On the use of atmospheric solar spectra for spectroscopy validation

Geoffrey Toon, Linda Brown, Bob Toth,
Jet Propulsion Laboratory, California Institute of Technology

Iouli Gordon, Larry Rothman,
Harvard Smithsonian Center for Astrophysics

It is usually impossible to fit IR solar spectra (high SNR) down to their noise level. Residuals are usually dominated by systematic errors arising from defects in:

- atmospheric T/P/VMR profiles
- instrumental response (e.g. ILS, zero-level-offsets, channel fringes, ghosts)
- spectroscopy

For spectra measured with a well-calibrated FTS under well-known atmospheric conditions, the first two systematic errors can usually be minimized, revealing the underlying spectroscopic problems.

I am not a spectroscopist. I am a user of spectroscopic data for the purpose of atmospheric remote sensing. I have never attended a HITRAN meeting before. I hope that this talk will give you the perspective of an end-user of your work.

Atmospheric Remote Sensing: Accuracy/Precision requirements

For long lived atmospheric gases (e.g. CO_2, N_2O , CH_4) , the variations of interest (e.g., seasonal cycle, inter-hemispheric gradient, sources, sinks) tend to be small in comparison with the accumulated total gas column.

So high fractional accuracy is needed for measurements to be scientifically useful.

For CO_2, column-averaged mole fraction need to be better than 0.3%
For N_2O, the tropospheric mole fraction needs to be better than 0.2%
For CH_4, the tropospheric mole fraction needs to be better than 0.4%

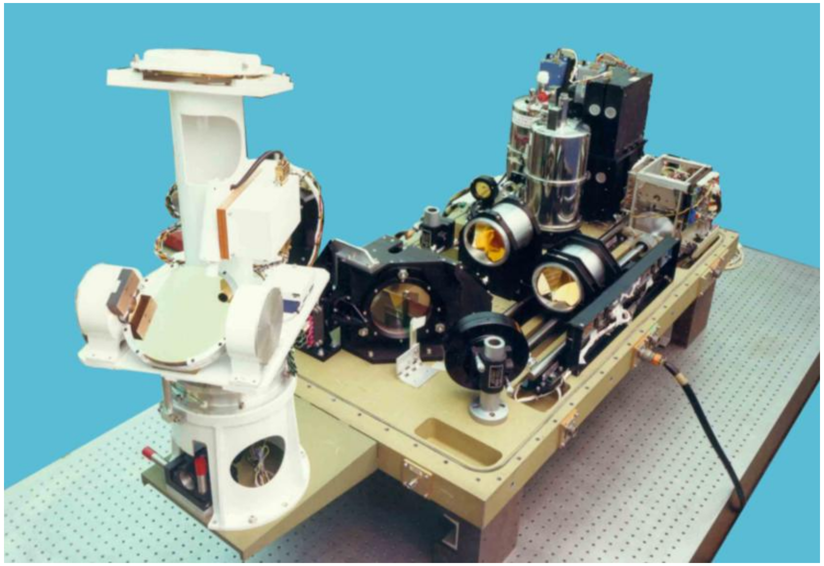
The OCO and GOSAT satellite sensors attempt to meet these goals, and ground-based validation netwotks (e.g. TCCON, NDACC) try to do even better.

Requirements will get stricter in the future as more is learned about these gases.

These goals impose strict requirements on the spectroscopy. Not only for these gases themselves, but also on other interfering absorbers (e.g. H_2O, HNO_3, etc.)

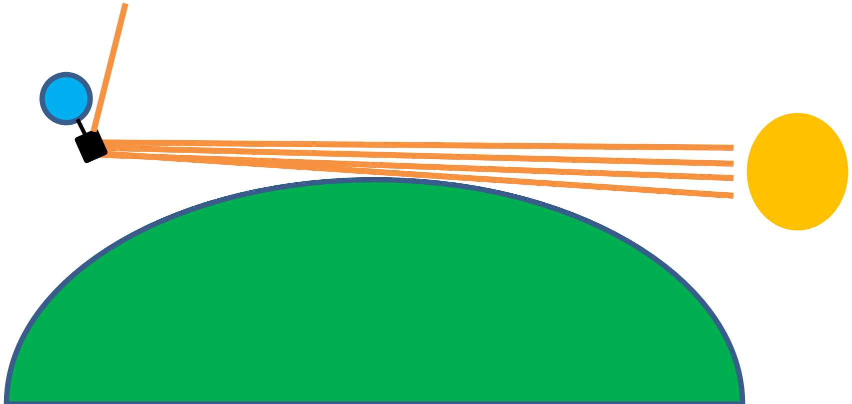
JPL MkIV interferometer

Most examples of atmospheric solar spectra shown in this talk come from the JPL MkIV. This solar absorption FTS was built at JPL in 1984 and has performed 22 balloon flights and over 1000 days of ground-based obs. Covers the entire 650-5650 cm^{-1} region simultaneously at up to 0.005 cm^{-1} resolution. MkIV is similar to ACE FTS, but is non-orbital.



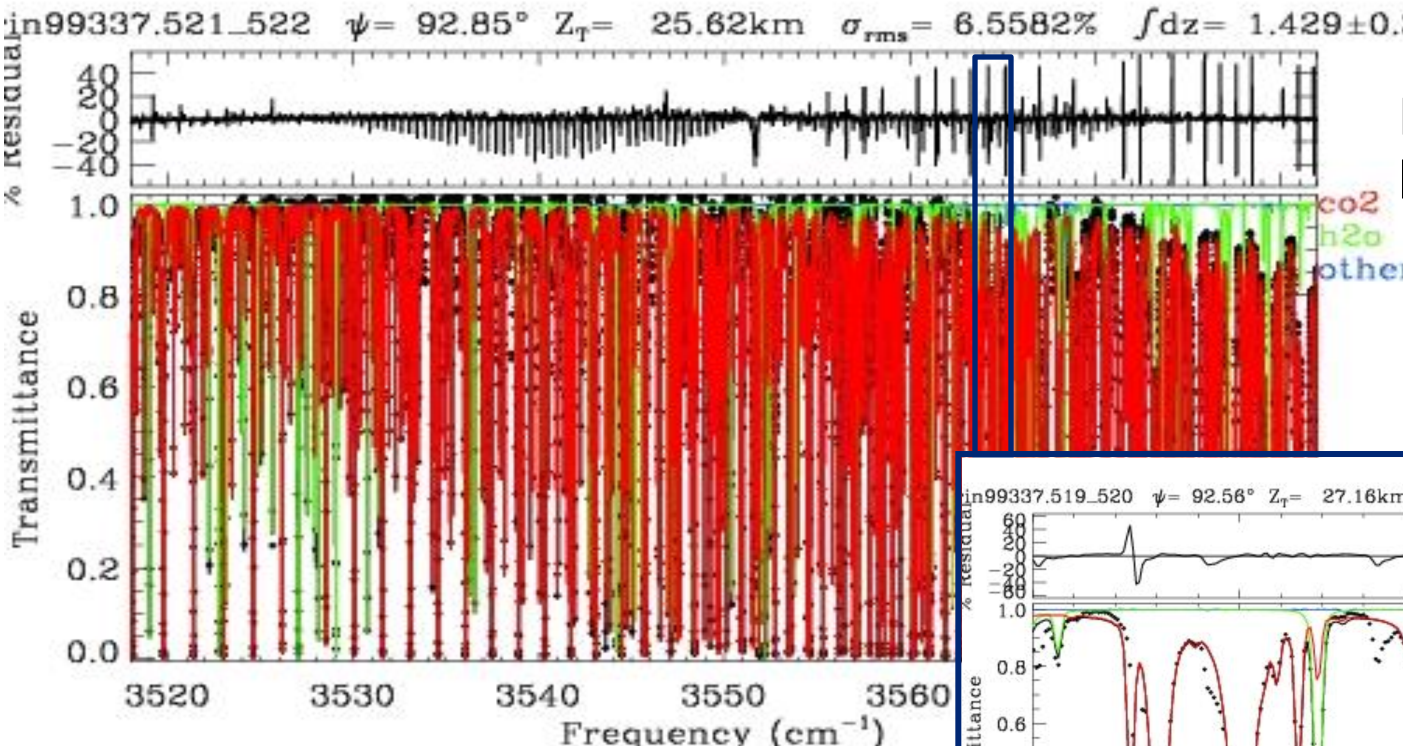
From the ground, data are acquired looking up at the direct sun. Absorption mainly from the lower troposphere. Ground-based MkIV spectra have been acquired from 0 to 4 km altitude, 0 to 91 deg SZA, and -40C to +40C.

From balloon, data are acquired at noon, sunset and sunrise from ~40 km altitude. Occultation viewing geometry provides long atmospheric path with high vertical resolution. Aimass changes by factor 10^4, allowing weak and strong lines to be evaluated.

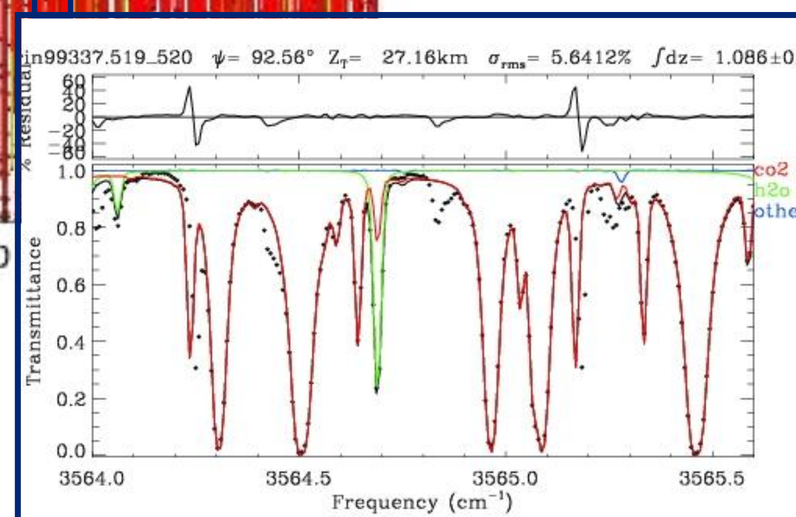


Case Study A: MkIV balloon HNO_3

Fit to a solar occultation spectrum measured at 26 km tangent altitude above Esrange, Sweden, in Dec 1999. Most absorption features are due to H_2O and CO_2. "% Residuals" are the differences between the measured and calculated spectra.



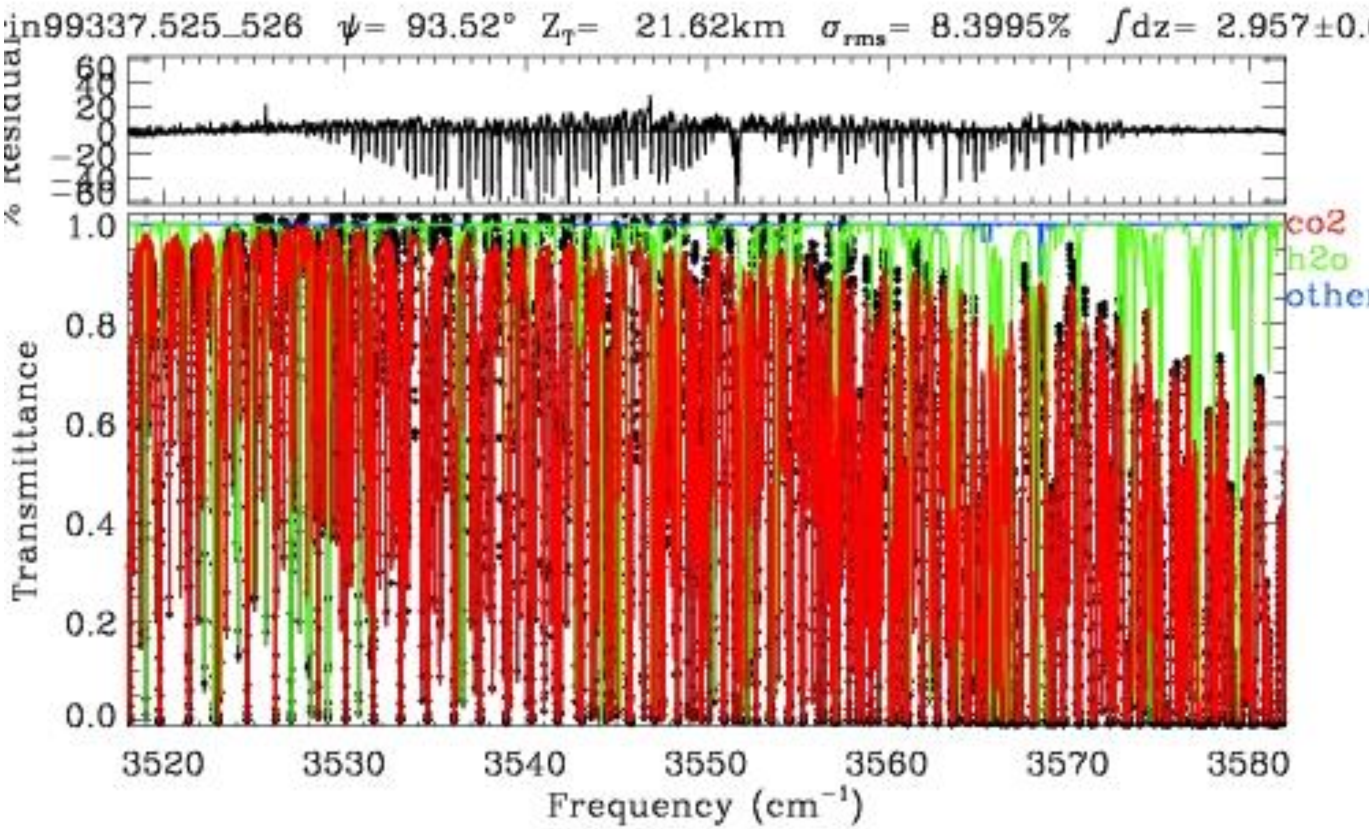
HITRAN 2008 linelist



Two problems are evident here:

