Evaluation of HITRAN 2016 O_3 linelist

Geoff Toon
Jet Propulsion Laboratory
2017-09-20

Recap of earlier work.

A candidate linelist, 03_hit16_beta.par (hit16), generated by the Reims group, was received from Iouli Gordon on Oct 27, 2016.

This was empirically modified in Jan 2017 by Geoff Toon based on fits to Kitt Peak lab spectra to create a new linelist named 03_hit16_beta_gct (hit16a). I wrote a report that compared these linelists with their predecessors (HIT08, HIT12, ATM16) in terms of the quality of the spectral fits to 194 Kitt Peak lab spectra and 38 MkIV balloon spectra, and in terms of the consistency of retrieved O_3 amounts from different windows:
http://mark4sun.jpl.nasa.gov/report/o3_spectroscopy_evaluation_20170117.pdf

In May/June 2017, an improved O_3 linelist was generated by the Reims group. And additional lab spectra have been made available. And I increased the number of fitted windows from 20 to 35. All of which seemed sufficient to warrant a further evaluation.

Recent Progress

Increased the number of fitted windows from 20 to 35. This was accomplished mainly by splitting the existing windows, but also by extending the analysis into regions of weak lines, previously ignored.

Received two new lab spectra covering the 3820-5010 cm^{-1} region, from Alain Barbe (U Reims) in Mar 2017. Deleted 6 previously-used Kitt Peak spectra that had large zero-level offsets and/or other artifacts. This brings the total number of lab spectra to 194-6+2 = 190.

Another candidate linelist was received from the Reims group (via Iouli) in June 2017, under the name 03_hit16.par (hit16b). This was better than the Oct 2016 Reims linelist (hit16) in most regions. But in general it was not as good as the empirically adjusted linelist (hit16a) in terms of fitting KP lab spectra.

So now we had two linelists (hit16a, hit16b) from the same common ancestor (hit16). To capture as many of the improvements as possible, I did a 3-way merge (using the kdiff3 tool). This identified 226 unresolved conflicts (the three files were all different) and 5316 automatically-resolved differences (2/3 files were identical). These 226 unresolved conflicts were all decided in favor of hit16b.

Developed a program to automatically adjust the position of lines in a linelist based on fits to a series of lab spectra. This was applied to the hit16b linelist. In most windows this improved the spectral fits. In a few windows it made the fits worse, and so in these cases the update was discarded. This work resulted in linelist hit16c. Wavenumber mis-calibration of some lab spectra is the likely reason for worsening spectral fits in some windows.

Finally, it was discovered that some residuals simply could not be eliminated by position adjustments only. It was necessary to adjust line intensities of some hot bands to obtain fits as good as HITRAN 2012 in certain regions. This resulted in the linelist hit16d.

Linelist Summary

The table below summarizes the attributes of the 7 linelists that were evaluated for this report.

In the evaluation, the non-O_3 gases were modeled using the ATM16 linelist, a greatest hits compilation based on earlier HITRAN versions. ATM16 was selected because it gave the best fits to MkIV balloon spectra (HITRAN 16 not yet being available when I did this evaluation), which allows the O_3-related difference to show up more clearly. So the only difference between the fits is the O_3 linelist.

I believe that the hit16b linelist was released as HITRAN 2016. But since that two improved linelists have been generated: hit16c and hit16d.

Linelist	File Name	Date	# Lines	Comments
hit08	03hit08.par	2009	409686	HITRAN 2008 O_3 linelist
hit12	03hit12.par	2013	422116	HITRAN 2012 O3 linelist
hit16	03_hitran16_beta.par	Oct 2016	449567	Calculation by Reims group
hit16a	03_hitran16_beta_gct	Jan 2017	449583	Manual position + shift adjustments, 16 new lines
hit16b	03_hit16.par	Jun 2017	449570	Improved calculation from Reims
hit16c	03_hit16c.161	Aug 2017	449589	Automated empirical position mods, 19 new lines
hit16d	03_hit16d.161	Sep 2017	449590	Intensity modifications, 1 new line

hit16d linelist Intensity Adjustments

Having adjusted the line positions, there were still some regions where the fits using the hit16x linelists produced poorer fits than HITRAN 2012, e.g. 1100-1220 & 1970-2040 cm^{-1}.

