Using AIRS CO to Improve Ground Based CO measurements in the Boundary Layer

Chris Wilson¹

1. NASA JPL / California Institute of Technology
Outline

• Motivation
• Data setup
• Retrieval Setup
• Results with a priori first guess
• Results with AIRS first guess
Introduction and Motivation

Reasons to study carbon monoxide

- Surface Air quality
- Climate Change
- Tracer of other pollutants

Ground and AIRS DATA Symbiosis

- AERI (Ground) based CO measurements have peak sensitivity near the surface
- AIRS CO measurements have peak sensitivity in the middle troposphere

*McMillan et. Al. 2011
\[ x_{ret} = R(y, b, x_{fg}, c) \pm \varepsilon_R \]
\[ y = F(x_{true}, b) \pm \varepsilon \]

**New Technique:** to use AIRS v5 retrieved CO profiles as a first guess in a ground based retrieval

- Department of Energy has a well instrumented radiation measurement site in Oklahoma (ARM)
- Unique Site in relation to instrumentation
- 57 Flights between 2007-2008 measuring profiles of CO
True Composite CO profiles

- 57 true Composite profiles made up of ground measurements, aircraft measurements, and satellite retrievals
- The forward model requires a finer grid than the true CO profile
Atmospheric Emitted Radiance Interferometer (AERI)

- Measures downwelling IR radiation spectra (3 – 18 μm) passively
- 1.037 cm Path length results in ~0.48 cm\(^{-1}\) resolution

- Calibrated spectrum every 7-8 minutes
- 178 total spectra per day
Brightness Temperature Contrast Cloud Flag (BT)

- Clouds emit radiation like a blackbody
- Clear: Max BT diff (2100 - 2200 cm⁻¹) > 40

Cons
- Identifies low thick clouds
- Used in version 1 CO retrieval
- Misses thin cirrus clouds
- Misses scenes where clouds only partially fill the field of view
- Previous uses did not quantify effect on radiation

Compared this with a recently developed neural network that identifies 98% of cloudy
Neural Network Cloud Flag (NN)

- Neural Network trained on AERI, micro-pulse LIDAR, and millimeter wave cloud radar
- Accurately identifies 98% of cloudy scenes and 93% of clear scenes
- Missclassifies optically thin cirrus when radiation is attenuated by large amounts of water vapor

- 800,000 cases from AERI-01 instrument
- June 1996 – May 2010
- Thick/thin 70 min average > 250 K
- Tri-modal behavior

From Turner et. al. 2010
Cloud Flag Analysis 2007

<table>
<thead>
<tr>
<th>Cloud Filter</th>
<th>Total Spectra</th>
<th>Spectra Analyzed</th>
<th>Clear spectra</th>
<th>Percent Clear</th>
</tr>
</thead>
<tbody>
<tr>
<td>BT</td>
<td>62964</td>
<td>57345</td>
<td>37319</td>
<td>65</td>
</tr>
<tr>
<td>NN</td>
<td>62964</td>
<td>58852</td>
<td>13735</td>
<td>23</td>
</tr>
</tbody>
</table>

2007 Window Channel (2133) NN Filter

2007 Window Channel (2133) BT Filter
Building the CO Retrieval
CO Jacobians

\[ K = \frac{\partial F_v(x)}{\partial x(p)} \]

- 180 x 90 - 100 Matrix
- \( m = 180 \) spectral channels
- \( n = 90 - 100 \) pressure layers
  (depends on surface pressure)

- Jacobians are calculated using finite difference
- CO jacobians are regularly spaced
- No vertical differences results in nearly linearly dependent Jacobians
- Large redundancy in the channels
Error Components

\[ y = F(x, b) \pm \varepsilon \]
Full Error Covariance Matrix

Wavenumber (cm$^{-1}$)

Full Spectral Error Covariance ($S_e$)

Spectral Error Variance (Radiance$^2$)

Wavenumber (cm$^{-1}$)

Full Spectral Error Covariance ($S_e$)

Spectral Error Variance (Radiance$^2$)

Wavenumber (cm$^{-1}$)

