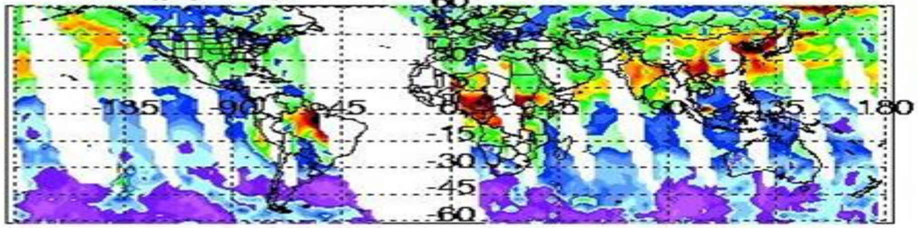
Ozone (ppbv), 20051201, at 6 - 10 km



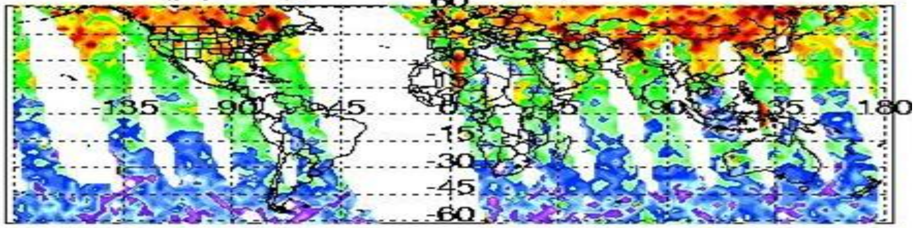
CO2 (ppmv), 20051201, at 6 - 10 km



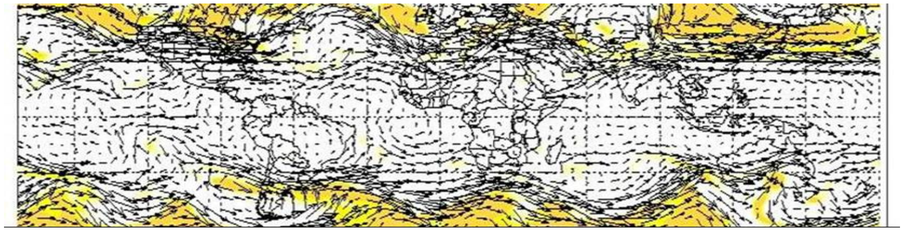
CO (ppbv), 20051201, at 6 - 10 km



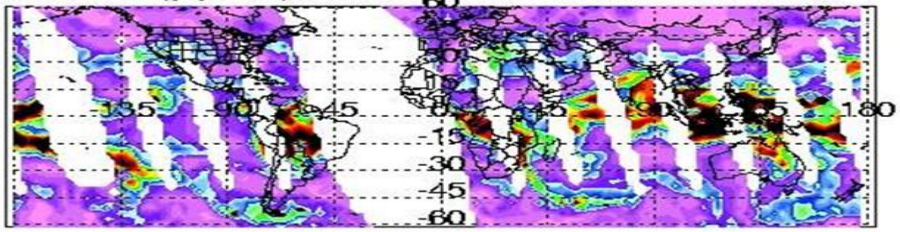
CH4 (ppbv), 20051201, at 6 - 10 km



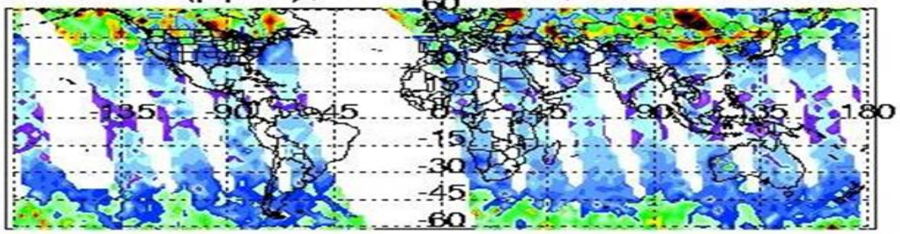
NCEP PV/Wind 20051201_18 at 300 hPa



H2O (ppbv), 20051201, at 6 - 10 km



HNO3 (pptv), 20051201, at 6 - 10 km



Stratospheric air masses (colored yellow in NCEP PV figure, where PVU ≥ 2) can be seen in AIRS upper tropospheric O3, CO, and HNO3 in the figures above. The H2O figure is scaled to show tropical convective features.

1DVAR versus AIRS Science Team Method

Simultaneous (1DVAR)	Sequential (AIRS method)
Solve all parameters simultaneously	Solve each state variable (*e.g.*, T(p)), separately.
Error covariance includes only instrument model.	Error covariance is computed for all *relevant* state variables that are held fixed in a given step. Retrieval error covariance is propagated between steps.
Each parameter is derived from all channels used (*e.g*., can derive T(p) from CO2, H2O, O3, CO, … lines).	Each parameter is derived from the *best* channels for that parameter (*e.g.,* derive T(p) from CO2 lines, q(p) from H2O lines, etc.)
A-priori must be rather close to solution, since state variable interactions can de-stabilize the solution.	*A-priori* can be less complex for sequential with well selected channels.
Regularization must include *a-priori* statistics to allow mathematics to separate the variables and stabilize the solution.	Regularization can be reduced (e.g., simple smoothing terms) and does not require *a-priori* statistics for most geophysical regimes.
This method has large state matrices (all parameters) and covariance matrices (all channels used). Inversion of these large matrices is computationally expensive.	State matrices are small (largest is 25 T(p) parameters) and covariance matrices of the channels subsets are quite small. *Very fast algorithm*. Encourages using more channels.
Has never been done simultaneously with clouds, emissivity(ν), SW reflectivity, surface T, T(p), q(p), O3(p), CO(p), CH4(p), CO2(p), HNO3(p), N2O(p)	*Can afford to repeat steps with improved knowledge of trace gas concentrations (i.e., repeated steps benefit from lower error estimates)*

Advantages of the AIRS Approach

- Sequential physical algorithm allows for a *robust* and *stable* system with minimal prior information
  - Sequential approach allows the more linear parameters to be solved for first -- can make the algorithm very stable
  - Can solve for all significant signals in the AIRS radiances.
- But … error from previous steps must be mapped into an error estimate from interfering parameters
  - A unique feature of this algorithm is that error estimates from previous steps are mapped into subsequent steps
    - Exploits *a-priori* information in forward model as a constraint
    - The observation covariance (S_ε in Rodgers 2000) contains both on- and off-diagonal terms composed of $(dR/dX)\cdot\delta x$ for all x's that are considered interference (including cloud clearing, correlation due to apodization, etc.).
  - Can be more robust than simultaneous retrieval because each step uses optimal sampling of channels (*i.e.*, low interference).

Advantages of optimal estimation

- O-E explicitly constrains the answer to lie within expectation of reasonable answers
  - Prior assumptions are always implicit in any retrieval approach
  - Note that "reasonable" can be in the *eye of the beholder* and sometimes that means a preference in the vertical null space.
- O-E explicitly derives the answer from prior information
  - in this sense, 1st guess can only speed up convergence
  - with enough iterations the same answer is usually achieved (up to non-linearity of Jacobians)
- Information content (or errors) in retrieval state can be partitioned between instrument and prior contributions
  - Averaging kernels or error covariance have more value
  - Modelers more likely to use product (rather than radiances)

Challenge #1: How to combine instruments with different characteristics

- AIRS/AMSU, IASI/AMSU/MHS, and CrIS/ATMS are processed with literally the same code.
  - Same underlying spectroscopy
  - Instrument specific items are file-driven
  - Code is backward and forward (as much as possible) compatible.
  - Operational code is a "filtered" version of the science code.
- Statistical *a-priori* for temperature and moisture derived from AIRS radiances
  - All channels used in *<u>constrained</u>* regression first guess
  - Captures high-vertical resolution content of T and q
- Physical-based approach
  - Avoid empirical corrections (including arbitrary *a-priori* constraints)
  - Use physical constraints (derived from both spectroscopy and geophysical variability) to regularize low information content domains
  - Avoid any unnecessary approximations that can induce systematic biases.

