Using Principal Component Analysis (PCA) to Speed up Radiative Transfer (RT) Computations

Vijay Natraj

Jet Propulsion Laboratory, California Institute of Technology

1st Workshop on Studies of Atmospheric Aerosol Using Multi-Angle Spectropolarimetry

July 17, 2012

Rationale

- **Multiple scattering RT calculations time-consuming**

- **Need a speed improvement of about 1000 (for OCO)!**

- **Solution: Make use of redundancies in spectra**

  – Correlated-k (Lacis and Wang, Lacis and Oinas, Goody *et al*, Fu and Liou)

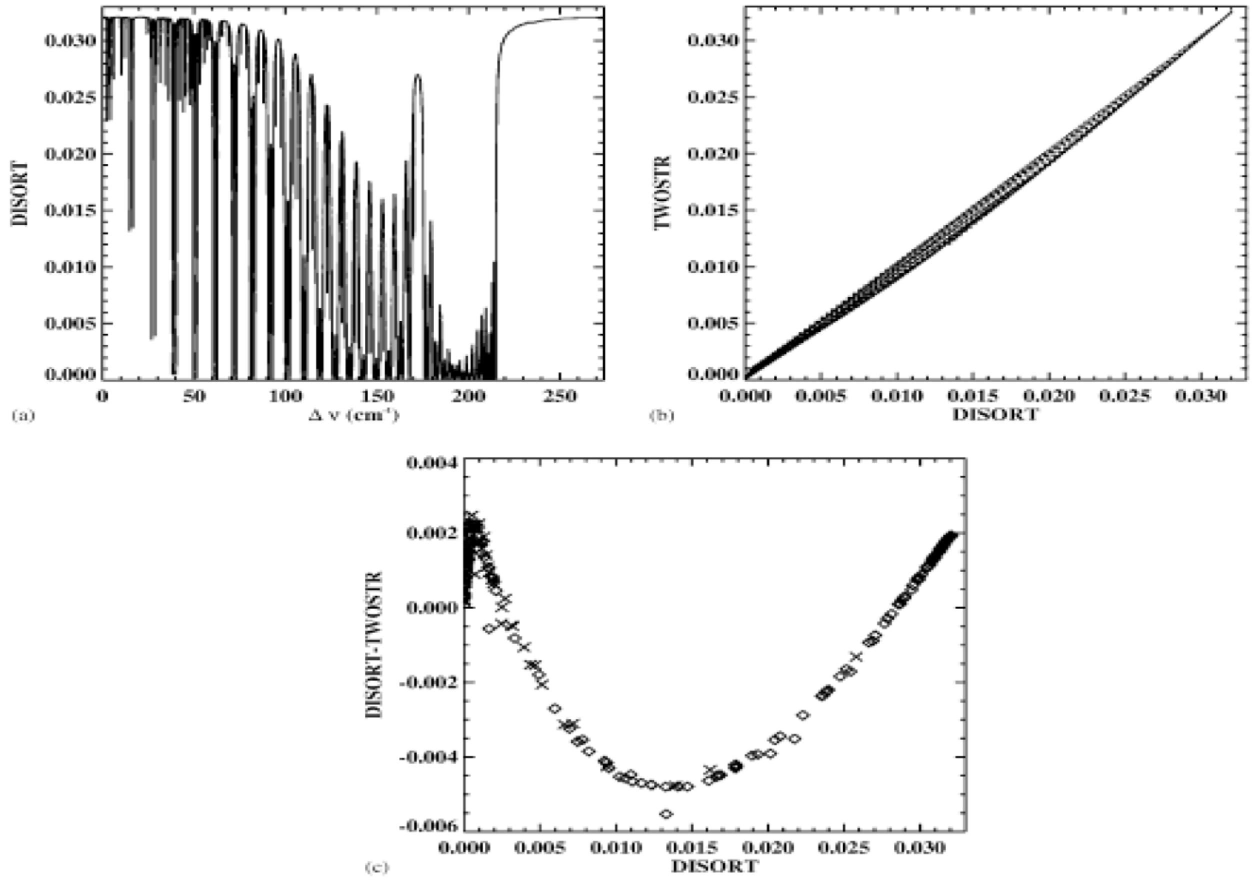
- **Problem**

  - Assume that spectral variation of atmospheric optical properties spatially correlated at all points along optical path

