SUPPLEMENT

Table S1. Vertical Scale for SBUV Partial Column data. SBUV Partial Column Ozone is reported in DU per layer within 21 defined pressure layers from the ground to the top of the atmosphere.

Pressure scale for SBUV partial column					
# of the layer	Layer bottom pressure, hPa	Layer top pressure, hPa	# of the layer	Layer bottom pressure, hPa	Layer top pressure, hPa
1	1013	639.16	11	10.13	6.3916
2	639.16	403.283	12	6.3916	4.03282
3	403.283	254.454	13	4.03282	2.54454
4	254.454	160.55	14	2.54454	1.6055
5	160.55	101.3	15	1.6055	1.013
6	101.3	63.916	16	1.013	0.63916
7	63.916	40.3283	17	0.63916	0.403283
8	40.3283	25.4454	18	0.403283	0.254454
9	25.4454	16.055	19	0.254454	0.16055
10	16.055	10.13	20	0.16055	0.1013
			21	0.1013	0.0

Table S2. Vertical scale for SBUV mixing ratio data. SBUV ozone mixing ratios are reported in parts per million by volume (ppmv) on 15 defined pressure levels between 0.5 hPa and 50 hPa.

Pressure scale for SBUV mixing ratios					
Level #	Pressure, hPa	Level #	Pressure, hPa	Level #	Pressure, hPa
1	0.5	6	3	11	15
2	0.7	7	4	12	20
3	1	8	5	13	30
4	1.5	9	7	14	40
5	2	10	10	15	50

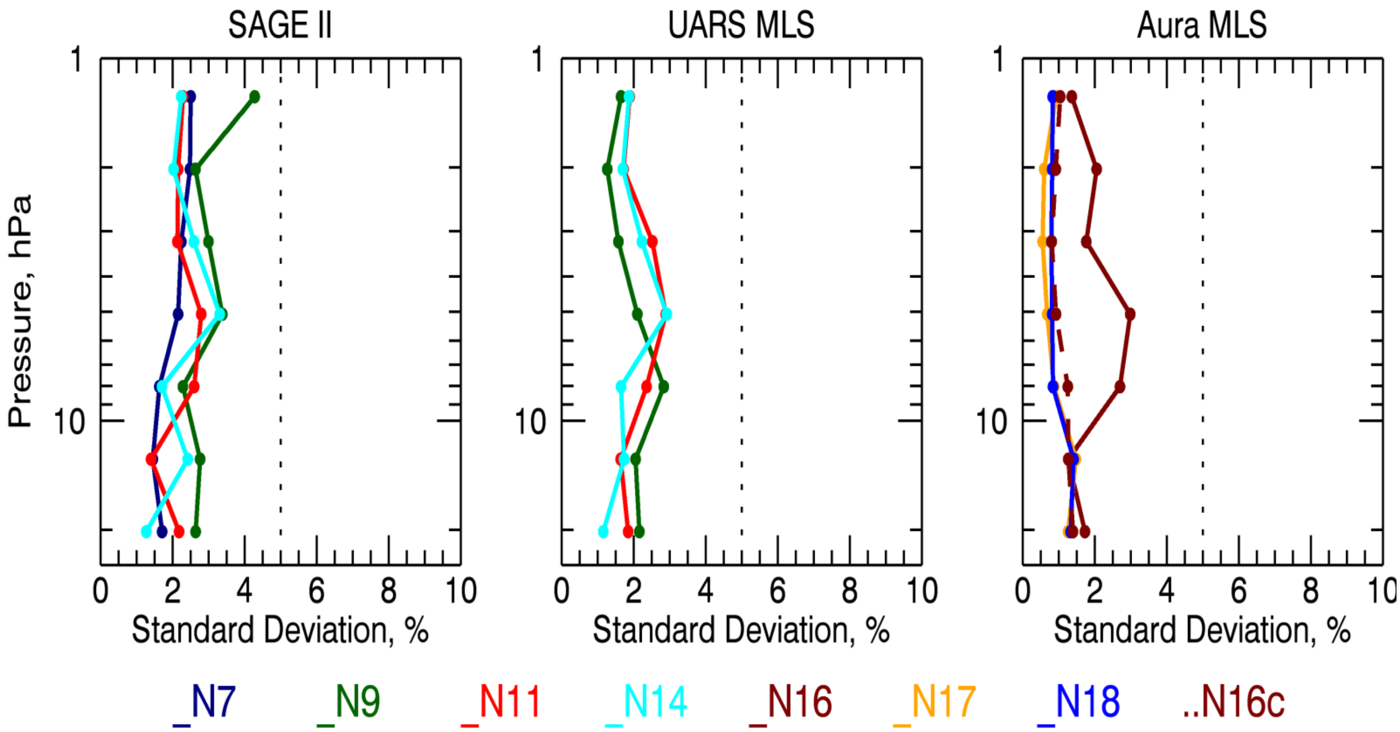

Fig. S1. Vertical profiles of the standard deviations of differences for the broad latitude band 50S - 50N. Colors correspond to individual SBUV instruments. The standard deviations vary from 1-4%. Lower standard deviations are found for N17 and N18 relative to Aura MLS. We found a significant drift in the N16 ozone time series after mid-2007 due to orbital drift, when the local equator crossing time of the N16 satellite orbit passes 4 pm. As a result the standard deviation for the full N16 record is larger compared with N17 and N18. The dashed maroon line on the right panel shows the standard deviation for N16 for the period from October 2004 to mid-2007. The N16 standard deviations for this time period are the same order as those for N17 and N18. We do not recommend using N16 ozone data after mid-2007.

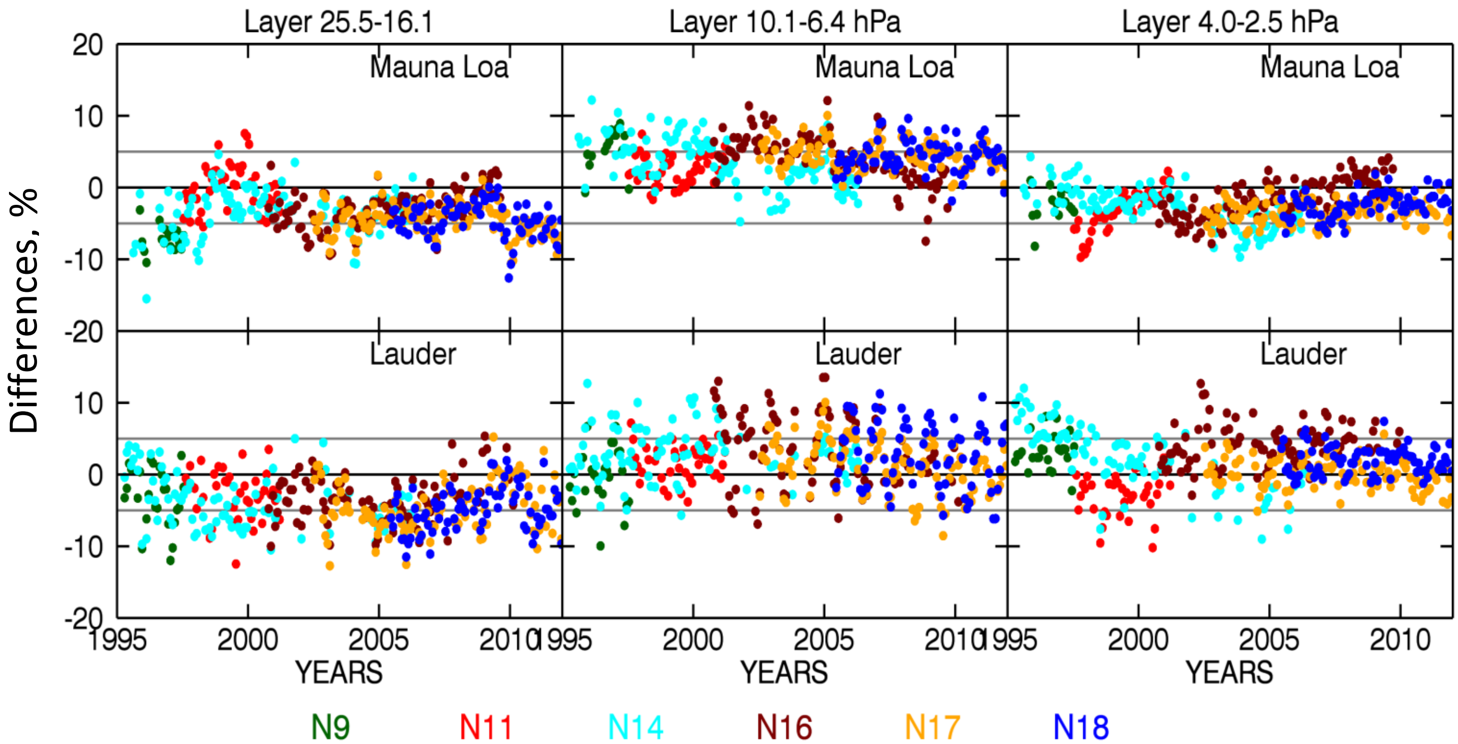

Fig. S2. Time series of monthly mean differences for individual SBUV instruments relative to ground-based microwave spectrometers. Columns on the panel show different layers, while rows correspond to different stations (Mauna Loa (20N) and Lauder (45S)). Colors correspond to individual SBUV instruments. Despite the mean offsets between microwave and SBUV measurements, the dispersion of the differences is small (10-15%), indicating good correspondence between the SBUV and microwave measurements.