- Missing ν_1 band of HNO_3.
- Position errors in a CO_2 sub-branch

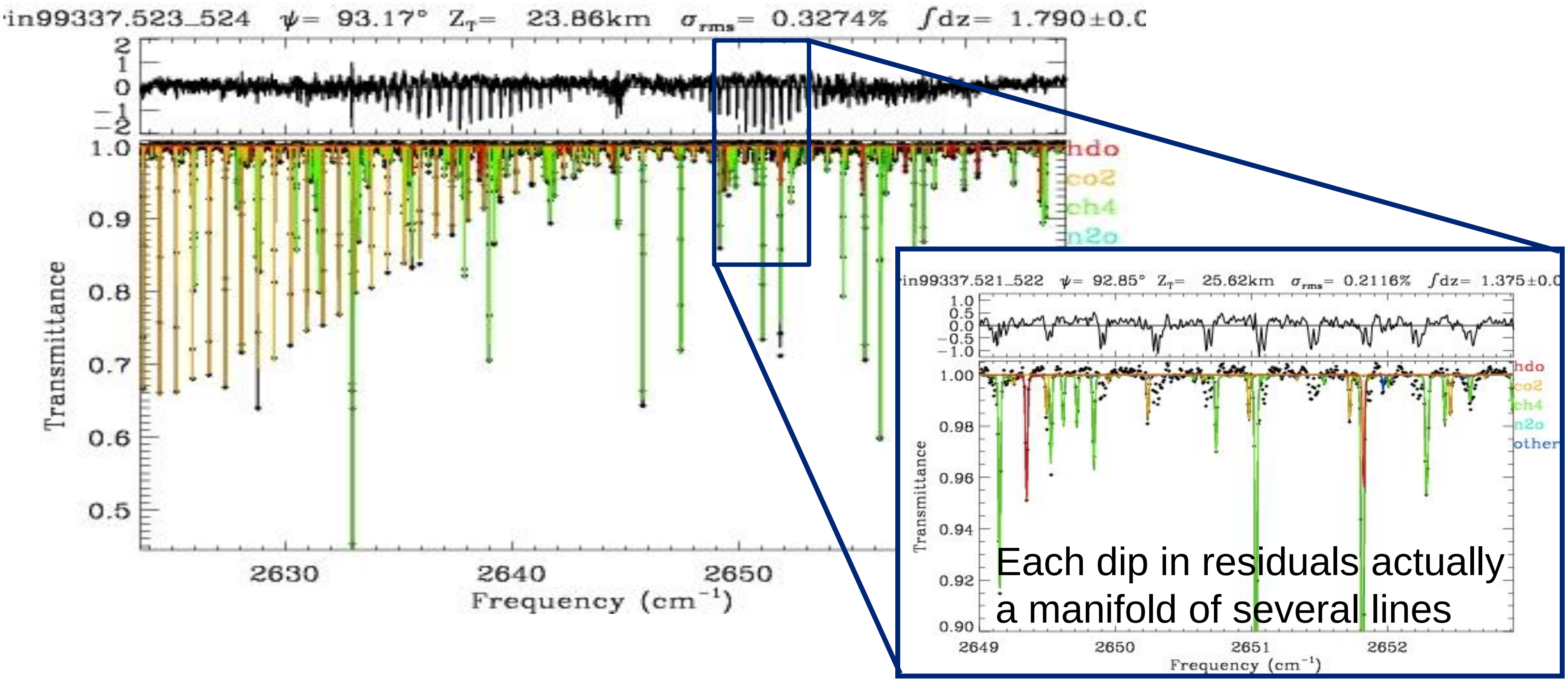
After empirically fixing CO_2 problem



The P, Q, and R-branch structures of the HNO_3 ν_1 band are now much clearer.
Missing HNO_3 features are up to 60% deep at 22 km tangent altitude.
The presence of such strong residuals make it impossible to quantify weakly absorbing gases (e.g. OH, HO_2) that are potentially observable in this region.

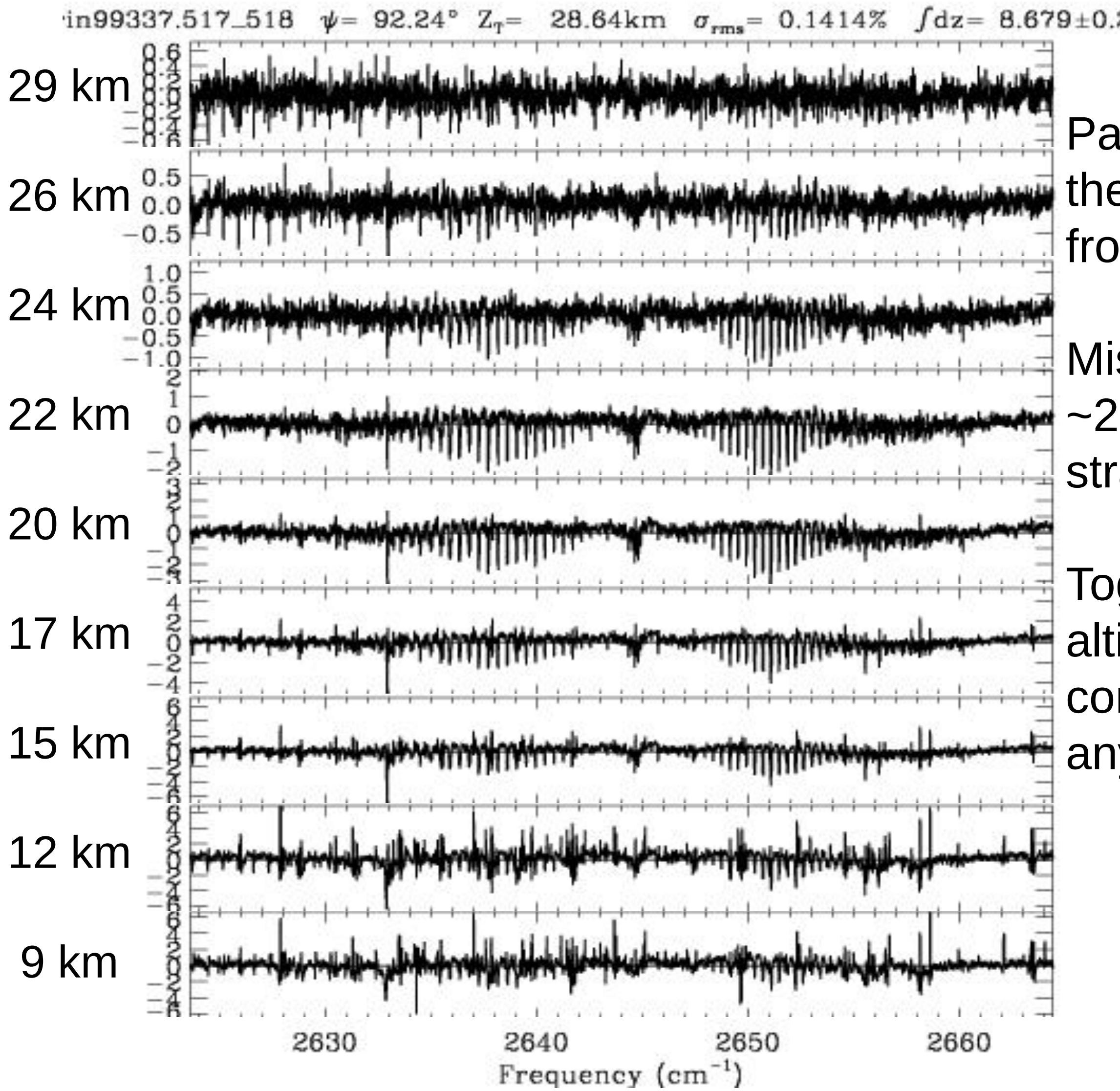
Another missing HNO_3 band

HNO_3 has dozens of other (weaker) bands missing from HITRAN, clearly visible in MkIV balloon spectra. Below, the missing $2\nu_3$ band of HNO_3 seen at 24 km altitude.



This band is not essential to quantify HNO_3. There are several other HNO_3 bands at $\nu<1750$ cm^{-1} that suffice. The problem is that the residuals prevent measurement of weakly absorbing trace gases (e.g. HBr) that absorb here.

Altitude variation of residuals

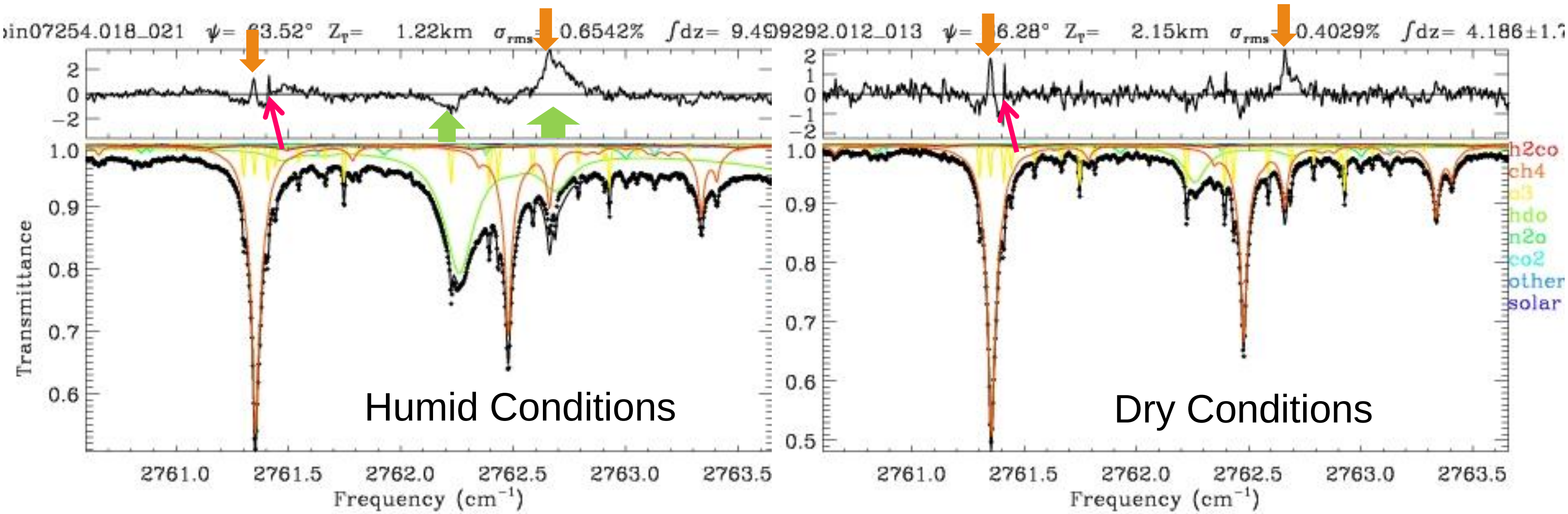


Panels show the altitude variation of the residuals seen on previous slide, from 29 km (top) to 9 km (bottom).

Missing absorption is maximum at ~20 km, confirming that it is a stratospheric gas, such as HNO_3.

Together with frequency signature, altitude variation of residuals helps confirm a unique identification of any missing absorbers.

Case Study B: Ground-based MkIV H_2CO

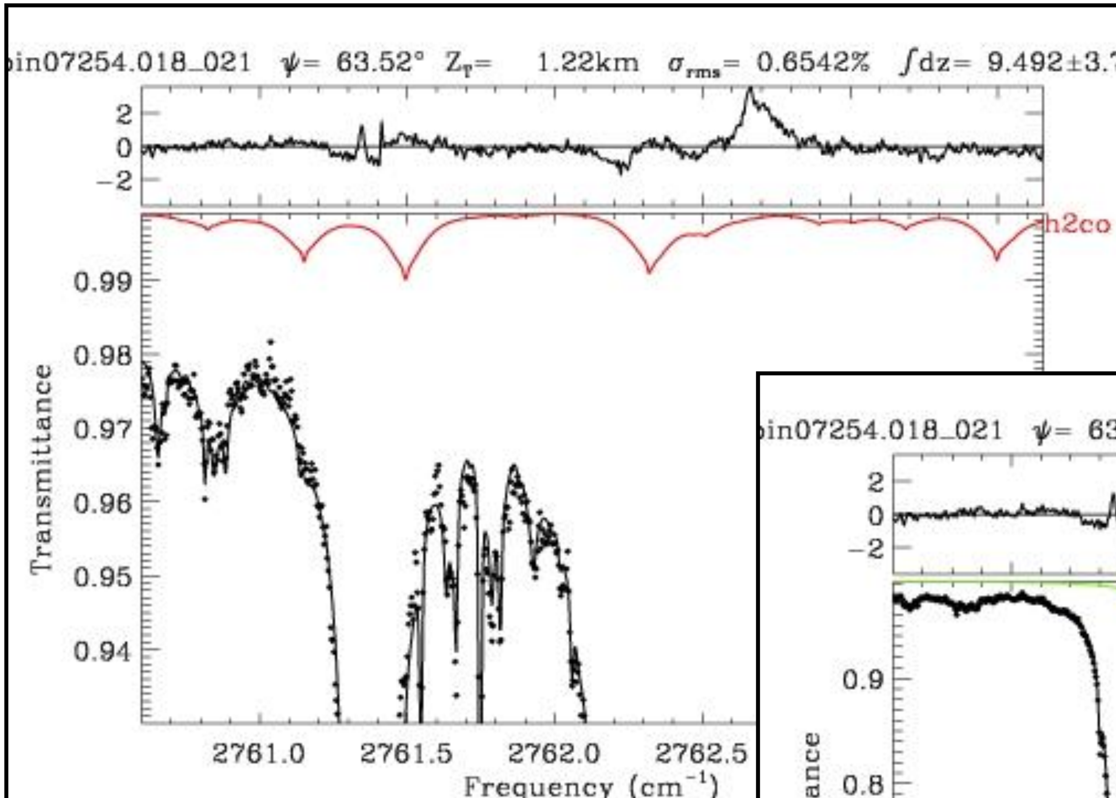


Fits to MkIV spectra measured under humid (left) and dry (right) conditions made using HITRAN 2008 linelist. This region used to measure H_2CO (formaldehyde). H_2CO absorptions are typically <1% deep, so 2-3% residuals are unacceptable.