Rationale (contd …)

- **Spectral Mapping (Crisp and West, Meadows and Crisp)**
  - No assumption about spatial correlation
  - Combine only spectral regions that remain in agreement at ALL points along optical path
  - Binning parameters: optical depth, single scattering albedo, asymmetry parameter, surface albedo
- **Problems**
  - Inadequate speed improvement for OCO precision constraints
  - Glitches in partial differentials

DISORT / TWOSTR Comparison

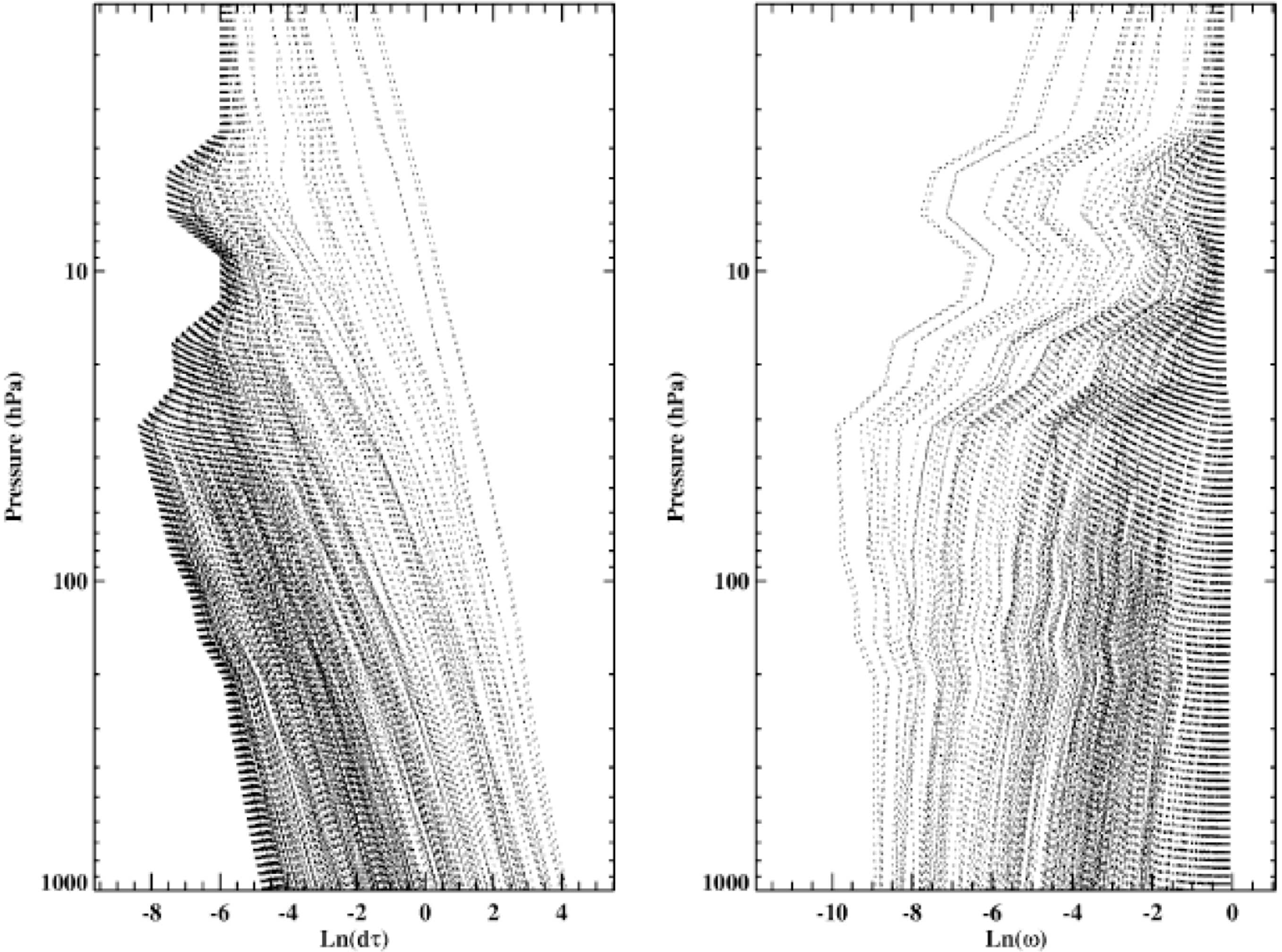


Natraj et al., 2005

PCA Technique

- **Data Set**
  - Optical properties in M atmospheric layers at N wavelengths
- **EOF**
  - Eigenvectors of covariance matrix of detrended (removed mean) data set
  - New basis to represent original data
  - No loss of information
- **PC**
  - Projection of original data set onto EOFs