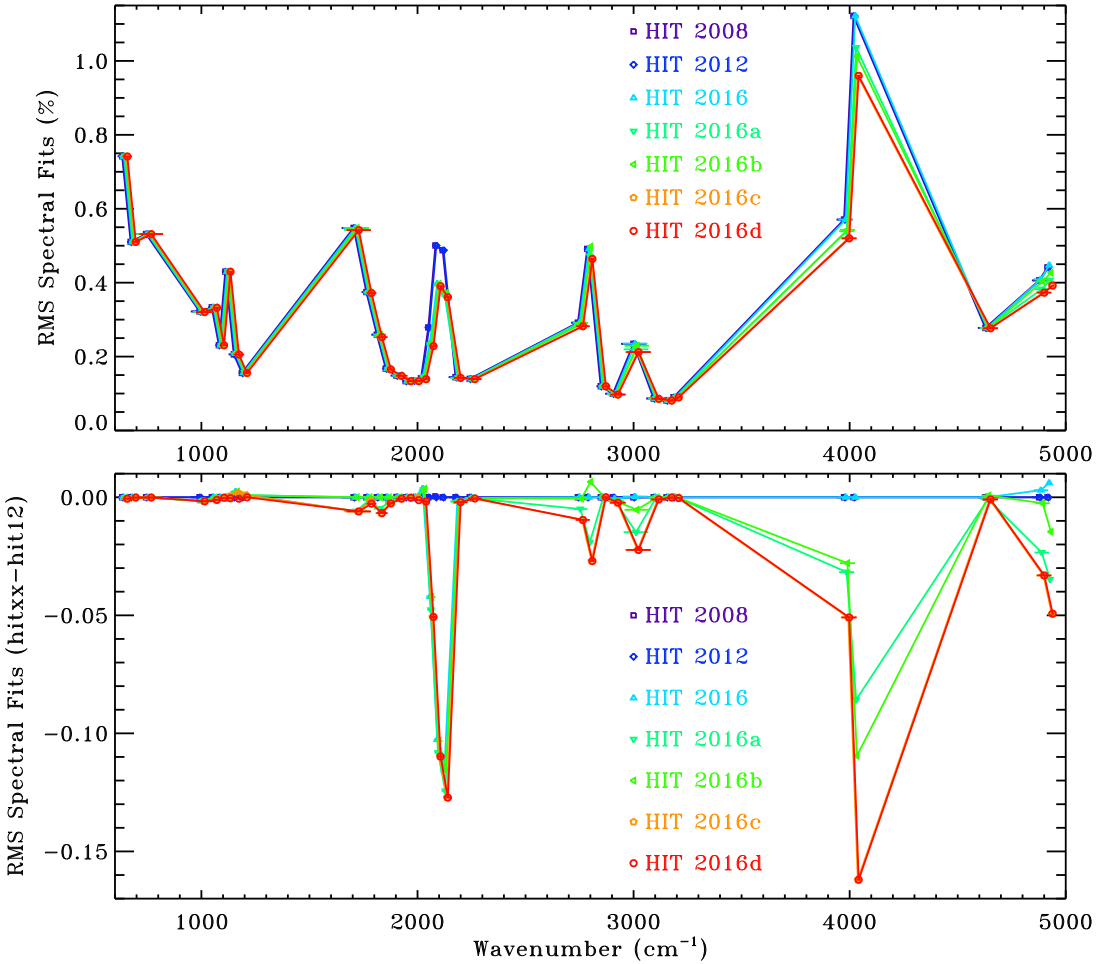
No amount of line position adjustment could remove some of the large positive residuals in the 1970 to 2040 cm^{-1} region. The offending lines belonged to the 1 0 2 ← 0 0 1 vibrational state. The lines were at the right frequencies, they were simply too strong.

It appears that the intensities of these lines were unchanged since HITRAN 2012, whereas the intensities of the overlapping 0 0 2 ← 0 0 0 overtone band had been reduced by 10-15% for improved consistency with the fundamental band at 1042 cm^{-1}. The resulting inconsistency between the intensities of these two overlapping bands at 1970 to 2040 cm^{-1} was driving up the rms residuals, as compared with HITRAN 2012. So I reduced the intensities of the strongest 1 0 2 ← 0 0 1 lines by 10%, resulting in a significant improvement in the rms fitting residuals.

In the 1100 1220 cm^{-1} region the 2 0 0 ← 0 0 1 transitions had been reduced in intensity since HITRAN 2012 by 5 – 25%, causing an inconsistency with overlapping lines that was driving up the fitting residuals. By restoring these intensities back to their 2012 values, this problem was fixed.

Kitt Peak laboratory spectra: RMS fits

Fitted 190 lab spectra of O_3 measured from 1984 to 1998, none of which covered the full 600-5000 cm^{-1} interval. Due to the limited spectral coverage of most lab spectra, only 35% of the potential 35x190 spectral fits could be performed. Lab spectra range from 210 to 300K and from 0.5 to 250 Torr.



Upper Panel: RMS fitting residuals. Their absolute values don't really matter. Their variation from linelist to linelist is more important.

HIT16x produces much improved residuals around 2100 cm^{-1} window.

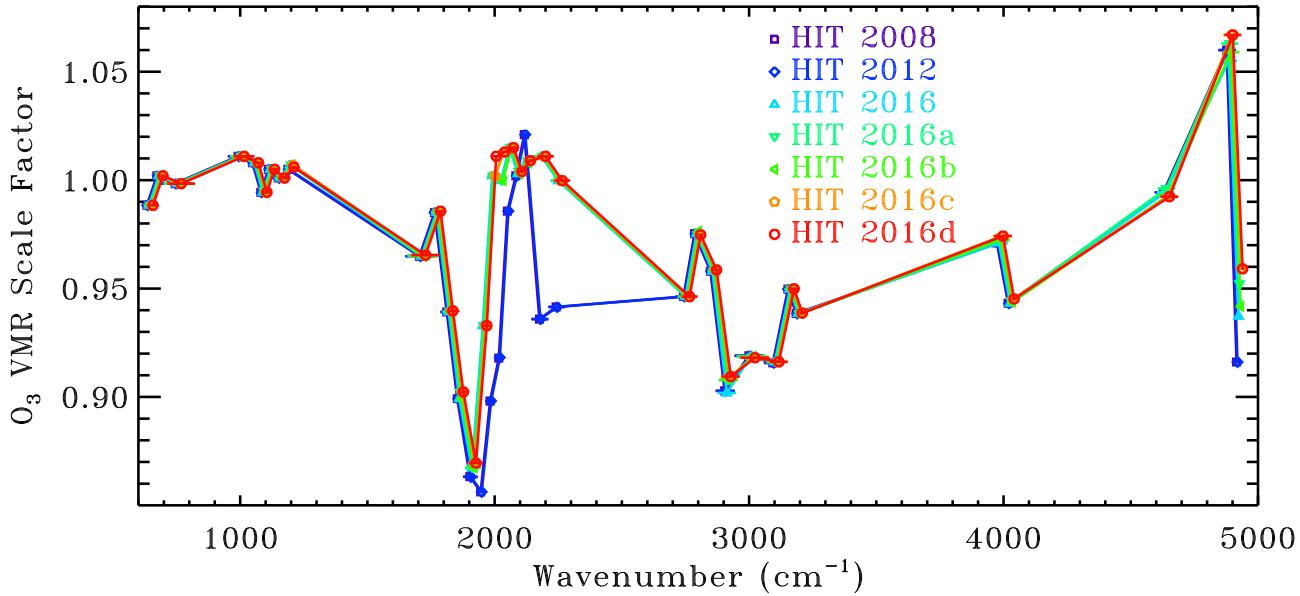
HIT16c,d produce improved residuals around 4000 cm^{-1}.