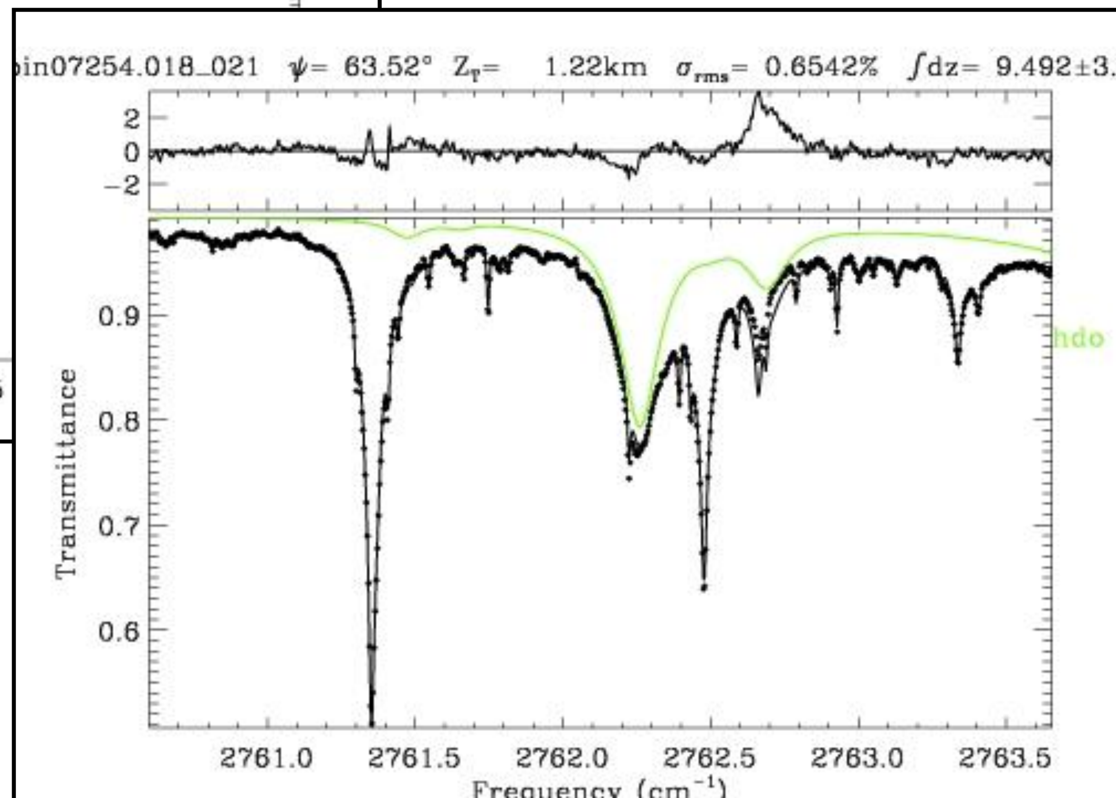
There are actually 3 spectroscopic problems visible here:

1) Broad positive residual around 2762.7 cm^{-1}, observed under humid conditions, due to intensity error in underlying HDO line.
2) CH_4 lines at 2761.34 and 2762.66 cm^{-1} are too narrow
3) O_3 line at 2761.41 cm^{-1} is in wrong position.

H_2CO and HDO contributions



Left: H_2CO absorptions of primary interest in this window are only ~1% deep. So 3% residuals are unacceptable. Need to improve interfering spectroscopy.



Right: Large residual at ~2762.704 cm^{-1} seems to be associated with an HDO line.

Toth 2003 Water Vapor Linelist

Available from: http://mark4sun.jpl.nasa.gov/data/spec/H2O/RAToth_H2O.tar
Non-HITRAN format containing calculated and measured intensities
All lines in the 2762 cm^{-1} region are HDO (gas=49)

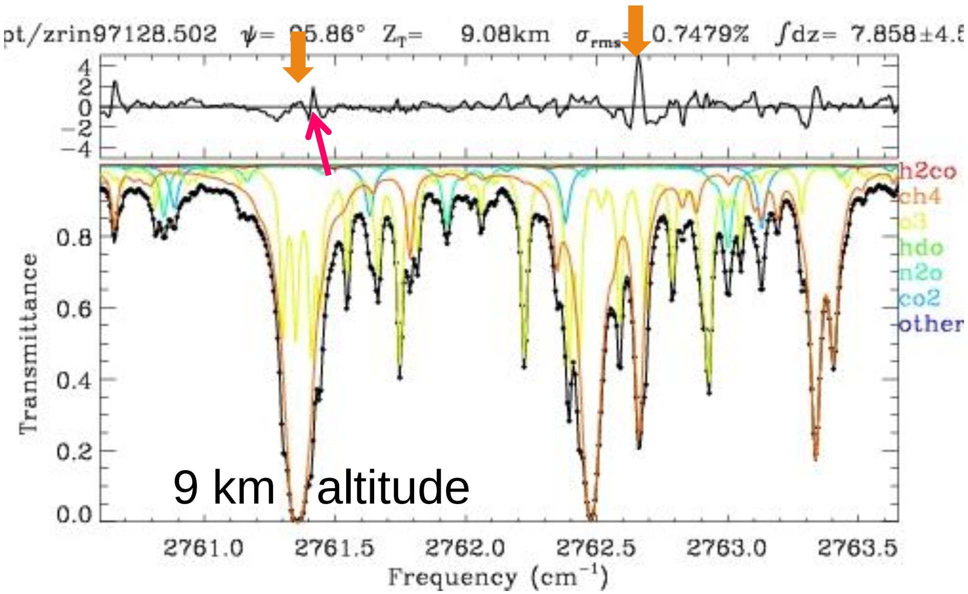
Calculated and Measured
Intensities agree well

Gas	I	Freq	Lower			Upper			Calculated Intensity	E”	ABHW	SBHW	SHIFT	?	?	Measured Intensity	
49	1	2761.66304	7	5	3	8	5	4	8.33E-08	942.53244	.0608	.303	.00000	25.		2.35E-07	
49	1	2762.26020	2	0	2	2	1	1	9.51E-06	66.18451	.1032	.430	.00000	3.	-20.	9.83E-06	2.
49	2	2762.50018	4	1	4	3	1	3	4.38E-07	99.58727	.0920	.496	.00000	6.			
49	1	2762.66663	5	1	4	5	0	5	9.11E-07	221.94609	.0934	.457	.00000	4.		2.22E-06	
49	1	2762.70437	3	2	1	3	1	2	1.58E-06	116.46133	.0964	.470	.00000	5.	-7.	7.15E-07	10.
49	1	2764.54270	3	1	3	2	1	2	1.92E-04	58.12689	.0959	.510	.00000	8.	4.	1.91E-04	3.
49	1	2765.11285	1	0	1	1	1	0	7.47E-06	32.49637	.1050	.462	.00000	2.	5.	7.51E-06	3.

Calculated & Measured
Intensities disagree by factor 2.2

HITRAN 2008 water positions and intensities 800-7900 cm^{-1} mostly based on Toth’s work.
Unfortunately the calculated intensities were adopted, not the measured intensities.
Changed intensity of 2762.704 cm^{-1} HDO line from the calculated to the measured value.

MkIV balloon spectra: 9, 12, 34 km

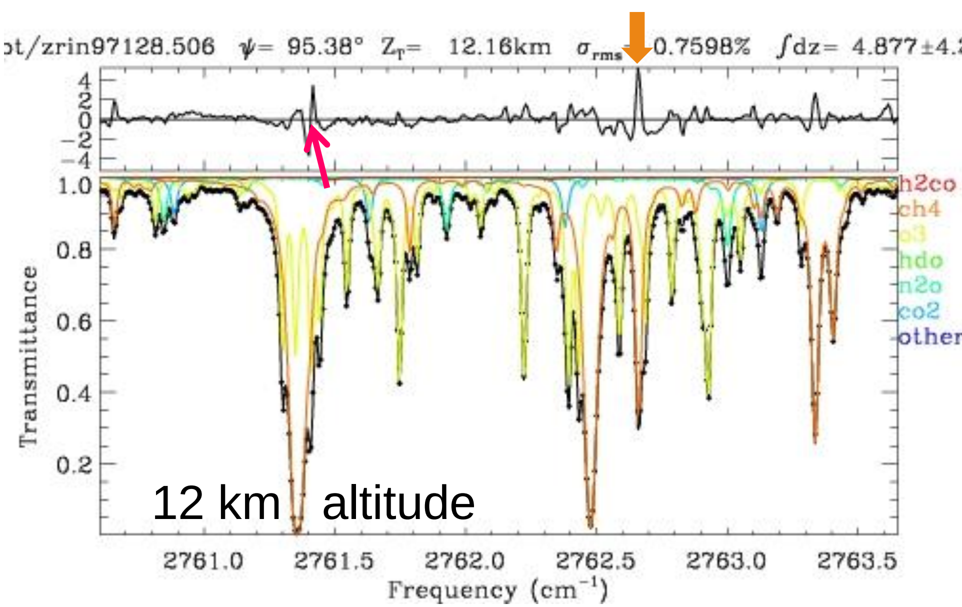


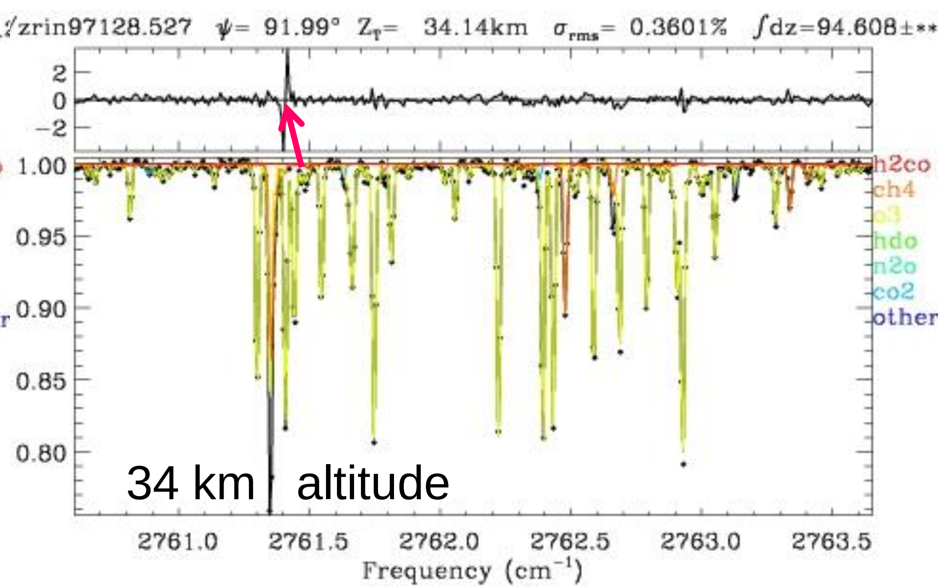
HITRAN 2008 linelist

Balloon spectra confirm fitting problems seen in ground-based spectra:

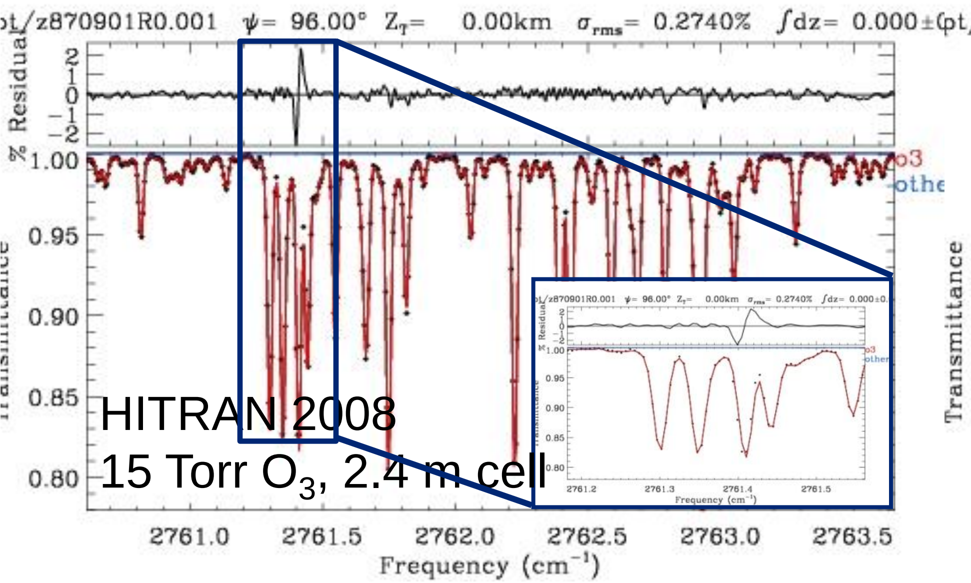
- CH_4 line at 2762.66 cm^{-1}: too narrow
- O_3 line at 2761.41 cm^{-1}: position error

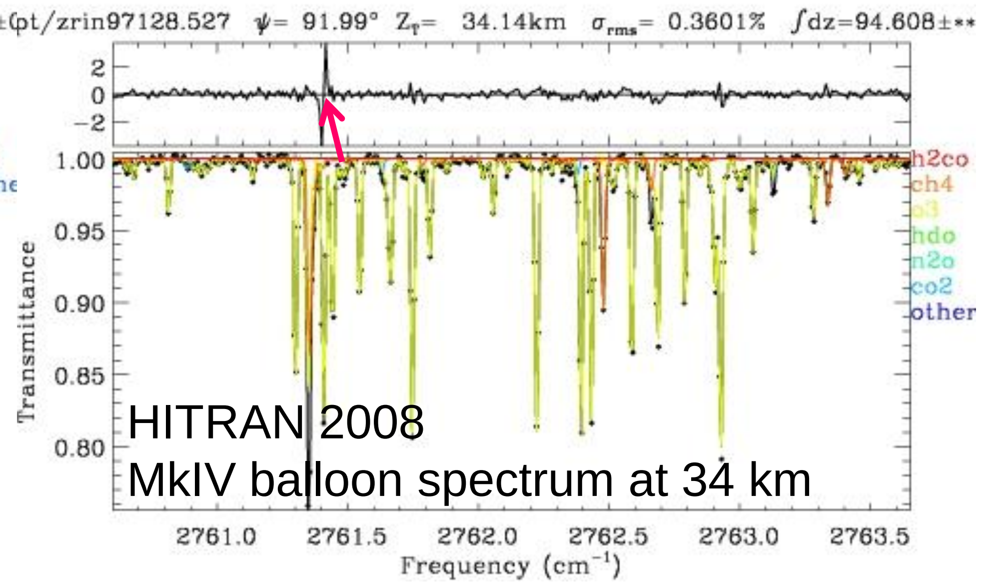
[HDO lines too weak to be observed above 9 km altitude in this region.]