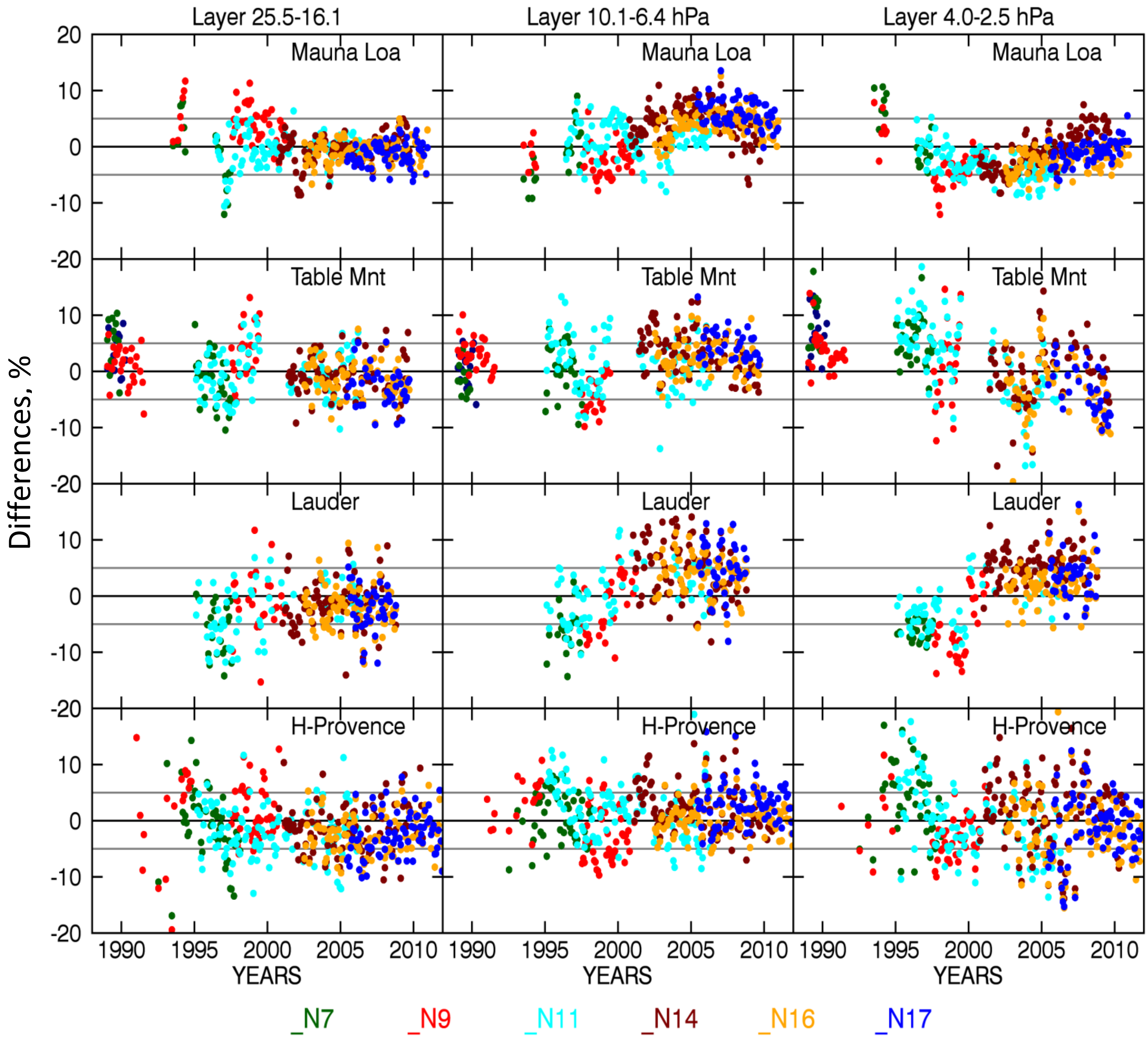

Fig. S3. Same as Fig. S2 only relative to ground-based lidar instruments at four stations: Mauna Loa (20N), Table Mountain (34N), Lauder (45S) and Haute Provence (44N). The dispersion of differences is larger relative to lidar observarions, and there are notable offsets in the long term data records at some layers and locations. At Table Mountain and Lauder we can see a transition around 2000, especially in the 4-2.5 hPa layer. At Mauna Loa the differences in the first half of the 1990s are slightly shifted relative to measurements after 1996. The time series of differences relative to lidar measurements at Haute-Provence do not show significant time dependence.

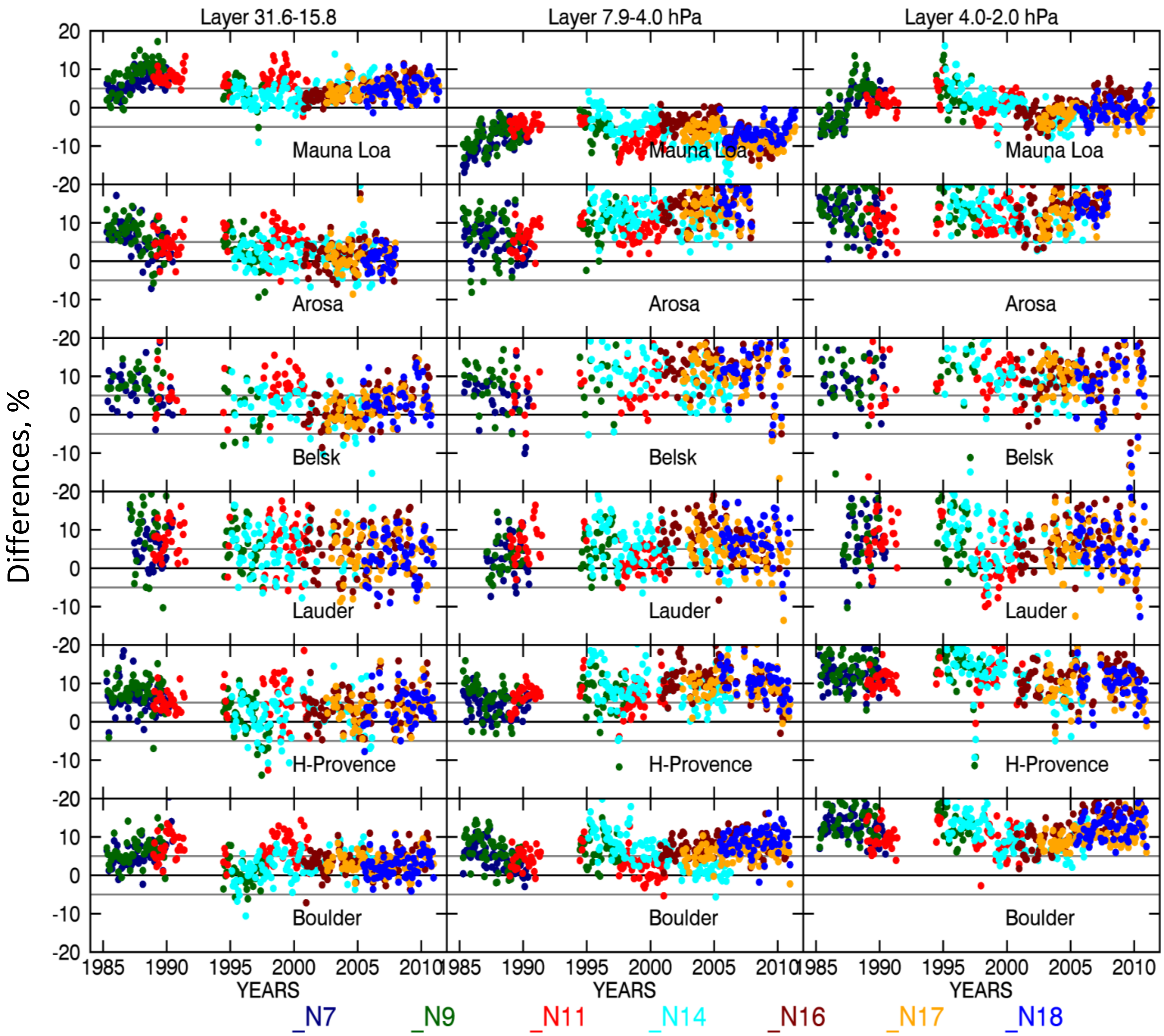

Fig. S4. Same as Fig. S2 only relative to ground-based Umkehr instruments at six stations: Mauna Loa (20N), Arosa (47N), Belsk (52N), Lauder (45S), Haute Provence (44N) and Boulder (40N). Records from all stations demonstrate that Umkehr instruments underestimate ozone amounts in the stratosphere above 31 hPa.

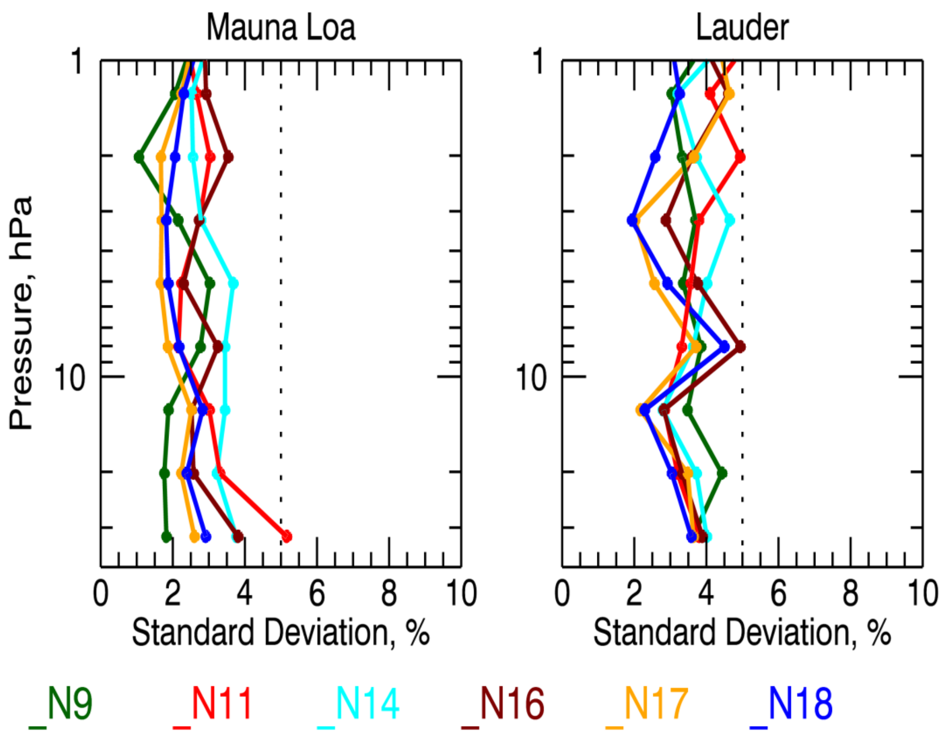

Fig. S5. Vertical profiles of the standard deviations for individual SBUV instruments relative to coincident ground-based microwave measurements at Mauna Loa and Lauder. Colors correspond to individual SBUV instruments. Standard deviations vary between 1-5% with slightly larger deviations at Lauder.

