CO_2 Spectroscopy Evaluation: 670 to 7000 cm^{-1}

Geoff Toon
Jet Propulsion Laboratory
California Institute of Technology

Several CO_2 linelists including HITRAN 2008, 2012, and 2016 (two versions), have been evaluated by fitting laboratory spectra (mainly Kitt Peak) and atmospheric solar absorption spectra (MkIV & TCCON).

The 670-7000 cm^{-1} region of interest was divided into 41 windows, most encompassing at least one complete CO_2 absorption band or sub-branch. Regions with no discernable CO_2 absorption were skipped.

The GFIT spectral fitting algorithm was used in all cases assuming a Voigt lineshape and no line-mixing. This evaluation focusses on the RMS fitting residuals that were achieved and the window-to-window consistency of the retrieved CO_2 amounts.

Between evaluations of the different linelists, only the CO_2 linelist was changed. So difference in the RMS fitting residuals or the retrieved CO_2 amounts is entirely attributable to the CO_2 linelist under evaluation.

A new "greatest hits" linelist (ATM18) was subsequently developed by selecting from the best predecessor linelists. Ad hoc manual adjustments were then performed to fix obvious errors (e.g. bad line positions, pressure shifts, inconsistent retrieved CO_2 amounts). To keep this report concise, the new ATM18 linelist is presented in parallel with the evaluation of earlier linelists, even though it was developed much later.

The CO_2 Linelists Evaluated

HITRAN 2008: 314,919 CO_2 lines

HITRAN 2012: 471,847 CO_2 lines

ATM 2016: 450,493 CO_2 lines
Based mainly HITRAN 2012. Uses Toth (2009) for the 5740-6500 cm^{-1} region because it gave better fits (and still does). Empirical adjustments have been made throughout to fix obvious errors (mainly line position errors).

HITRAN 2016a: 554,183 CO_2 lines
Based on files the linelist that Iouli Gordon sent me June 27, 2017 (O2_hit16_first-9iso) and on June 30, 2017 (hit838corr). Before using, I fixed 19 lines with ABHW=0 and one with an intensity of zero (used HIT 2012 value).

HITRAN 2016b: 554,879 CO_2 lines
Downloaded from HITRAN-Online website on Nov 28, 2017 (5a1de32a.par). Includes isotopologs 11 & 12. A format of "f5.3" had been enforced for SBHW, which changes some lines that were previously "f5.4", e.g.

25 3.681760 1.943E-33 1.457e-12.0865.1155 in HIT16a

became

25 3.681760 1.943E-33 1.457e-12.08650.116 in HIT16b

ATM 2018: 524,724 CO_2 lines (new linelist)
Mostly HIT 2016b, except in regions where ATM 2016 or HIT 2008 were better. Some ad hoc empirical adjustments.

The Laboratory Spectra of CO_2

Kitt Peak CO_2 lab spectra are available covering 600 to 12,000 cm^{-1}, although here we investigate 670 to 7000 cm^{-1}.

There are 148 spectra: 136 from Kitt Peak and 12 from JPL (Keeyoon Sung). Pressures range from 0.1 to 700 Torr.

All at room temperature (291-303 K) except for two Kitt Peak spectra:

- One at 268K and 14.2 Torr in a 30 cm cell covering 600 -1400 cm^{-1}
- One at 235K and 12.8 Torr in a 30 cm cell covering 600 -1400 cm^{-1}

23 Kitt Peak spectra are enriched in ^{13}C, giving the lab spectra a much higher sensitivity to spectroscopic errors in isotopologs 2, 5, 6, 10, 11, 12, than the other lab spectra (or atmospheric spectra).

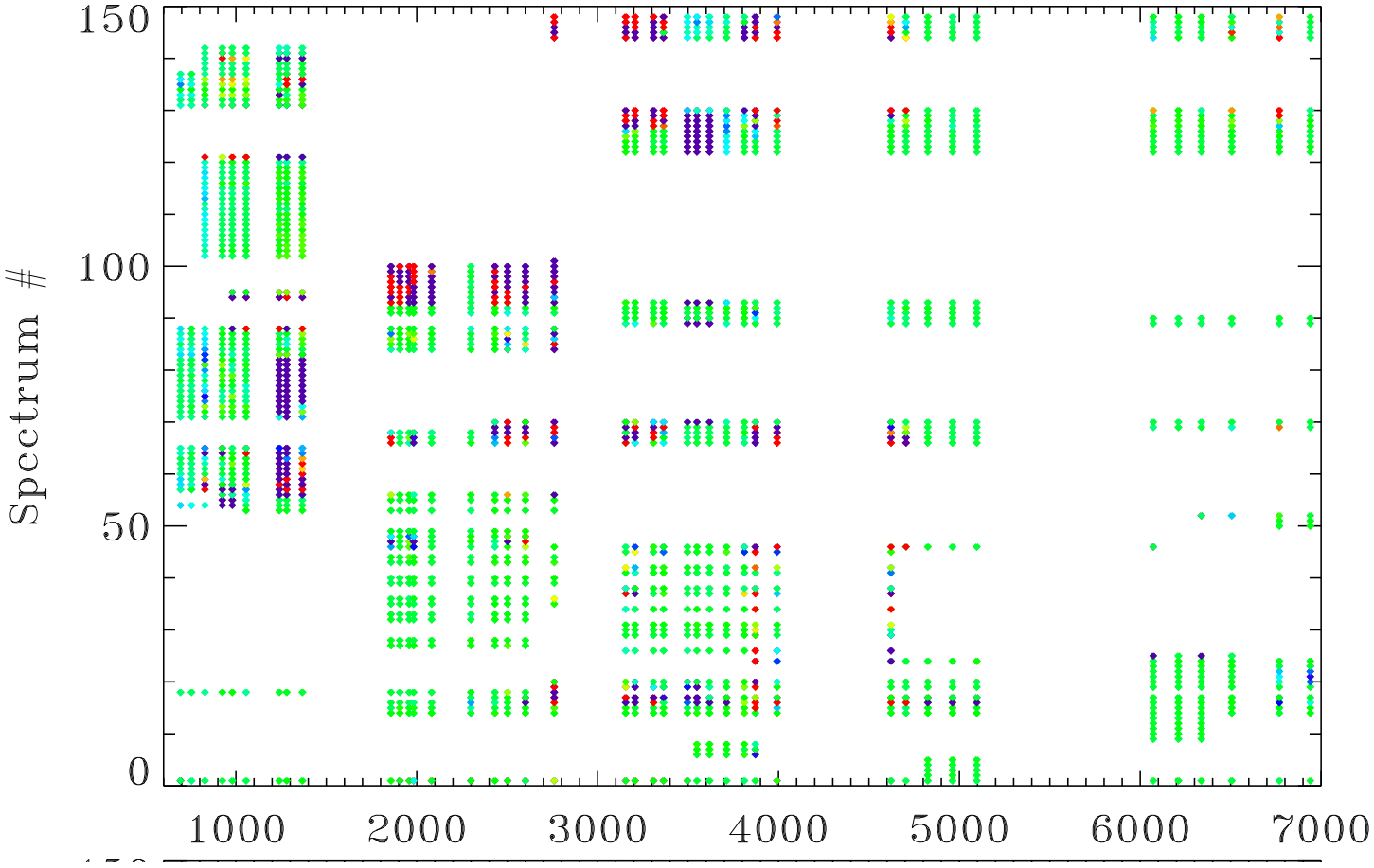
Of the 6068 potential spectral fits (41 windows x 148 spectra), only 1816 (29.9%) could actually be performed for each linelist due to the limited spectral coverage of the individual spectra, most of which have < 1000 cm^{-1} of useful coverage.

This makes it difficult to compare intensities measured at low wavenumbers with those from high wavenumbers because these are seldom in the same spectrum. And on the rare occasions when they are, the SNR is poor.

Retrieved CO_2 VMR Scale Factors

VMR Scale Factors (VSFs) are the ratio of the retrieved gas amount to that expected based on the measurement condition (cell length, T, P, VMR). In a perfect case, the VSFs should all be 1.0.

Upper Panel. The retrieved CO_2 VSFs values, for each window and each lab spectrum are color-coded (blue=0.5, green=1.0, red=1.5) and are plotted versus the window center wavenumber and an arbitrary spectrum #. Spectra 1-12 are from JPL (Sung), the remainder from Kitt Peak. Only one spectrum (#1) covers the full spectral range. Most cover less than 1000 cm^{-1}. Gaps in the wavenumber coverage (e.g. 4000-4600, 5200-6000 cm^{-1}) imply weak, undetectable CO_2 lines.

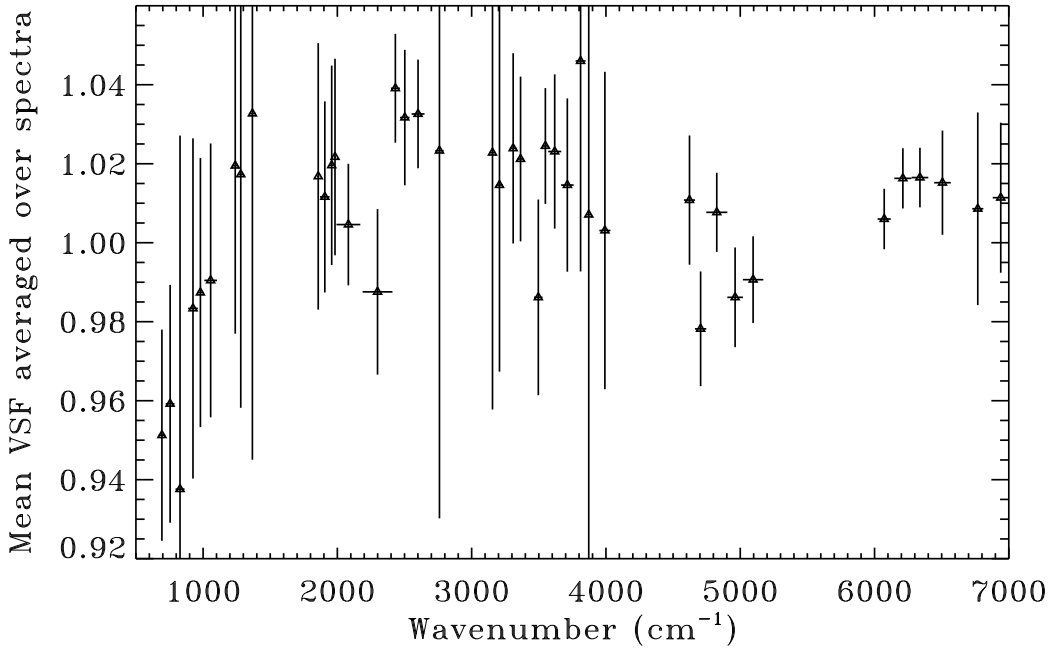


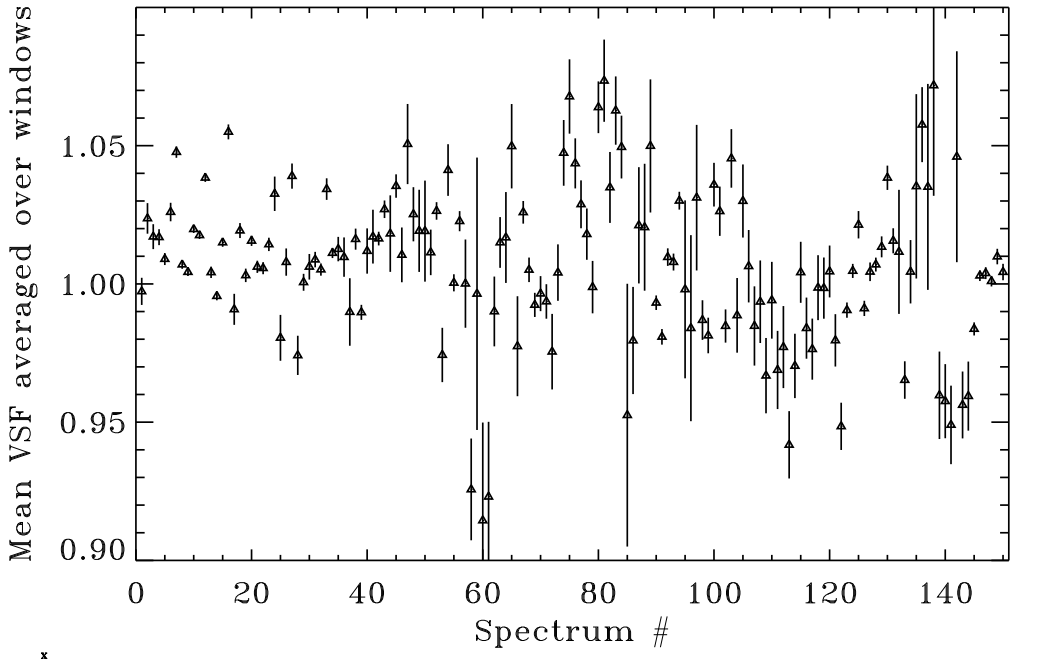
CO_2 VSFs averaged by window (top) and by spectrum (bottom)

Top Panel: VSF values obtained using the ATM18 linelist, averaged over the different lab spectra fitted in a particular window and plotted versus its center wavenumber. This exposes windows in which the retrievals are wrong due to factors common to the majority of the fitted spectra, e.g., spectroscopy.