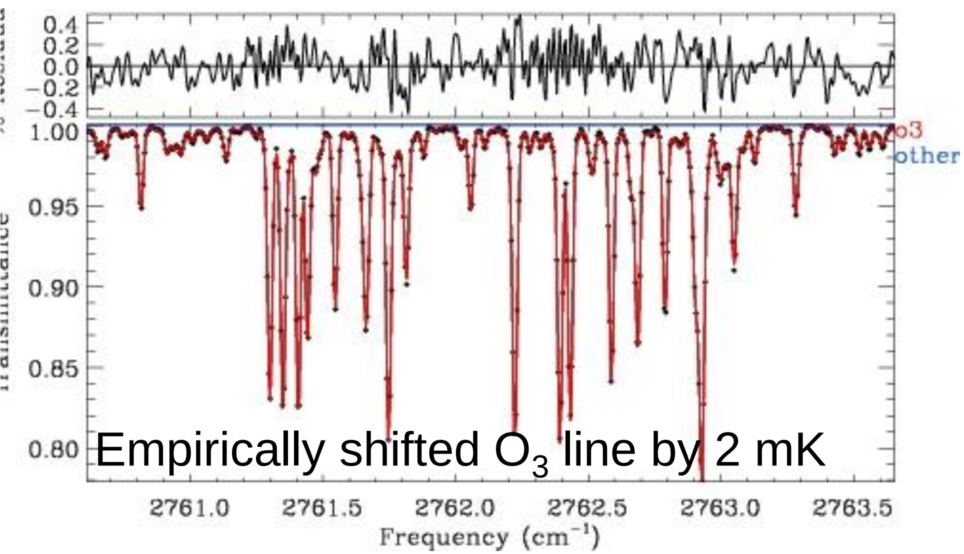




Fitted Kitt Peak laboratory O_3 spectra







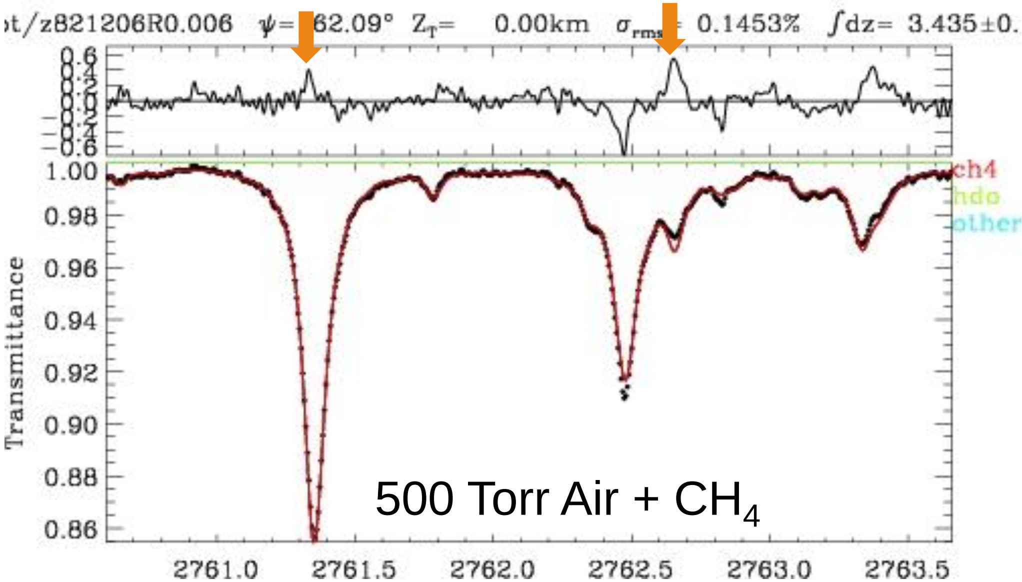
The Kitt Peak O_3 spectral fitting residuals are remarkably similar to those in the 34 km MkIV balloon spectrum.

After empirically correcting the position of the wayward O_3 line at 2761.4 cm^{-1}, the KP fitting residuals improved by a factor 5

To me, it's surprising how one line can be so poor, when all the others are so good.

31 2761.4095 5.586E-23 8.155E-02 .0749 0.102 243.0333 0.73 0.00 1 1 1 0 0 0 22 2 21 23 2 22 4655502130 5 2 3 0

Fits to Kitt Peak Laboratory CH_4 Spectra

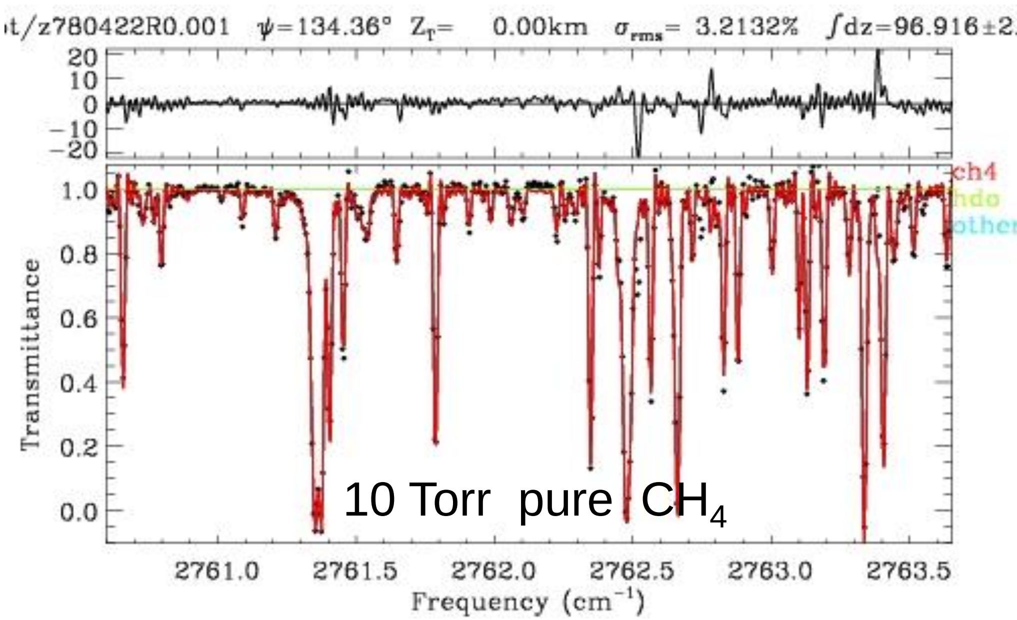


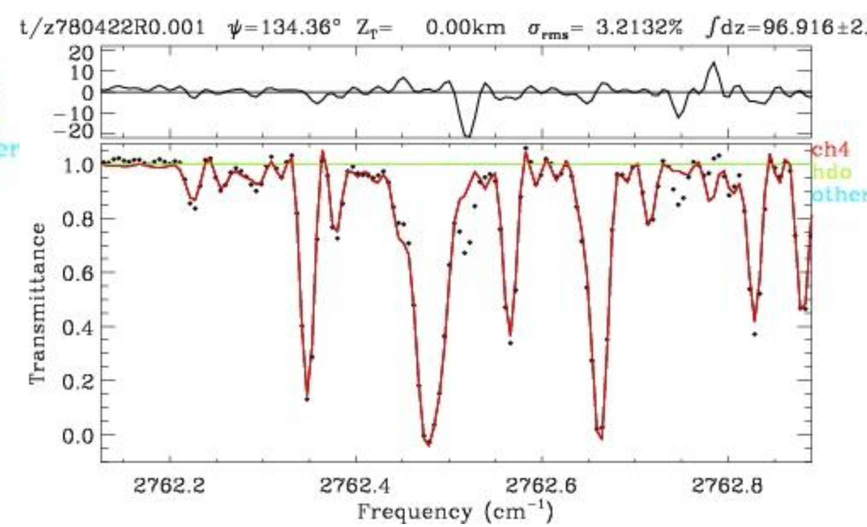
HITRAN 2008 Linelist

Very similar residuals to those seen in MkIV balloon spectra

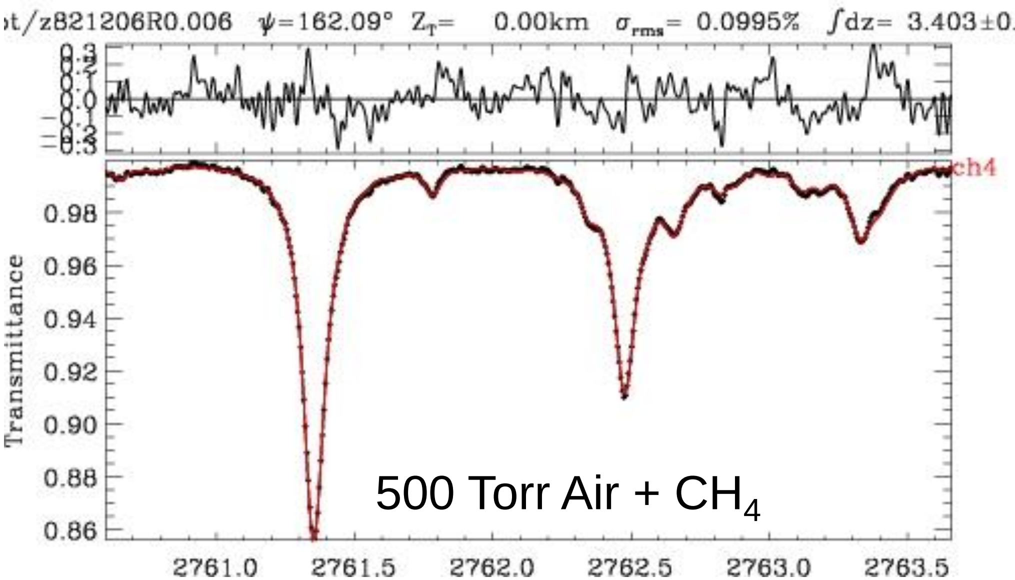
Problems seen in high pressure lab spectrum, not low pressure - width.

Also, missing or misplaced line at 2762.521 cm^{-1}



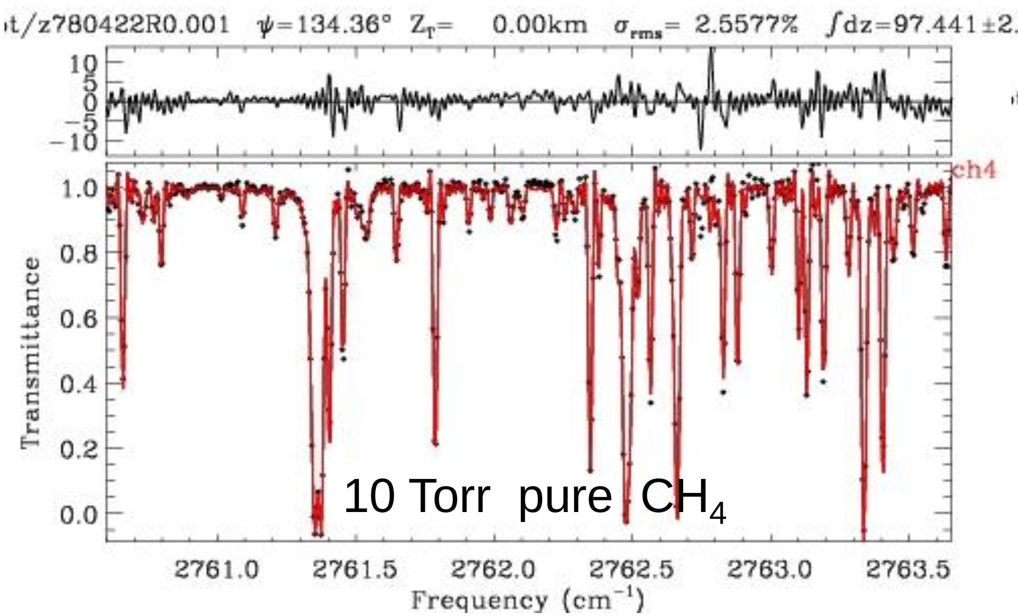


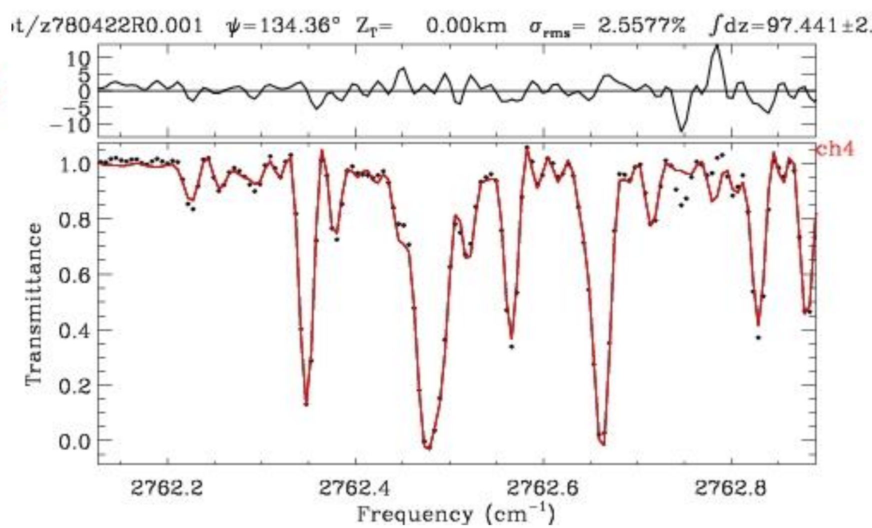
Fits to Kitt Peak Laboratory CH_4 Spectra



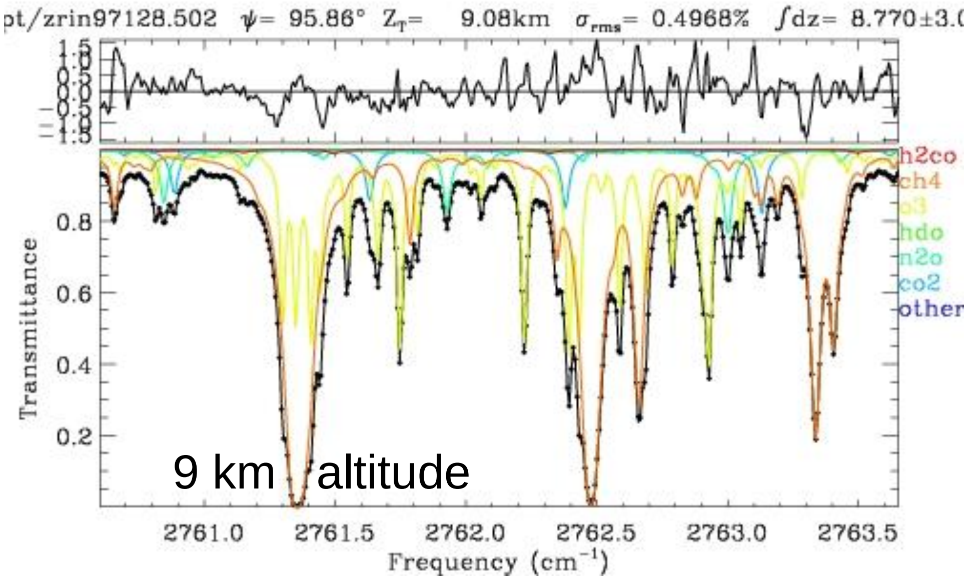
Empirically Adjusted Linelist

After manually adjusting the widths of the stronger CH_4 lines and adding the missing line at 2762.52 cm^{-1}, fits to lab spectra are much improved, both at low and high pressure.