Full Spectral Error Covariance ($S_e$)

Spectral Error Variance (Radiance$^2$)

Wavenumber (cm$^{-1}$)

Full Spectral Error Covariance ($S_e$)

Spectral Error Variance (Radiance$^2$)

Wavenumber (cm$^{-1}$)
### Information Content of the Retrieval

**Rank of the Problem = 1**

- Number of independent pieces of information able to be determined greater than measurement error
- Determined by singular values > 1: 
  \[ S_e^{-1/2} K S_a^{1/2} \]
  (Signal to noise term)

\[ \Lambda_1 = 93 \quad \Lambda_2 = .5891 \]

### Averaging Kernel

- Describes the sensitivity of the retrieval to the true atmospheric state: 
  \[ \frac{dx_{\text{ret}}}{dx_{\text{true}}} \]
- Determined by \( A = GK \)
- Primarily sensitive to changes in first 2 layers

### Shannon Information Content

- Measurement process reduces entropy of a system (improves knowledge)
- Used in a selection process to determine 19-21 best measurement channels
- **New** selection process accounts for correlated error

\[ H_s = -\frac{1}{2} | I_n - A | \]
Non-linear Retrieval solution

• Retrieval solution finds \( x \) for the relation: \( y = F(x, b) + \epsilon \)

\[
x_{i+1} = x_1 + (K^T S_e^{-1} K + H)^{-1} K^T S_e^{-1} [(y - F(x_i)) + K(x_i - x_1)]
\]

From Rodgers 2000

• Constraining to the first guess profile (\( x_1 \))
• Level of constraint is defined by \( H \) (inverse of the variance for a layer)
• CO signal: \( y - F(x_i) \)
• \( Se \) removes sensitivity to noise inside \( y - F(x_i) \)

### Operational Solution

<table>
<thead>
<tr>
<th>Practical Problems with Solution</th>
<th>Fix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inverse is unstable for large numbers (( K \sim 1 \times 10^7 ))</td>
<td>Scale Jacobians to radiances (Retrieve a percent change)</td>
</tr>
<tr>
<td>( x ) has a large null space</td>
<td>Retrieval is calculated on retrieval grids</td>
</tr>
<tr>
<td>Step size must be linear</td>
<td>factor in a dampening factor (step &lt; 15%) (( a = 0.01:1 ))</td>
</tr>
</tbody>
</table>

\[
x_{i+1} = I \left( \frac{z_1}{z_i} + \alpha \right) x_i
\]

(I stands for interpolation function) 

\[
x_{i+1} = \beta x_i
\]
Retrieval Results with *a priori* first guess

28 of 57 days determined to be cloud free
One Layer Results

NEW

Version 1

No Bias spectrum is used in either plot above

Conclusions

- Centering data improves results from 11% to 6%
- Spectral Bias and not clouds causes version 1 retrieval bias

One Layer Summary

<table>
<thead>
<tr>
<th>Bias Spectrum</th>
<th>Slope</th>
<th>Intercept</th>
<th>$R^2$</th>
<th>RMSE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Bias</td>
<td>0.87</td>
<td>0.06</td>
<td>0.89</td>
<td>11</td>
</tr>
<tr>
<td>BT Bias</td>
<td>0.99</td>
<td>0.01</td>
<td>0.89</td>
<td>6</td>
</tr>
<tr>
<td>NN bias</td>
<td>1.02</td>
<td>-0.05</td>
<td>0.89</td>
<td>6</td>
</tr>
</tbody>
</table>
Profile Results