Bottom Panel: VSF values from a particular spectrum averaged over the fitted windows and plotted versus spectrum#. This exposes spectra in which the retrievals are wrong due to factors specific to that particular spectrum, e.g. the assumed VMR, Pressure, Temp, or path length may be wrong. Or the ILS might be mis-aligned. Or a large zero-offset is present.

These plots summarize the information presented 2 slides ago by averaging the columns and the rows. In general the error bars in the lower panel are smaller than those in the upper panel, which implies that spectrum-to-spectrum uncertainties in retrieved CO_2 are larger than window-to-window variations.





RMS Residuals from fits to laboratory spectra

Showing the RMS spectral fits for 41 windows, averaged over the 148 lab spectra. This is done for 6 different linelists. These are the same data tabulated on the previous slide.

Upper panel shows absolute RMS residuals. Lower panel shows differences from HIT12.

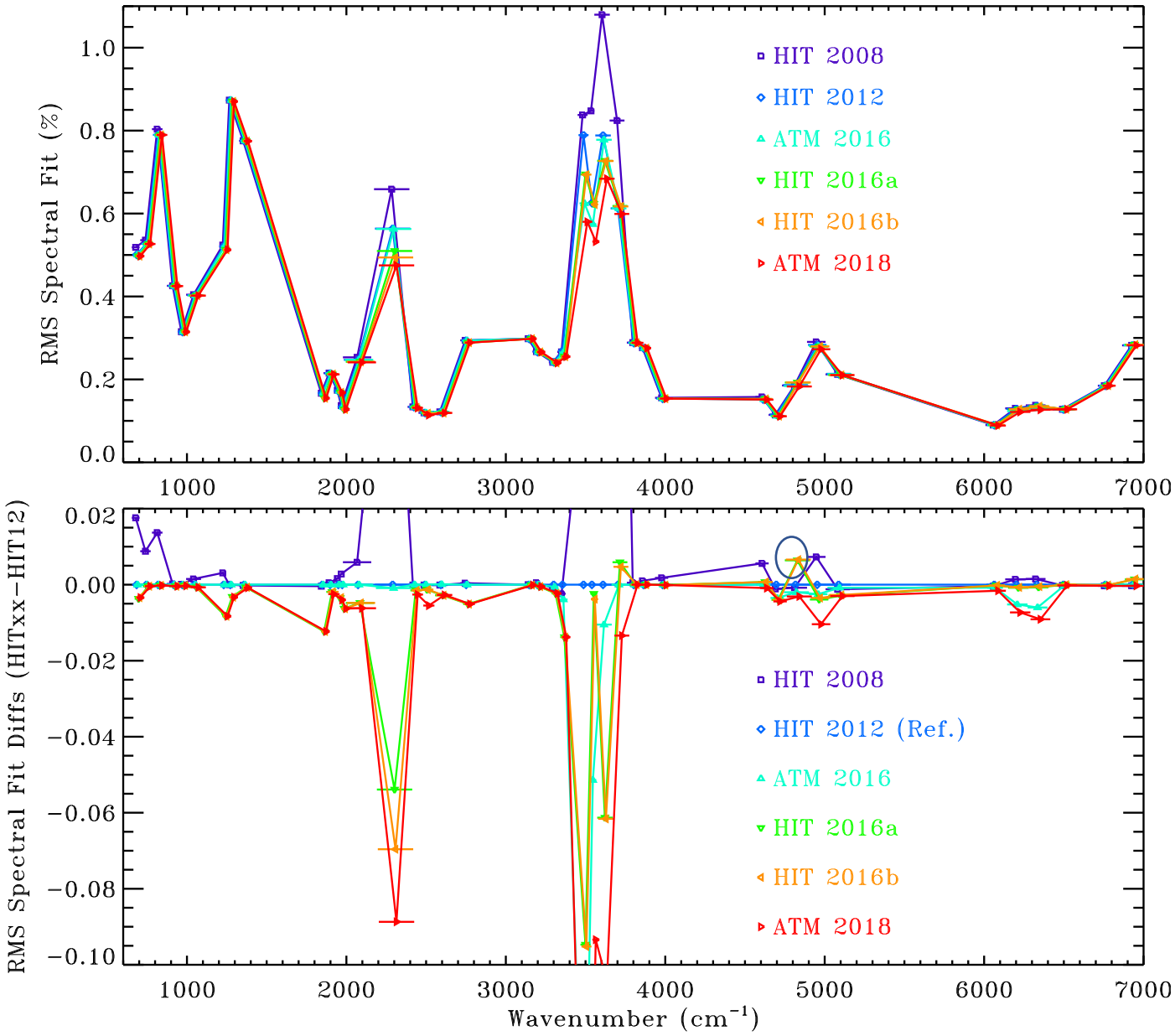
The absolute value of the RMS fit is unimportant. This is generally dominated by instrumental issues and interfering absorptions (e.g. H_2O).

Variation of RMS from linelist to linelist is entirely due to the CO_2 spectroscopy, since nothing else has been changed.

Big improvements are apparent for HIT16a,b as compared with HIT12 at 2300 cm^{-1} & 3600 cm^{-1}.

HIT16a and HIT16b produce similar results. The largest difference is seen at 2300 cm^{-1} due to inclusion of isotopologs 11 & 12 into HIT16b.

In the 4825 cm^{-1} window, HIT16 produces the worst fits (circled).



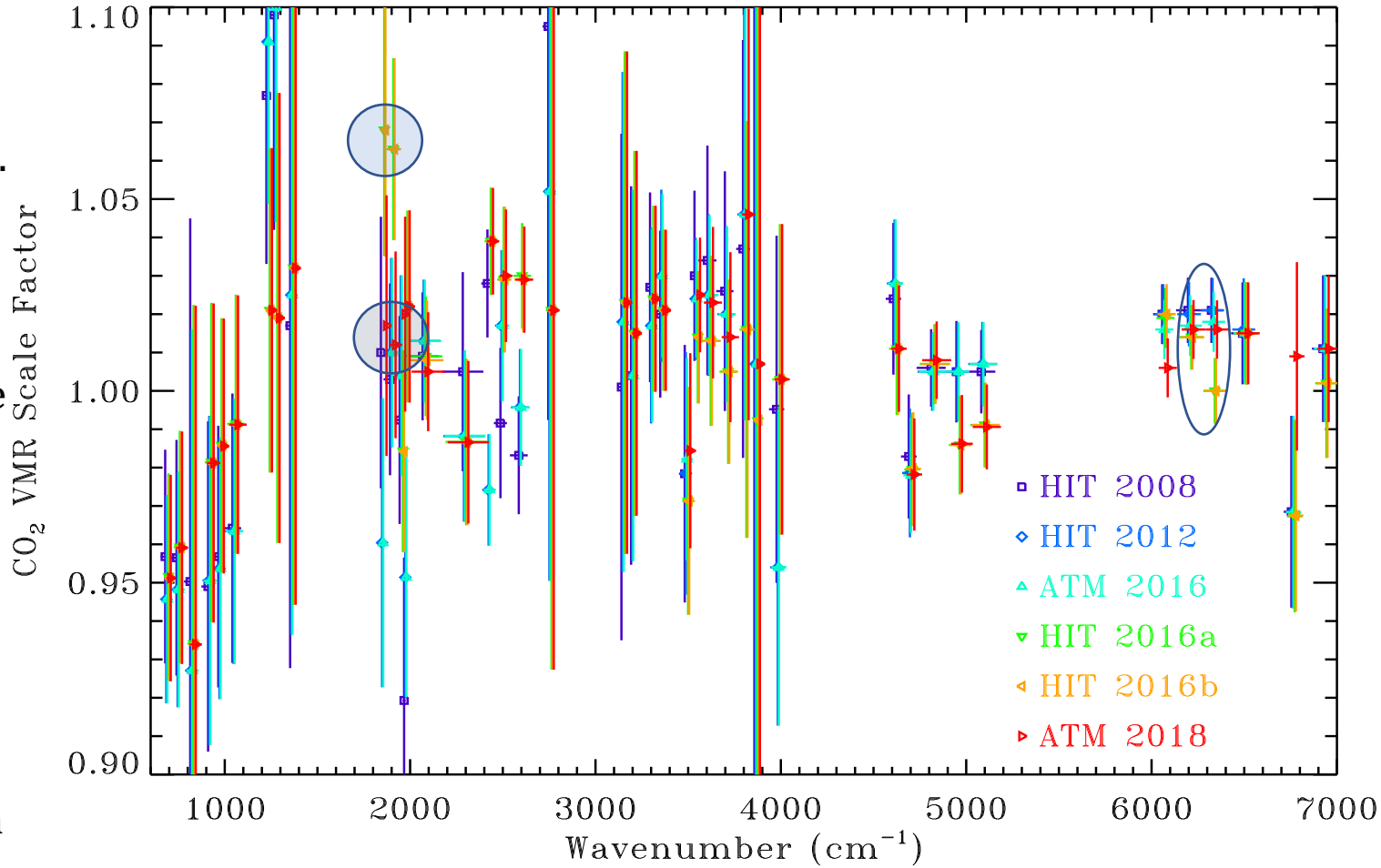
CO_2 VMR Scale Factors retrieved from lab spectra

Comparison of VSF values for all 6 linelists using the data tabulated in the previous slide, together with their (untabulated) uncertainties. Red points (ATM 2018) are identical to those shown in top panel of slide # 5.

VSF values greater than 1 mean that the line intensities, or the absorber amounts, need to be multiplied by the VSF value.

Width and line position errors can also contribute to an incorrect retrieved CO_2 amount, but in this case the relationship between the VSF value and the width/position error is more complicated.

A discrepancy is apparent between retrievals in the 6220 and 6338 cm^{-1} bands using HIT16



Chris Boone reported 5-10% larger retrieved ACE CO_2 amounts from 1915 cm^{-1} band than from 2050 cm^{-1} band using HIT16 with ACE data. Kitt Peak lab spectra confirm this (upper circle). This bias didn't exist with earlier linelist editions. It was fixed in ATM18 (lower circle). A 4% reduction to ATM18 CO_2 intensities in the 6740 cm^{-1} band is also apparent.

RMS Residuals from fits to lab spectra

HITRAN 2008 is clearly the worst overall.
Of the pre-2018 linelists, HIT16b is the best overall.
ATM18 is of course the best overall, being cherry-picked from the best parts of the earlier linelists.