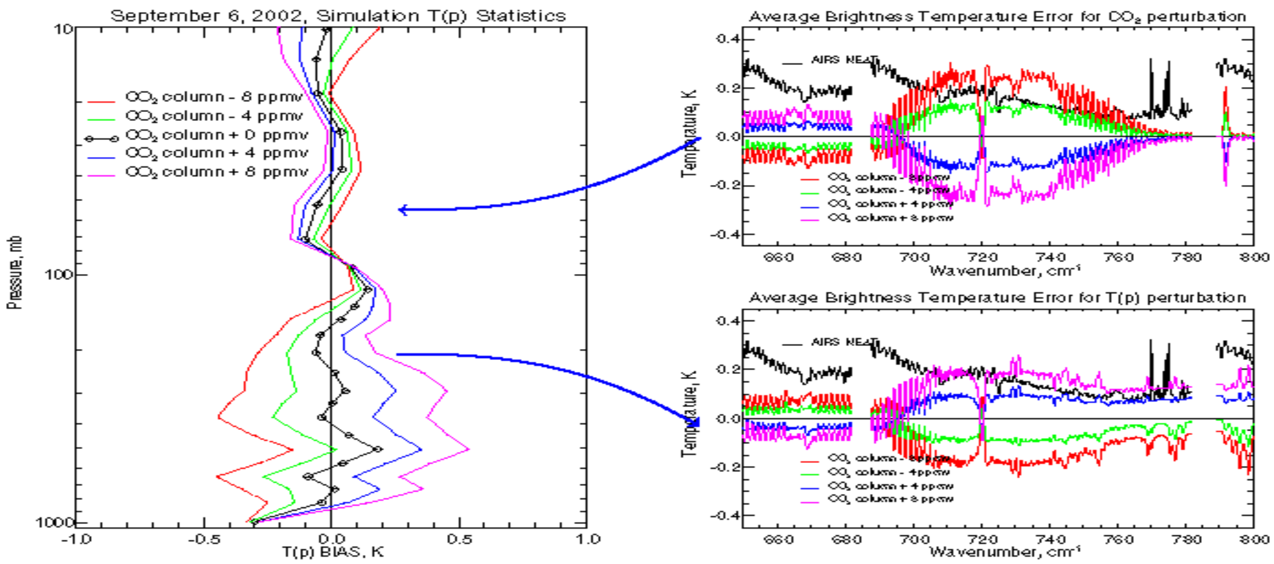
Challenge #2: Separating effects of clouds/ surface, T/CO2, q/CH4, etc.

- Physical, 1st principles based algorithm using knowledge of radiative transfer to identify unique spectral "fingerprints." Problem areas are:
  - Sensitivity of temperature sounding channels to CO_2
  - Sensitivity of temperature sounding channels to N_2O.
  - Sensitivity of cloud clearing to surface gradients and low clouds.
- Approach the problem as a physics problem.
  - With trace gases *a-priori* information is limited and there are many geophysical correlations (T/CO2, Ts/CH4, CO/CH4/O3) and spectral correlations (CO2,O3,H2O in 15 um band)
  - Solve problem sequentially (Taylor expansions) solving for most linear (including cloud clearing) or high S/N parameters first.
- Also, inversion solutions are not unique
  - In information limited regions (cold scenes, low lapse rate, uniform clouds) we must use statistics as a constraint.
  - Product becomes more difficult to use – need to convey variable information content (*e.g.,* vertical averaging functions) to users.

CO$_2$ and temperature Jacobians are similar

(see Maddy et al. 2005 OSA)

- When retrieval is told the wrong CO2 it results in a vertically biased T(p) and radiance bias.

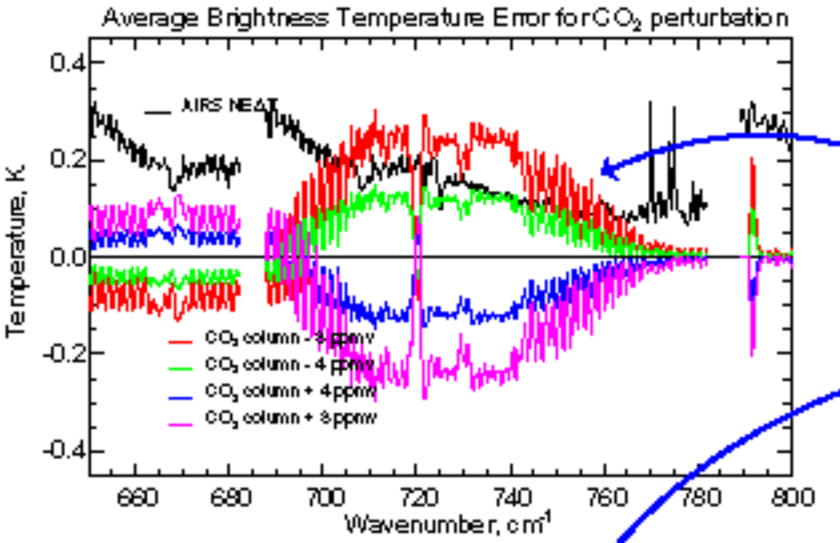


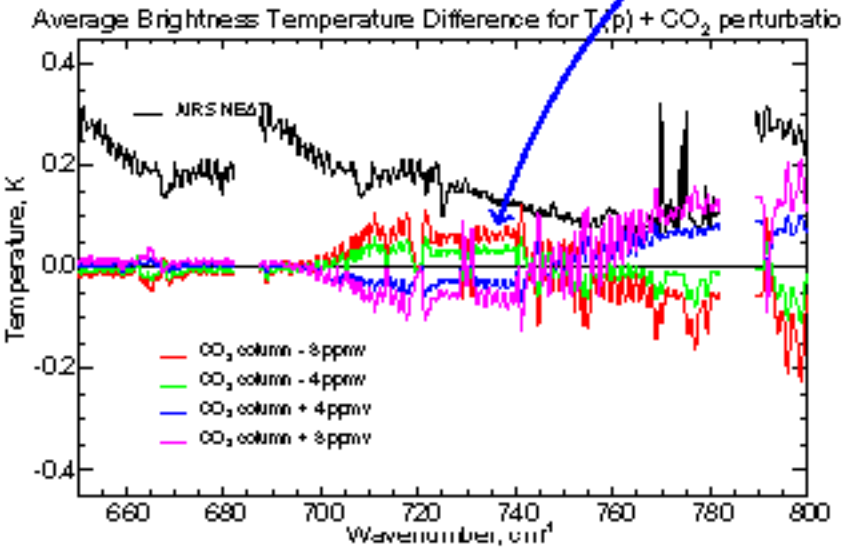
To 1st order, CO2 and T Jacobians are mirror images of each other

One exception is 790 cm^{-1} channel

- AMSU 57 GHz and multi-spectral IR (15 and 4.3 μm and *a-priori* information must be used to separate T and CO2.

To first order, T(p) and CO2 biases can cancel making the separability difficult





- Average BT error due to ±4 and ±8 ppm CO$_2$ perturbations
- Sum of average BT error resulting from T(p) bias plus CO$_2$ perturbations
- In essence, this is the signal, not the Jacobian above, that tells us the CO$_2$ prior is wrong.

Why are CO_2 averaging functions broad while T(p) functions have profile information?