Optical Depth and Single Scattering Albedo Profiles



4 EOFs capture > 99.99% of variance

Bin Selection (Original)

- **Reference: Natraj *et al.*, JQSRT, 95, 539-556, 2005**
  - http://www.gps.caltech.edu/~vijay/Publications/Natraj-etal-05.pdf
- **Grouping Criteria**
  - $c_1 < \ln(2\tau_2) < c_2$; τ_2: total OD of lower half of atmosphere
  - $c_3 < \omega_1 < c_4$; ω_1: ssa of top layer
  - Criterion 2 accounts for vertical structure of gas absorption
- **Surface albedo and phase function assumed not to vary with wavelength**

Mapping to Radiance

- **Compute DISORT-TWOSTR differences**
  - I_d: mean properties
  - $I_d^{+/-}(k)$: mean+/-EOF for k^{th} EOF

- **First and second order differences**

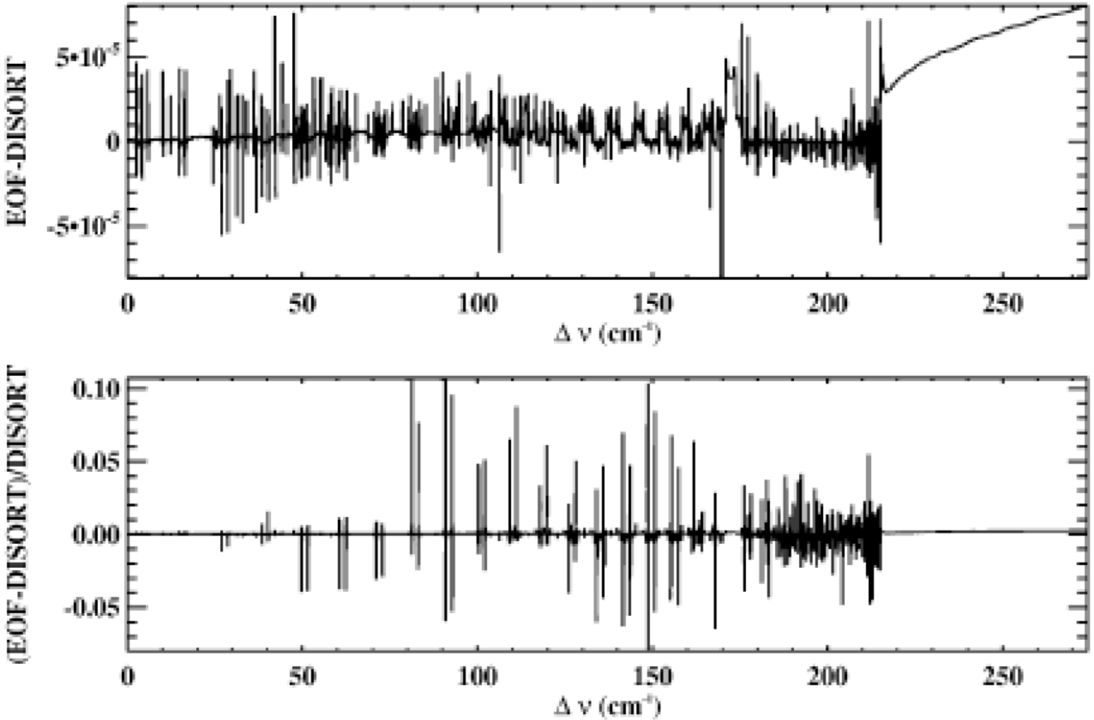
$$\delta I_k = \frac{I_d^+(k) - I_d^-(k)}{2}$$