Comparing HITRAN 2012 with ATM 2016: They produce similar results

- In 14 windows ATM 2016 is better
- In 27 windows they produce equally good fits
- In 0 windows HITRAN 2012 is better

It is no surprise that HITRAN 2012 is nowhere better than ATM 2016. If it had been, I would have replaced the offending lines in ATM 2016 with those from HITRAN 2016. Additionally, empirical adjustments have been performed to the ATM linelist to fix obvious deficiencies (e.g., position errors).

Comparing HITRAN 16a and 16b: They produce very similar results. Improvements seen in strong bands at 2300 cm^{-1} and ~3600 cm^{-1} in ^{13}C-enriched lab spectra due to addition of isotopologs 11 and 12 to HIT16b

- In 10 windows HIT16a is slightly better
- In 24 windows they produce the same rms fit
- In 7 windows HIT16b is slightly better

Since only 2/148 lab spectra used here was below 290K, these results don't really validate the T-dependent parameters.
In the 4825 cm^{-1} window the HIT16 linelists produces significantly poorer residuals than any predecessor.

MkIV Balloon Spectra: RMS Spectral Fitting Residuals

Top Panel: Plot of the data tabulated on the previous slide. Shows RMS residuals for 35 windows using 6 different linelists. The absolute fitting residuals are dominated by interfering atmospheric absorptions, especially H_2O.

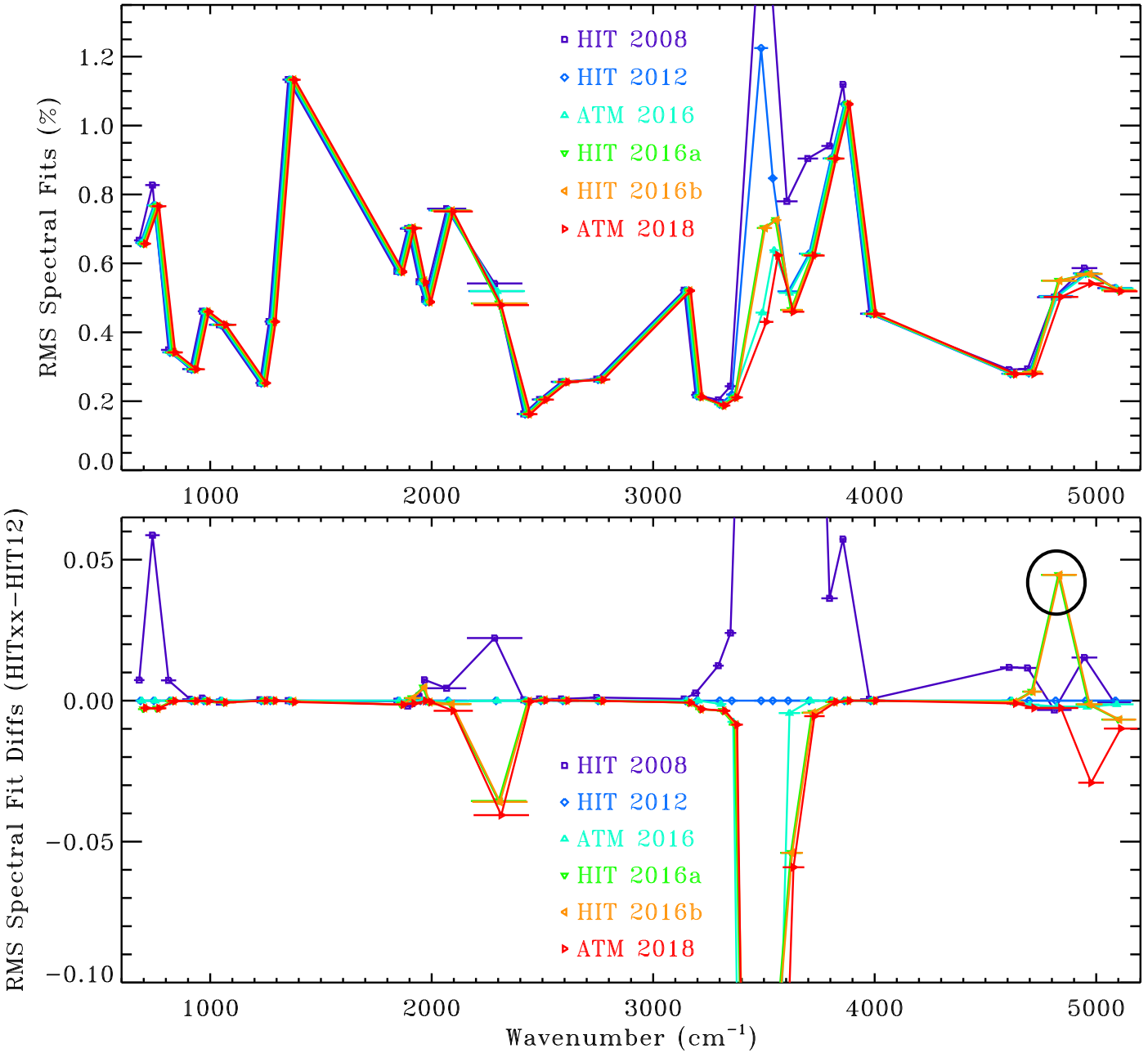
Bottom Panel: Differences from HIT12.

Difference between HIT16a and HIT16b are tiny because isotopologs 11 & 12 are not discernable in atmospheric spectra and because the rounding of the SBHW values doesn't matter in air-broadened spectra.

HIT16 shows improvements over HIT12 below 900 cm^{-1}, and in windows centered at 2290, 3496, 3548, 3623, 4962 and 5096 cm^{-1}.

In the 4825 cm^{-1} window, used by OCO & GOSAT, the HIT16 linelist achieves the worst fits (circled) and HIT08 the best, as for lab spectra.

The ATM18 linelist is always best, or close to.



MkIV Balloon: Retrieved CO_2 VMR Scale Factors

Plotting the VSF values tabulated in the previous slide, along with their error bars.

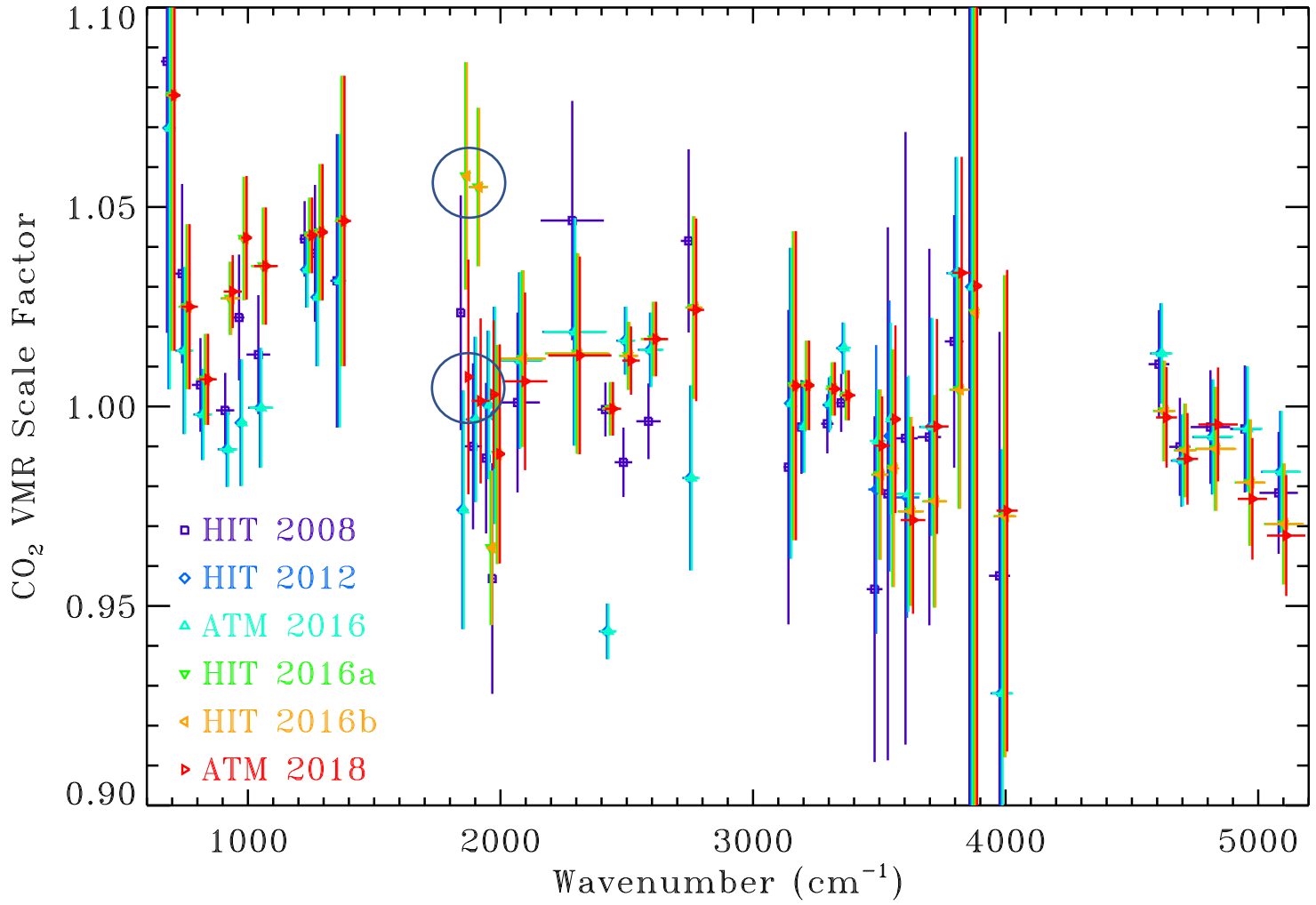
HIT12 points mostly buried under the ATM16 points, except around 3600 cm^{-1}.

HIT16a points mostly buried beneath HIT16b points. X-values offset for clarity.

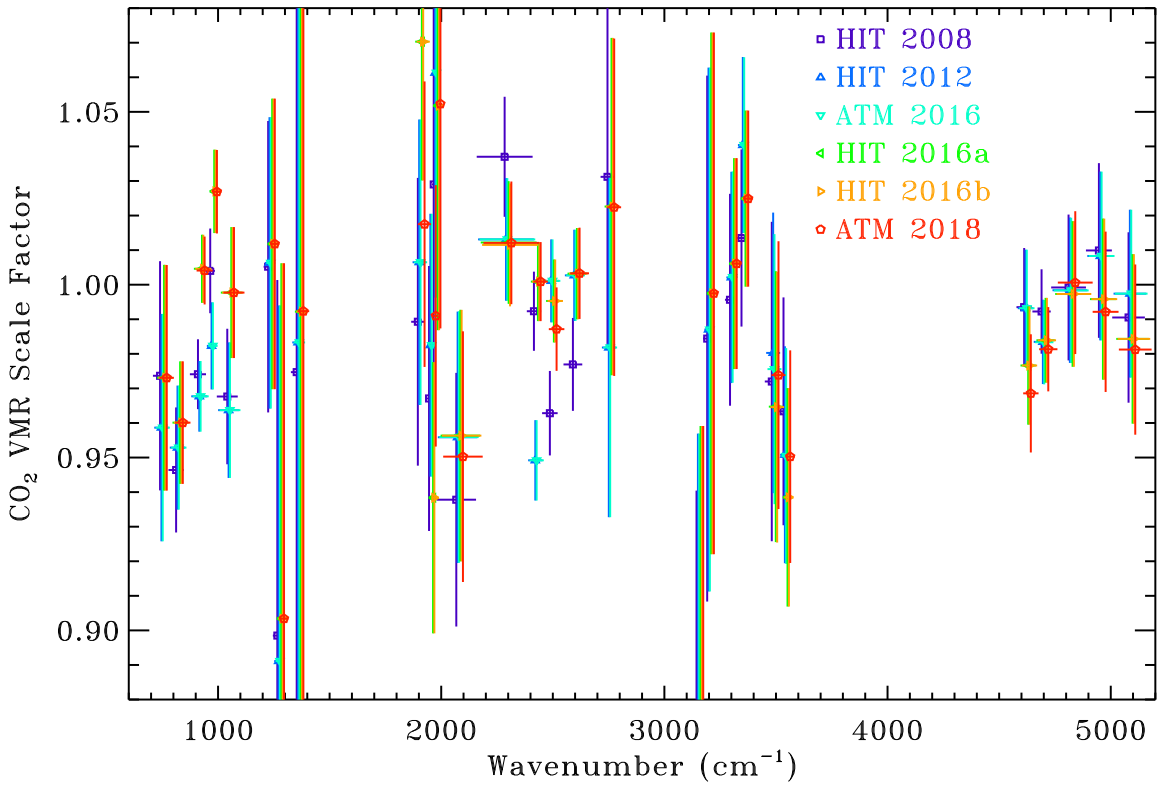
MkIV instrument records 600-5650 cm^{-1} simultaneously, so derived VSFs should have good window-to-window consistency.