- Spectroscopy: The CO2 lines are strong narrow lines. Temperature affects the width of line while # of CO2 molecules, N_i, affects the strength. Once the line is saturated (near the surface, where p is large) we loose sensitivity.

$$\kappa_i(\nu, p, T, \theta) \simeq \sum_{j=1}^{J} \frac{N_i \cdot S_{ij}}{\pi} \frac{\gamma_{ij}}{(\nu - \nu_{ij})^2 + (\gamma_{ij})^2} \cdot \sec(\theta) \qquad \gamma_{ij} \simeq \gamma_{ij}^0 \cdot \frac{p}{P_0} \cdot \sqrt{\frac{T}{T_0}}$$

- Radiative transfer: The temperature enters both in the absorption coefficient (above) and in the Planck function.

$$R_n(\vec{X}) \simeq \int_{\nu} \Phi_n(\nu) \int_{p} B_\nu(T(p)) \cdot \frac{\partial \exp\left(-\int_{z'=\infty}^{z(p)} \sum_i \kappa_i(\vec{X}, p, \ldots) dz'\right)}{\partial p} \cdot dp \cdot d\nu$$

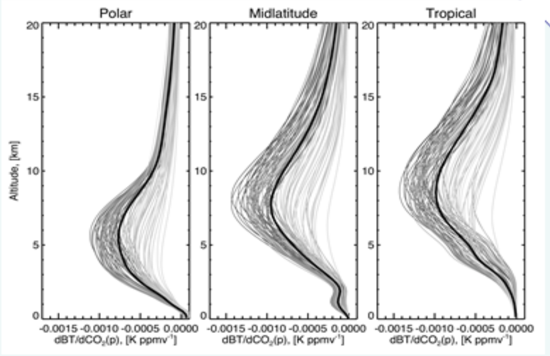
- Change in concentration is more uniform vertically as it affects all channels in a proportional way
- Individual channels select vertical layers by strength of the lines, S_{ij}, and is enhanced due to the Planck function sensitivity

Validation of CO2

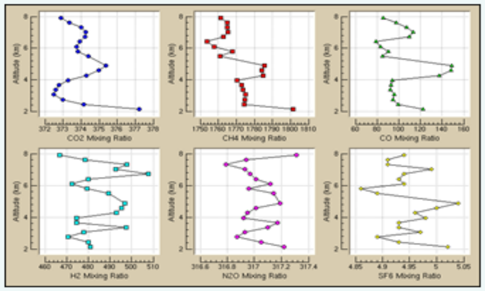
(see Maddy 2008 JGR, doi:10.1029/2007JD009402)

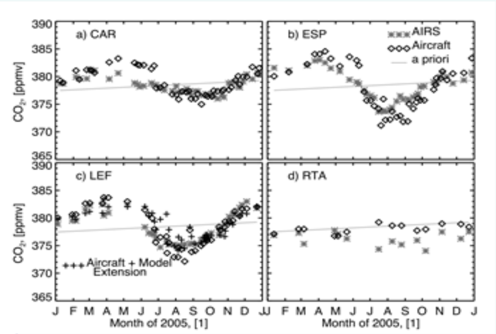
Validation

Thermal sounders, such as AIRS, IASI, and CrIS, measure traces in a thick tropospheric column averages. The vertical region sounded is a function of the atmospheric state as shown at right for CO2



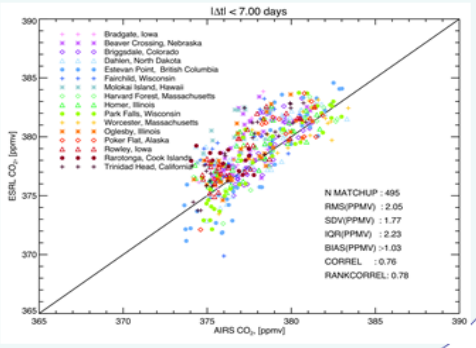
The best in-situ validation products are gas flask samples taken during aircraft flights. The NOAA/ESRL monitoring network provides high precision vertical profiles for a number of locations



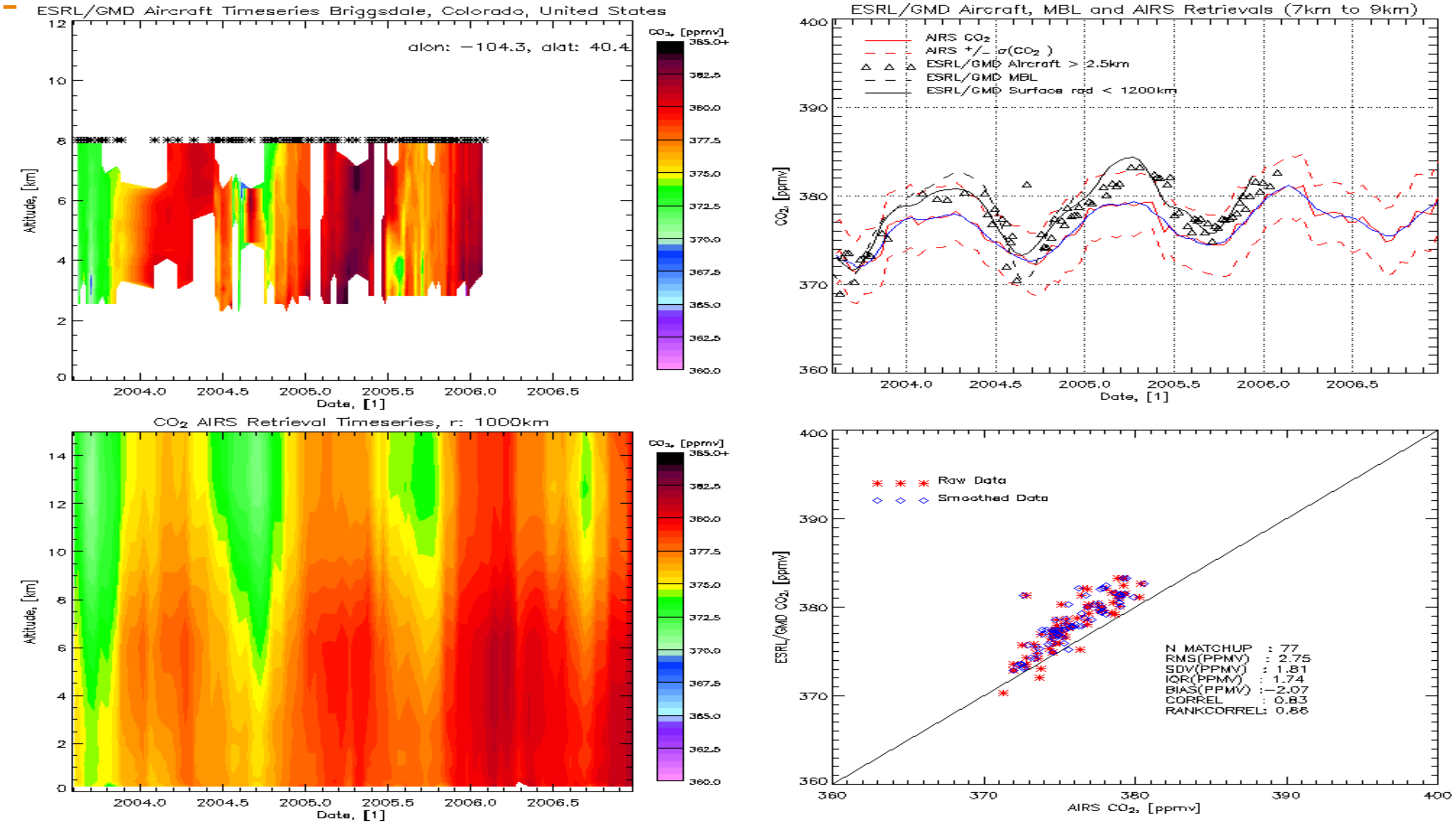


Comparison of AIRS CO2 product with aircraft measurements at 4 ESRL sites shown as a function of time (left)

and as a scatter plot for all ESRL sites (right)