$$\delta^2 I_k = I_d^+(k) - 2I_d + I_d^-(k)$$

- **Radiance**

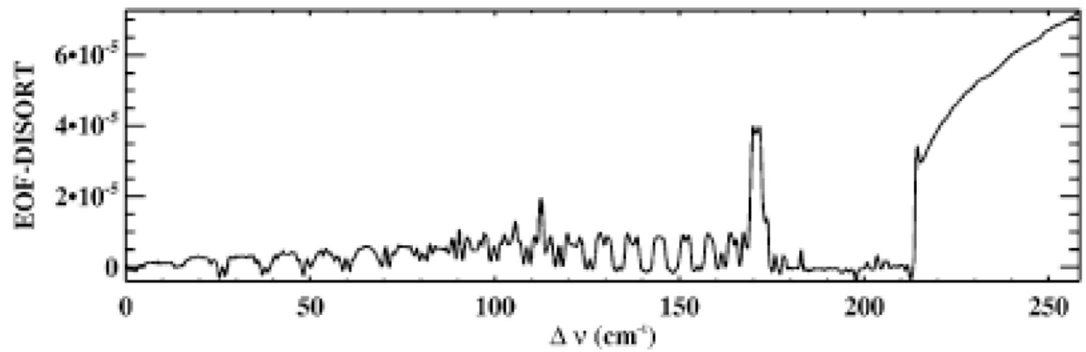
$$I_l = I_l^{TS} + I_d + \sum_{k=1}^{4} \delta I_k P_{kl} + \frac{1}{2}\sum_{k=1}^{4} \delta^2 I_k P_{kl}^2$$

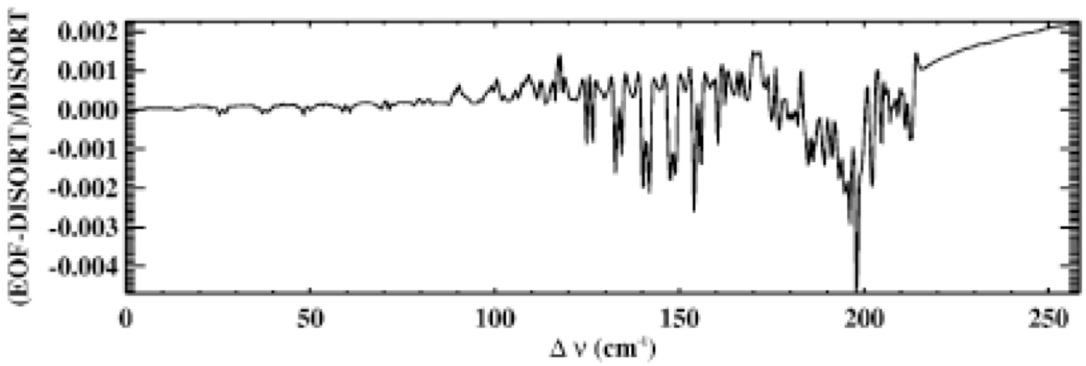
Results: High Resolution



Order of magnitude improvement in speed

Results: Convolved





Residuals smaller after convolution => PCA errors mostly random

Bin Selection (Modified)

- **Reference: Natraj *et al.*, JQSRT, 111, 810-816, 2010**
  - http://www.gps.caltech.edu/~vijay/Publications/Natraj-etal-10.pdf
- **Grouping Criteria**
  - $c_1 < \ln(\tau_g) < c_2$; τ_g: total gas OD
  - $c_3 < \omega_1 < c_4$; ω_1: ssa of top layer
- **Rationale for change in Criterion 1**
  - No justification to use OD of lower half of atmosphere
  - Aerosol OD can vary substantially from scene to scene
  - Gas OD has the most significant spectral variation