Lower Panel: RMS fitting residuals relative to HIT2012. Negative values mean that the linelist produces better fits than HIT2012

The hit16d linelist is lowest or equal lowest in all windows. And it should be because the adjustments were performed to make it so.

Kitt Peak laboratory spectra: Retrieved O_3

Used 190 Kitt Peak lab spectra of O_3 measured from 1984 to 1998, although none covered every window. Due to the limited spectral coverage, only 39% of the potential 35x194 spectral fits could be performed. These lab spectra range from 210 to 300K and from 0.5 to 250 Torr.



The VMR Scale Factor represents the retrieved O_3 amount divided by the expected amount, based on nominal cell conditions (Length, T, P, VMR).

In the 2029 and 2180 cm^{-1} windows the HIT16x linelists yield 8% larger retrieved O_3 amounts.

The 1850-1900 cm^{-1} region produces 10-14% low retrieved O_3 amounts (line intensities are too large) for all linelists. In this region the O_3 absorption lines are quite weak.

The 2900 - 3120 cm^{-1} windows also produce 8-9% low O_3 amounts in all linelists.

RMS Spectral fits by window (KP lab)

	Center (cm^{-1})	Width (cm^{-1})	N_{spec}	RMS Spectral Fits (%) HIT08	HIT12	HIT16	HIT16a	HIT16b	HIT16c	HIT16d
1	645.50	21.51	105/190	0.7420	0.7420	0.7419	0.7419	0.7419	0.7414	0.7414
2	684.50	16.51	105/190	0.5103	0.5103	0.5102	0.5102	0.5102	0.5102	0.5102
3	756.00	55.01	105/190	0.5315	0.5315	0.5314	0.5314	0.5314	0.5314	0.5314
4	1003.40	40.51	162/190	0.3221	0.3221	0.3205	0.3205	0.3205	0.3205	0.3204
5	1060.00	16.26	162/190	0.3331	0.3331	0.3331	0.3321	0.3330	0.3320	0.3320
6	1092.86	16.68	162/190	0.2307	0.2307	0.2307	0.2305	0.2307	0.2305	0.2305
7	1122.68	13.14	162/190	0.4299	0.4299	0.4299	0.4295	0.4299	0.4295	0.4295
8	1199.25	11.10	162/190	0.1556	0.1556	0.1565	0.1565	0.1565	0.1565	0.1556
9	1161.90	24.30	162/190	0.2060	0.2060	0.2084	0.2079	0.2084	0.2079	0.2054
10	1716.40	56.02	60/190	0.5482	0.5482	0.5482	0.5423	0.5482	0.5422	0.5422
11	1775.44	22.46	60/190	0.3746	0.3746	0.3746	0.3734	0.3746	0.3732	0.3719
12	1823.85	25.85	60/190	0.2593	0.2593	0.2593	0.2545	0.2593	0.2526	0.2526
13	1865.00	14.70	60/190	0.1677	0.1677	0.1677	0.1653	0.1677	0.1651	0.1651
14	1914.62	27.45	60/190	0.1486	0.1485	0.1485	0.1485	0.1485	0.1479	0.1479
15	1957.12	14.13	60/190	0.1336	0.1336	0.1336	0.1335	0.1336	0.1334	0.1334
16	1994.00	19.50	60/190	0.1346	0.1346	0.1347	0.1344	0.1346	0.1344	0.1334
17	2028.22	13.53	60/190	0.1413	0.1407	0.1444	0.1439	0.1443	0.1389	0.1389
18	2062.00	17.01	60/190	0.2789	0.2789	0.2369	0.2312	0.2366	0.2282	0.2282
19	2095.15	16.20	60/190	0.5006	0.5002	0.3976	0.3920	0.3970	0.3904	0.3904
20	2128.57	17.23	60/190	0.4878	0.4878	0.3724	0.3634	0.3722	0.3606	0.3606
21	2188.00	32.20	60/190	0.1443	0.1443	0.1426	0.1424	0.1425	0.1422	0.1422
22	2253.60	31.40	60/190	0.1396	0.1396	0.1393	0.1392	0.1393	0.1391	0.1391
23	2754.00	31.51	67/190	0.2917	0.2917	0.2817	0.2867	0.2911	0.2821	0.2821
24	2797.50	10.51	67/190	0.4913	0.4912	0.4912	0.4721	0.4977	0.4643	0.4642
25	2860.00	20.01	23/190	0.1197	0.1197	0.1197	0.1197	0.1197	0.1197	0.1197
26	2916.00	36.01	11/190	0.0994	0.0994	0.0988	0.0978	0.0977	0.0971	0.0971
27	3011.00	57.01	11/190	0.2344	0.2344	0.2347	0.2196	0.2291	0.2121	0.2121
28	3105.00	35.01	11/190	0.0865	0.0865	0.0864	0.0864	0.0864	0.0856	0.0856
29	3164.50	22.51	11/190	0.0810	0.0810	0.0810	0.0810	0.0810	0.0809	0.0809
30	3197.00	10.01	11/190	0.0899	0.0899	0.0899	0.0899	0.0899	0.0896	0.0896
31	3985.88	34.88	12/190	0.5708	0.5708	0.5708	0.5391	0.5429	0.5199	0.5199
32	4029.75	8.99	12/190	1.1218	1.1218	1.1218	1.0360	1.0123	0.9598	0.9598
33	4640.00	35.01	5/190	0.2775	0.2775	0.2775	0.2767	0.2792	0.2766	0.2766
34	4888.85	33.35	6/190	0.4062	0.4062	0.4092	0.3827	0.4036	0.3730	0.3730
35	4927.35	5.17	6/190	0.4414	0.4414	0.4475	0.4066	0.4268	0.3921	0.3921