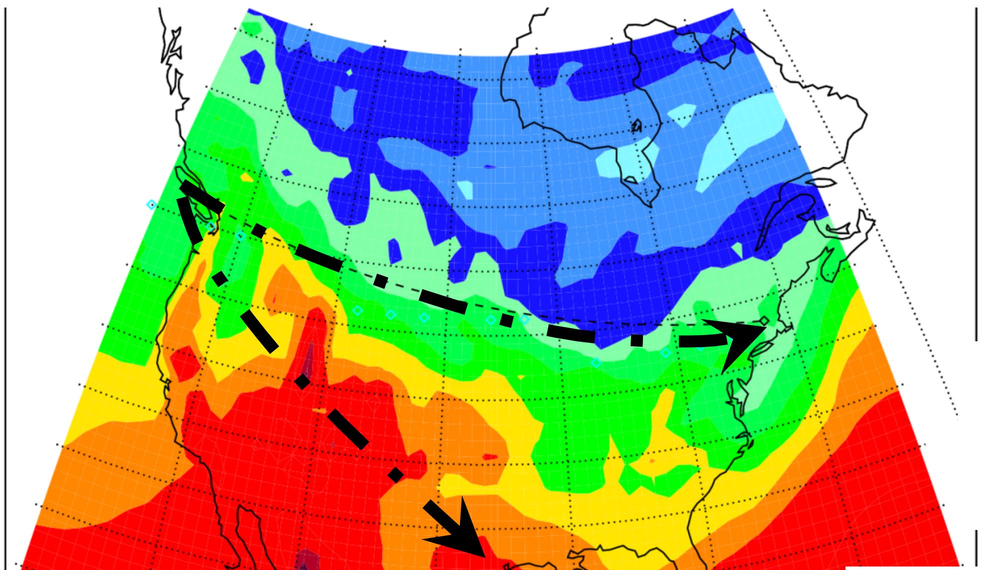


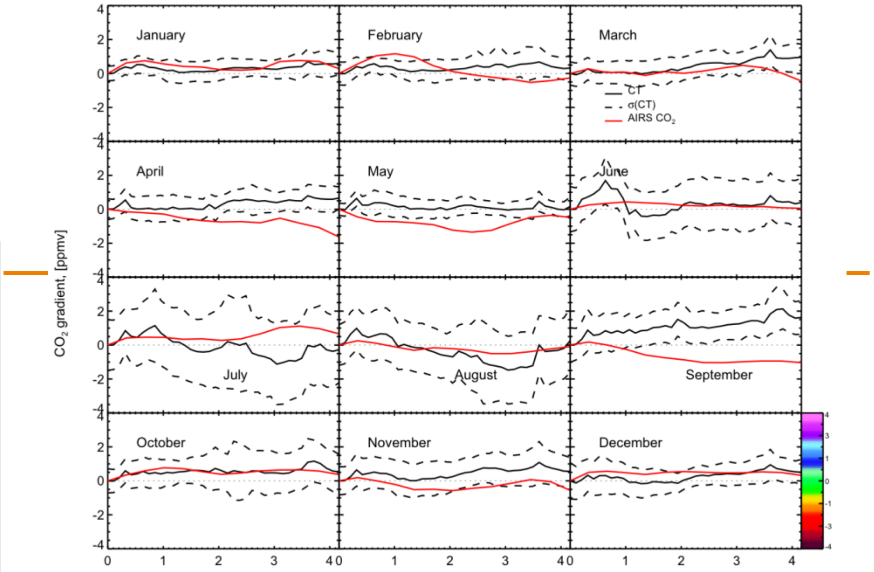
Comparison of NOAA CO_2 product with *in-situ* aircraft at Carr, CO



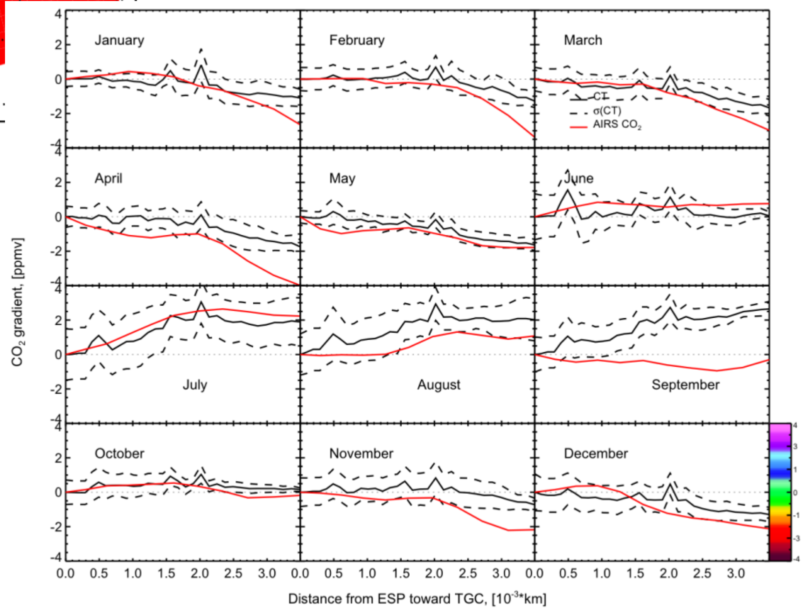
NOTE: Currently we are using a 1:24 spatial sampling for these comparisons from our “3x3 global grid” reprocessing dataset.

NOAA CarbonTracker Upper Troposphere CO2 Gradient Over N. America Relative to NW US For July





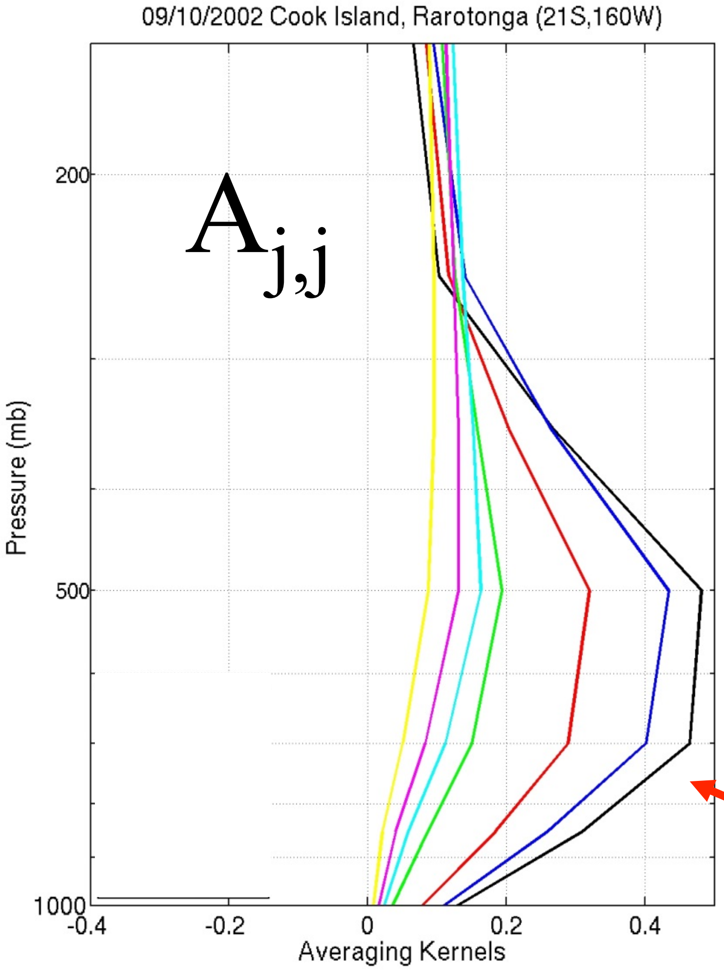
(Above) Comparison of Monthly Average **AIRS** and **CarbonTracker** Gradients From Northwest US to Massachusetts

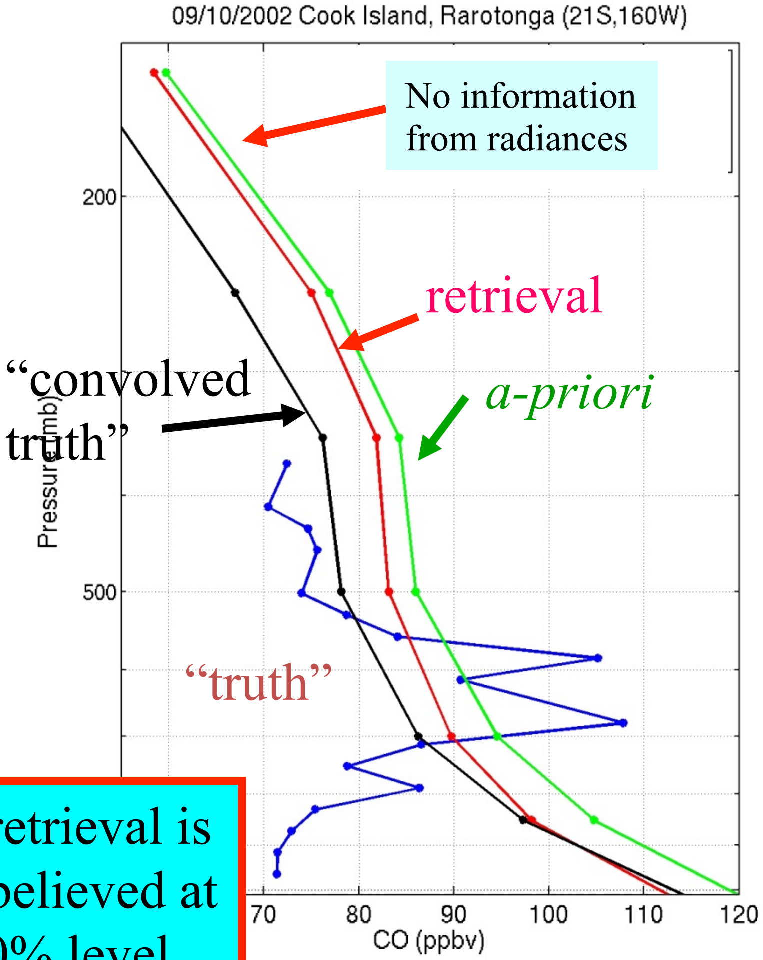


(Right) Monthly Average **AIRS** and **CarbonTracker** Gradients From Northwest US to Texas

Validation has been used to identify and mitigate problems in the AIRS forward model.