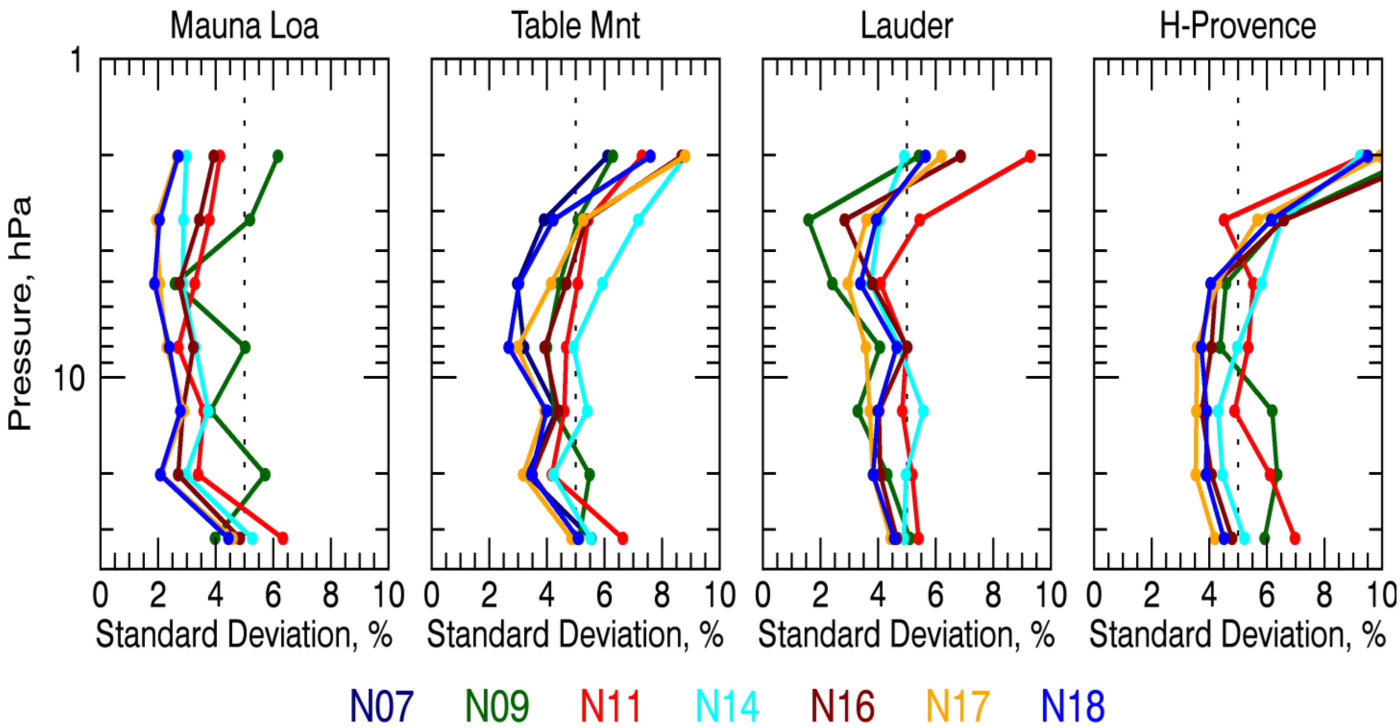

Fig. S6. Same as Fig. S5 but relative to ground-based lidar measurements at four specified stations. Between 40 and 2 hPa standard deviations are within 2-6%, and significantly increase above 2 hPa up to 10-12%. The larger deviations are related to the reduced number of lidar observations at high altitudes after applying the 10% lidar precision screening.

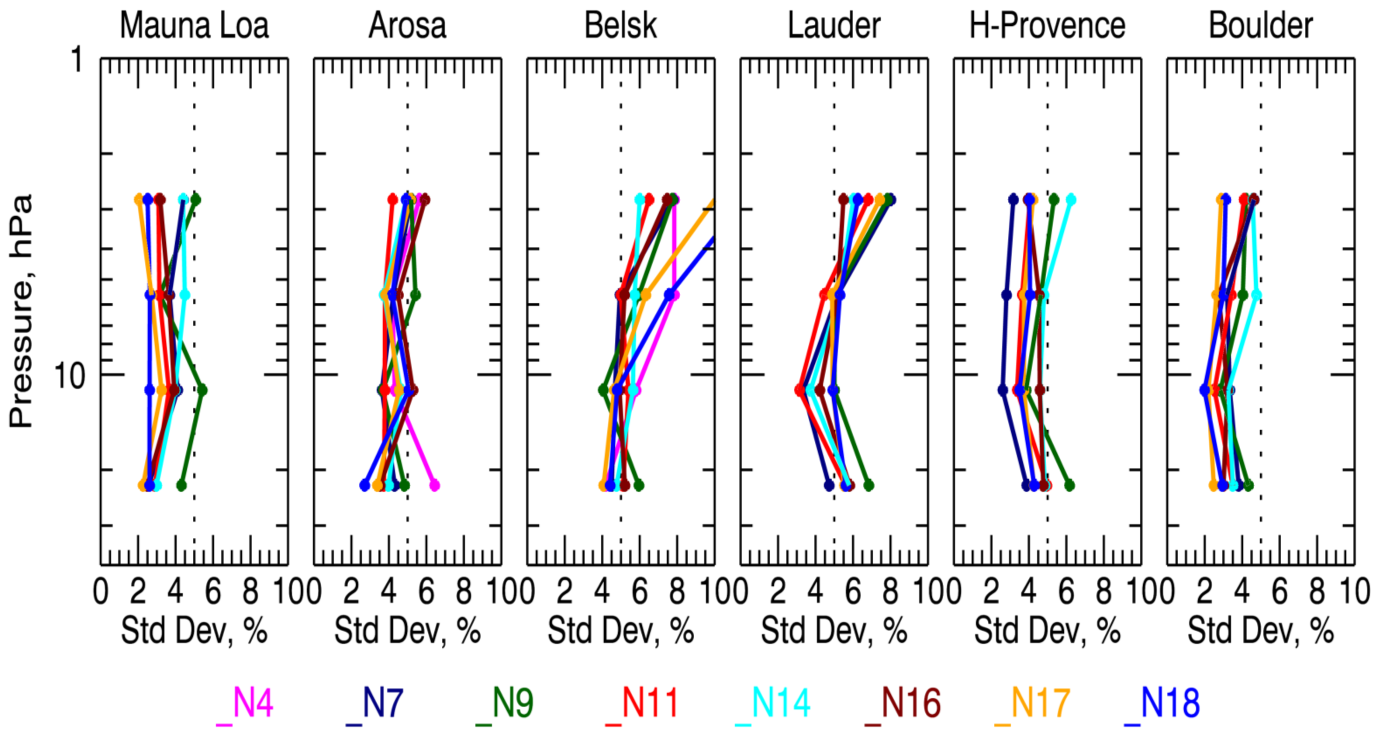

Fig. S7. Same as Fig. S5, but relative to ground-based Umkehr measurements at six specified stations. Standard deviations for comparisons with Umkehr measurements are mostly within 2-6%. However, we find larger deviations relative to the instrument at Belsk (up to 10-14%) above 4 hPa. The lowest deviations are found at Mauna Loa and Boulder.