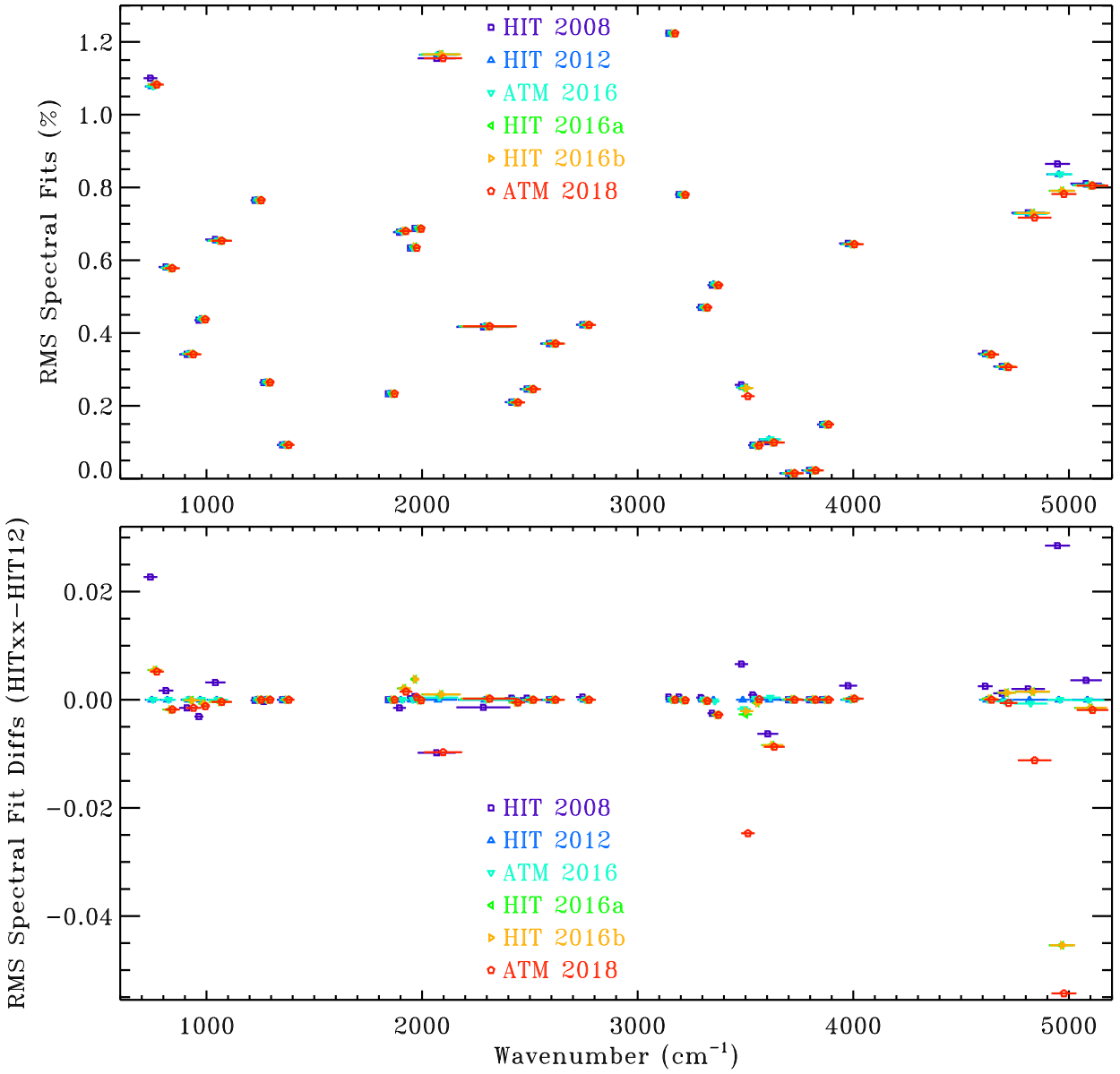
Upper circle indicate anomalously high HIT16 values in the 1800-2000 cm^{-1} region, which were also seen in lab spectra. Lower circle shows pre-2016 HITRAN & ATM18.

Somewhat high (~1.04) VSFs for all linelists are seen in 1200-1400 cm^{-1} region containing the v_1 band (symmetric stretch) of the ^{17}O and ^{18}O isotopologs. These are not definitive enough to warrant fixing.



MkIV Ground-based: RMS Fitting residuals (right) VMR Scale Factors (below)