Red values indicate largest RMS. Black values indicate intermediate RMS. Blue values indicate smallest RMS.

Discussion of Lab Results

Empirical adjustment of the line positions of the hit16b linelist was performed to improve rms residuals in fits to Kitt Peak lab spectra. A few intensities were also adjusted. And 36 additional O_3 lines were added. All of this produced significant improvements in the fits to KP lab spectra.

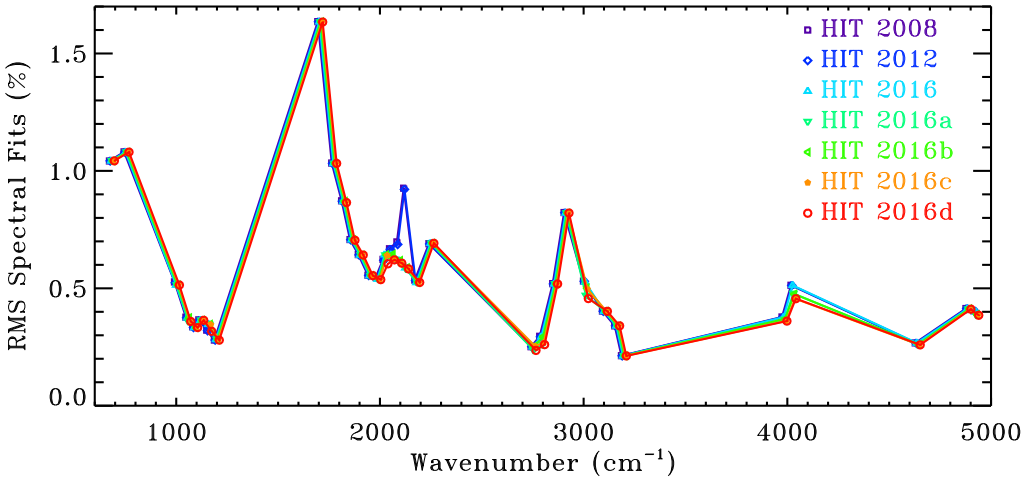
The question is: are these improvements real, or are we simply fitting the noise and artifacts (e.g., contaminants) in the Kitt Peak spectra?

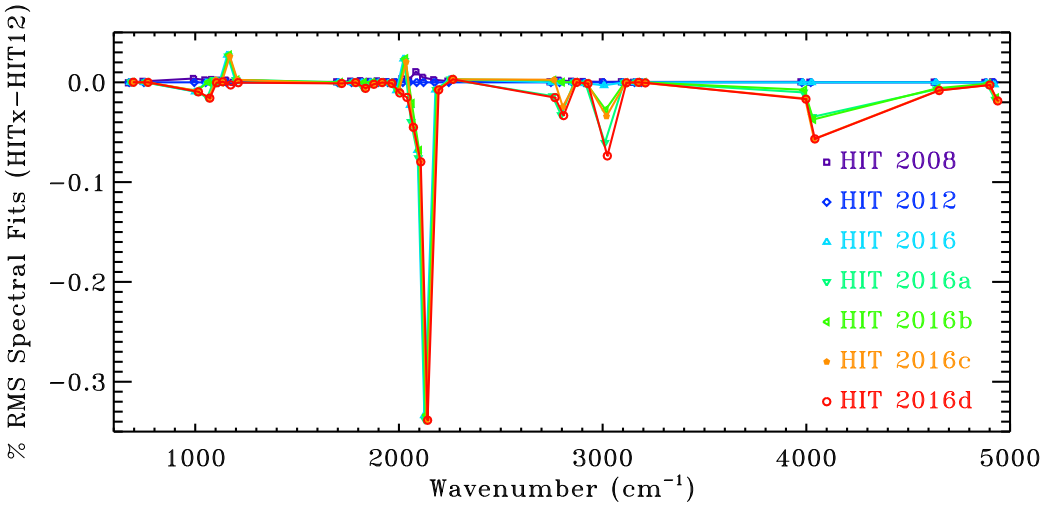
To find the answer, look at atmospheric spectra (MkIV balloon) and see if those fits also improve. If so, the empirical adjustments are likely valid.