Mapping to Radiance

$$I_d^0 = \ln(I^0/I_2^0),$$

$$I_{d,k}^+ = \ln(I_k^+/I_{2,k}^+),$$

$$I_{d,k}^- = \ln(I_k^-/I_{2,k}^-),$$

$$Q_d^0 = Q^0 - Q_1^0,$$

$$Q_{d,k}^+ = Q_k^+ - Q_{1,k}^+,$$

$$Q_{d,k}^- = Q_k^- - Q_{1,k}^-.$$

$$\delta I_k = \frac{I_{d,k}^+ - I_{d,k}^-}{2},$$

$$\delta^2 I_k = I_{d,k}^+ - 2I_d^0 + I_{d,k}^-,$$

$$\delta Q_k = \frac{Q_{d,k}^+ - Q_{d,k}^-}{2},$$

$$\delta^2 Q_k = Q_{d,k}^+ - 2Q_d^0 + Q_{d,k}^-.$$

First and second order differences

$$I_l = I_2 \exp\left[I_d^0 + \sum_k \delta I_k P_{kl} + \frac{1}{2}\sum_k \delta^2 I_k P_{kl}^2\right],$$

$$Q_l = Q_1 + Q_d^0 + \sum_k \delta Q_k P_{kl} + \frac{1}{2}\sum_k \delta^2 Q_k P_{kl}^2,$$

Radiance

- **Logarithm better than differences**
  - Avoids negative intensities
  - Scattering is basically a perturbation to Beer's Law

Basics of Speed-up Technique

- **High accuracy (HI) and 2-stream (2S) calculations have high correlation**

- **Single scattering (SS) computations highly scenario-dependent, but not time consuming**

- **Perform SS and 2S calculations at every wavelength**

- **Perform small number of HI computations**

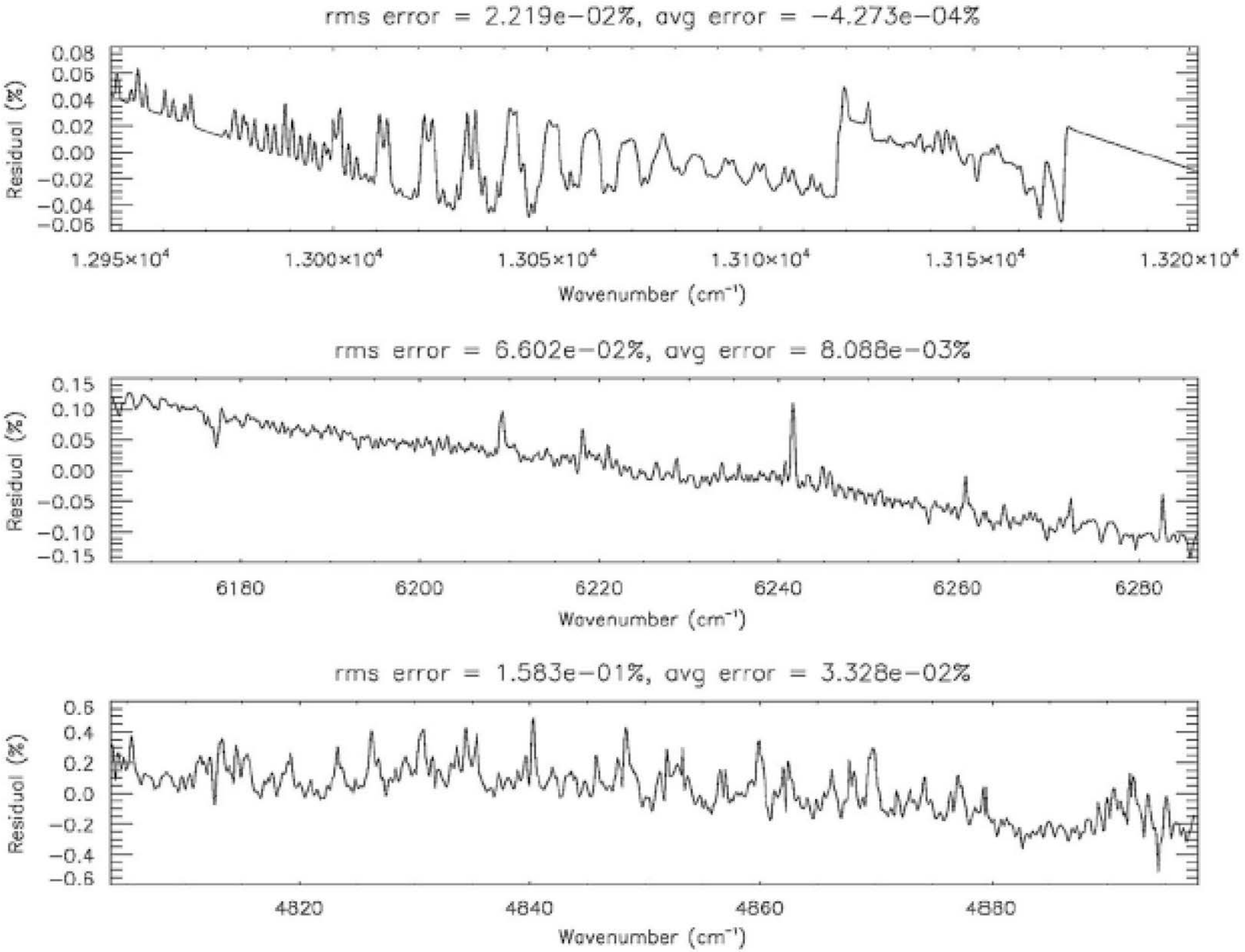
- **Need to compute correction factor *B* at every wavelength**