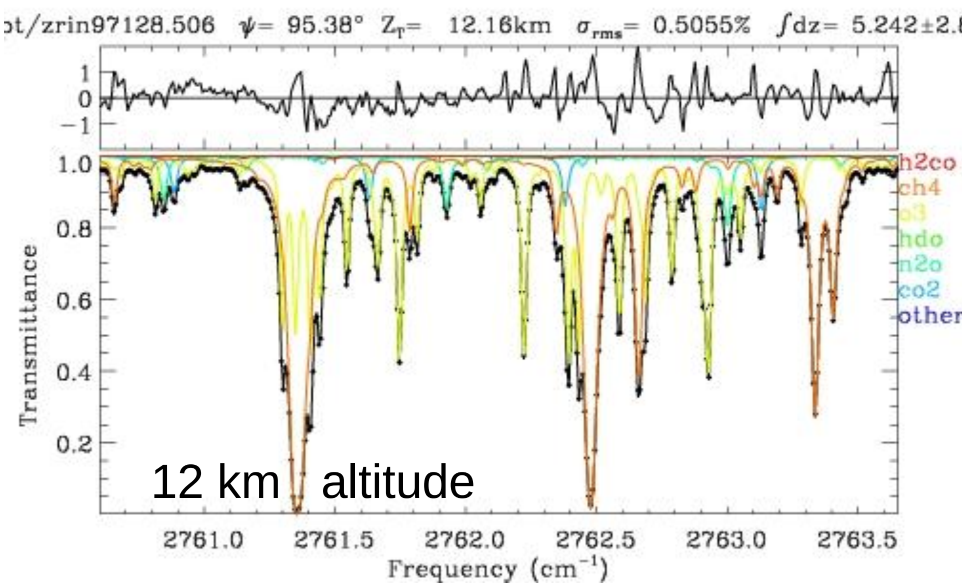
Re-fitted MkIV solar spectra: 9, 12, 34 km

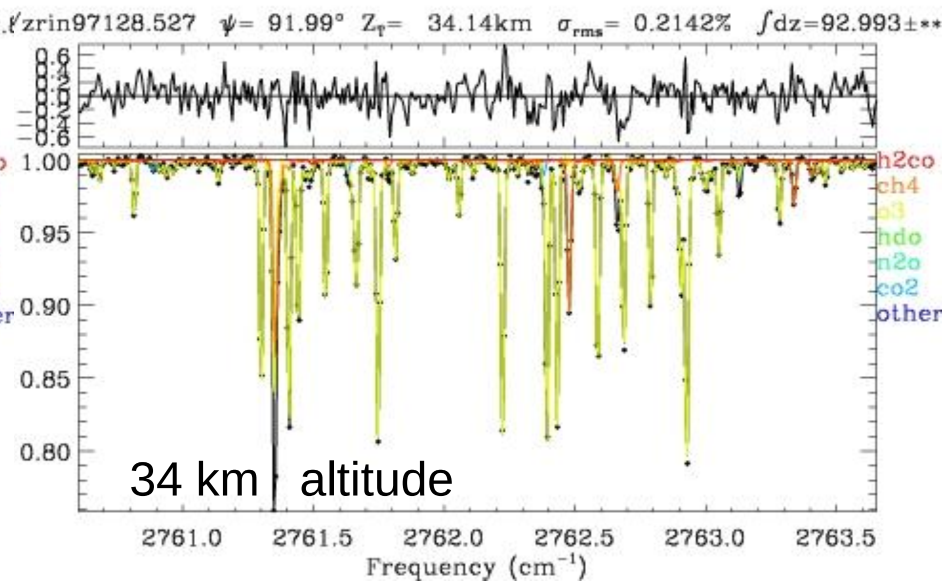


Using Empirically Adjusted Linelists

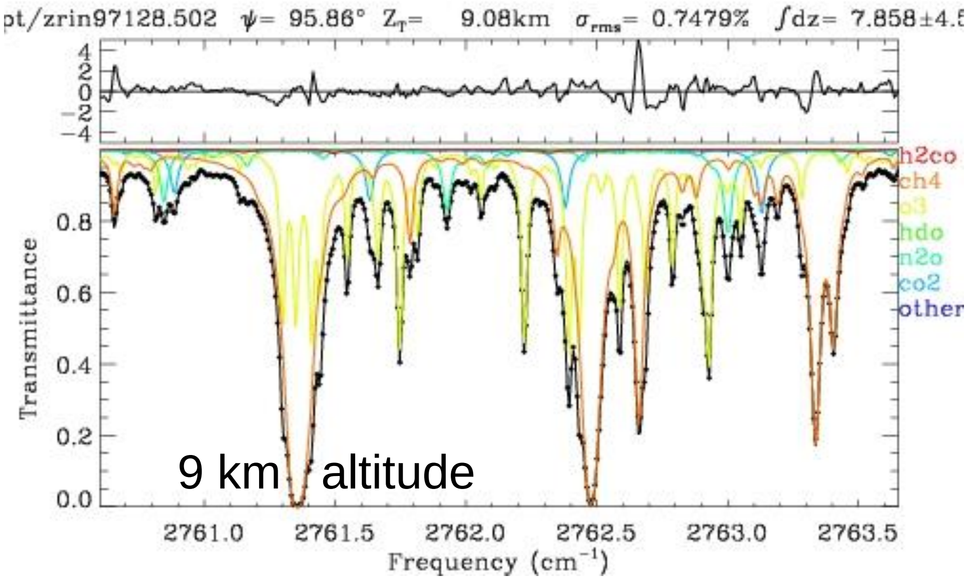
Peak residuals are 2-5 times smaller than using HITRAN 2008

RMS residuals are reduced by 30%



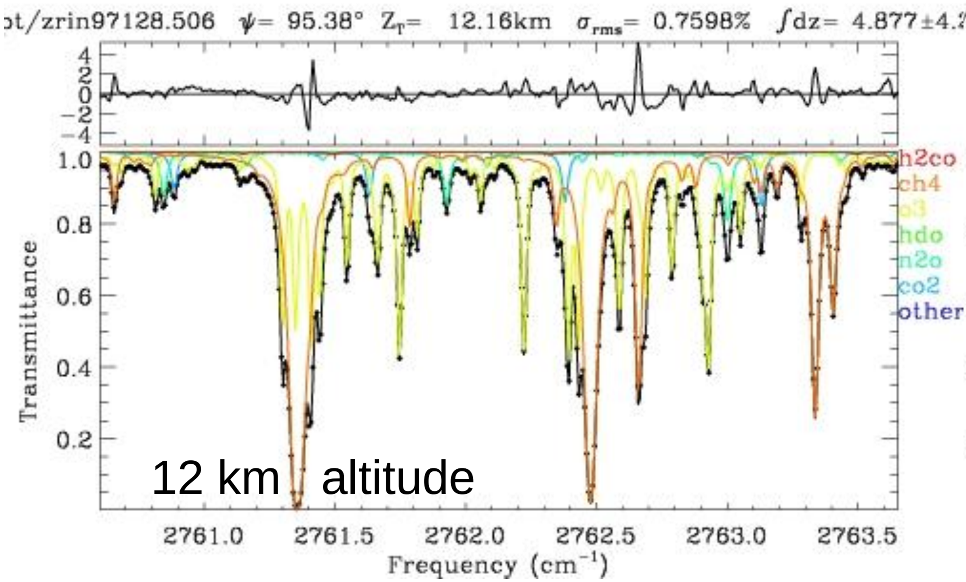


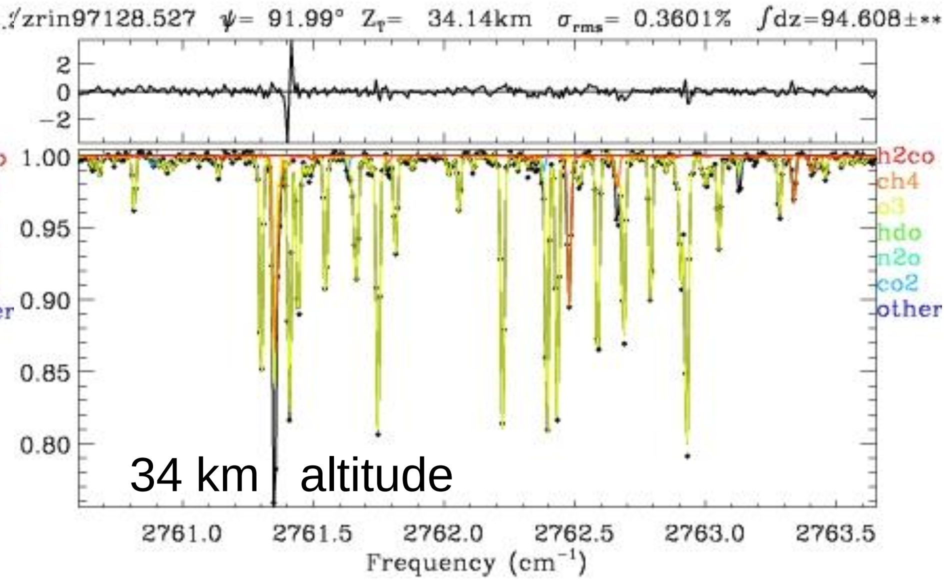
MkIV balloon spectra: 9, 12, 34 km



HITRAN 2008 linelist

Reminder of what it looked like
Before empirical adjustments.



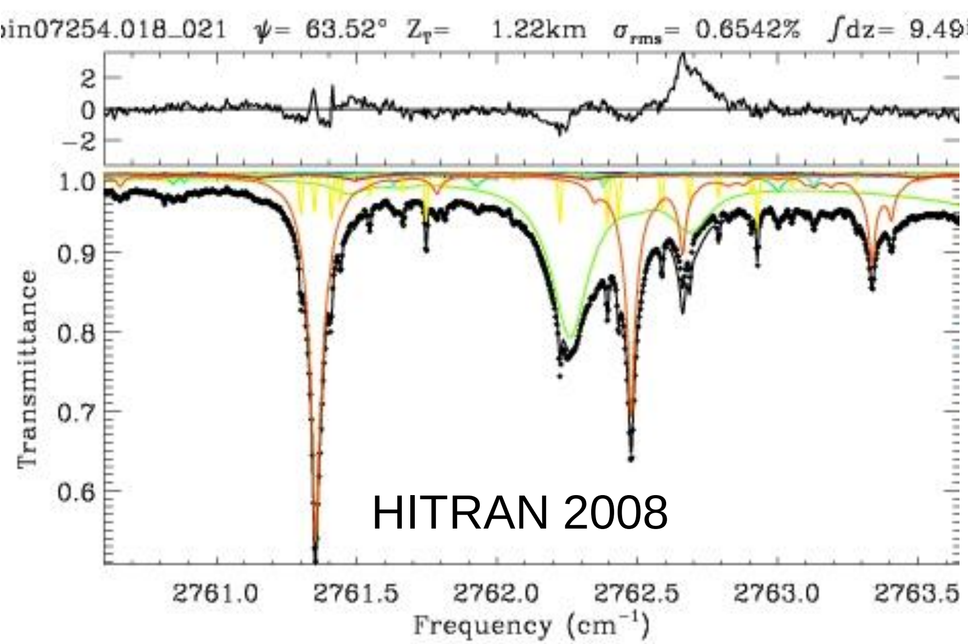


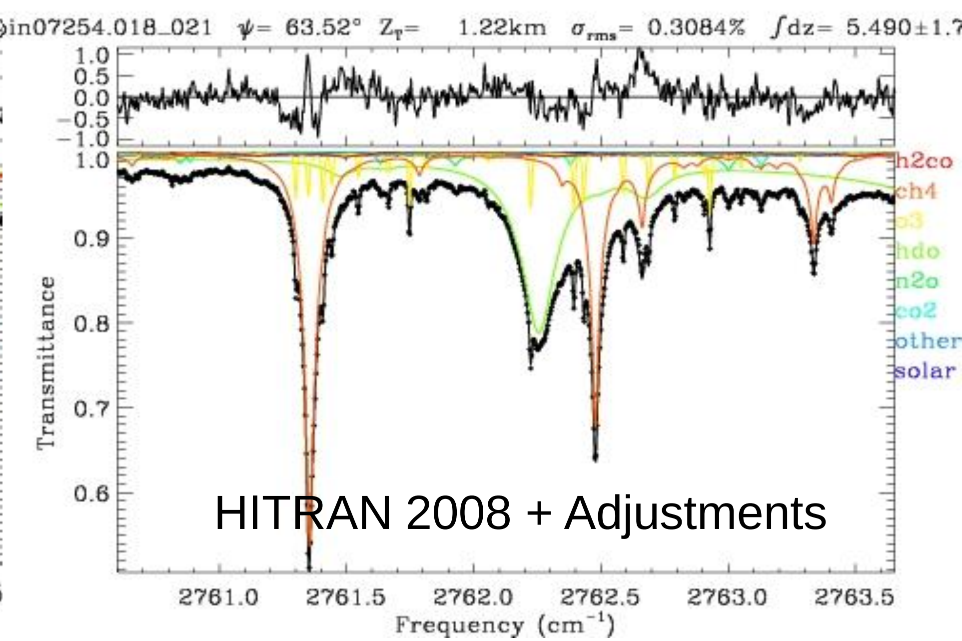
Summary: Ground-based MkIV H_2CO

Main linelist adjustments were:
- Factor 2.2 reduction in strength of HDO line at 2762.70 cm^{-1}, consistent with Toth
- Increased width of the CH_4 line at 2762.662 cm^{-1}, from 0.056 to 0.070 cm^{-1}/atm
- Added missing CH_4 line at 2762.521 cm^{-1}, consistent with Kitt Peak lab spectra
- Shifted position of the O_3 line at 2761.4095 to 2761.4074 cm^{-1}