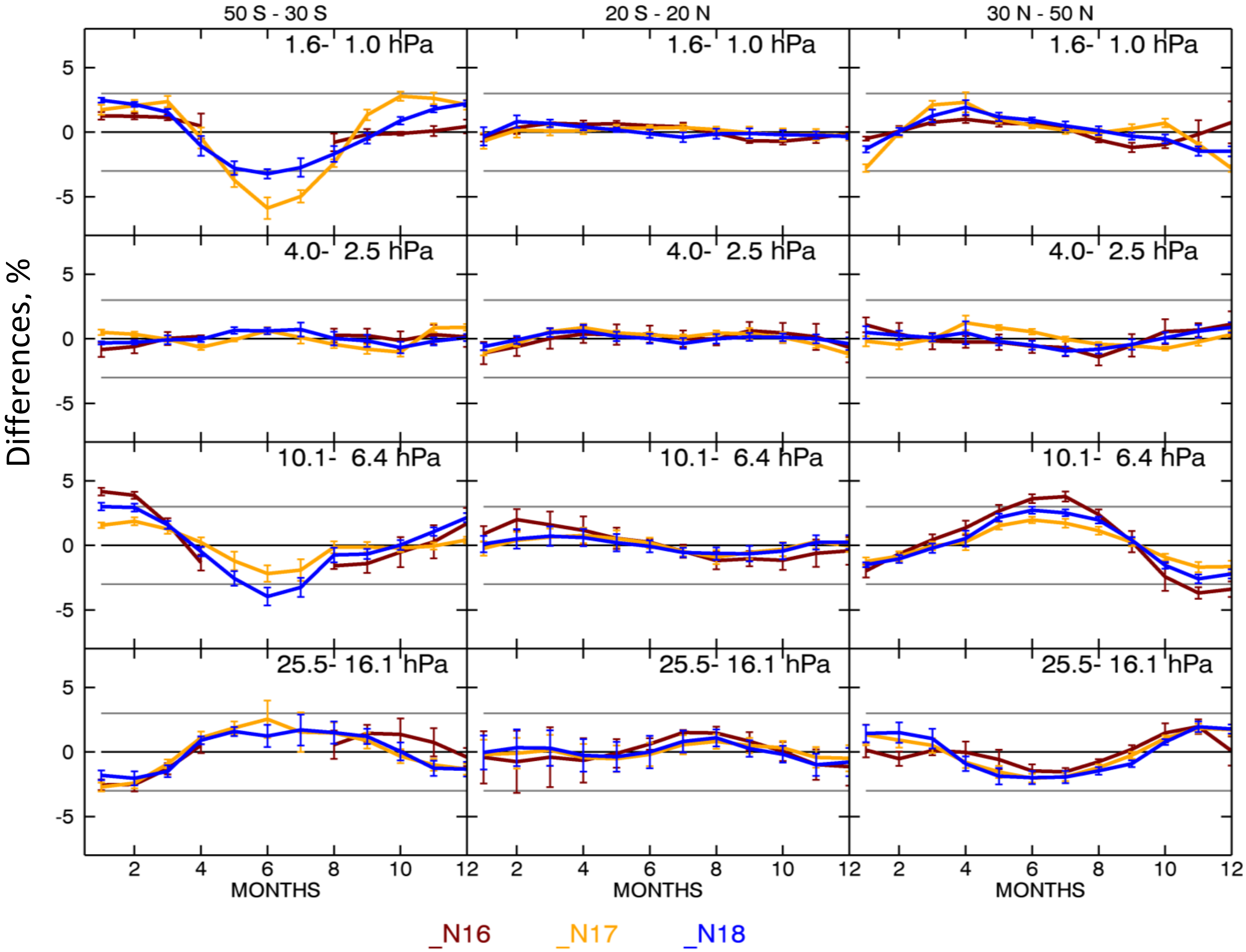

Fig. S8. Seasonal biases for N16, N17 and N18 SBUV relative to AURA MLS for four layers and 3 latitude bands. The vertical lines define the standard error of the mean. We calculate seasonal biases as the differences between datasets at each month after removing the mean bias. We require at least two measurements from two different years in each specified month to calculate the seasonal bias. The seasonal biases are smaller over the tropics, where they are mostly insignificant and less than 2%. We find clear seasonal patterns in both extratropics, with a 6-month lag between the two hemispheres. The seasonal biases are mostly within ±2-3%, but at some layers the amplitude is as high as 5-6%. We find larger seasonal biases in the middle and lower stratosphere with the maximum in the 10-6 hPa layer.

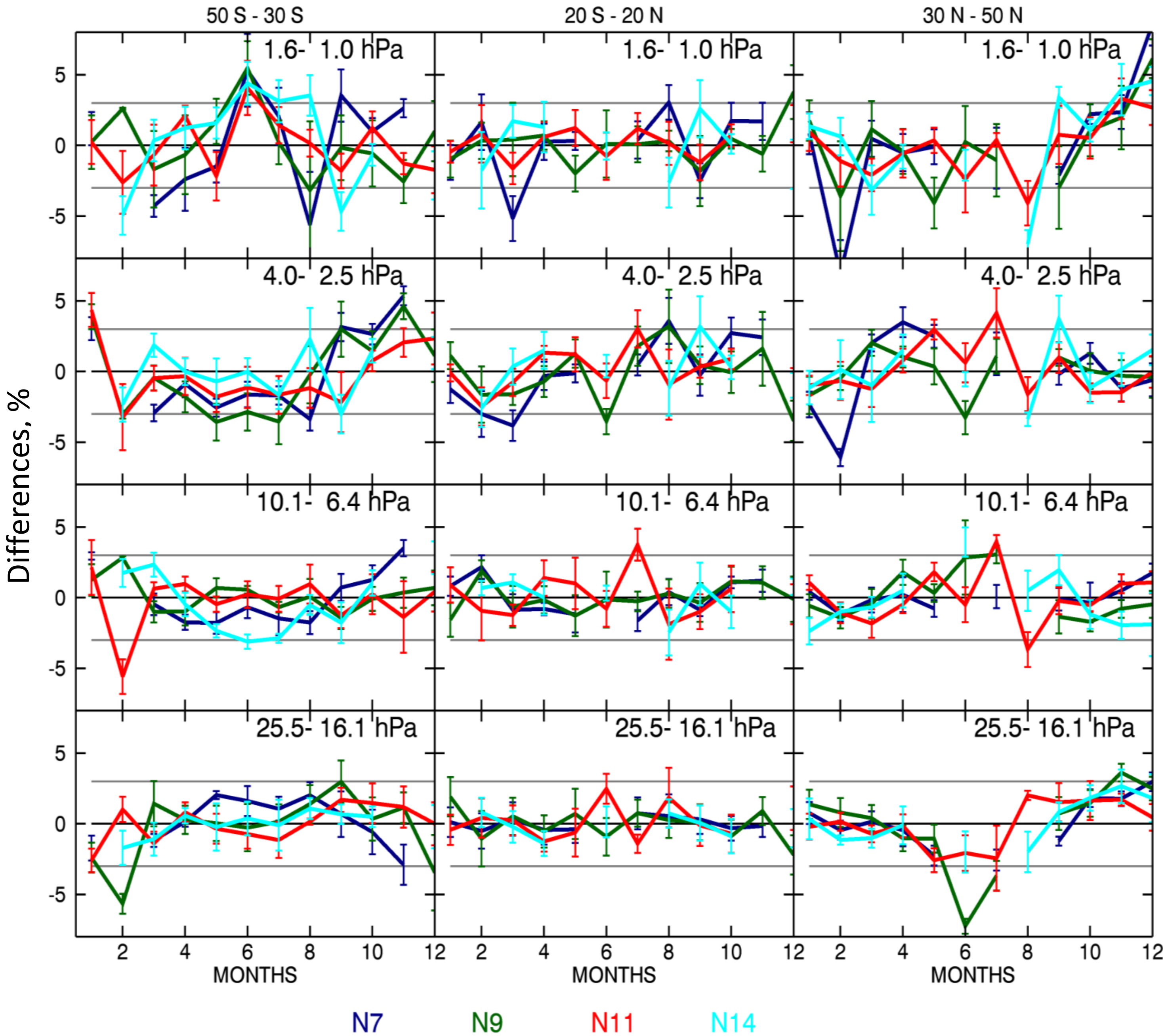

Fig. S9. Same as Fig. S8 but for N7, N9, N11 and N14 SBUV relative to SAGE II. We did not find seasonal patterns from the comparisons with SAGE II. However, it is important to note that we have limited overlap time periods and sparse SAGE II sampling. These factors complicate the analysis of the seasonal biases.

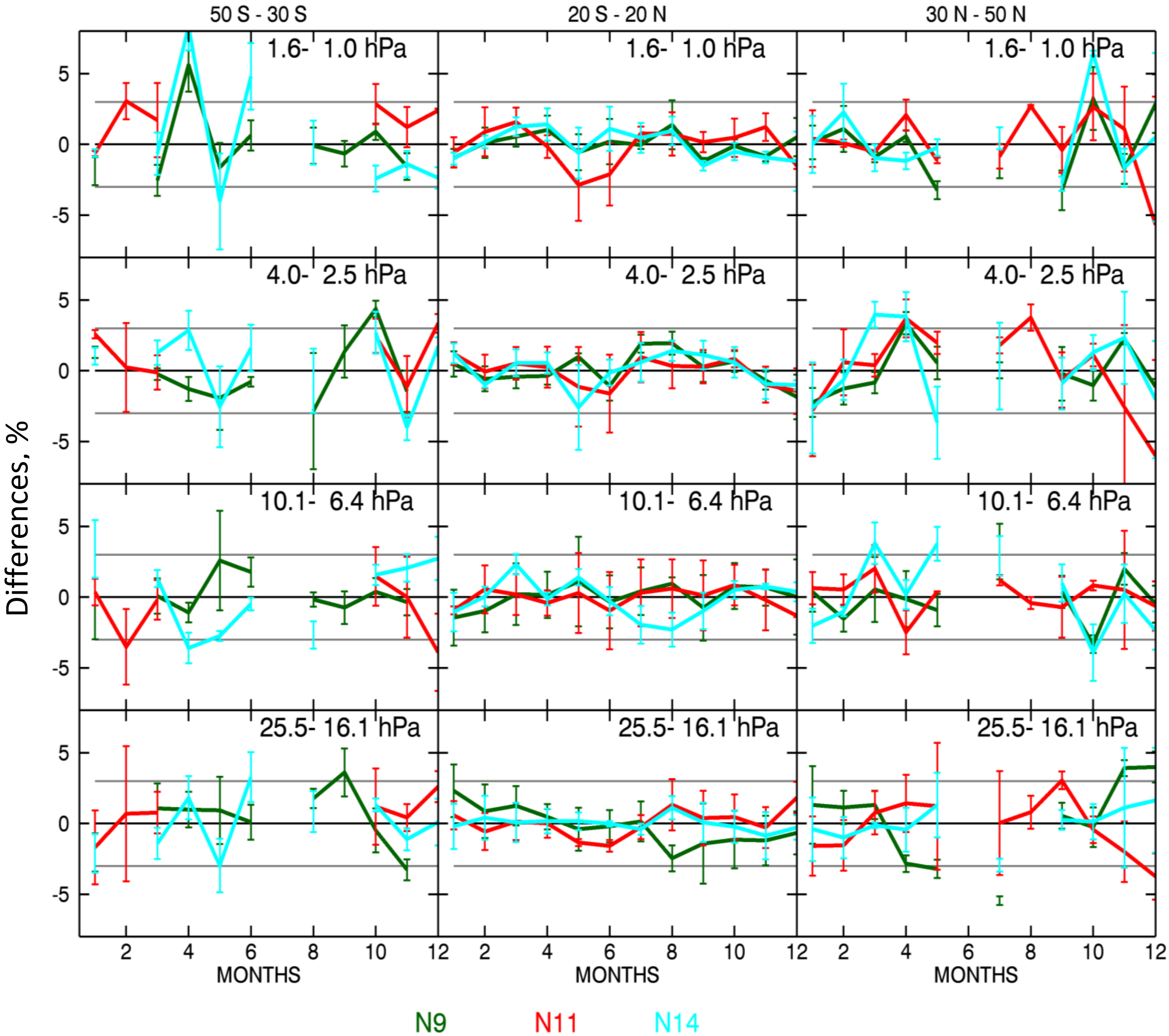

Fig. S10. Same as Fig. S8 but for N9, N11 and N14 SBUV relative to UARS MLS. There are not enough coincident measurements in the extra-tropics to estimate biases for all months. Coincidences are limited by short overlap time periods between SBUV instruments and UARS MLS and by the specific orbital track of the UARS satellite, which scans middle and high latitudes of each hemisphere only every other month. The biases in the tropics are small and mostly less than 2-3%.