$$I_{Approx}(\lambda_i) \cong [I_{2S}(\lambda_i)+I_{FO}(\lambda_i)]B(\lambda_i)$$

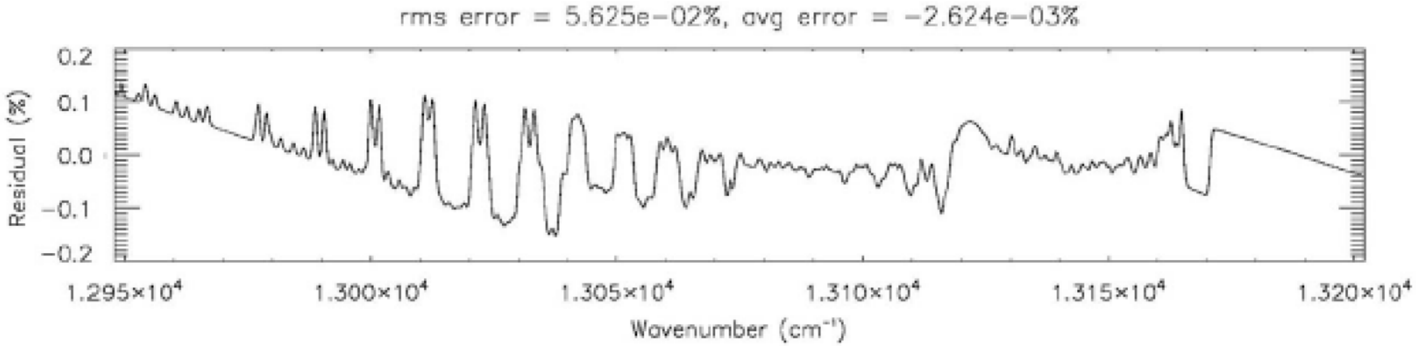
Scenarios

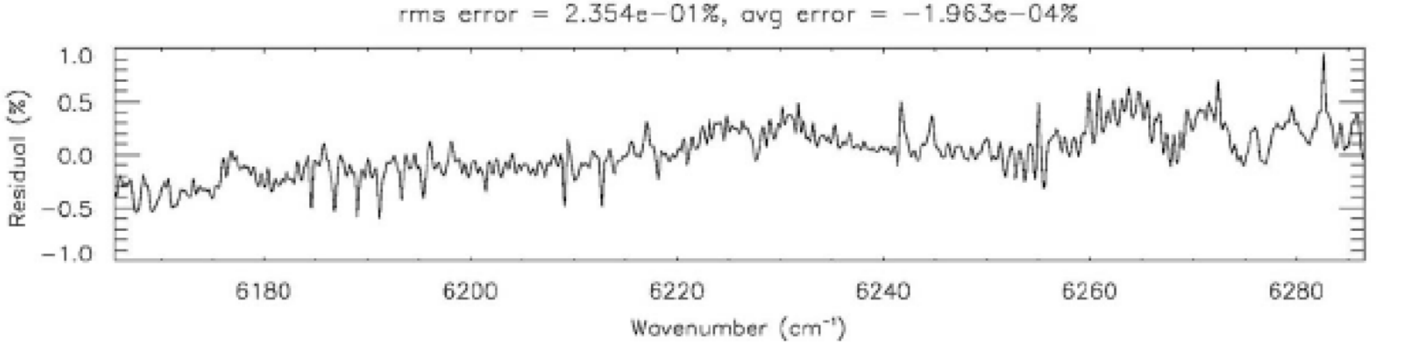
Profile Index	Cloud + Aerosol OD	Water Cloud OD	Ice Cloud OD	Aerosol OD
174	0.27358E+00	0.00000E+00	0.20848E+00	0.65099E-01
2558	0.14782E+00	0.00000E+00	0.00000E+00	0.14782E+00
256	0.27129E+00	0.20605E+00	0.48814E-03	0.64759E-01
28	0.64800E-01	0.00000E+00	0.00000E+00	0.64800E-01