The atmospheric balloon spectra were not used to guide the empirical linelist improvements. They are used simply to evaluate the linelist consistency, that had been previously adjusted on the basis of KP lab spectra.

Balloon Fitting Results (da_fai97rat.brl)

An occultation of MkIV balloon spectra was fitted, covering altitudes from 9 to 37 km, temperatures from 210K to 250K, and pressures from 3 mbar to 300 mbar. Every MkIV spectrum covered every window, so 100% of potential spectral fits were performed





Upper Panel: RMS fitting residuals. Absolute values are dominated by interfering gases (e.g. H_2O & HNO_3 at 1730 cm^{-1}). The linelist-to-linelist variations that pertain to O_3.

Lower Panel: Difference in RMS residuals between HITx and HIT2012. Negative values imply fits are better than with HIT2012.

In the 2050 to 2150 cm^{-1} region fits are much better with hit16 and later linelists than with HIT08 or HIT12.

Smaller improvements are also seen in other bands.

% RMS Spectral fitting residuals (balloon)

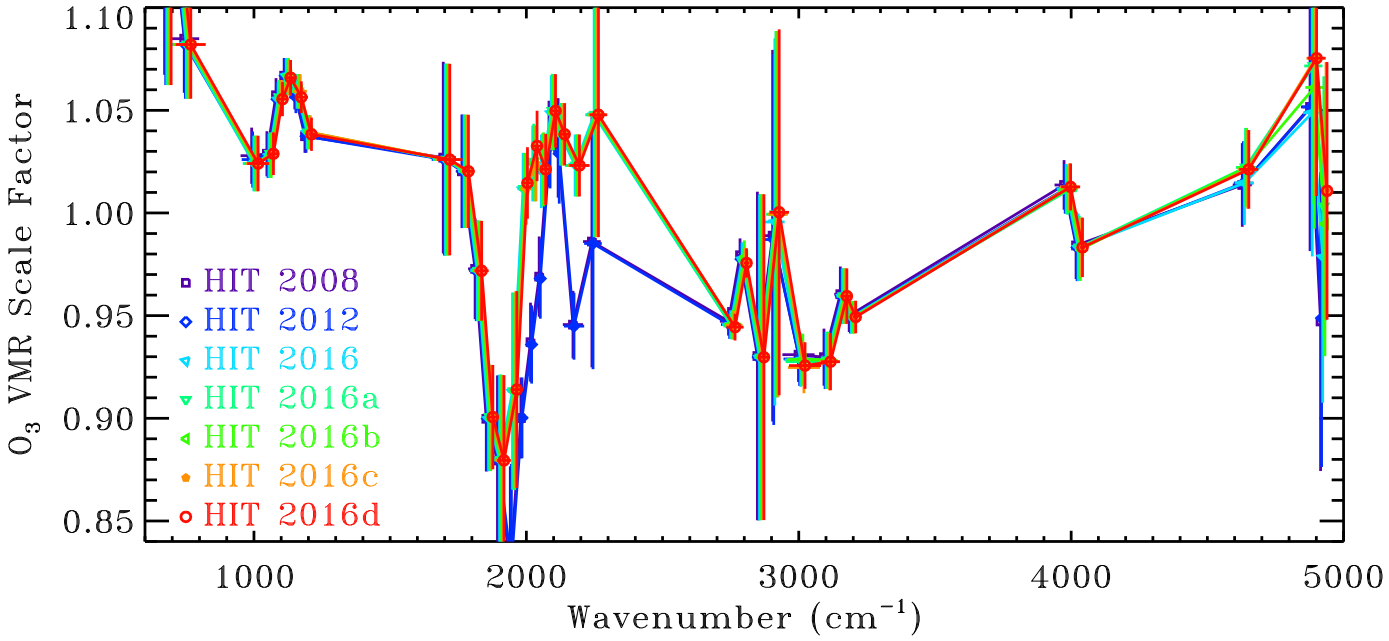
	Center (cm^{-1})	Width (cm^{-1})	N_{spec}	RMS Spectral Fits (%) HIT08	HIT12	HIT16	HIT16a	HIT16b	HIT16c	HIT16d
1	645.35	21.36	0/38	-	-	-	-	-	-	-
2	684.00	17.01	38/38	1.0426	1.0426	1.0429	1.0429	1.0429	1.0428	1.0428
3	756.00	55.01	38/38	1.0800	1.0800	1.0800	1.0800	1.0800	1.0800	1.0800
4	1003.40	40.51	38/38	0.5234	0.5234	0.5141	0.5141	0.5148	0.5148	0.5139
5	1059.40	16.30	38/38	0.3446	0.3746	0.3744	0.3593	0.3749	0.3588	0.3588
6	1092.86	16.68	38/38	0.3333	0.3333	0.3333	0.3358	0.3333	0.3357	0.3329
7	1122.68	13.14	38/38	0.3618	0.3618	0.3621	0.3623	0.3621	0.3623	0.3618
8	1161.90	24.31	38/38	0.3196	0.3196	0.3471	0.3456	0.3471	0.3456	0.3170
9	1199.25	24.31	38/38	0.2797	0.2797	0.2822	0.2822	0.2822	0.2822	0.2792
10	1706.90	46.52	38/38	1.6352	1.6352	1.6352	1.6341	1.6352	1.6341	1.6341
11	1775.44	22.46	38/38	1.0321	1.0321	1.0321	1.0320	1.0321	1.0319	1.0319
12	1823.85	25.85	38/38	0.8710	0.8710	0.8710	0.8654	0.8710	0.8650	0.8650
13	1865.00	14.70	38/38	0.7067	0.7067	0.7068	0.7050	0.7068	0.7049	0.7049
14	1905.82	25.25	38/38	0.6423	0.6423	0.6422	0.6422	0.6422	0.6422	0.6422
15	1953.12	18.13	38/38	0.5551	0.5551	0.5544	0.5543	0.5544	0.5541	0.5541
16	1992.00	20.50	38/38	0.5476	0.5476	0.5395	0.5390	0.5393	0.5389	0.5370
17	2027.22	14.53	38/38	0.6225	0.6199	0.6436	0.6414	0.6436	0.6403	0.6047
18	2059.00	17.01	38/38	0.6667	0.6667	0.6438	0.6270	0.6455	0.6237	0.6215
19	2095.15	16.20	38/38	0.6936	0.6872	0.6194	0.6115	0.6193	0.6076	0.6076
20	2128.57	17.23	38/38	0.9222	0.9222	0.5886	0.5837	0.5887	0.5836	0.5836
21	2183.00	37.20	38/38	0.5327	0.5327	0.5253	0.5252	0.5253	0.5252	0.5252
22	2252.60	32.40	38/38	0.6888	0.6888	0.6919	0.6919	0.6919	0.6919	0.6919
23	2754.00	31.51	38/38	0.2511	0.2510	0.2510	0.2372	0.2531	0.2536	0.2357
24	2796.30	10.51	38/38	0.2955	0.2936	0.2936	0.2608	0.2938	0.2691	0.2603
25	2860.00	20.01	38/38	0.5188	0.5188	0.5188	0.5188	0.5188	0.5188	0.5188
26	2916.00	36.01	38/38	0.8219	0.8219	0.8210	0.8208	0.8208	0.8208	0.8207
27	3011.00	59.01	38/38	0.5300	0.5300	0.5271	0.4695	0.5026	0.4961	0.4564
28	3104.50	34.51	38/38	0.4033	0.4024	0.4022	0.4022	0.4022	0.4020	0.4020
29	3164.50	22.51	38/38	0.3403	0.3402	0.3402	0.3402	0.3402	0.3403	0.3402
30	3197.00	10.01	38/38	0.2129	0.2126	0.2126	0.2126	0.2126	0.2120	0.2121
31	3985.88	34.88	38/38	0.3772	0.3772	0.3772	0.3672	0.3698	0.3611	0.3604
32	4029.75	8.99	38/38	0.5126	0.5126	0.5126	0.4780	0.4755	0.4560	0.4560
33	4640.00	35.01	38/38	0.2673	0.2673	0.2673	0.2604	0.2614	0.2591	0.2591
34	4888.85	33.35	38/38	0.4138	0.4138	0.4136	0.4114	0.4125	0.4112	0.4112
35	4927.35	5.17	38/38	0.4039	0.4039	0.4017	0.3865	0.3884	0.3853	0.3853