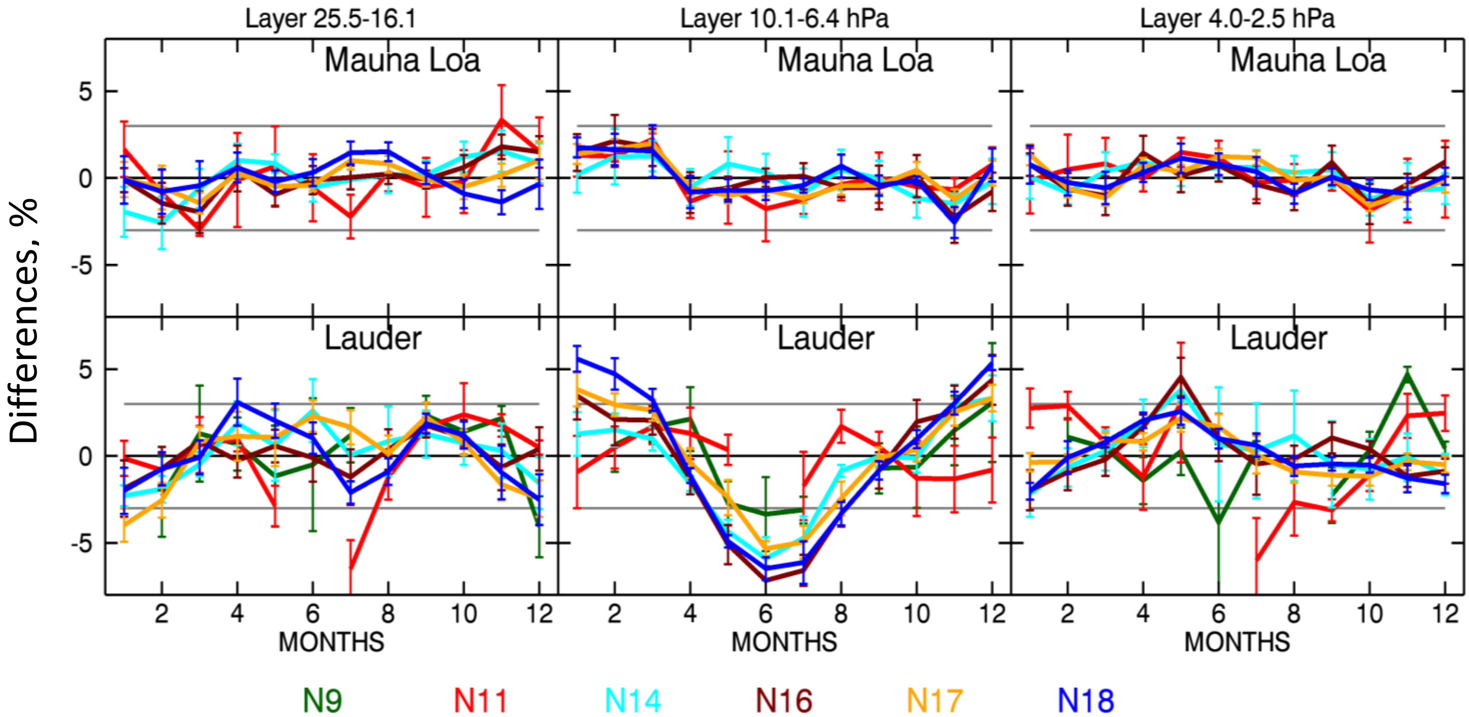

Fig. S11. Seasonal biases for individual SBUV instruments relative to ground-based microwave spectrometers at Mauna Loa and Lauder. Colors correspond to individual SBUV instruments. Seasonal biases are very small at Mauna Loa and consistent with the satellite comparisons which also show smaller seasonal biases over the tropics. Biases are larger at Lauder with amplitudes up to 7% in the 10-6 hPa layer. The seasonal pattern found at Lauder is similar to the one detected from comparisons with Aura MLS (see Fig. S8).

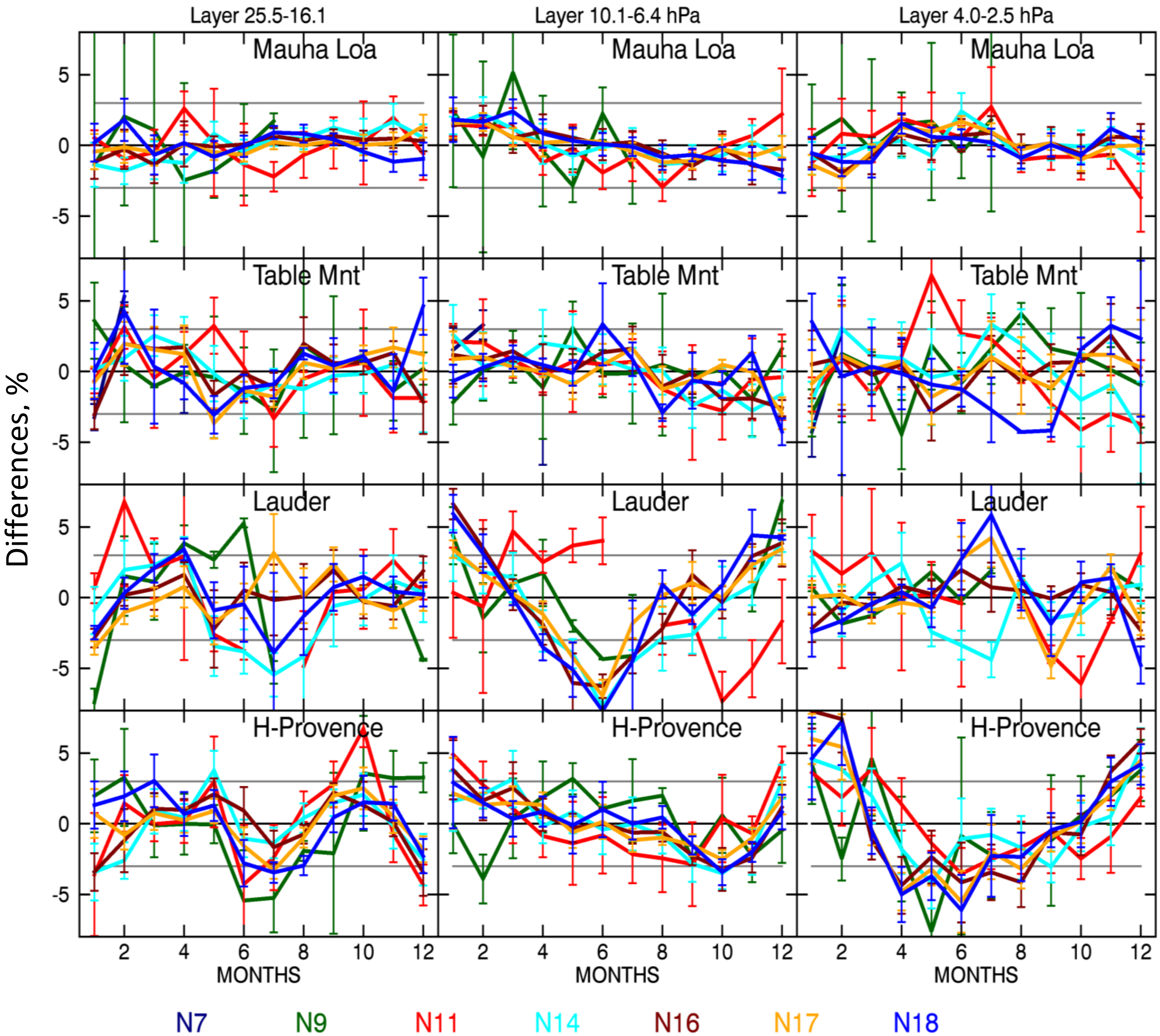

Fig. S12. Same as Fig. S11 but relative to ground-based lidars at four stations. Seasonal biases for N9 and N11 are substantially greater with large error bars. Again seasonal biases are smaller in the tropics at Mauna Loa. In the layer 4-2.5 hPa, the seasonal pattern at Table Mountain shows inconsistent differences that are not seen in Mauna Loa and Haute Provence. At Lauder the seasonal pattern in the 10-6 hPa layer is consistent with that detected from comparisons with the Lauder microwave spectrometer and Aura MLS. Seasonal structures of biases relative to the Haute Provence lidar are not consistent with Aura MLS. For example, in the 2-4 hPa layer the amplitude of the seasonal biases is about 8%, while comparisons with Aura MLS showed no sign of seasonal biases at that altitude.