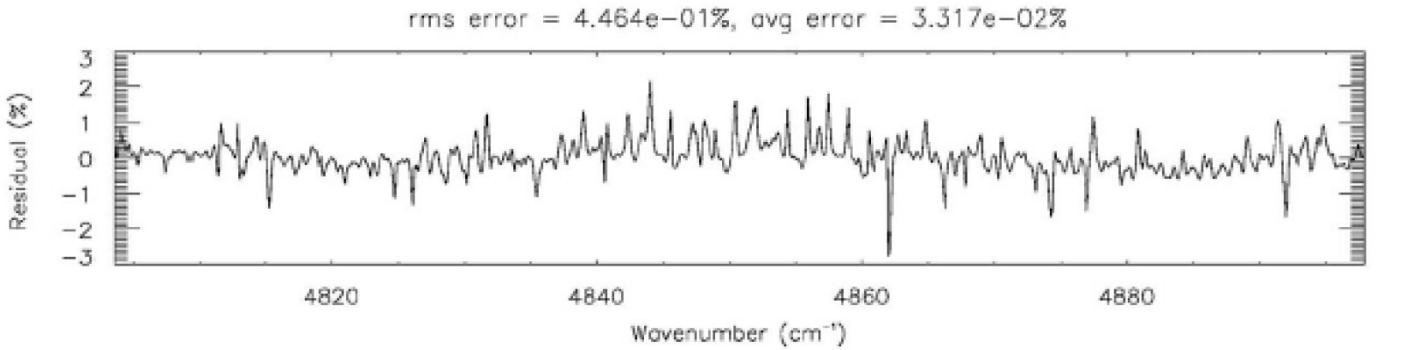
Results (I): Profile 256



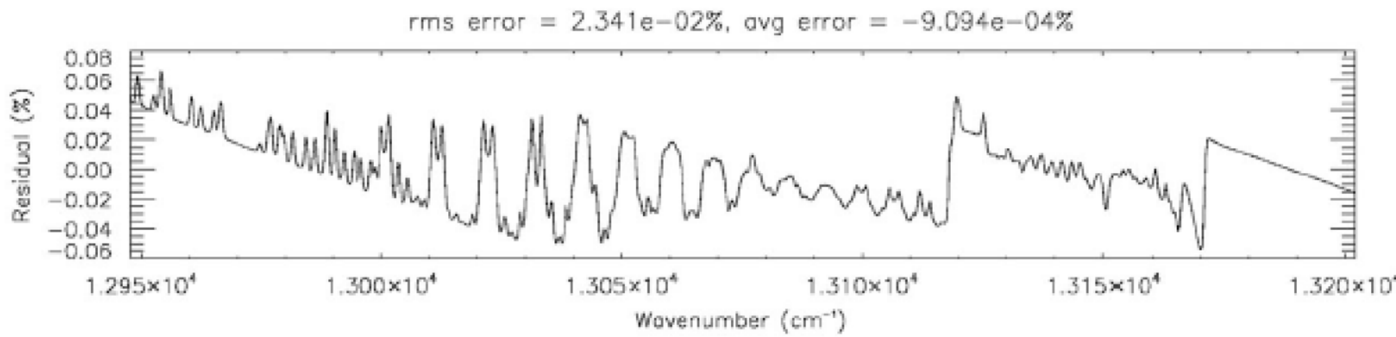
Results(Q): Profile 256

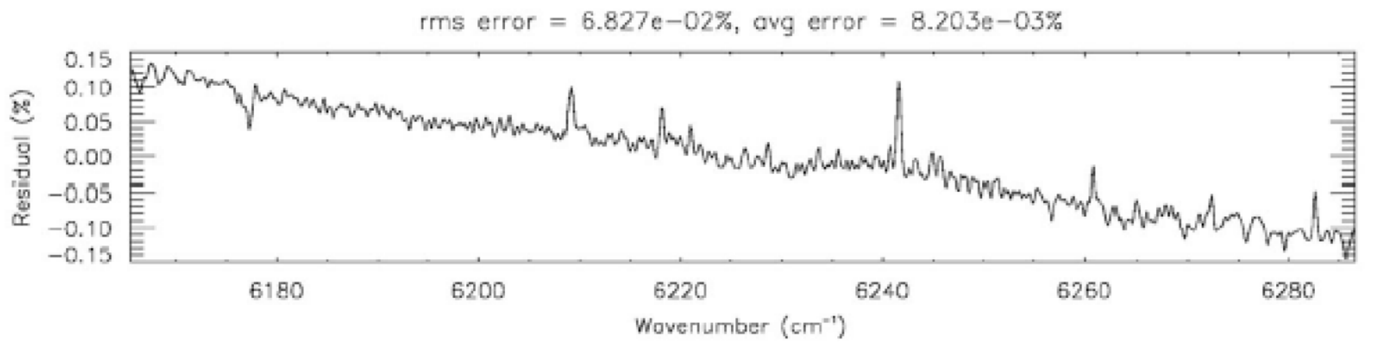


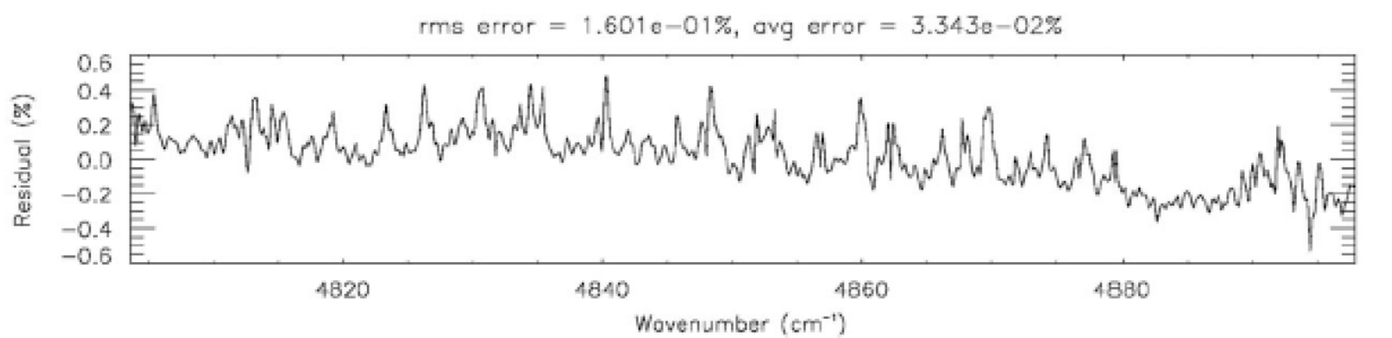




Results(I-Q): Profile 256







RMS Errors and Timing Details

Scenario	% RMS error (O_2 *A* band)	% RMS error (1.61 μm CO_2 band)	% RMS error (2.06 μm CO_2 band)
28	0.003, 0.015, 0.003	0.001, 0.021, 0.001	0.005, 0.071, 0.005
174	0.009, 0.176, 0.011	0.017, 0.043, 0.019	0.022, 0.183, 0.019
256	0.022, 0.056, 0.023	0.066, 0.235, 0.068	0.158, 0.446, 0.160
2558	0.007, 0.030, 0.011	0.004, 0.094, 0.005	0.003, 0.052, 0.002

Scenario	RT time (s) (LBL)	RT time (s) (PCA)	Total time (s) (LBL)	Total time (s) (PCA)
28	4858.12	61.15	4875.28	83.45
174	5067.82	63.96	5096.67	96.08
256	5160.10	68.99	5193.31	105.34
2558	4875.67	61.99	4891.94	83.57

Further Improvements

- **Automated binning for** τ_g
- **Introduced binning for** τ_{sca}
- **Linearized PCA**
- **Testing PCA for Jacobians**
- **Expanding PCA to work from UV to TIR**