Red values indicate largest RMS. Black values indicate intermediate RMS. Blue values indicate smallest RMS.

Retrieved O_3 amounts: Balloon

An occultation of 38 MkIV balloon spectra was fitted, covering altitudes from 9 to 37 km at 0.01 cm^{-1} spectral resolution (57 cm Max OPD). Every MkIV spectrum covered every window, so 100% of potential spectral fits were performed.

The figure below shows the O_3 vmr scale factors averaged over all 38 spectra. Since we don't actually know the atmospheric O_3 amount, the absolute value of the scale factors is unimportant. But the variation from window to window is trustworthy. If fact, balloon spectra are the best way of checking the consistency of the various windows because they are all measured simultaneously in every spectrum by the same instrument.



Discussion of Balloon Results

The linelist adjustments, made on the basis of KP lab spectra, also improve fits to MKIV balloon spectra.

In nearly all cases, the hit16d linelist produces the best fits to MKIV balloon spectra.

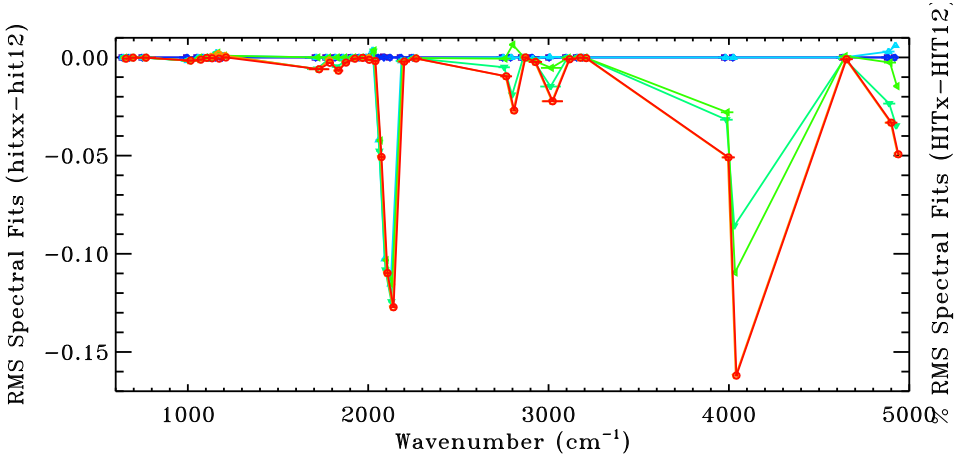
The RMS spectral fits show a similar window-to-window pattern of improvements to the KP lab spectra