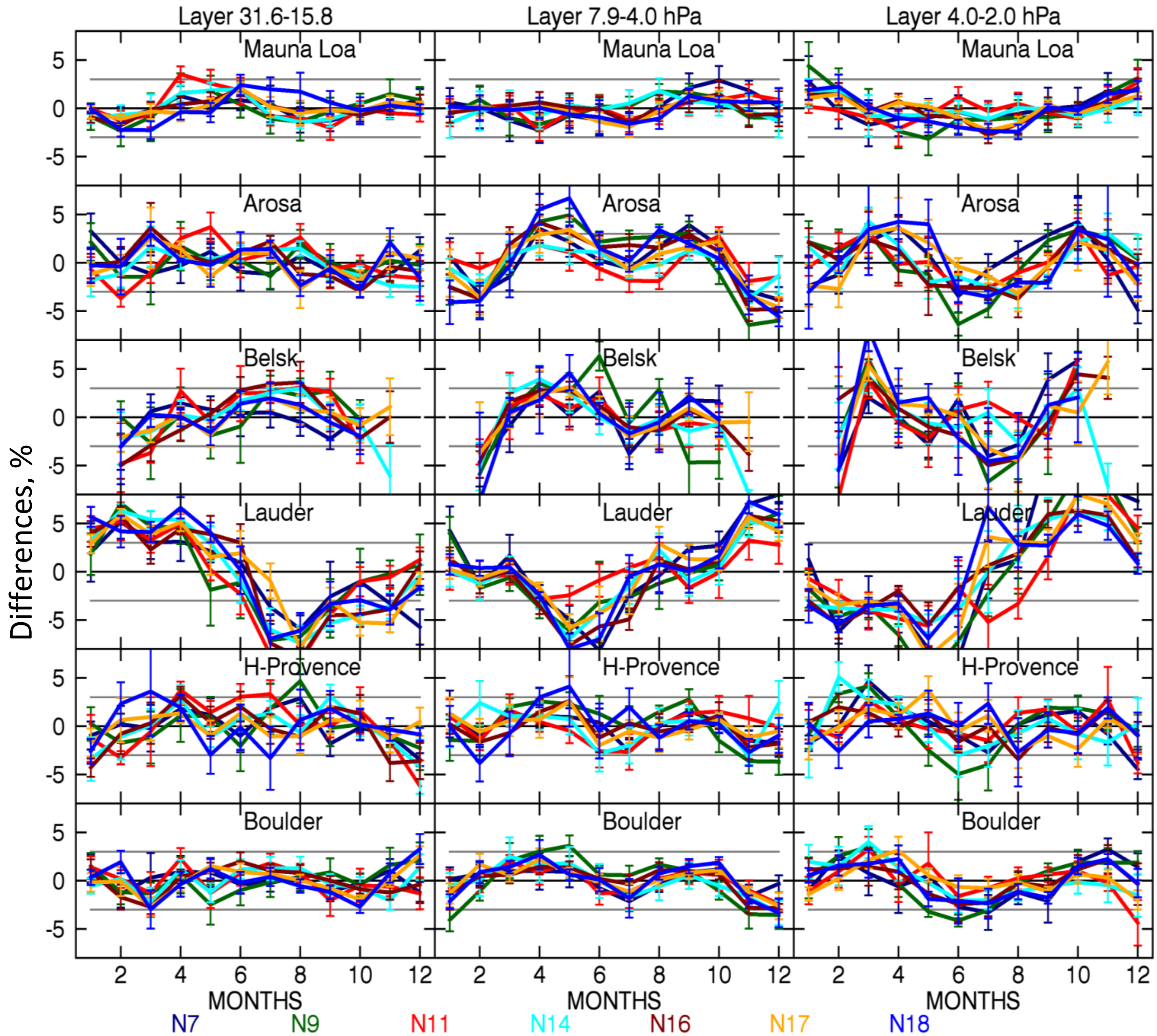

Fig. S13. Same as Fig. S11 but relative to ground-based Umkehr instruments at six stations. Note that the vertical scale for validation with Umkehr is different. Seasonal biases at Mauna Loa are less than 2-3%. Biases relative to two northern hemisphere instruments (Arosa and Belsk) have similar structures in the 8-4 hPa and 4-2 hPa layers. In the 8-4 hPa layer the biases are positive in summer and negative in winter. These results are consistent with Aura MLS comparisons (see Fig. S8), but the two other northern hemisphere stations, Haute Provence and Boulder, show small seasonal biases within ±5% in all layers. At Lauder the seasonal pattern of biases in the 4-8 hPa layer is consistent with other comparisons (in the 10-6 hPa layer for the Lauder microwave (Fig. S11) and lidar (Fig. S12) comparisons), but large seasonal biases in other layers are not consistent with other observations.

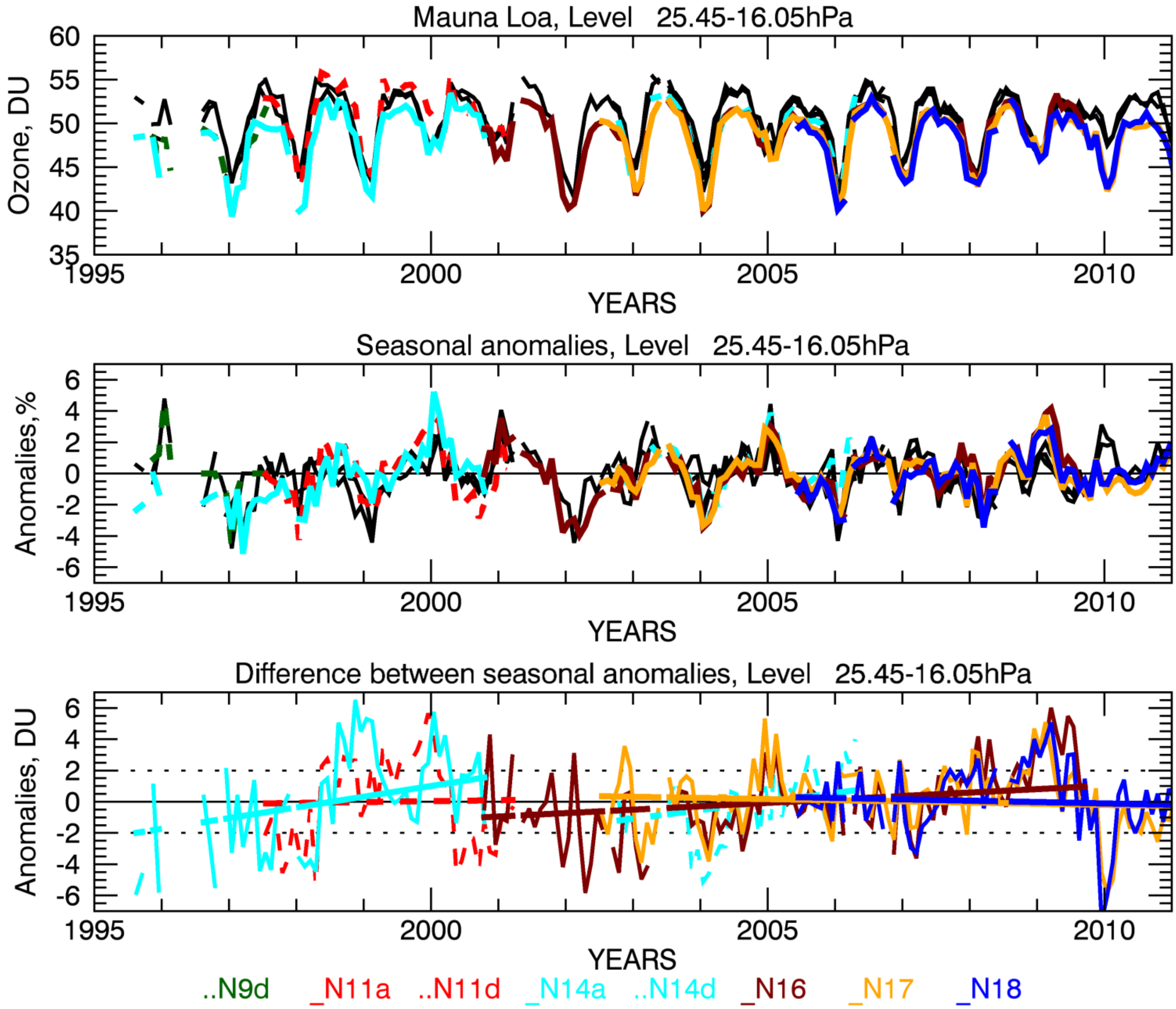

Fig. S14. (top) Time series of the coincident microwave (black line) and SBUV measurements at 25-16 hPa layer. Different colors correspond to individual SBUV instruments, and dashed lines correspond to descending orbit modes.

(middle) Time series of the deseasonalized ozone anomalies obtained from individual SBUV instruments (different colors) and microwave spectrometer (black line). Note the good correspondence between the deseasonalized time series.

(bottom) Time series of the differences between the deseasonalized microwave and SBUV anomalies (thin color lines) along with the linear drifts (thick color lines). This figure demonstrates a typical example of the differences time series and associated linear drifts.