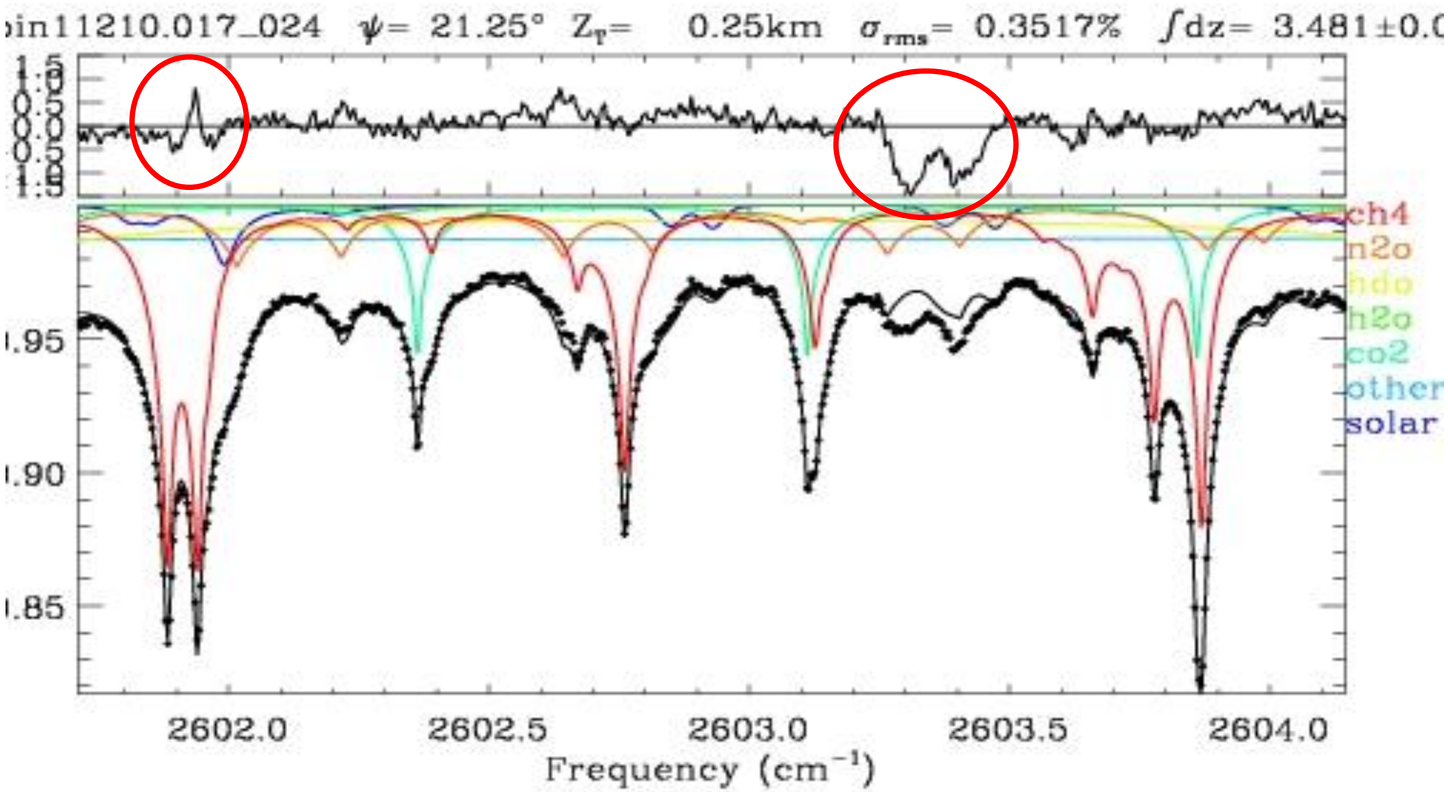
Adjustments reduce rms residuals by a factor 2 for wet case, and 20% for dry case.

More importantly, for wet spectrum, linelist adjustments reduce H_2CO by 40%



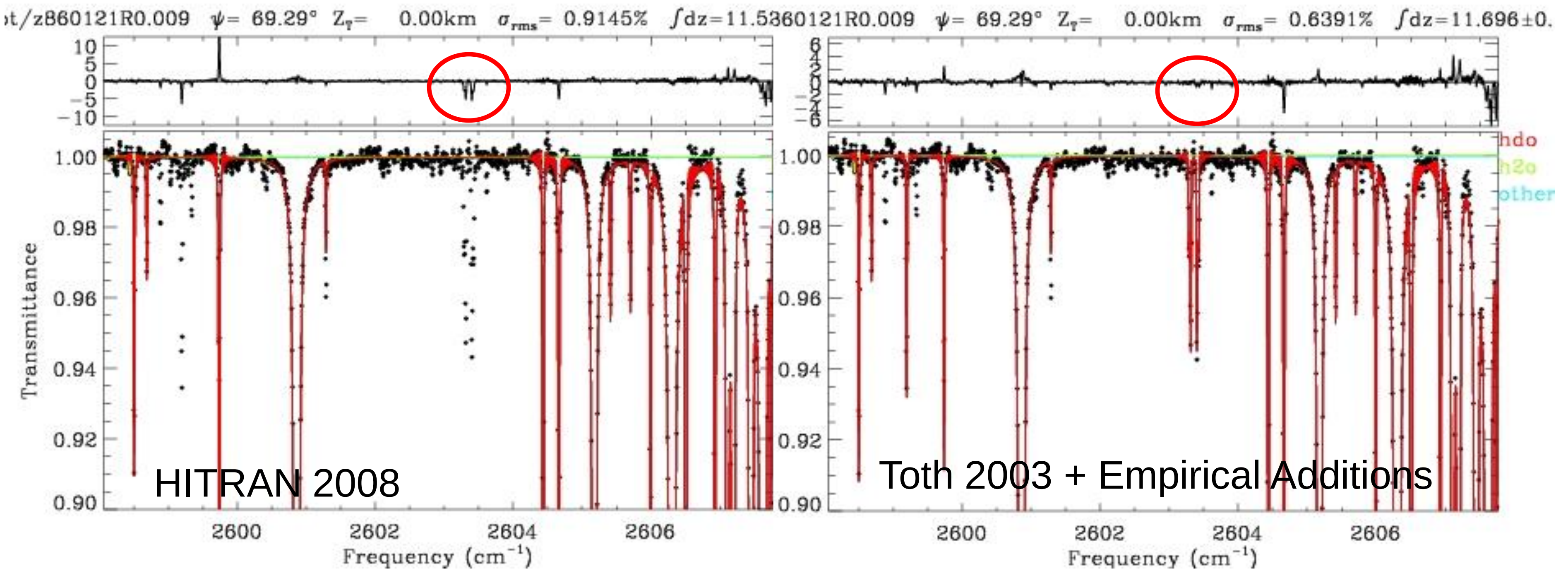


Case Study C. Ground-based CH_4



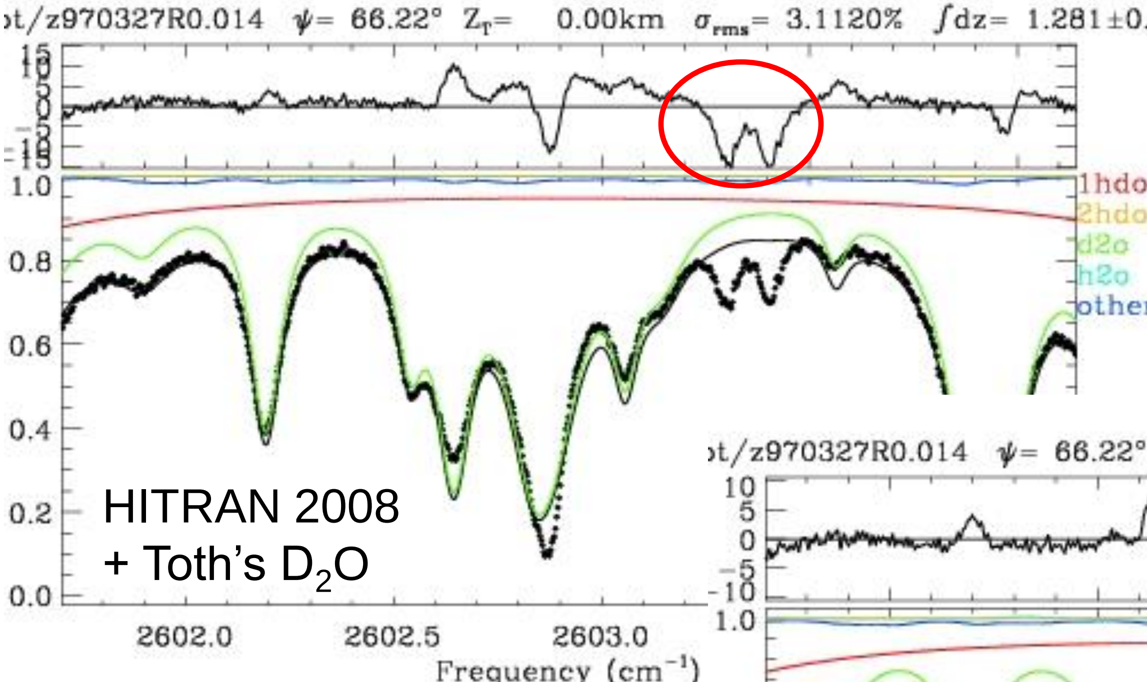
One of three mid-IR windows used to retrieve CH_4 from the ground
Large residuals seen at 2603.3 and 2603.4 cm^{-1} under humid conditions.
Residuals can reach 1.5% at low airmass and 10% at high airmass.
Shape of residuals, and their variability suggest water vapor.
Also, CH_4 line at 2601.93 cm^{-1} has wrong shape (width problem?)

Kitt Peak HDO Laboratory spectrum



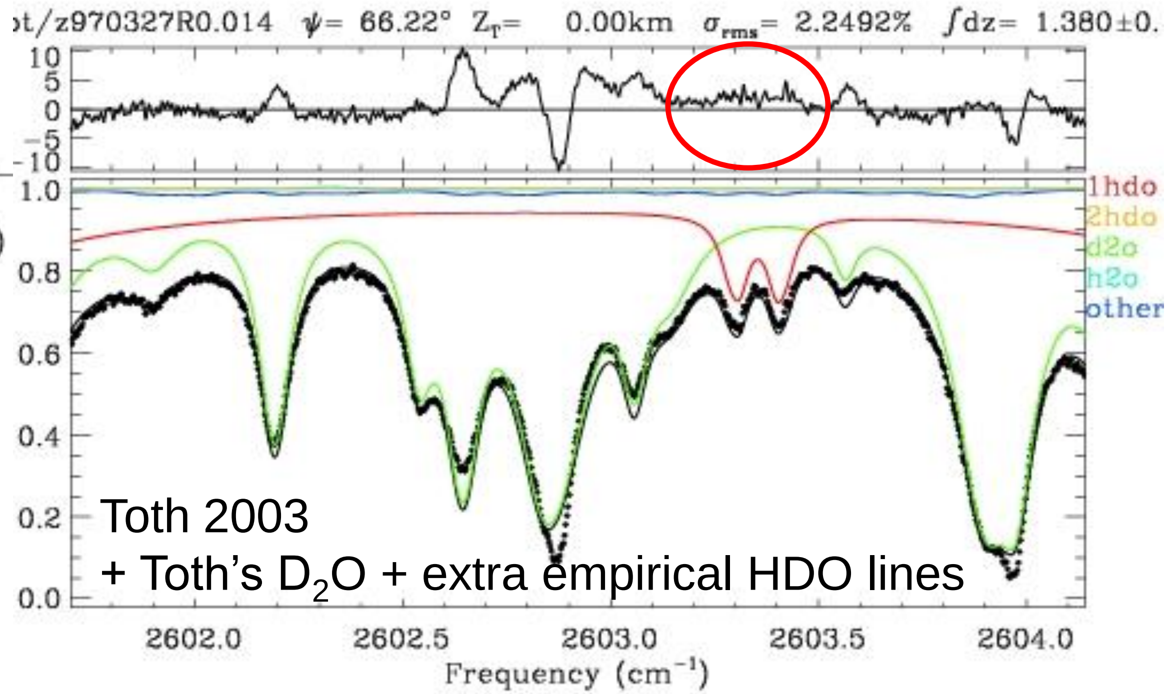
Fitted a Kitt Peak lab H_2O spectrum measured in Jan1986 at 4 Torr and 24C.
Left panel shows that are several lines missing from HITRAN 2008 in this region.
Right panel shows fits using HDO linelist empirically adjusted and supplemented with several new lines. Note change in residual scale from 12% peak to 6 % peak.

500 Torr KP D-enriched lab spectra



HITRAN 2008
+ Toth's D_2O

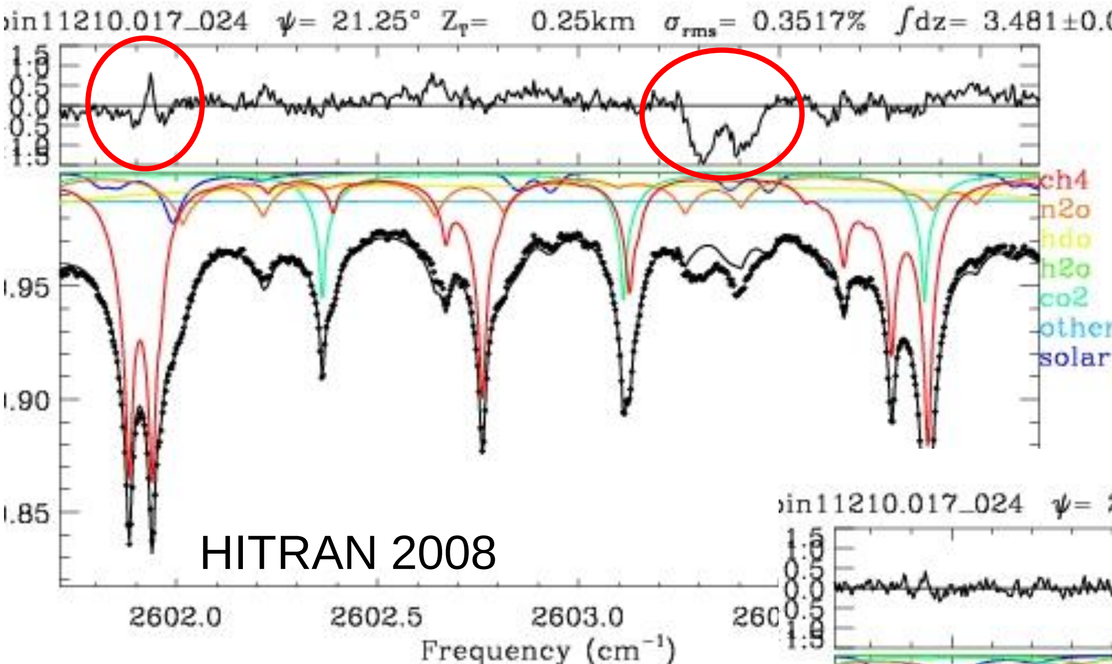
Residuals seen in fits of atmospheric spectra at 2603.3 and 2603.4 cm^{-1} also seen in D-enriched lab spectra, implying that missing lines are HDO.