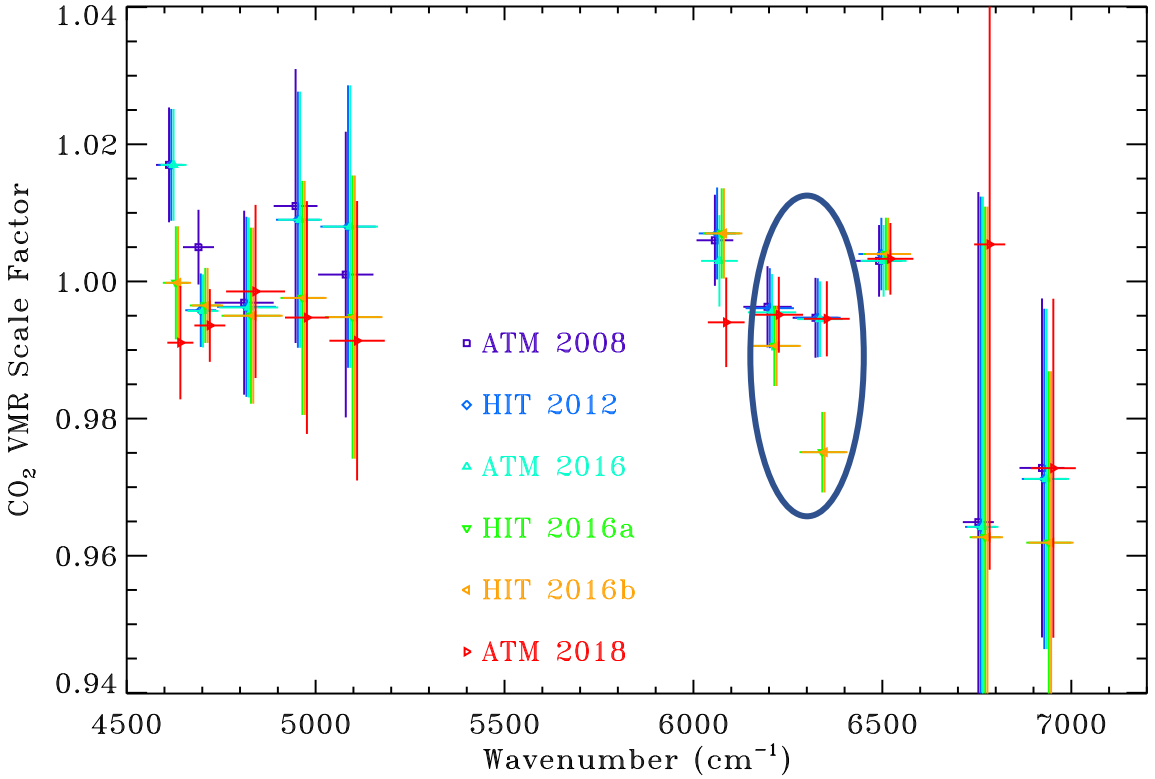


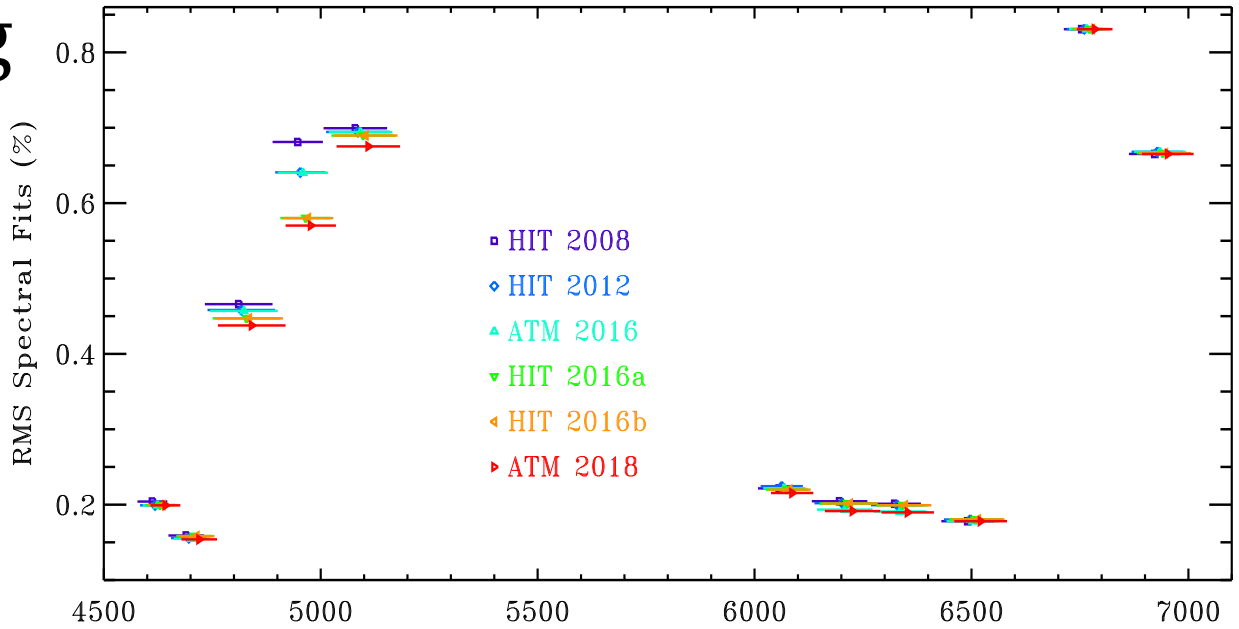


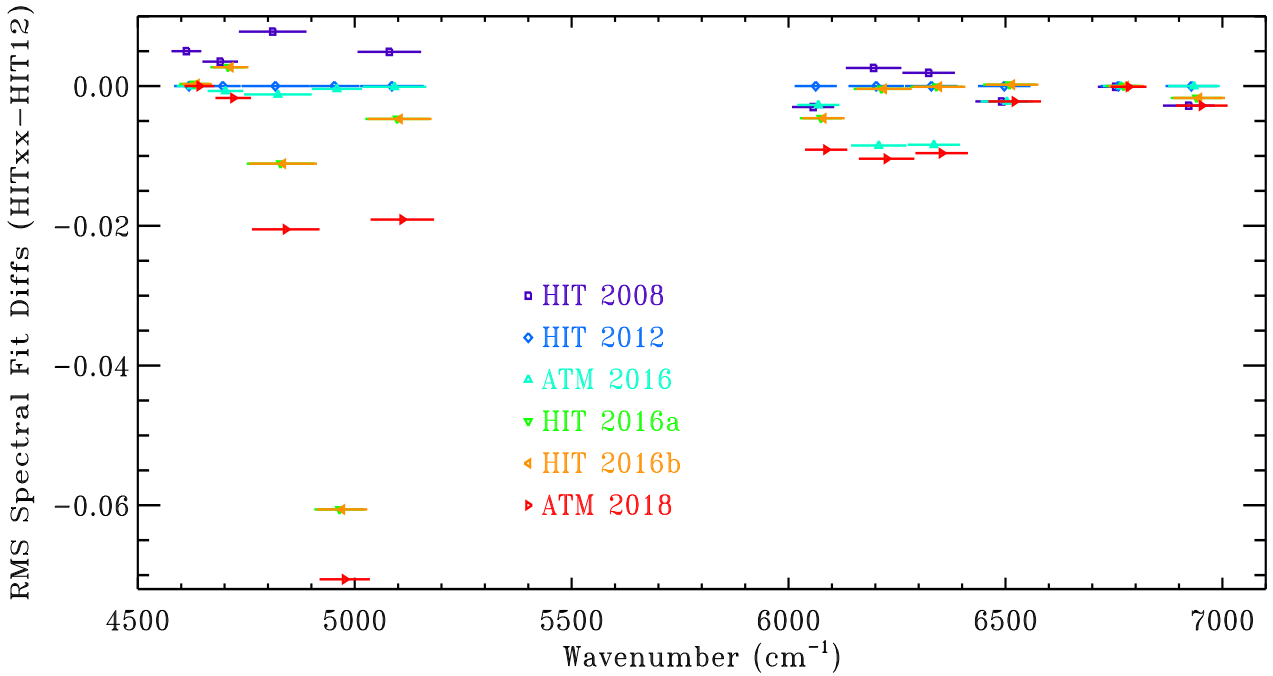
The MkIV ground VSFs (above) are generally close to 1. For the 3150 cm^{-1} window, however, the VSF values are around 0.8 for all linelists and therefore off the bottom of the plot. But the tops of the error bars are visible, not quite reaching 1. In this window, there is strong absorption from H_2O and CH_4 which dwarfs that of CO_2.

TCCON Ground-Based RMS fitting residuals (right) & VSFs (below)

TCCON CO_2 windows circled below. The HIT16 linelists reduce the CO_2 retrieved from the 6221 and 6338 cm^{-1} windows by 0.5% and 1.5% respectively, as compared with the other linelists, introducing a new 1.5% inconsistency.







Usefulness of Spectra for Spectroscopy Evaluation

Type	Pros	Cons
Laboratory	• Well-known cell conditions (Leng, T, P, VMR) • VMR up to 1 are possible, testing SBHWs • Large isotopic enrichments possible	• Dim source, so narrow spectral coverage or poor SNR • Isotopic composition often uncertain • Mostly at room-T
Occultation MkIV Balloon	• Bright source (sun) allows simultaneous coverage 650-5650 cm^{-1} at high resolution • Colder temperatures (210-250K) • Wide range of slant columns • Solar and instrumental features removed • Long path lengths (~400 km)	• Inhomogenous atmospheric path • No control over P/T or VMR (400 ppm) • Interferences from other gases • CO_2 used to determine tangent altitude so no info on absolute CO_2 amounts
Ground-based MkIV / TCCON	• Bright Source (sun) • Broad simultaneous coverage • Long path lengths (~100 km) • Sensitive to lineshape (e.g. width, shifts, LM) • Accurate knowledge of airmass	• Inhomogeneous atmospheric path • No Control over P/T or VMR (400 ppm) • Wide regions blacked out: - H_2O (1350-1900; 3350-4000 cm^{-1}) - CO_2 (650-700; 2280-2390 cm^{-1})

Atmospheric spectra have better-known isotopic composition than lab spectra, unless the lab samples have been independently measured, e.g., by mass spectrometry. For example, for atmospheric CO_2, the $^{13}C/^{12}C$ ratio can be predicted anywhere to 0.1%. Atmospheric spectra contain no information on the SBHW.

In ground-based geometry, airmass is known to 0.1%, given a surface pressure measurement of 1 mbar accuracy.

Ad Hoc Linelist Corrections

After selecting lines from the best predecessor linelist for each spectra region, there would typically still be some large spectral fitting residuals. Sometimes these had an obvious cause (e.g. line position errors, pressure shifts).

Even though these defects were usually discovered in fits to atmospheric spectra (because I look at more atmospheric spectra than lab fits), their correction was always performed while fitting lab spectra. But the atmospheric spectra still play a role in deciding which fitting residuals to investigate.

Since the ad hoc correction process is highly error-prone, it is important that the spectra be refitted after the corrections have been made to ensure that the expected benefits materialized.

Are RMS fitting residuals a useful metric of linelist quality?

The RMS fitting residuals tell us about the consistency of the spectroscopy within a given window. They don't tell us whether it is right or wrong. For example, if all the CO_2 line intensities within a given window were 50% too high, we would still get a good spectral fit, after GFIT scales the assumed CO_2 vmr. Similarly, if all the CO_2 line positions within a given window were in shifted by 0.1 cm^{-1}, we would again get a good spectral fit because the GFIT code retrieves a frequency shift. In fact, these retrieved frequency stretches are a useful diagnostic, and are looked at but not reported here. Of course, if there are multiple gases present in a sample (e.g. CO_2 and H_2O) and the CO_2 lines have a position errors, but not the H_2O, then the position inconsistency will cause an increase in the RMS residuals because GFIT retrieves a single shift for the entire window, not one per gas.

Retrieved CO_2 VSFs: All 4 datasets

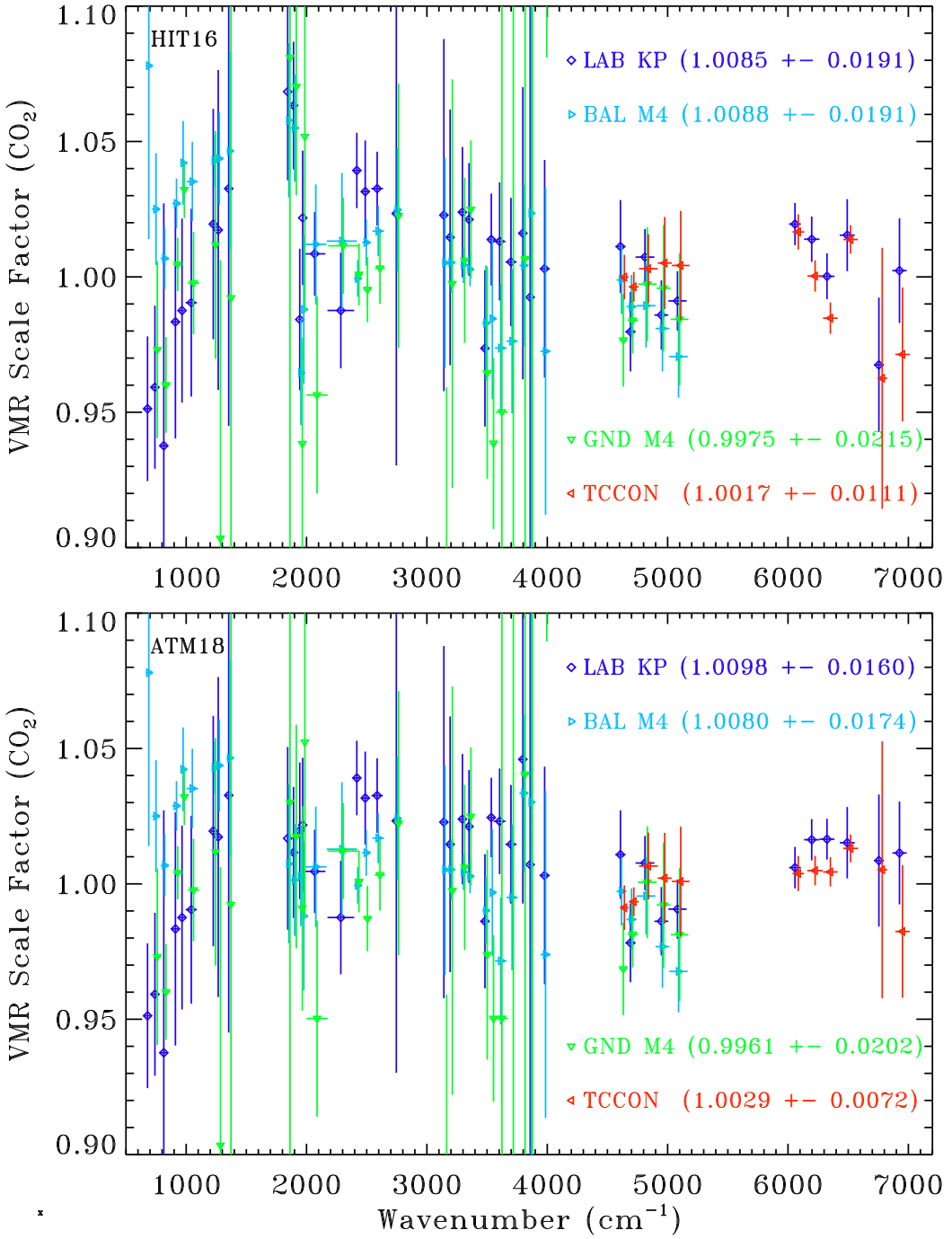
Plot show CO_2 VSFs for the HIT16b linelist (top panel) and the ATM18 linelist (bottom panel). The data are color-coded by the measurement data type (not the linelist). The ATM18 has better window-to-window consistency for all four measurement types, due mainly to the explicit adjustments made in the 1800-1993 cm^{-1} and 6720-6800 cm^{-1} regions.

In general, the error bars overlap between the four different datasets, in terms of the bias in the retrieved CO_2 amounts.

Lab data cover the entire wavenumber range. The M4_GND and M4_BAL datasets cover 700 to 5500 cm^{-1}. The TCCON covers 4000+ cm^{-1}. Only in the 4000 to 5600 cm^{-1} interval do all four datasets overlap.

MkIV GND VSFs are generally lower than those from the other datasets. Below 1200 cm^{-1} the lab data and MkIV GND show VSFs below 1.0, but the MkIV BAL shows values above 1.0 and with smaller error bars.

MkIV ground also have the worst window-to window consistency (2.02%). TCCON has the best (0.72%) but only covers well-behaved, unsaturated windows.



Summary & Conclusions

Four spectral datasets (Kitt Peak lab, MkIV balloon, MKIV ground, and TCCON ground) have been used to evaluate six different CO_2 linelists (HIT08/12/16a,b and ATM 16/18) over 670 to 7000 cm^{-1}.

Spectral fitting was performed with the GFIT code using a Voigt lineshape. The linelists were evaluated in terms of: the rms fitting residuals; and the window-to-window consistency of the retrieved gas amounts. There was no analysis of separate isotopologs. They were all lumped together as CO_2 which makes it important to know the fractionation. Analyzing the twelve CO_2 isotopologs separately is beyond the scope of this work.

RMS Spectral Fitting Residuals

Results show progressive overall RMS fit improvements in each HITRAN version, but there have been some regions where the HITRAN 2016 fits have regressed. For example, in the 4825 cm^{-1} window used by OCO-2 and GOSAT, HIT16 produces the worst fits to lab and MkIV balloon spectra (low-P) but the best fits to ground-based spectra (high-P), suggesting the positions and/or relative intensities in HIT16 are worse than predecessors, but that the widths/shifts are better.

Window-to-Window Consistency of Retrieved CO_2 Amounts

Retrieved CO_2 in the 1900 cm^{-1} region with HITRAN 2016 is biased 5% larger than in the 2050 cm^{-1} region, as pointed out by Chris Boone from ACE data. In previous HITRAN version the 1900 cm^{-1} window produced no significant bias. This problem was fixed for ATM18.

In the 6200-6400 cm^{-1} region used by TCCON the consistency of the retrievals between the 6220 and 6338 cm^{-1} bands has degraded from better than 0.1% to 1.5%. This is a serious problem because TCCON performs an weighted average of the CO_2 retrieved from these two windows. With the existence of a bias, anything that affects the uncertainties of one window relative to the other will perturb the weighted average.