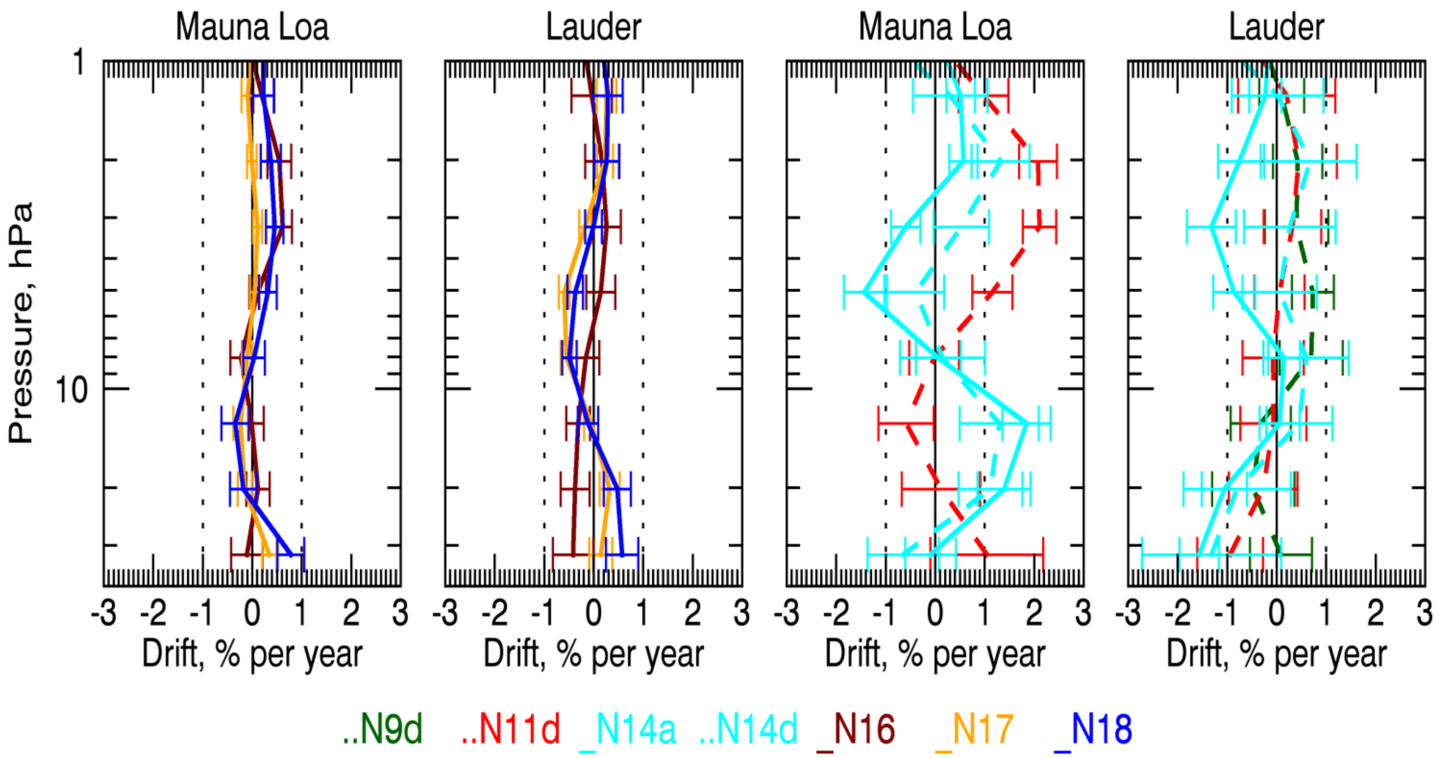

Fig. S15. Vertical profiles of drifts for individual SBUV instruments relative to ground-based microwave instruments at Mauna Loa and Lauder. Drifts for ascending and descending orbit modes are calculated independently. The drifts for N16, N17 and N18 are small, reflecting the stability of the most recent instruments. Drifts are larger for previous SBUV instruments. The vertical pattern of drifts for N14 ascending is similar for both stations. This consistency will allow us to account for the N14 drift in the merged data set uncertainty analysis.

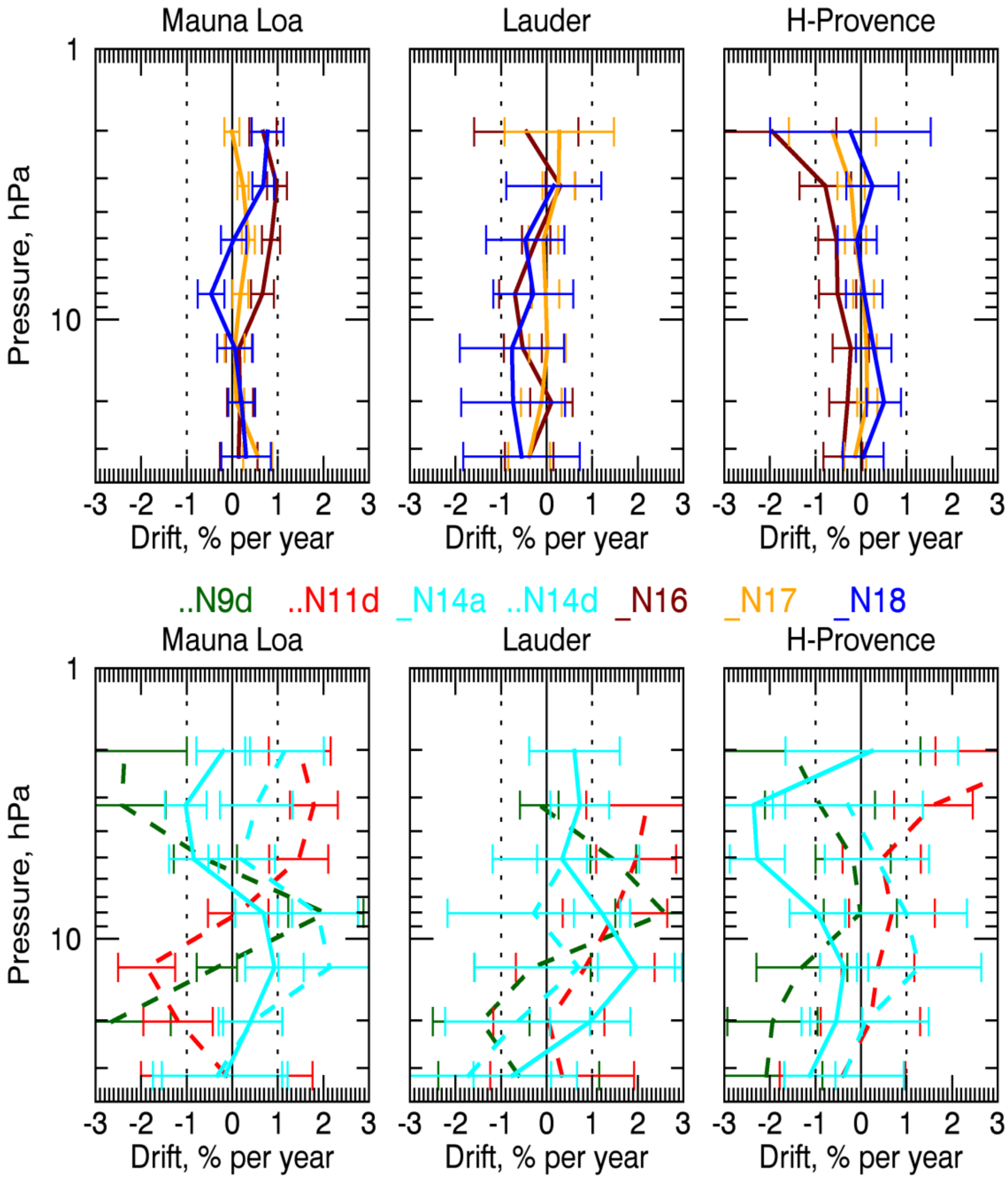

Fig. S16. Same as Fig. S15 but relative to lidar instruments at Mauna Loa, Lauder and Haute Provence. Drifts for three recent instruments shown on the top panel are small (less than 0.5% per year) at all locations, with slightly larger drifts and corresponding uncertainties for the upper layer at Haute Provence. We found significantly larger drifts (1-3% per year) for earlier instruments.