Toth 2003
+ Toth's D_2O + extra empirical HDO lines

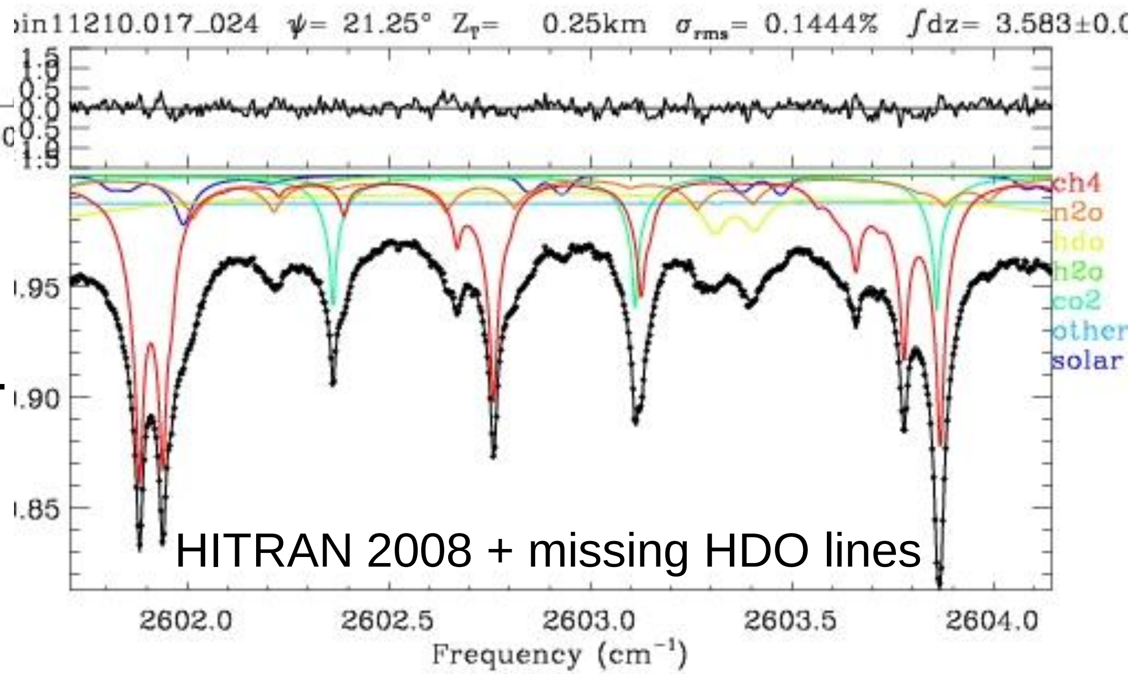
Adding extra HDO lines fixes the problem. Other large residuals due to D_2O, which are not a problem for atmospheric observations.

Summary – Ground-based MkIV CH_4



In fits to solar spectra, addition of the missing HDO lines completely fixes residuals in fits to atmospheric spectra

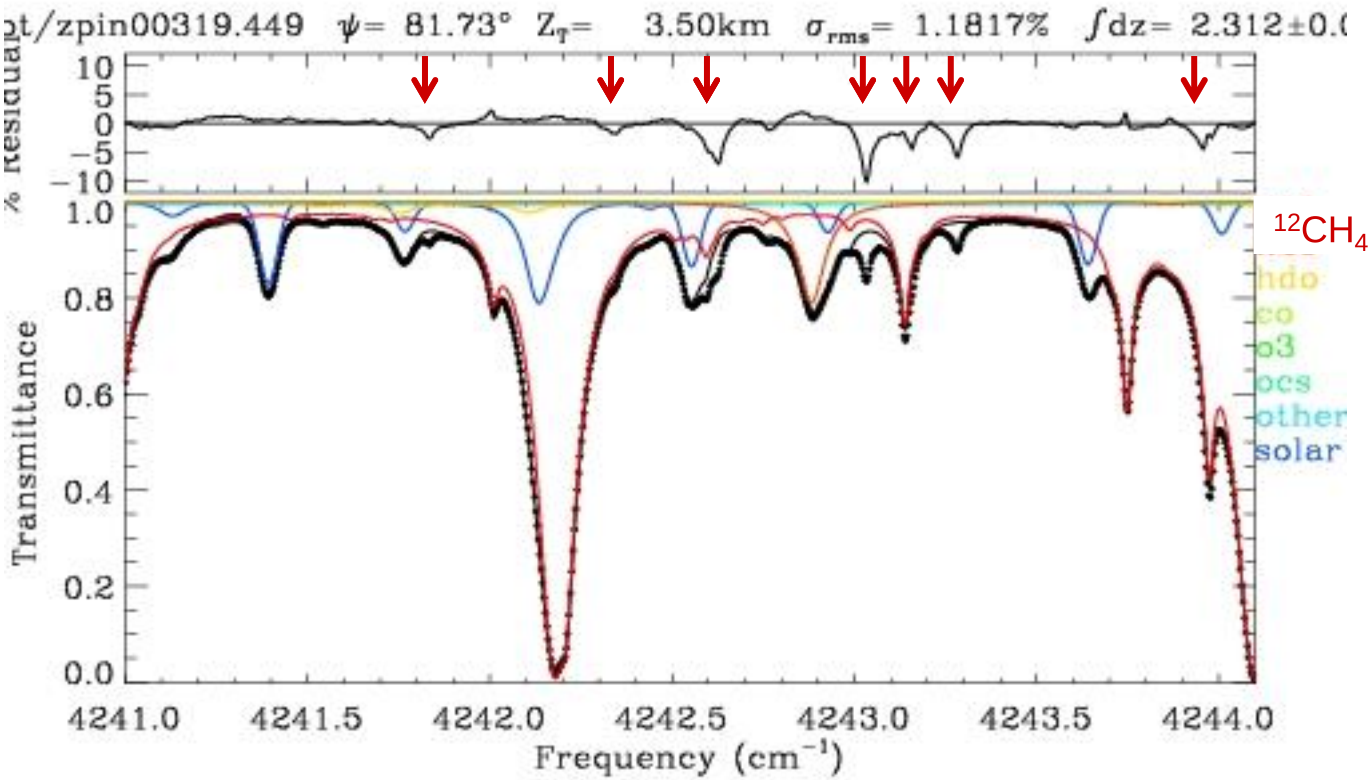
Addition of missing HDO lines improves spectra fits by factor 2.5, and changes retrieved CH_4 by 3%. Since atmospheric HDO is highly variable, the missing lines make the retrieved CH_4 seem noisy.



Case Study D: Missing $^{13}CH_4$ lines

HITRAN 2008 contains no $^{13}CH_4$ lines between 3364 and 5898 cm^{-1}

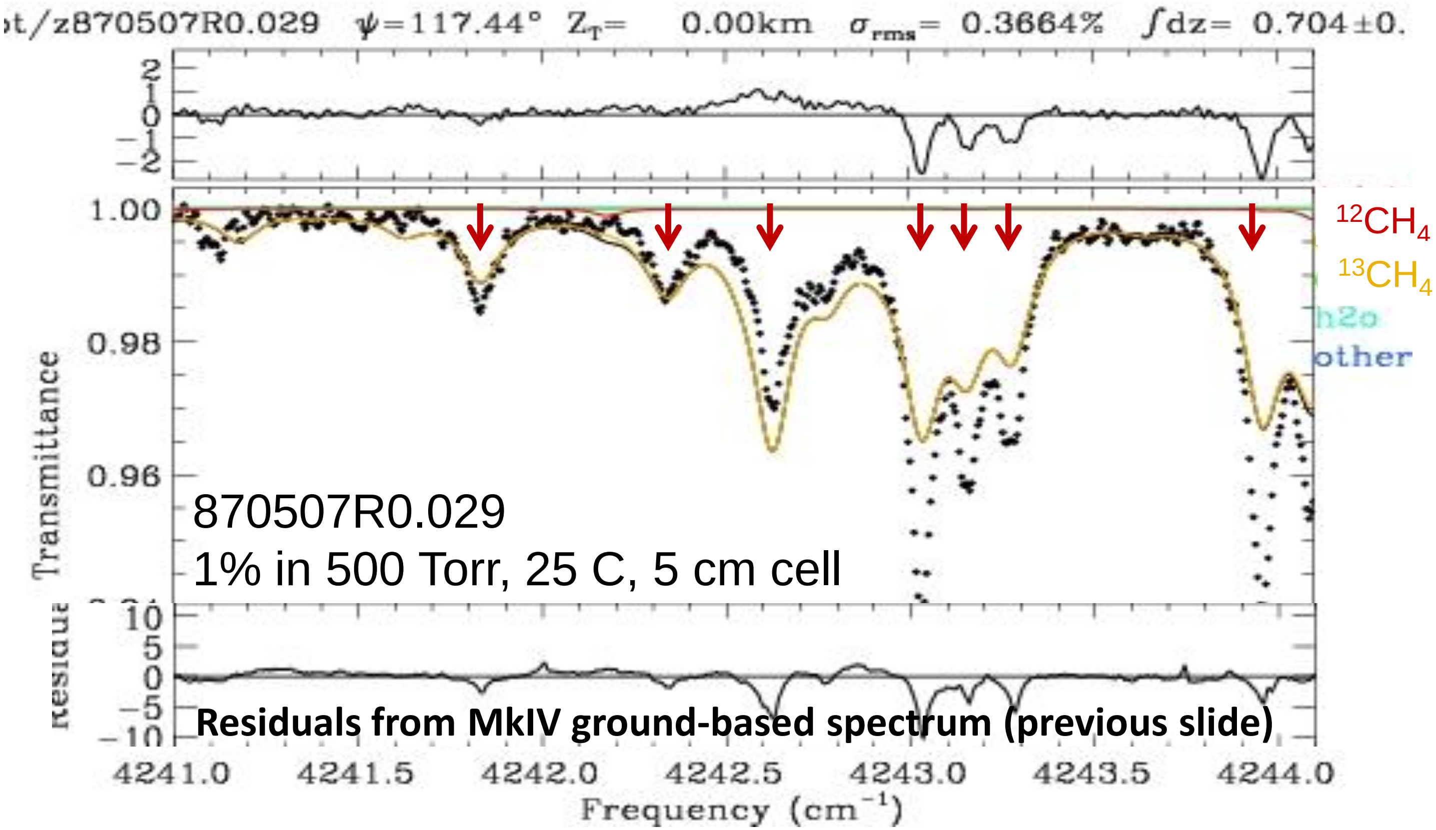
Since ^{13}C is only 90 times less abundant than ^{12}C, and $^{12}CH_4$ lines are saturated, the missing $^{13}CH_4$ lines should be several % deep.



Above is a fit to a MkIV ground-based spectrum using HITRAN 2008.

Could $^{13}CH_4$ be causing the dips in the residuals which reach 10% ?

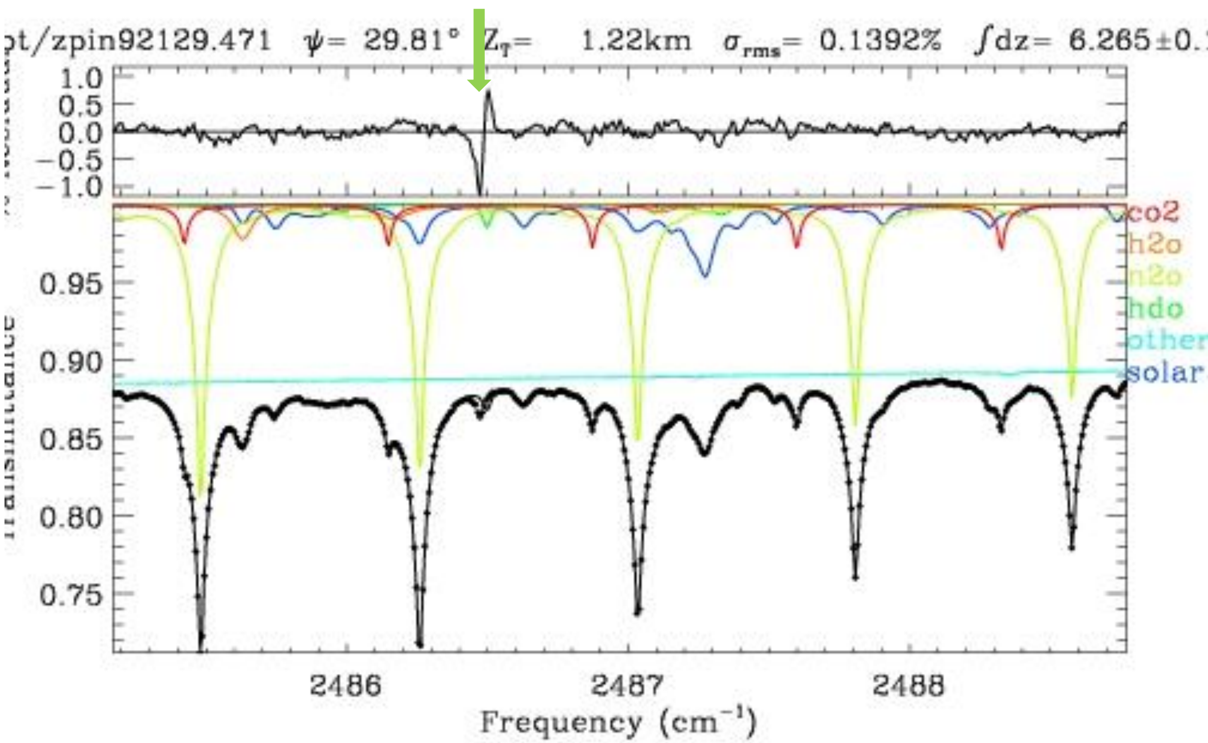
Fit to Kitt Peak $^{13}CH_4$ lab spectrum

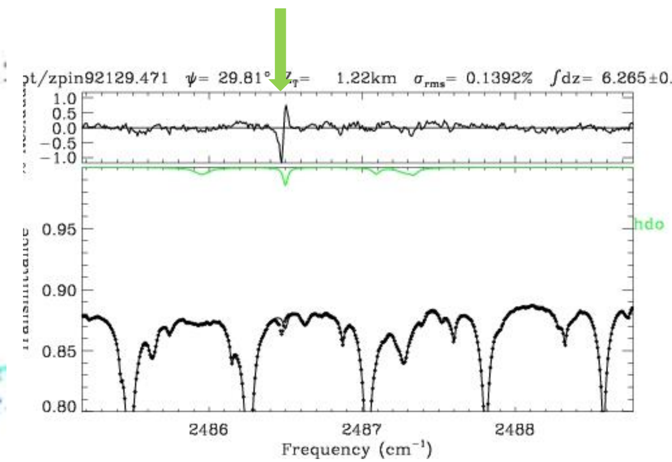


Almost perfect correlation between the MkIV fitting residuals from the previous slide and the $^{13}CH_4$ absorption features in the Kitt Peak spectrum (black dots). This confirms that the major residuals in previous slide are due to $^{13}CH_4$. Yellow line is $^{13}CH_4$ calculation using linelist downloaded from "Methane@Titan" website: http://www.icb.cnrs.fr/titan/

Case Study E: Ground-based N_2O & CO_2

To be useful, column-average CO_2 and N_2O measurements must be better than 0.2%.
This represents a <0.04% transmittance error to a N_2O line that is 20% deep.
So systematic residuals exceeding 0.1% are unacceptable, even though they don't overlap the N_2O or CO_2 features.