Example of AIRS CO product





No information from radiances

retrieval

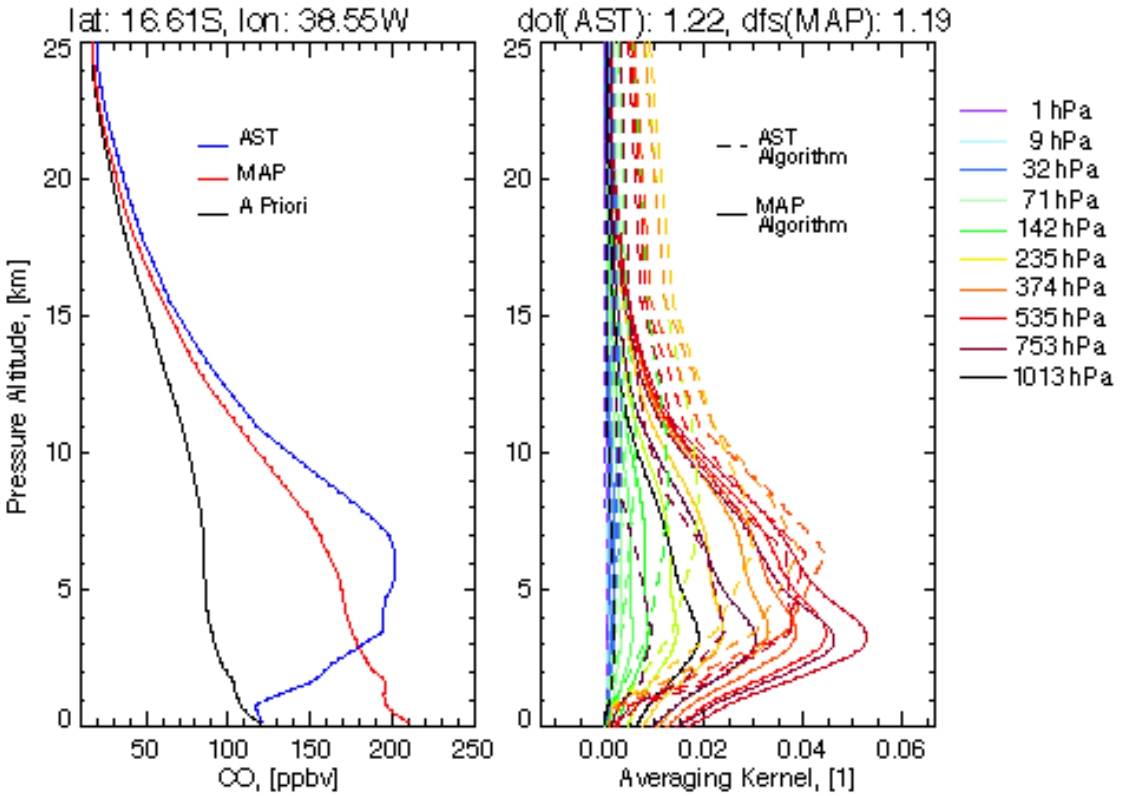
a-priori

“convolved truth”

“truth”

This retrieval is only believed at the 50% level

Comparison of NOAA/STAR optimal estimate and AIRS science team algorithm’s CO product (Maddy 2009 IEEE Geosci. V.6 p.802)



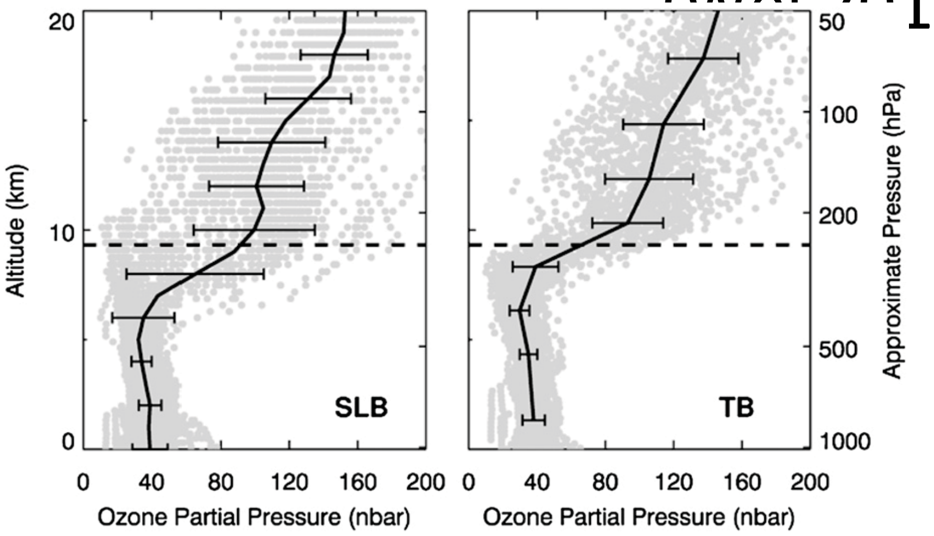
- AIRS science team approach uses regularized least squares without a prior constraint.
- This impacts the averaging kernels in the sense that better information can be acquired if profile shape is more realistic (less errors in Jacobian, *K*)

Above left: Optimal estimation (red) and AIRS science team (blue) methods produce similar total column amounts. Both profiles have no-skill in lower 3-km, but O-E profile is statistically. more realistic.

Above right: O-E averaging kernels (solid) are slightly lower, therefore, O-E allows more lower tropospheric sensitivity (in this case).

Tropopause based *a-priori* improves O_3 retrieval results near the tropopause

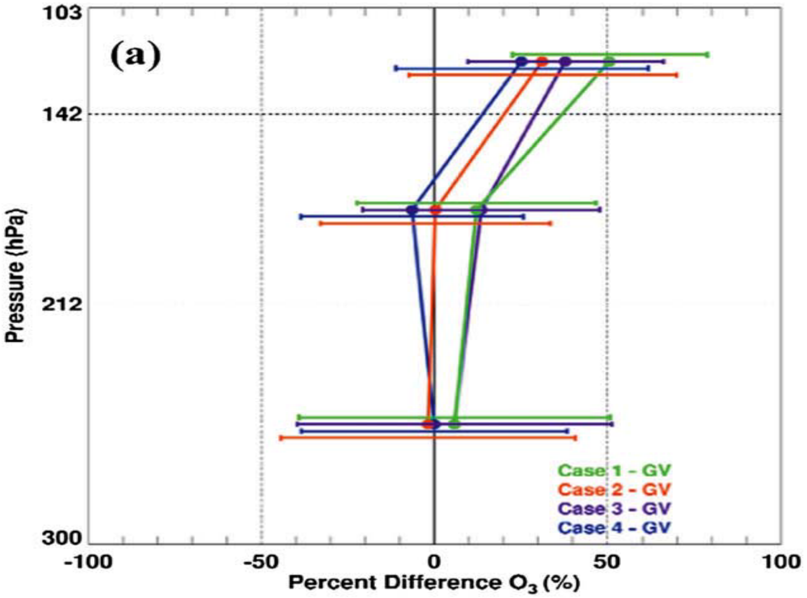
(Wei 2010 JAOT 27 1123)



Sea-level based (SLB, left) versus Tropopause Based (TB, right) ozone climatology