RMS Difference on Retrieval Grid

<table>
<thead>
<tr>
<th>Pressure (mb)</th>
<th>First Guess</th>
<th>Retrieval</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>200</td>
<td>0</td>
</tr>
<tr>
<td>100</td>
<td>400</td>
<td>100</td>
</tr>
<tr>
<td>200</td>
<td>600</td>
<td>200</td>
</tr>
<tr>
<td>300</td>
<td>800</td>
<td>300</td>
</tr>
<tr>
<td>400</td>
<td>1000</td>
<td>400</td>
</tr>
<tr>
<td>500</td>
<td>1200</td>
<td>500</td>
</tr>
<tr>
<td>600</td>
<td>1400</td>
<td>600</td>
</tr>
<tr>
<td>700</td>
<td>1600</td>
<td>700</td>
</tr>
<tr>
<td>800</td>
<td>1800</td>
<td>800</td>
</tr>
<tr>
<td>900</td>
<td>2000</td>
<td>900</td>
</tr>
<tr>
<td>1000</td>
<td>2200</td>
<td>1000</td>
</tr>
</tbody>
</table>

RMS Difference (percent error)

2 Layer

- Interpolation correlates layers
- Higher layer degrades the accuracy of bottom layer
- 5 layer is “well behaved” (relaxes back to first guess profile)

3 Layer

5 Layer
Single Day (20070126)

070126LSQwAP\textsubscript{TT} Mixing Ratios on 5 Layer RET. Grid

- New retrieval overcompensates by perturbing lower layer by 30 ppbv
- Displays the ambiguity as to the sensitivity of the version 1 CO retrieval
Convolved Truth

- Smoothes the true profile with retrieval vertical sensitivity
- Difference between convolved and retrieval represents effect of error

\[ z_{\text{conv}} = A z_{\text{true}} + (I_n - A) z_1 \]

- Strongest perturbation is in first 2 layers
- Some error covariance is not strong enough in this case

New vs. Version 1

- Averaging kernels allow characterization of vertical sensitivity with each retrieval
- Effect of noise is reduced
- Error analysis is possible with each retrieval
- Still dominated by first 2km of the atmosphere
Retrieval Results with AIRS CO profiles used for first guess

Figure 2.13: (a) AIRS v5 AK with the effective pressures for each row of the AK listed in the legend. This plot comes from [86]. (b) SAAC averaging kernel. The components for this retrieval use a diagonal constraint matrix called $H_{TT}$ and a 5 layer retrieval grid. These components are defined in Chapter 4.
**Channel Selection Comparison**

- An algorithm has been developed to select the optimal channels for use in this CO retrieval.
- Incorporates the channel to channel correlation from apodization.
- Selected 19-21 channels (depending on # of grid layers).

---

**Optimal Channels**

- Optimal channels is better than the whole band on every level (2-6%).
- Places most channels in R band where effect of water is less.
First Guess Comparison

• The first guess only affects parts of the profile in the null space of the measurement
• Constraint matrix is defined by a 30 percent variation

AIRS first guess is better above 700 mb and worse below compared with a priori
• Retrieval plus AIRS is better constrained in the lower atmosphere
• Still relaxes back to the correct solution in the top portion
First guess Results

- RMS is calculated between retrieval solutions and true composite CO profiles for 13 cases
- RMS is reported as a percent error on the retrieval grid

First guess = a priori

First guess = AIRS

• 1\textsuperscript{st} layer represents majority of AERI measurement: 8 vs 11\% for AIRS vs. a priori
• 2\textsuperscript{nd} layer represents the merge point of AERI and AIRS: 4 vs. 13\% for AIRS vs. a priori
• Subsequent layers merge back to the AIRS retrieval
Conclusions and Future Applications

• New AERI retrieval can retrieve total columns to better than 10% accuracy in boundary layer
• Including AIRS as the first guess improves the first 2 layers in the 5 layer grid system
• Research should be done to identify how long AIRS can represent upper atmosphere CO
• More satellites can be used in this method
Extra Slides
Error Simulation

Estimated Error of Various Terms

- Total
- Modeled
- Input Water
- Input Temp
- Instrument
- Ret indice

Radiance \([\text{mW/(m}^2 \text{str cm}^{-1})]\)

Wavenumber (cm\(^{-1}\))
kCARTA Simulation Results

- Uncertainty in water is the dominating term in modeling spectra
Another Comparison

First guess = AIRS

080227LSQ TT Mixing Ratios on 5 Layer RET. Grid

First guess = a priori

080721LSQ TT Mixing Ratios on 5 Layer RET. Grid