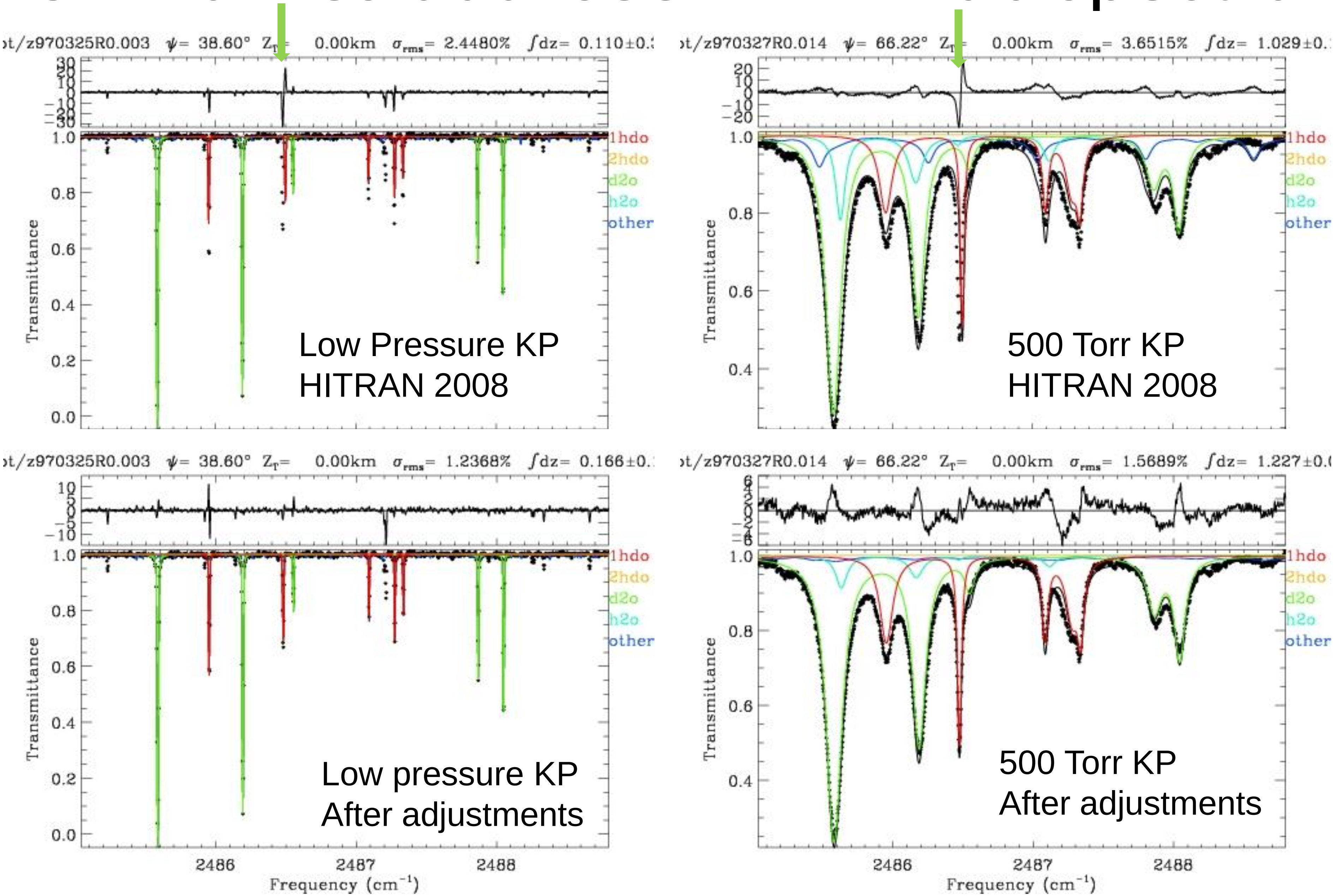


Large fitting residual due to error in position of narrow HDO line at 2486.49 cm^{-1}

Other HDO lines seem okay

Position error or shift error?

Similar residual seen in KP lab spectra



Adjusting the problematic HDO lines

Problem HDO lines tabulated in HITRAN:

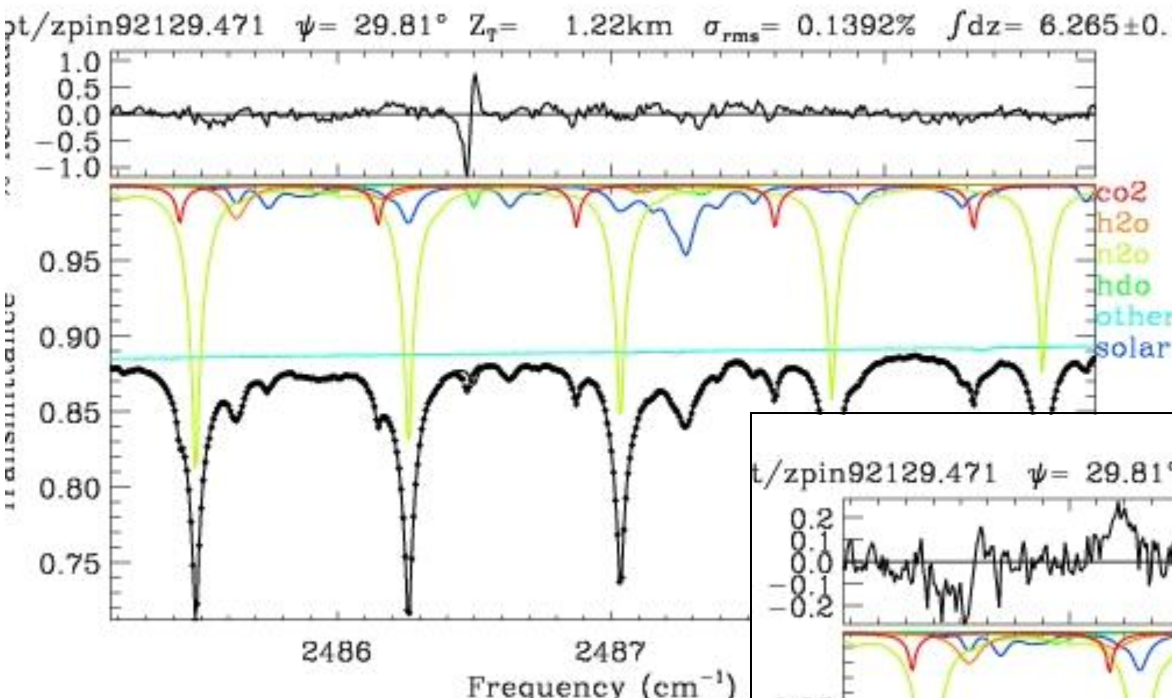
HDO 49 1 2486.496120 1.307E-26 1839.7539 .0181 .1330 .38 0.000 1 0 0 0 0
0 15 0 15 16 0 16 355540301824 3 2 0 186.0 198.0

HDO 49 1 2486.501000 1.307E-26 1839.7570 .0178 .1400 .38 0.000 1 0 0 0 0
0 15 1 15 16 1 16 355540301824 3 2 0 186.0 198.0

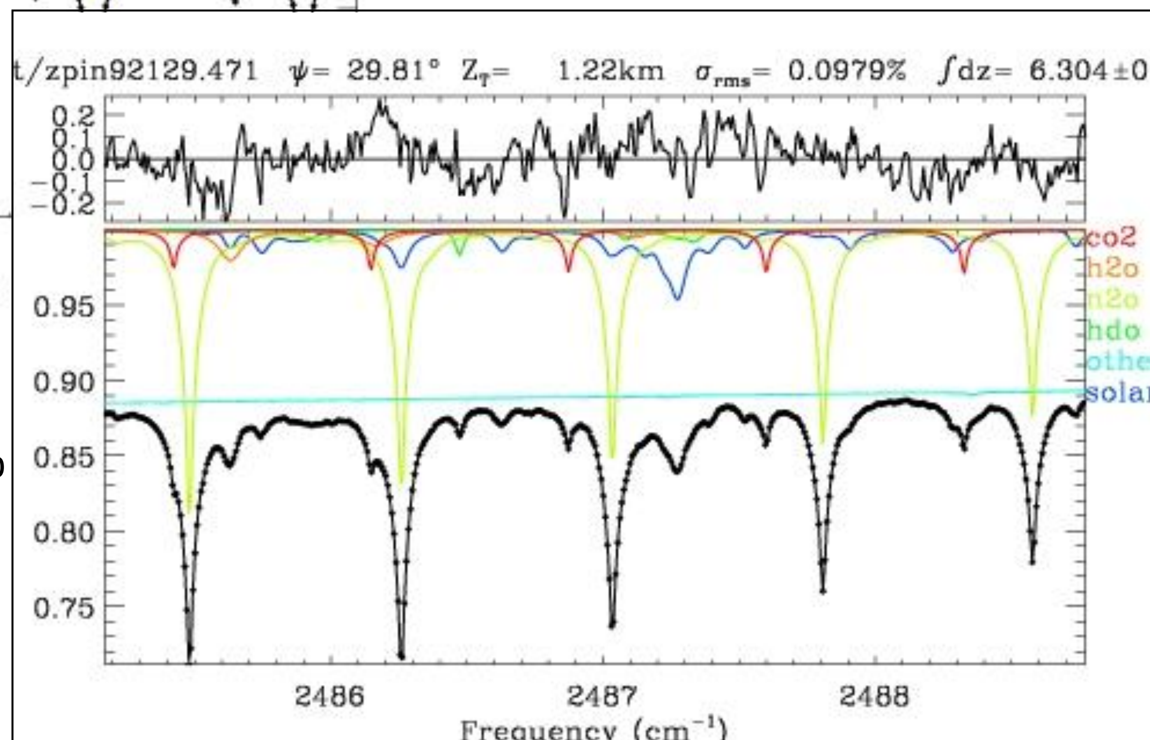
Used low pressure KP spectra to adjust line positions by -0.017 cm^{-1}
Used high pressure KP spectra to adjust the pressure shift to -0.010 cm^{-1}/atm
Both are necessary to fit both low- and high-pressure KP lab spectra

BTW, In HITRAN 2008, above 2000 cm^{-1} all HDO pressure shifts are zero.

Summary: Ground-based N_2O & CO_2



Fits to ground-based MkIV Spectra before (left) and after (below) adjusting problematic HDO lines. Note residual scale change.



After fixing HDO lines, peak residual drops from 1.2% to 0.3%

More importantly, the retrieved CO_2 column changed by 0.7%

Conclusions (1/2)

In almost any region in the mid-IR, systematic residuals arising from spectroscopic inadequacies can be seen in solar atmospheric spectra.

Non-linear least-squares spectral fitting techniques, generally used in remote sensing, give undue influence to the largest residuals. So need to fix them.

In nearly all cases, the same systematic residuals seen in fits to solar spectra can also be found in fits to lab spectra with similar conditions.

So analysis of relatively few solar spectra measured under a wide, but well known, range of conditions (cold/warm, low/high airmass) can provide a good test of spectroscopy of nearly all gases important for Earth remote sensing, indentifying:

- Important missing lines
- Errors in existing lines

Solar spectra also have the benefit (over lab spectra) of high SNR and broad spectral coverage because the sun is a very bright source. Allows band-to-band consistency checks.

Conclusions (2/2)

Atmospheric spectra not only allow identification of spectroscopic problems, they also provide diagnostic information, and allow evaluation of proposed solutions (e.g. new/adjusted linelists) in terms of:

- Quality of the spectral fits
- Impact on retrieved gas abundances

Acknowledgements

NASA Upper Atmosphere Research Program who have funded the JPL MkIV task for many years.

Kitt Peak National Observatory, whose library of over 30,000 high-quality solar and laboratory spectra acquired over the past 35 years is a gold-mine. [*Warning: do not trust the measurement conditions in Kitt Peak file headers. Instead use published values or contact people who made the measurements*]

I thank the various people who made preliminary linelists available to me.

Finally, I thank the entire international spectroscopy community. Without their dedicated work, atmospheric remote sensing would be impossible.

Comment:

I hope my talk didn't sound like I was complaining all the time about HITRAN.
There was a lot of progress between HITRAN 2004 and 2008.
But the improvements are less newsworthy than the deficiencies.
Spectral fits showing random residuals are not so interesting.