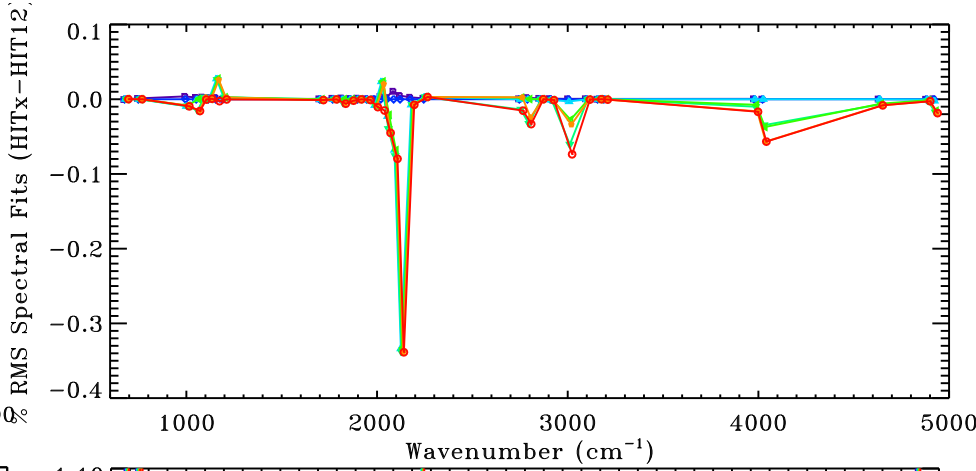
The retrieved gas amounts show a similar window-to-window pattern of biases to the KP lab spectra.

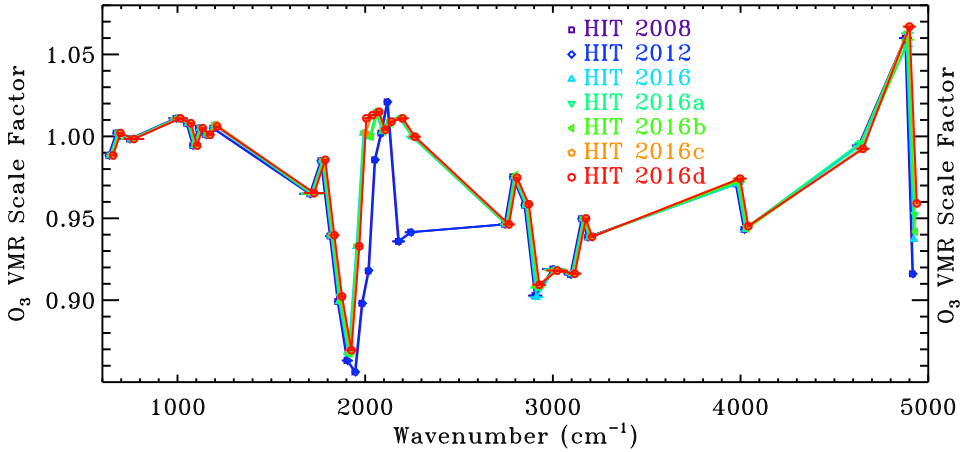
RMS improvement seen in balloon spectra seen around 4000 cm^{-1} and 4900 cm^{-1} is smaller than in lab spectra. This is because O_3 absorption is weak in balloon spectra compared with interferers, whereas in lab spectra the O_3 absorption is fairly strong and dominant.

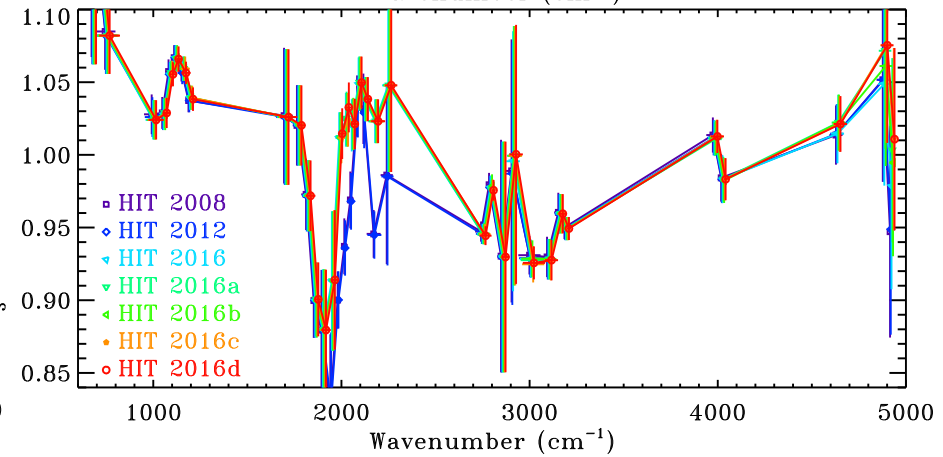
Laboratory (left) and Balloon (right) results

Upper panels show average RMS spectral fitting residuals relative to hit12. The large dips at 2050 to2150 cm^{-1} and 3950-4100 cm^{-1} indicate improvements.