ATM18 CO_2 Linelist

A new linelist (ATM18) was generated, based primarily on HITRAN 2016b, except for:

- Replacing the 3419 - 3923 cm^{-1} and 5750 - 6598 cm^{-1} sections with ATM16.
- Replacing the 6715 - 7000 cm^{-1} section with HIT08
- For isotopologues 10,11,12, using HIT16b throughout.
- Scaling all CO_2 line intensities in the 1800 to 1993 cm^{-1} interval by 1.05
- Scaling all $^{12}CO_2$ line intensities in the 6720 to 6800 cm^{-1} interval by 0.96
- Scaling all $^{12}CO_2$ line widths by 0.99 over 960 to 1000 cm^{-1}

On top of this, ad hoc corrections (mainly position adjustments) were applied, where beneficial.

There are very few windows where the ATM18 linelist doesn't produce the best (or equal best) RMS fits. In all four datasets, the ATM18 linelist produces the best average rms fits. In 2/4 datasets the ATM18 linelist produces the best window-to-window consistency in retrieved CO_2 amount, the exceptions being that the **HIT08** linelist produces the best consistency for the MkIV balloon and TCCON ground datasets.

The main weakness of this evaluation is that there were very few low-temperature lab measurements. So the selection of predecessors lines for inclusion into ATM18 might be different with more low-T lab spectra. In future, obtain additional low-T lab spectra (air-broadened) to test T-dependence of ABHW.

I recommend the ATM18 linelist for use by the NDACC and TCCON FTIR networks.

H_2O Spectroscopy Evaluation 700-12,000 cm^{-1}

Geoff Toon
Jet Propulsion Laboratory
California Institute of Technology

Water vapor is extremely variable in the atmosphere. So multiple windows needed to encompass its dynamic range. Water vapor is also major interferent in retrievals of other gases. Hence the need for consistent and accurate spectroscopy.

Four different H_2O linelists were evaluated: HITRAN08, HITRAN12, ATM16, and HITRAN16.

This is done by fitting lab spectra (mainly Kitt Peak), MKIV balloon spectra, MkIV ground-based spectra, and TCCON ground-based spectra.

Generating a fifth linelist, ATM18, by "cherry-picking" from the predecessor linelists and by correcting obvious problems.

HITRAN 16 has 7/9 possible water vapor isotopologs, all except $D_2{}^{18}O$ and $D_2{}^{17}O$. Although D_2O can never been seen in the Earth's atmosphere, it can be seen in D-enriched lab spectra.

The H_2O Linelists

HIT08: 69,201 lines
54,177 H_2O lines covering 0 to 25,232 cm^{-1}
15,024 HDO lines covering 0 to 22,708 cm^{-1}

HIT12: 224,515 lines
209,492 H_2O lines covering 0 to 25,710 cm^{-1}
15,023 HDO lines covering 0 to 22,708 cm-1

ATM16:169,134 lines
147,636 H_2O lines covering 0 to 25,711 cm^{-1}
18,713 HDO lines covering 0 to 22,708 cm^{-1}
2,785 D_2O lines covering 2198 to 4255 cm^{-1}

HIT16: 304,225 lines
207.277 H_2O lines covering 0 to 25,711 cm^{-1}
73.460 HDO lines covering 0 to 19,935 cm^{-1}
23.488 D_2O lines covering 0 to 12,797 cm^{-1}

ATM18
Under construction. Based on previous linelists with empirical adjustments.

H_2O is represented by isotopologs 1-3, HDO by isotopologs 4-6, and D_2O by isotopologs 7-9.

D_2O cannot be seen in the Earth's atmosphere. The main purpose of the D_2O linelist is to facilitate analysis of highly D-enriched lab HDO spectra, in which D_2O absorptions can be strong. HIT16 includes a linelist for $D_2{}^{16}O$ (isotopolog #7) for the first time. Bob Toth had one in 2006 that never got into HITRAN, but was included in ATM16.

Whereas the HDO linelist in HIT08 and HIT12 extends to 22,708 cm^{-1}, in HIT16 it extends only to 19,935 cm^{-1}, despite containing 4x more lines.

The Fitted Spectra

Laboratory (650-11,000 cm^{-1})

143 from Kitt Peak (1983-1996)

- 58 of which are D-enriched
- 7 of which are ^{18}O-enriched
- 3 of which are ^{17}O-enriched

13 from Manfred Birk (2014-2015)

MkIV Balloon-borne (650-5650 cm^{-1})

38 spectral pairs (HgCdTe & InSb) covering 9 to 38 km altitude

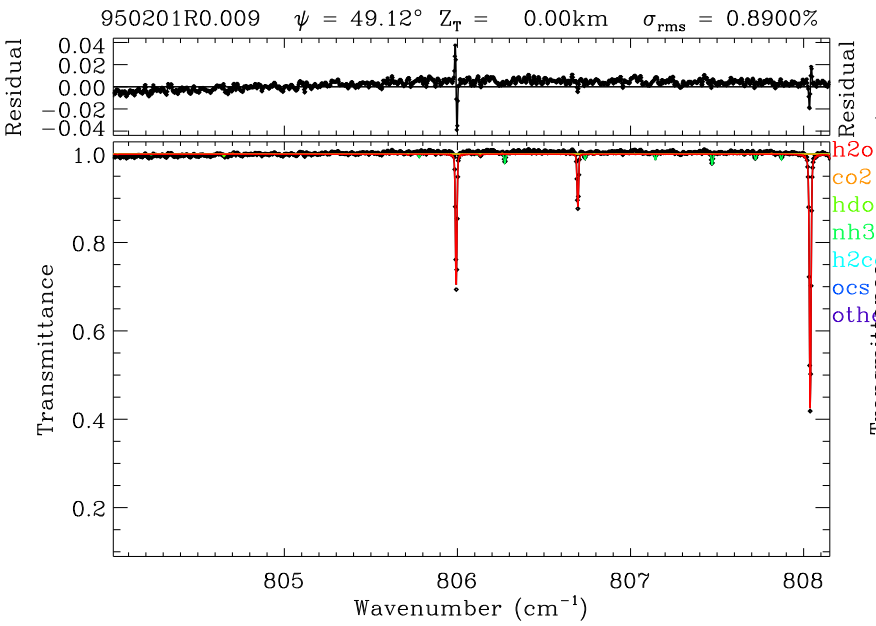
MkIV Ground-based (650-5650 cm^{-1})

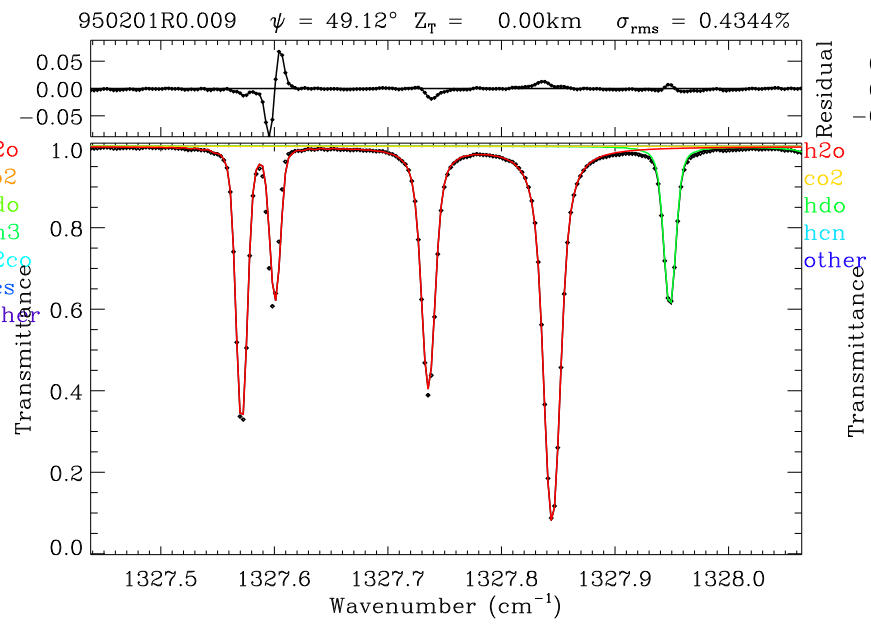
124 spectral pairs (HgCdTe & InSb) covering 0 to 3.8 km altitude.

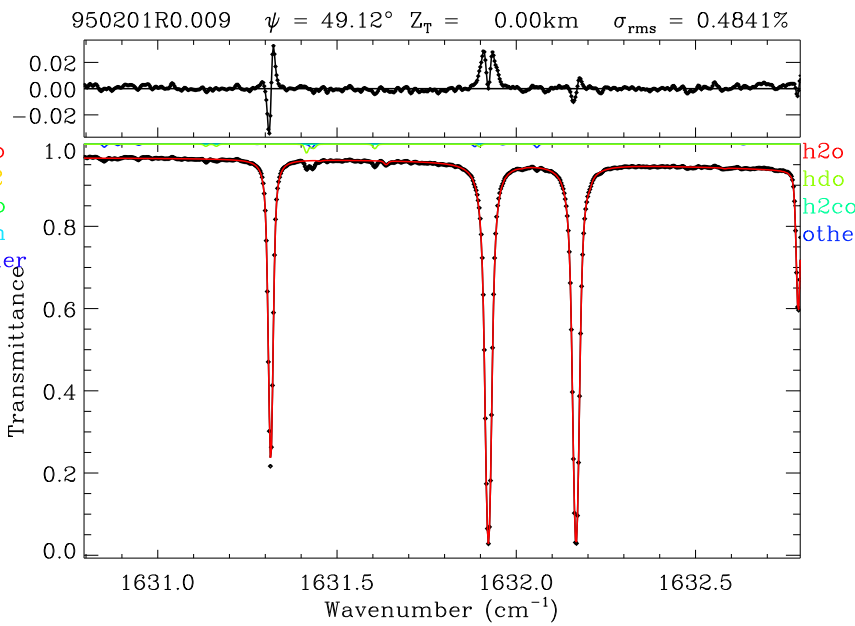
TCCON Ground-based (4000-15,500 cm^{-1})

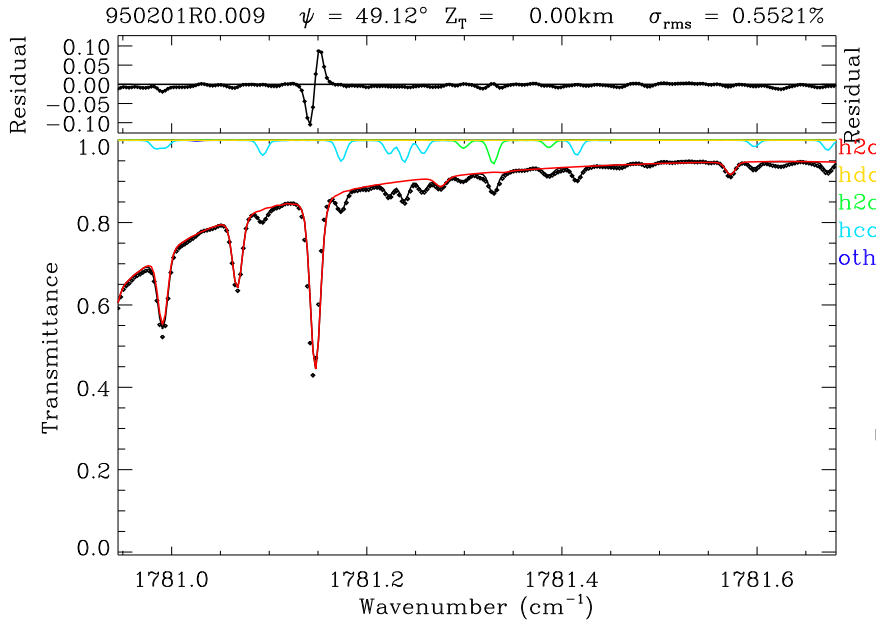
26 spectral pairs (InGaAs & Si)