Near the tropopause the TB climatology provides a better shape

- This is region where IR hyperspectral has most sensitivity

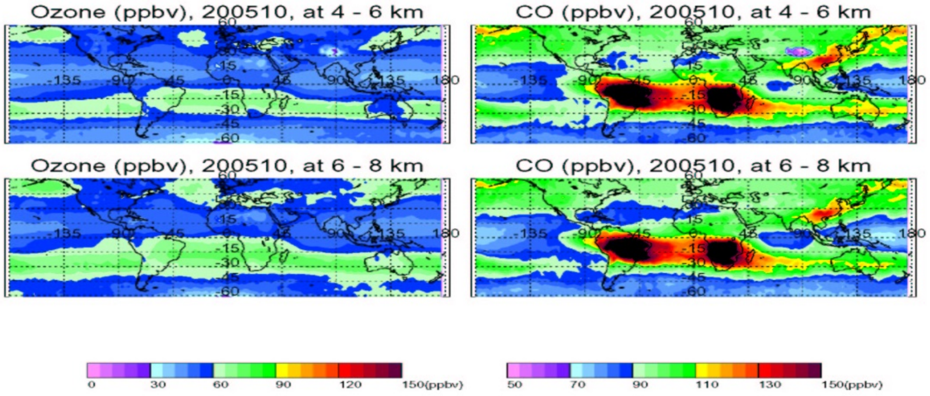


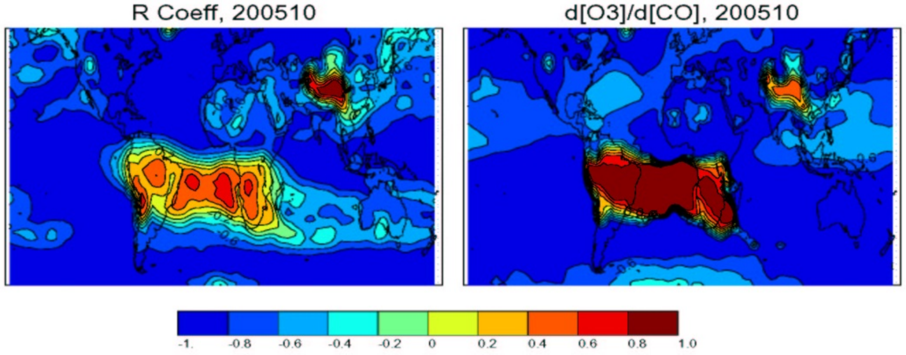
Case 1: AST w/ SLB   Case 3: O-E w/ SLB
Case 2: AST w/TB   Case 4: O-E w/ TB
GV = Start-2008 Gulfstream-V measurements.

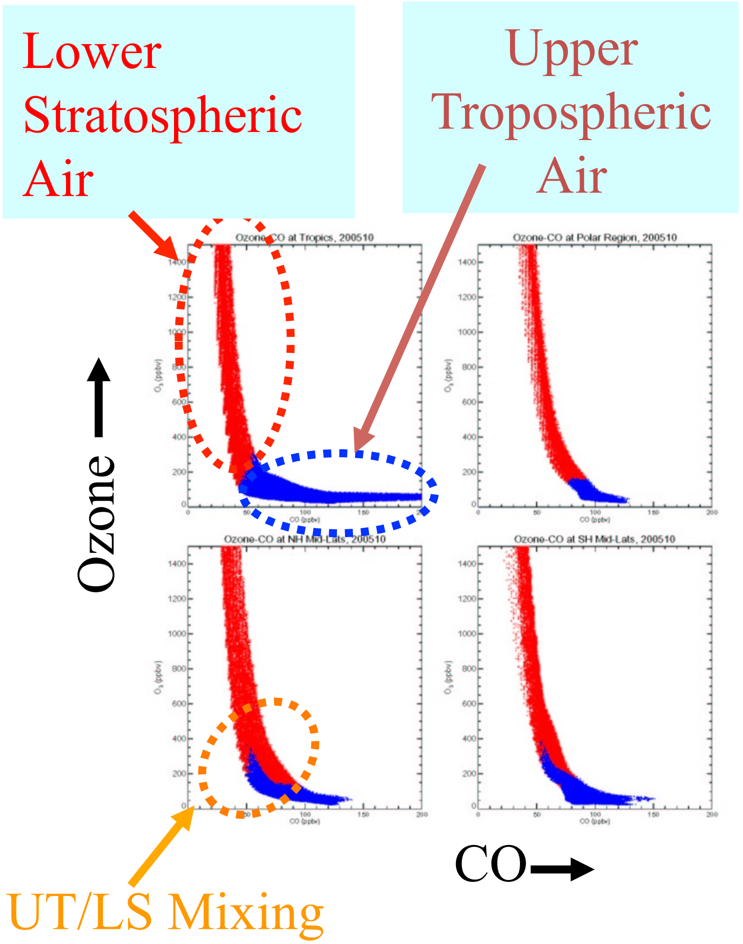
Shape preserving retrievals (perform better with TB (Case 2 and 4)

Tracer-Tracer correlations can define regions (AIRS v5.0 O_3 and CO products)

Production of O_3 in biomass burning regions (high CO production)







Lower Stratospheric Air

Upper Tropospheric Air

UT/LS Mixing

See L. Pan et al. JGR 2007 doi:10.1029/2007JD008645

Future Work

- Last time I gave a talk on hyperspectral trace gas products was April 2009
  - Nothing has really changed with the algorithm since then
- Need to properly propagate errors from upstream steps
  - Focus of my recently funded NASA-NPP proposal
- Upstream T/q and downstream trace gas steps need to be converted from SVD to O-E
  - CO has been implemented already
  - CO2 and O3 versions exist, but were not implemented
- Need to utilize tropopause relative methodology for O_3
  - May consider similar ideas for CH_4, HNO_3
- Need to explore derived tracer-tracer index products.

Biggest challenge – funding

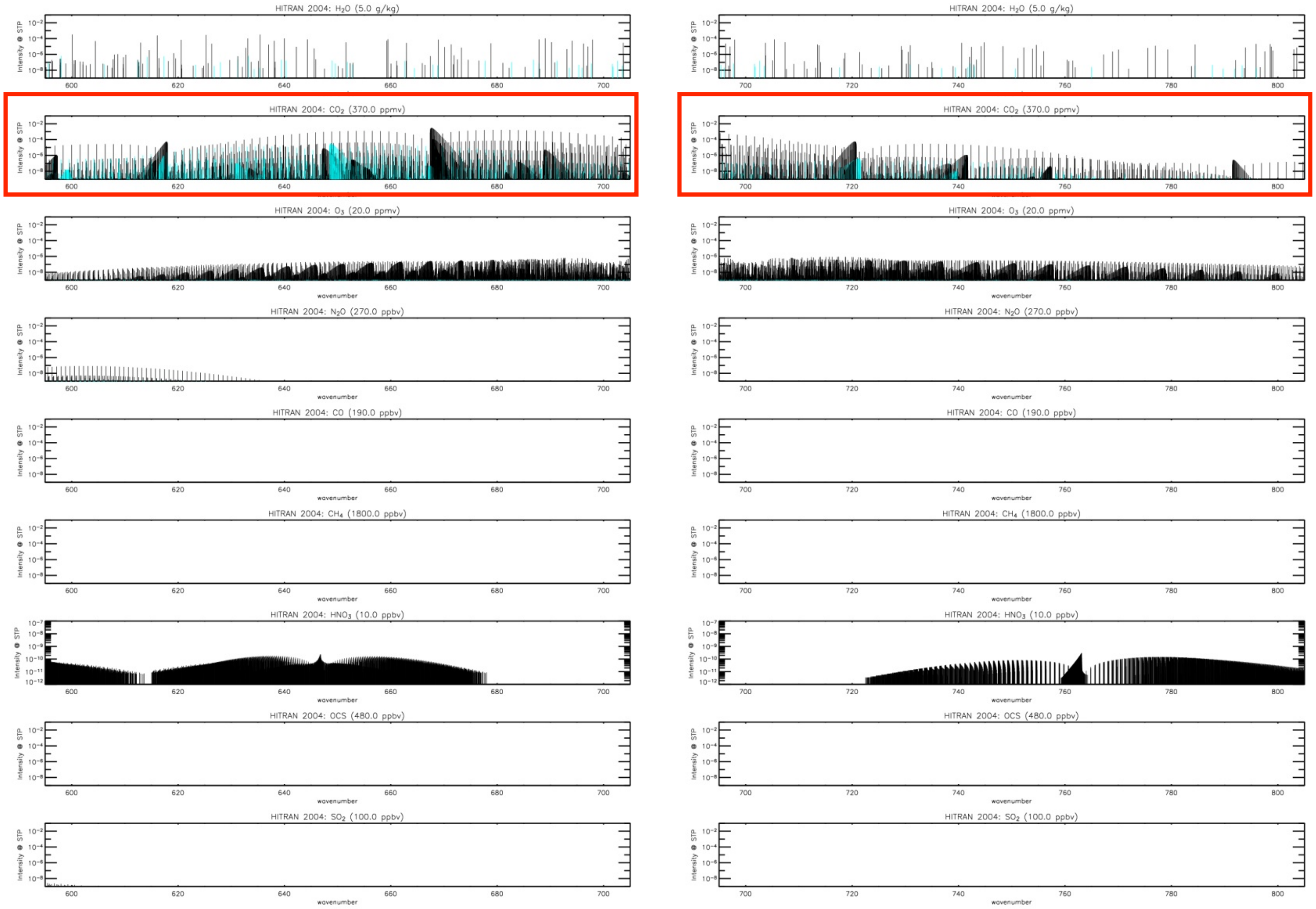
"I was close to a breakthrough when the grant money ran out."