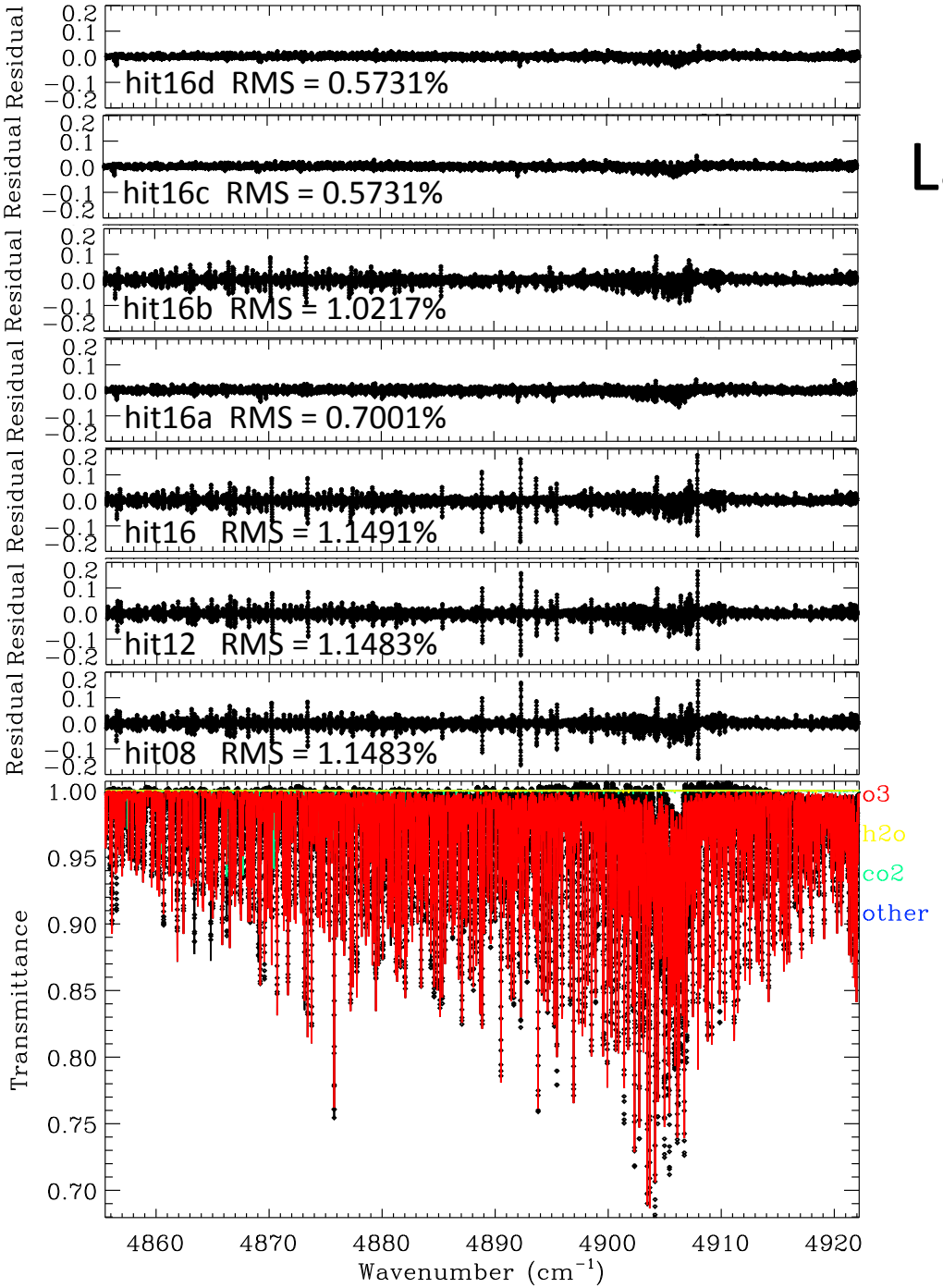








Lower panels show average vmr scale factors (VSF). Above 2000 cm^{-1} there is good consistency between balloon and laboratory, but below 2000 cm^{-1} the balloon results are ~5% larger.



Example: Fits to a Reims Lab spectrum at 4900 cm^{-1}

Figure shows fits to a Reims lab spectrum, kindly provided by Alain Barbe, using the 7 different linelists

These spectra have larger O_3 slant columns than any of the Kitt Peak spectra that I found, which is important for validating the weaker O_3 bands.

The lowest panel shows a spectral fit to 68 cm^{-1} wide portion of a Reims lab spectrum. The main absorbers are O_3 and CO_2. The residuals obtained for the 7 evaluated linelists are shown in the panels above.

hit08 and hit12 produces identical fits in this region, as do hit16c and hit16d.

Summary and Conclusions

Seven different O_3 linelists (HIT08, HIT12, HIT16, hit16a, hit16b, hit16c, hit16d), together with the non-O_3 lines from ATM16, were used to fit Kitt Peak lab spectra of O_3 and MkIV balloon spectra in 35 windows covering significant O_3 absorption bands over 600-5000 cm^{-1}.

The largest residuals seen in fits with HIT16 or hit16c could be fixed simply by adjusting line positions, typically by less than 1mK. In two cases, a single line was split into two weaker lines with the same total intensity, which were adjusted.

The position (and intensity) adjustments that improved fits to lab spectra also improved fits to MkIV balloon spectra.

Around 2000 and 2200 cm^{-1} the HIT16x linelists produces 5-10% larger retrieved O_3 amounts than earlier linelists, bringing these regions into better consistency with the others.

The 3000 cm^{-1} region yields low retrieved O_3 amounts, about 5-8% lower than average, both in fits to Kitt Peak lab and MkIV limb spectra. This indicates errors in line widths or intensities, since the errant line positions were corrected.

Pressure shifts need updating. Most O_3 lines have shifts of zero, even some strong lines. The inconsistency between the lines with and without pressure shifts drive up the residuals in fits to higher-pressure lab spectra.