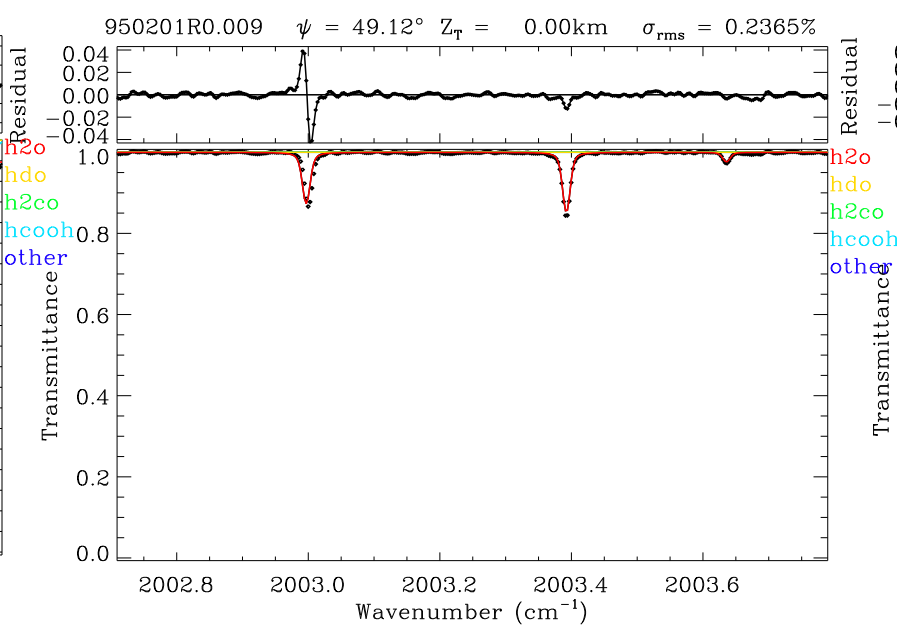
Examples of Line position error in HIT16 linelist in Kitt Peak lab spectra

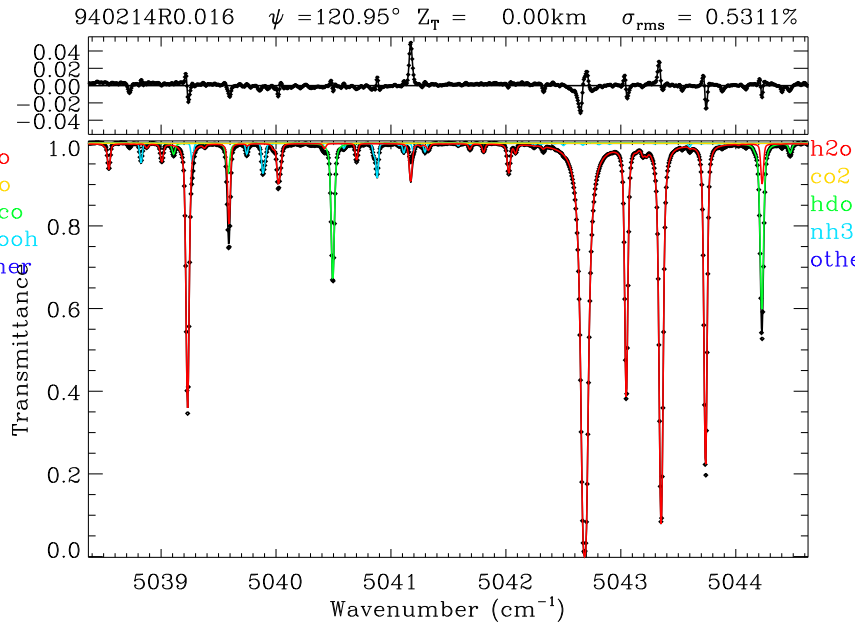












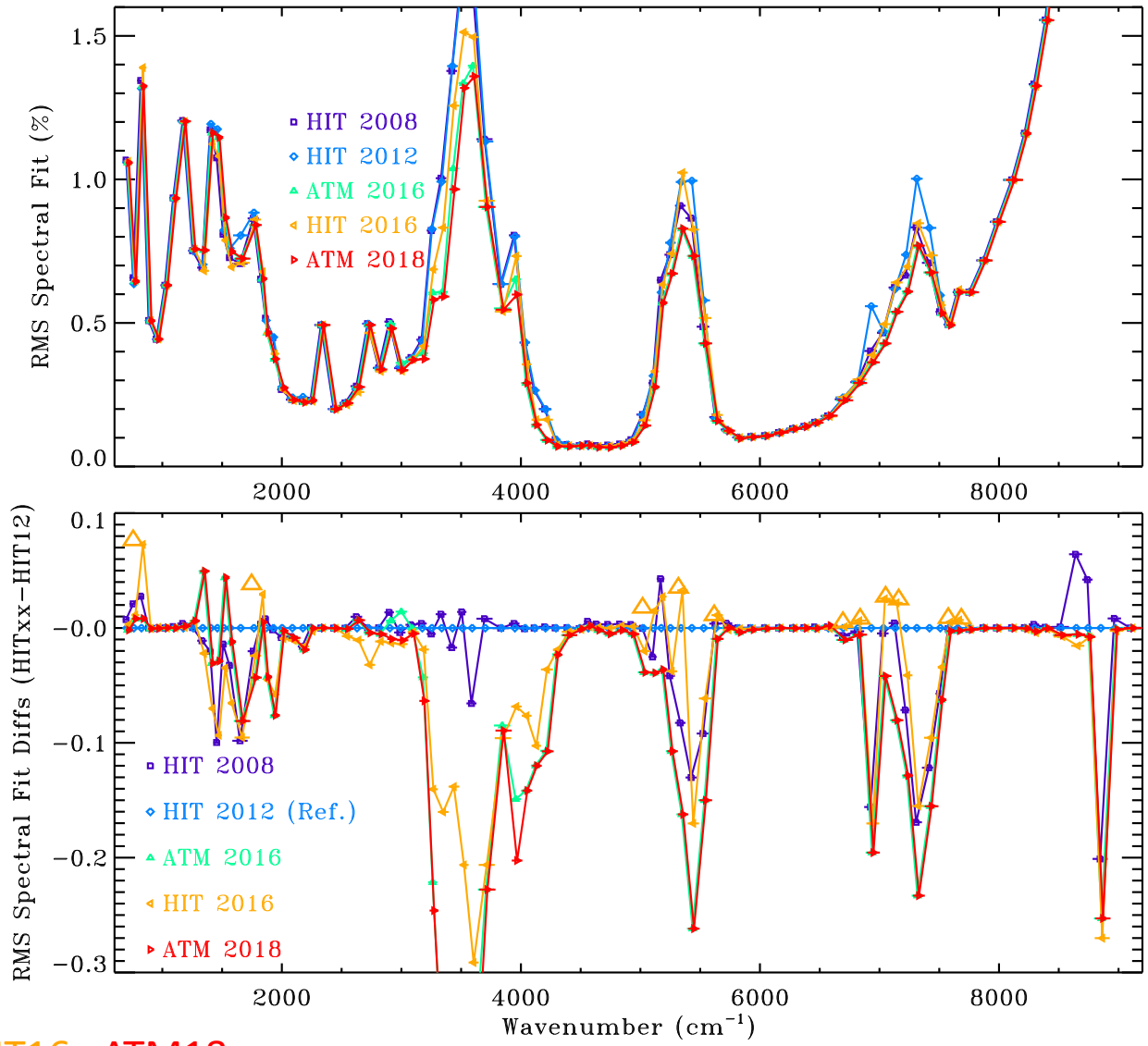
Kitt Peak Lab Spectra – RMS Fitting Residuals

Top panel shows the absolute RMS Spectral fitting residuals. These tend to be large in regions with strong H_2O absorption and smaller in regions of weak absorption. Above 6000 cm^{-1} the Kitt Peak lab spectra that I have available start to fall in signal, so the RMS expressed in gas transmittance, increases.

Lower panel shows the difference in RMS relative to HIT12. In general the ATM16 and ATM18 linelists give the best fits. This situation is partly a result of ad hoc adjustments having been made to these linelists in the past to improve their fits to a large subset of these same lab spectra -- an unfair advantage. Hence the need to look at other spectra.

Overall, HIT12 is the worst linelist. HIT08 was substantially better than HIT12 in several regions (1400-1900, 2000-2200, 5300-5600, 7200-7500 cm^{-1}. HIT12 was better than HIT08 at 600-800 and 8500-8800 cm^{-1}.

HIT16 is the worst of all linelists in 11 windows (depicted by orange triangles in lower panel), mostly above 5000 cm^{-1}.



Linelist	HIT08	HIT12	ATM16	HIT16	ATM18
Average % RMS over all windows:	1.3323	1.3461	1.2968	1.3161	1.2939

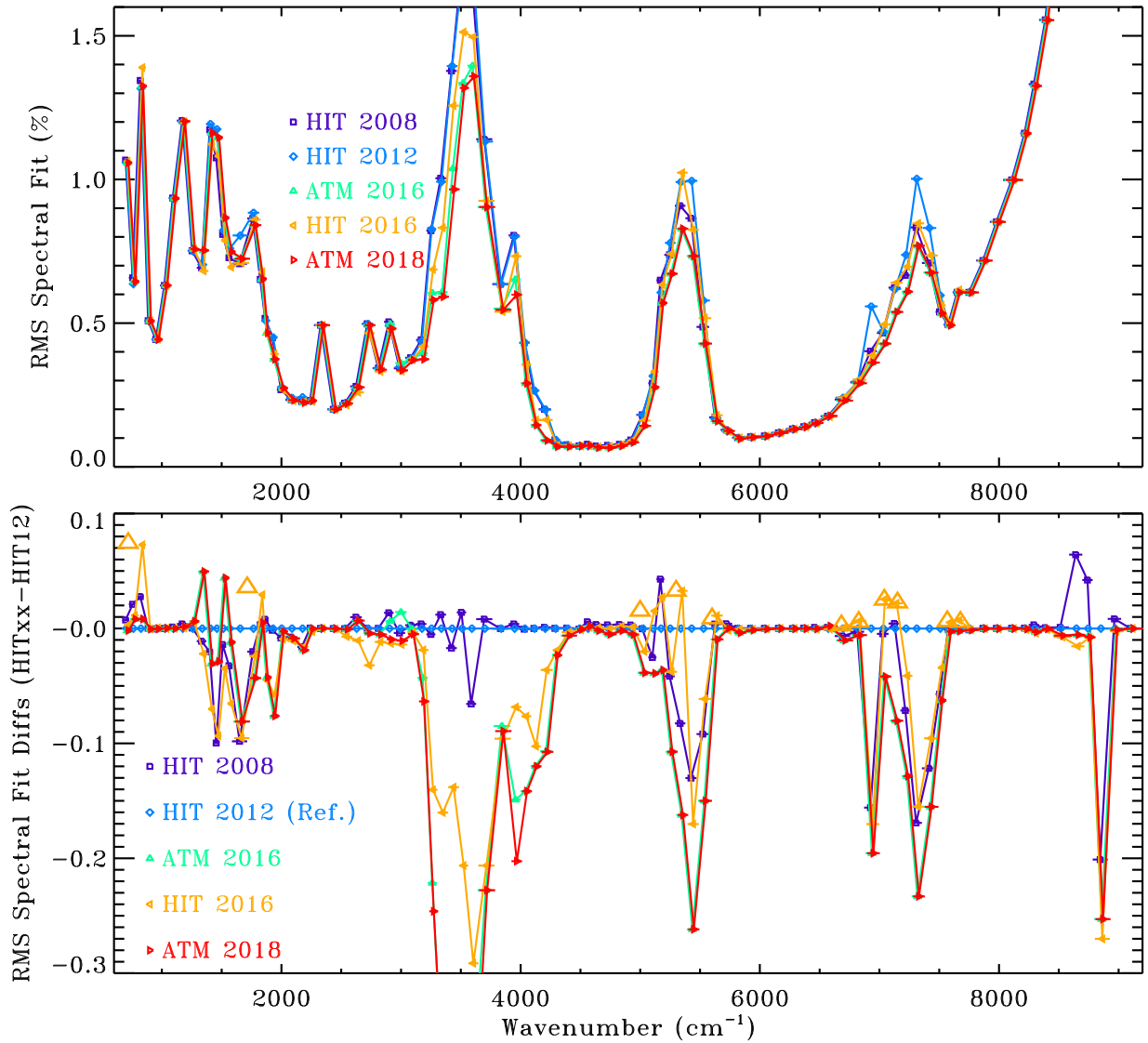
Kitt Peak Lab Spectra – RMS Fitting Residuals

Top panel shows the absolute RMS Spectral fitting residuals. These tend to be large in regions with strong H_2O absorption and smaller in regions of weak absorption. Above 6000 cm^{-1} the Kitt Peak lab spectra that I have available start to fall in signal, so the RMS expressed in gas transmittance, increases.