QUESTIONS?

Example of vibration rotational line strengths in 15 μm band region

600 to 700 cm^{-1}

700 to 800 cm^{-1}



H2O

CO2

O3

N2O

CO

CH4

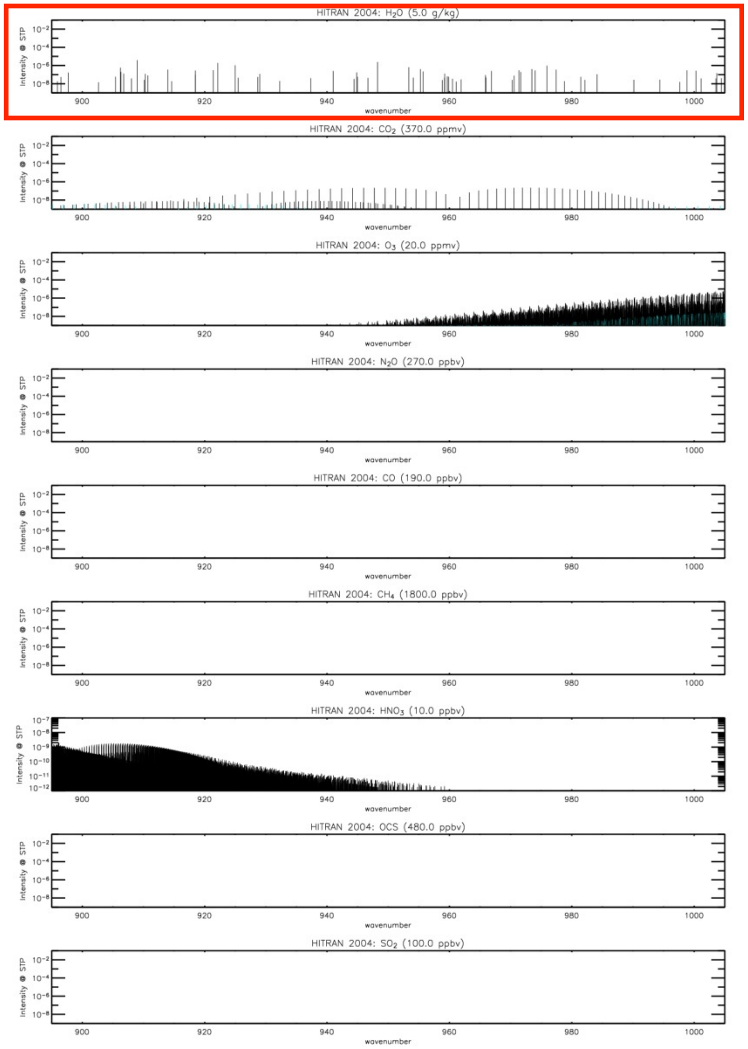
HNO3

OCS

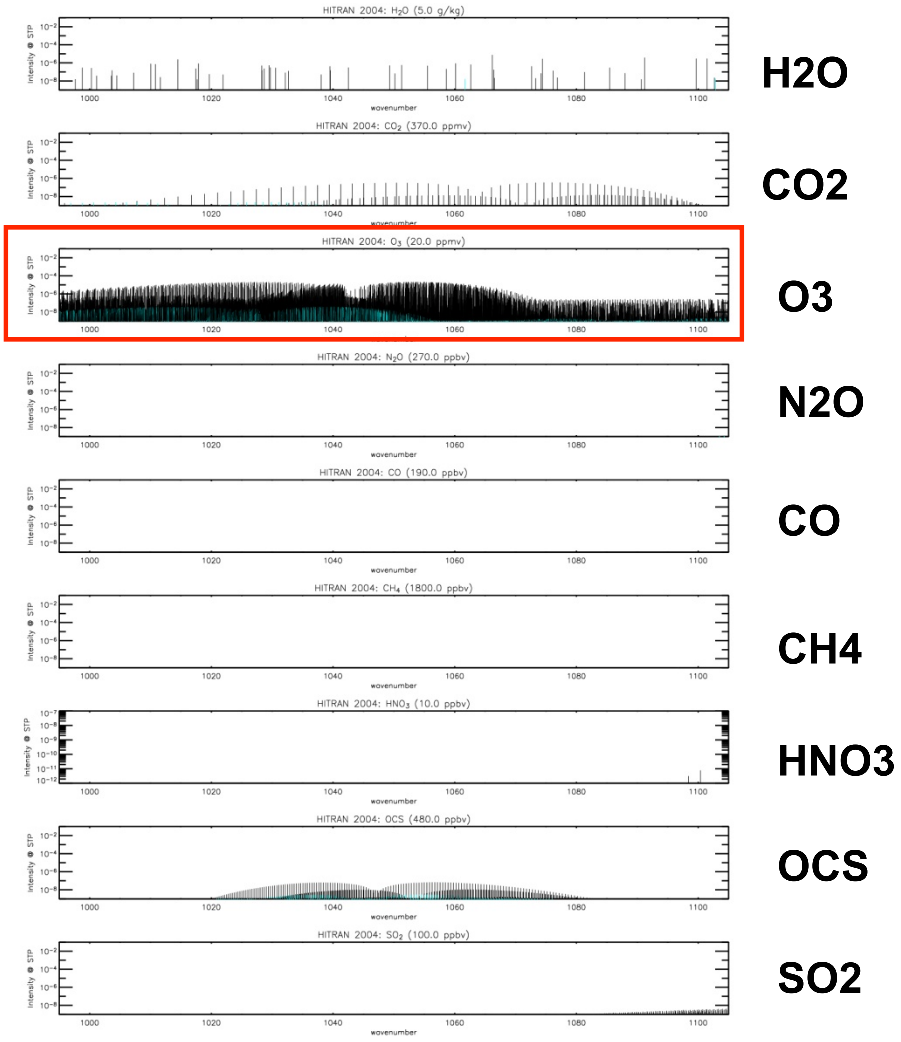
SO2

Example of vibration rotational line strengths in 10 μm band region

900 to 1000 cm^{-1}

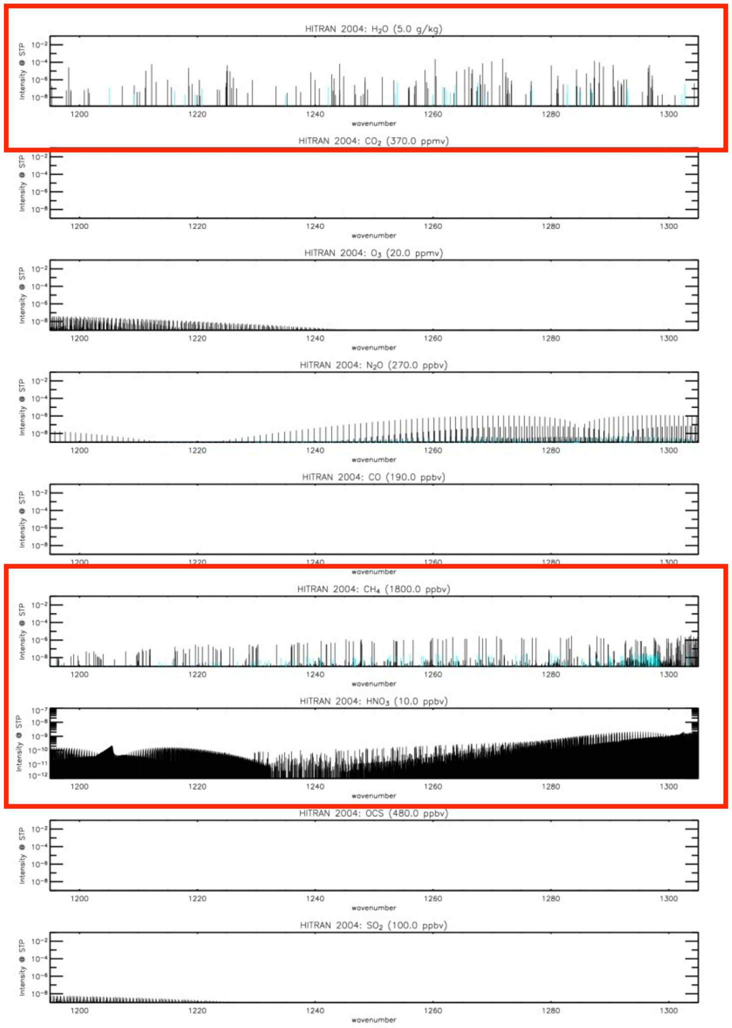


1000 to 1100 cm^{-1}