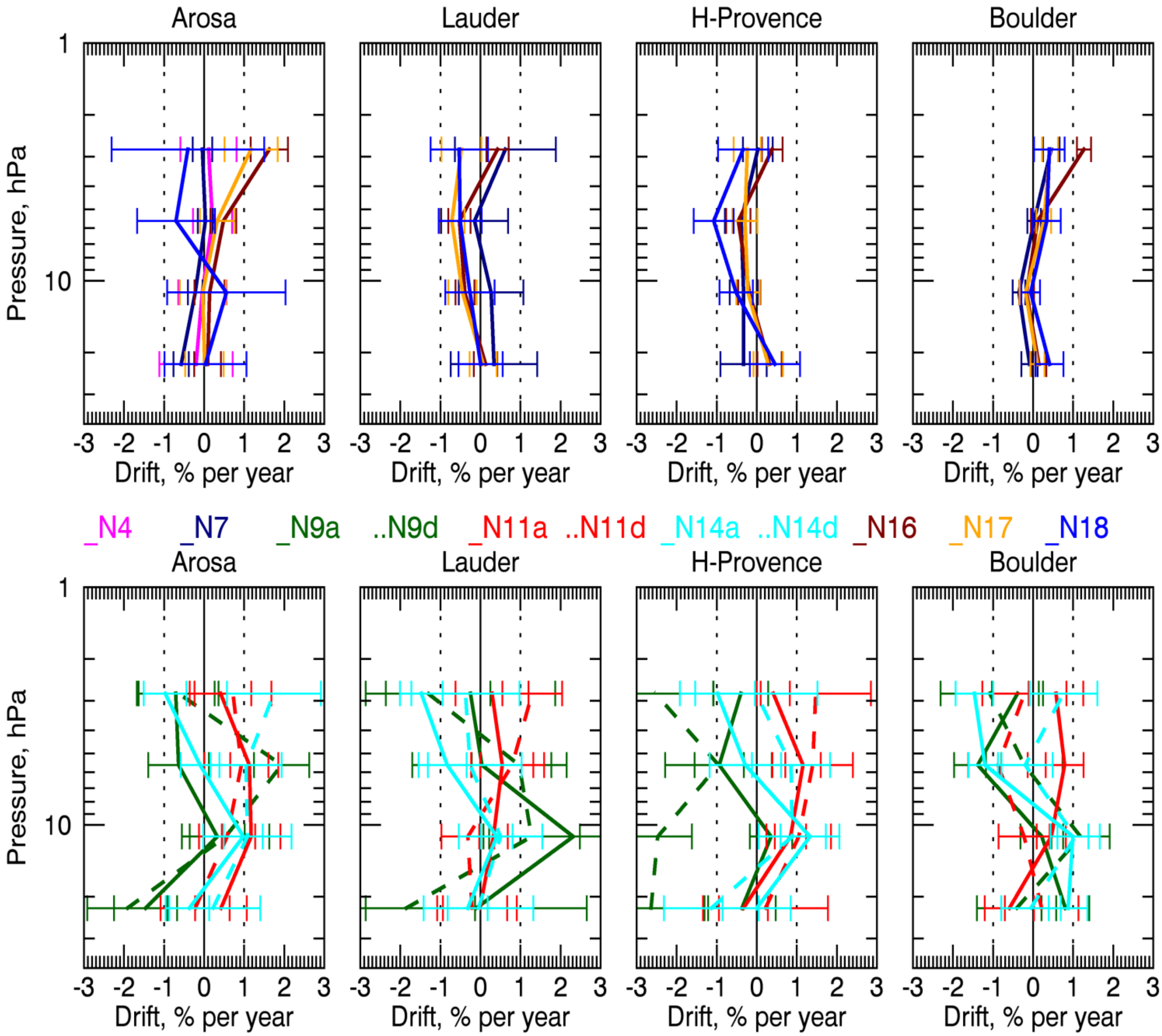

Fig. S17. Same as Fig. S15 but relative to ground-based Umkehr observations at Arosa, Lauder, Haute Provence and Boulder. The top panel shows drifts for N4, N7, N16, N17 and N18 relative to Umkehr observations at four locations. Drifts are small and mostly less than 0.5% per year. However we found slightly larger drifts for N16, N17 and N18 relative to Arosa, most likely due to a shorter overlap time period. The bottom row shows drifts for ascending (solid lines) and descending (dashed lines) orbital modes of N9, N11 and N14. Drifts are larger and vary within ±3%. We note a stable pattern of drifts at all stations for the ascending mode of N11 and the ascending mode of N14.

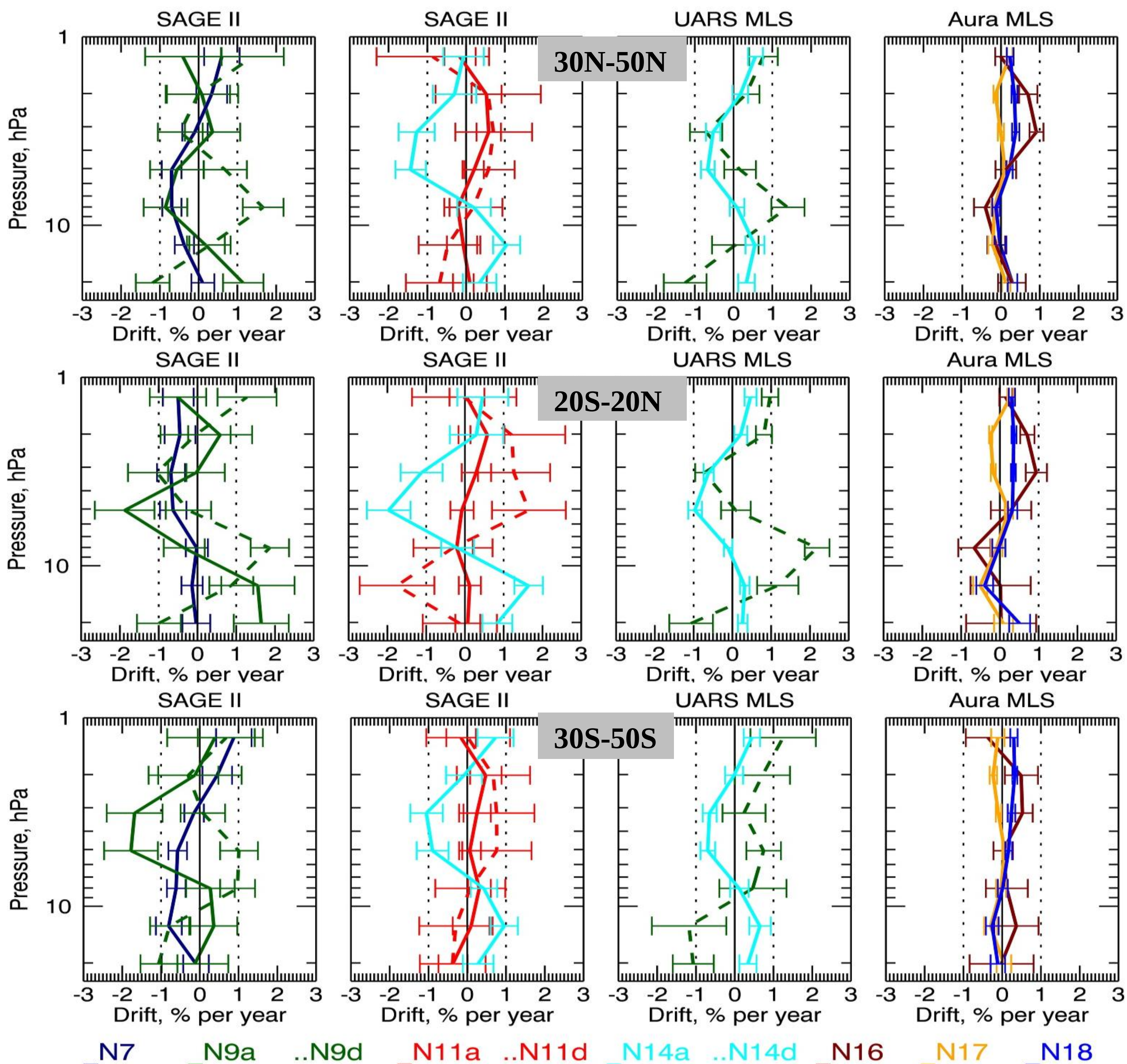

Fig. S18. Vertical profiles of drifts for individual SBUV instruments relative to independent satellite measurements for three latitude bands: northern mid-latitudes 30N-50N (top row), tropics 20S-20N (middle row) and southern mid-latitudes 30S-50S (bottom row). Drifts for ascending and descending orbit modes are calculated independently. The horizontal error bars indicate two times the standard deviation of the slope. Column 1: Drifts for N7 and ascending and descending modes of N9 relative to SAGE II; Column 2: Drifts for ascending and descending modes of N11 and ascending mode of N14 relative to SAGE II; Column 3:Drifts for descending mode of N9 and ascending mode of N14 relative to UARS MLS; Column 4: Drifts for N16, N17 and N18 relative to Aura MLS.

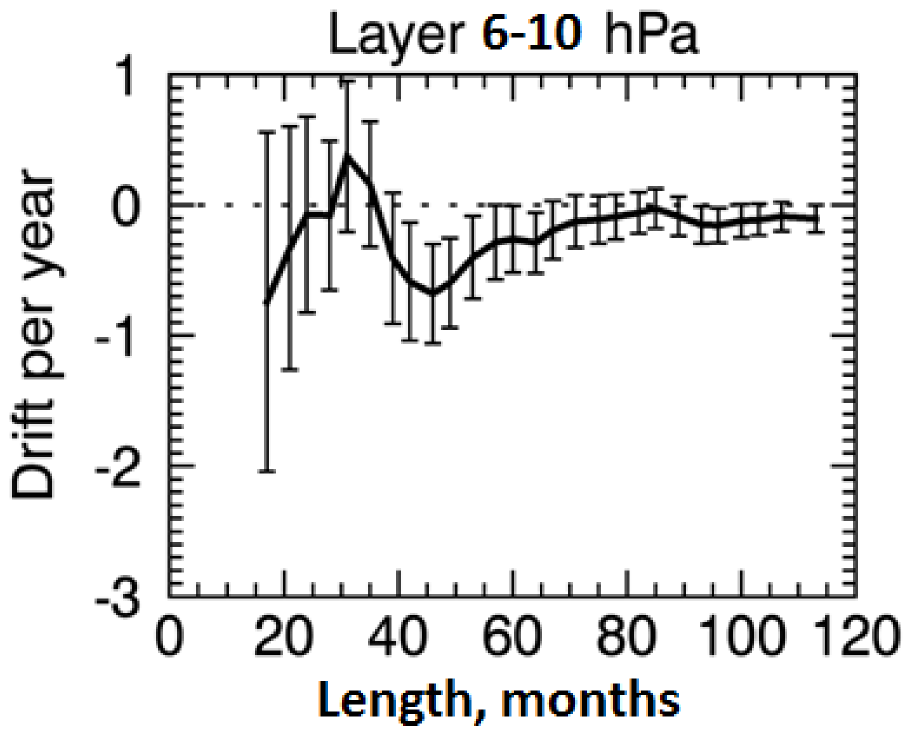

Fig. S19. Drift between N17 and ground-based microwave at Mauna Loa as function of the overlap length. The vertical error bars indicate 2σ standard deviation of the slope. The overlap between these two instruments is 9.5 years (July 2002 and December 2011). We calculated drifts for shorter periods starting with the 1.5-year overlap between July 2002 and Dec 2003 (18 months), then adding three months of overlap at a time, ending with the full period of 9.5 years (114 months). This figure demonstrates the importance of having long overlapping time periods and sufficient sampling to accurately estimate a relative drift and to reduce the standard deviations of the slope. The error bars significantly decrease with increasing overlap time. It is also interesting to note that for overlap of about 40-50 months (between July 2002 and November 2005-May 2006) the relative drift is negative (-0.7 yr^{-1}) and significant at the 2σ level, while the drift over the whole time period is only about -0.1 yr^{-1} and is not significant at the 2σ level. This results from the statistical characteristics of the difference time series, with anomaly variations that are correlated over many months.