Lower panel shows the difference in RMS relative to HIT12. In general the ATM16 and ATM18 linelists give the best fits. This situation is partly a result of ad hoc adjustments having been made to these linelists in the past to improve their fits to a large subset of these same lab spectra -- an unfair advantage. Hence the need to look at other spectra.

Overall, HIT12 is the worst linelist. HIT08 was substantially better than HIT12 in several regions (1400-1900, 2000-2200, 5300-5600, 7200-7500 cm^{-1}. HIT12 was better than HIT08 at 600-800 and 8500-8800 cm^{-1}.

HIT16 is the worst of all linelists in 11 windows (depicted by orange triangles in lower panel), mostly above 5000 cm^{-1}.



Linelist	HIT08	HIT12	ATM16	HIT16	ATM18
Average % RMS over all windows:	1.3323	1.3461	1.2968	1.3161	1.2939

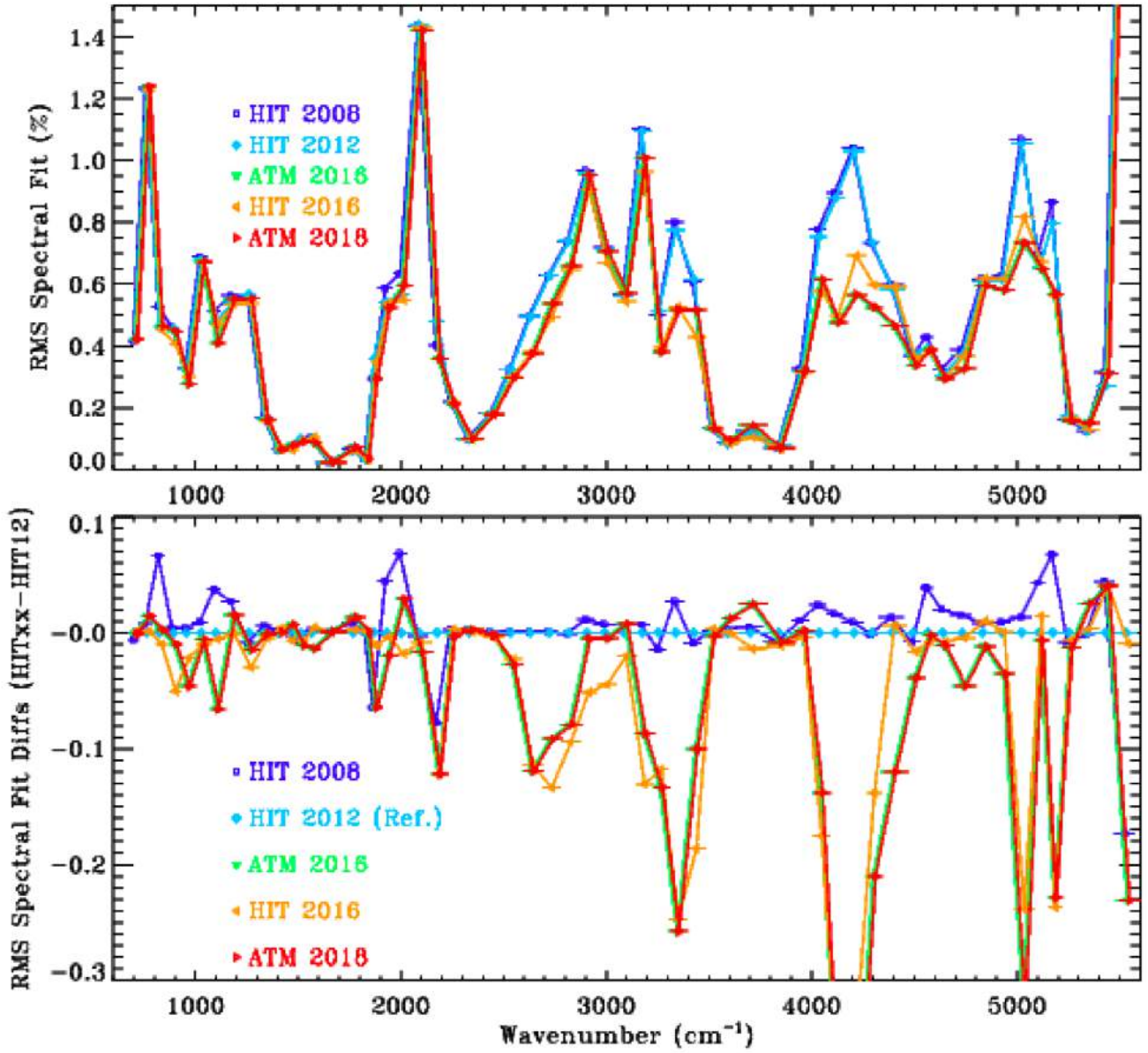
MkIV ground-based: RMS Fitting residuals

Residuals tend to be small in blacked out regions: by strong H_2O absorption (e.g. 1300-1900 cm^{-1}; 3700-4000 cm^{-1}), or strong CO_2 absorption (2200-2400 cm^{-1}; 3500-3700 cm^{-1}).

Above 5500 cm^{-1} the MkIV response falls off and so the RMS, expressed in transmittance, increases.

Only 1 windows in which HIT16 is worst, and many where it is best (e.g. 2600-3200 cm^{-1} – HDO?)

HIT12 slightly better than HIT08



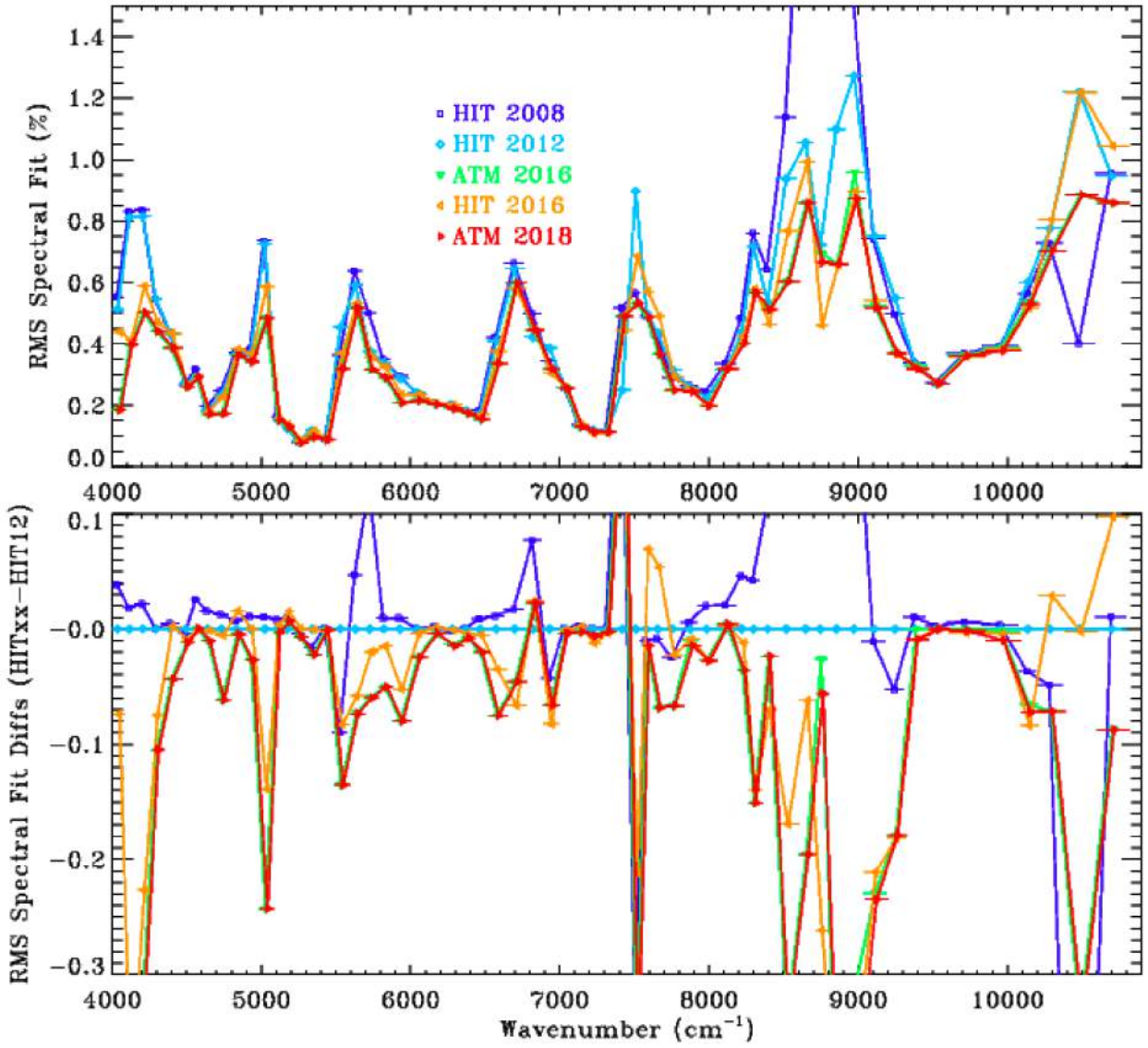
Linelist	HIT08	HIT12	ATM16	HIT16	ATM18
Average % RMS over all windows:	0.5213	0.5161	0.4585	0.4658	0.4585

TCCON Ground-based - RMS Fitting Residuals

Overall HIT08 is the worst linelist, except at 10,500 cm^{-1} where is is by far the best.

HIT12 is the best linelist at 6950 cm^{-1}

HIT16 is the worst linelist at 4850, 5200, 7600-7700,10200-10800 cm-1. HIT16 is the best linelist at 6950, 8750 and 10,100 cm^{-1}.



Linelist	HIT08	HIT12	ATM16	HIT16	ATM18
Average % RMS over all windows:	0.5034	0.4567	0.3743	0.4049	0.3718

Water Spectroscopy: Summary and Conclusions

HIT16 contains much weaker HDO lines than HIT12. So for humid conditions where weak lines are observable, it tends to do better than previous linelists, which are missing those weak lines. The stronger lines In HIT16 are empirically-based and therefore haven't changed much since HIT2012

The ranking of the linelists in terms of RMS and VSF varies substantially from window to window and from dataset to dataset. Lab spectra are generally at lower pressure than atmospheric and so the results are sensitive to line strengths, positions and, for pure gas samples, the SBHW. Ground-based observations are relatively more sensitive to lineshape (ABHW, pressure shifts, and Line Mixing) and to the weaker lines. Balloon measurements encounter low temperatures and provide a good validation of the T-dependent spectroscopic parameters of the stronger lines.

Over the 2500-3000 cm^{-1} region HIT16 provides the best RMS fits (improved HDO) in ground-based spectra. Above 4000 cm-1 HIT16 is worse than existing ATM linelists.

For D_2O, necessary to analyze D-enriched lab spectra of HDO, HIT16 is also an improvement.