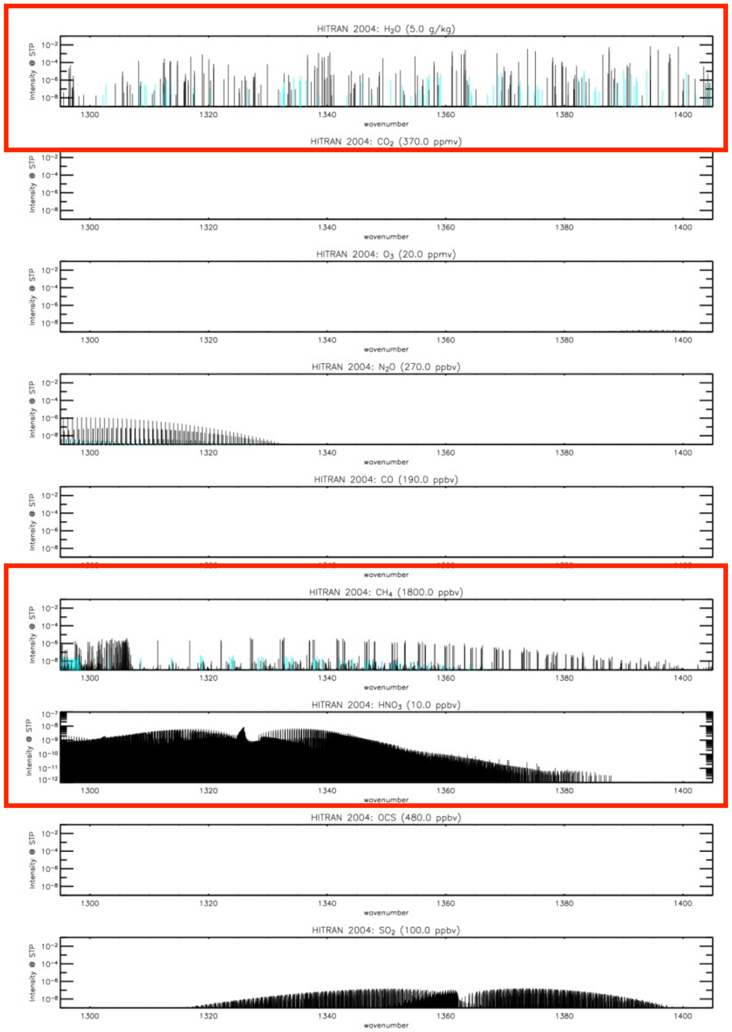


Example of vibration rotational line strengths in 6 μm band region

1250 to 1350 cm^{-1}



1350-1450 cm^{-1}



H2O

CO2

O3

N2O

CO

CH4

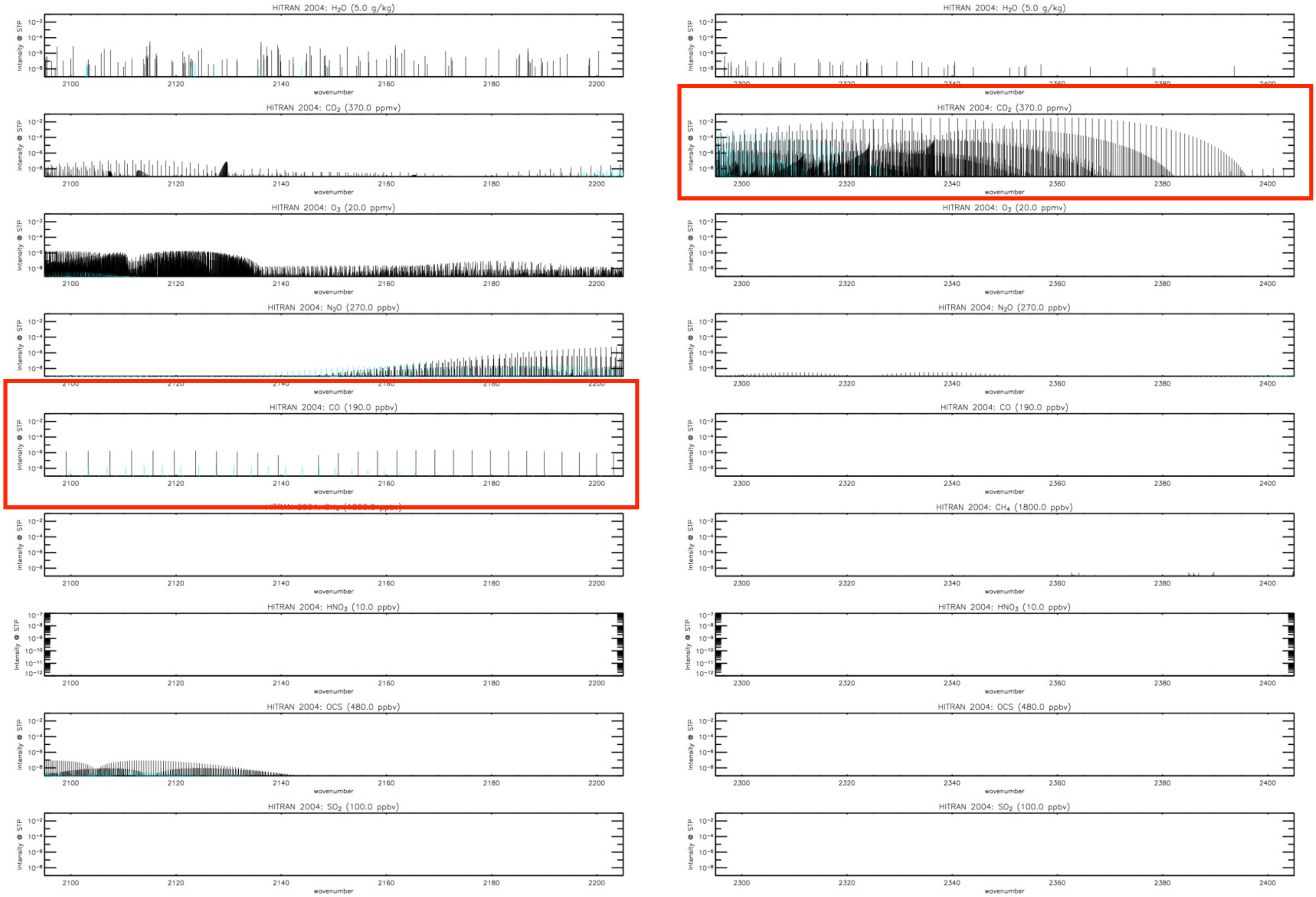
HNO3

OCS

SO2

Example of vibration rotational line strengths in 4 μm band region

2100 to 2200 cm^{-1} **2300 to 2400 cm^{-1}**



H2O

CO2

O3

N2O

CO

CH4

HNO3

OCS

SO2