Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2018-275-RC2, 2019

Atmospheric Measurement Techniques Discussions
Open Access

Interactive comment on "Merging of ozone profiles from SCIAMACHY, OMPS and SAGE II observations to study stratospheric ozone changes" *by* Carlo Arosio et al.

Anonymous Referee #1

Received and published: 7 January 2019

Review of 'Merging of ozone profiles from SCIAMACHY, OMPS and SAGE II observations to study stratospheric ozone changes', Arosio et al, AMTD, 2018

1 Short resume

The construction of long-term ozone profile data records for trend studies has been a lively line of research these past few years. The objective is simple, yet has proven hard to realise: obtain a climate data record which is sufficiently stable (better than a few % per decade) over multiple decades and which ideally covers (most of) the globe at

Printer-friendly version

Discussion paper

high spatial resolution. Arosio et al. explore two established, complementary merging methods to combine measurements by two dense limb samplers SCIAMACHY (2003-2012) and OMPS-LP (2012-now) using a third limb sensor as transfer standard, Aura MLS (2005-now). Contrary to most earlier efforts by other groups, the authors attempt to preserve longitudinal information. The resulting data record is then analysed for trends over the 2003-2018 period using a widely used regression model. The authors discuss the spatial structure of the trends and they claim that longitudinal patterns are discerned that are indicative of changes in the Brewer-Dobson circulation. The paper concludes by extending the SCIAMACHY-OMPS data record to earlier decades using similar techniques and SAGE II measurements. Zonally averaged trends from the longer record confirm earlier findings for the 1985-1997 and 1998-2018 periods.

2 Recommendation

This paper fits the scope of AMT and would be suitable for ACP as well, since equal shares of the manuscript are devoted to merging methods and trend analysis results. I would recommend publication as long as the authors are willing to improve the discussion of several topics.

3 Major comments

3.1 Demonstrate that the longitudinal structures are realistic

My main criticism on this paper is that there is no substantial proof of the robustness of the reported longitudinal structure of the time series and derived trends.

A much more profound discussion is needed about the validity of the longitudinally-resolved results, especially since this is one of the central points of the paper. Constructing a lon-resolved data record is one thing, but the authors need to show that the longitudinal information in the data record is reliable and stable. This should have been the cornerstone of this paper, but it is entirely missing from the paper.

As an illustration (p.12, l.2): "A plot of the longitude-resolved drift values is shown in the Supplements, Fig. S1". But no discussion of key results follows: is there longitude structure in the drift field, or not? Another check would be to compare lon-resolved maps of trend results at neighbouring vertical levels to demonstrate their stability in the vertical domain. Once this validation step is over with, you could gain additional confidence by discussing how the derived trend fields compare to what is expected.

3.2 Elaborate discussion of merging technique

The authors present two merging methods and the resulting difference time series with respect to Aura MLS (Figs. 3-4). Unfortunately, they miss the opportunity to discuss merits and weaknesses of each of the methods and in what way one or the other method can correct for specific issues. I feel such a discussion in Sect. 3 would improve the paper a lot. In the end, readers of this paper will be interested in what you recommend as merging approach: plain-debiasing or anomalies? The answer to this naive question may depend on the use case, of course, but this should be part of the discussion.

For instance, Fig. 4 shows a discontinuity of in the anomaly-merged time series between 10°N-10°S at 31.5 and 34.8 km. What is the cause of the feature and why is it not present in the plain-debiased time series (Fig. 3)? The trends in the tropics (Fig. 6) are, surprisingly, not very different using both data records. How can this be? On the other hand, p.14, l.14 claims "The general picture in the two panels is very similar, even though trend values in panel (b) are slightly larger". Can this observation be linked to the merging strategy?

3.3 Absolute vs relative offset corrections

The adopted plain-debiasing method (Eqs. 1-2) removes additive biases but not the multiplicative biases. And vice versa, the adopted anomalies method (Eqs. 3-4) removes multiplicative biases but not the additive biases. Can the authors clarify the statement on p.8, l.8-9: "Through the merging process biases will be subtracted, whereas the discrepancies in the shape of the SC are accounted for when calculating anomalies (subtraction of the SC)."?

3.4 Substantiate claim about stability MLS seasonal cycle

p.8, l.3-4: "In addition, we notice that MLS SC may vary within the instrument life time, as shown at 34.8 km in the tropics with change of up to 5–7 % between the two periods." This is quite a bold statement which may worry the users of Aura MLS data. But this claim is not really substantiated by the authors. Should a reader be really worried about the stability of MLS data, while you mentioned earlier on that it is stable? When looking at Fig. 2 only one panel indicates that MLS 2005-2012 deviates clearly from MLS 2012-2016. I would like to see more proof/discussion if you want to keep the statement that MLS SC varies over its life time.

3.5 Collapse of longitude dimension

p.7, l.15: "In this paper we only describe the analysis of the longitudinally resolved ozone profile product". If you do not consider the zonally averaged data in this Section, why mention the binning at all? This is confusing as most of the plots in Sect. 3 are latitudinal cross sections. More importantly, the authors do not clarify in what order and how the different dimensions are collapsed from the underlying alt-lat-lon-time resolved data, perhaps because it has no importance but -in that case- it should be stated somewhere. For e.g. Figs. 3 and 4, did you collapse longitude dimension before computing the difference to MLS, or, first compute difference to MLS then average over longitude?

3.6 Diurnal variation

p.9, l.24-25 reads "Furthermore, at these altitudes diurnal variation of ozone have to be accounted for (Sakazaki et al., 2013), which was not done in this study". This message is repeated on p.15, l.5-6. The correction scheme Eq. 1-2 removes (additive) biases between data records, irrespective of the nature of the bias. Biases due to diurnal variation are part of the total bias. Hence, I infer that diurnal variations are accounted for contrary to what the authors claim. Can the authors respond to this reasoning, and incorporate their answer in the manuscript?

3.7 Impact of using ERA-Interim data to convert Aura MLS data?

The authors mention that the Aura MLS data record is stable (p.6, l.18-19). However, it is not clear from the paper whether this holds for converted Aura MLS data as well. Please elaborate on how the ERA-Interim data may impact the converted Aura MLS data. Can it induce the change in seasonal cycle reported on p.7, l.6-7? Can it lead to the drift above 50 km reported in p.9, l.22-23?

3.8 Correlation between solar and trend term

The MLR regression model contains a term for the 11 year solar cycle and a linear trend term (p.12, Eq.6). The analysis period (2003-2018) contains one and a half solar cycle, which triggers the question as to how independent the two said low-frequency terms are. Can the authors elaborate on this? Could the change in derived trend for different starting times (p.15, l.7-10) be related to interference between the solar and trend term, this is exactly the region where solar influence should be large.

This concern may even be more important for the results shown in Figs. 7 and 9 where even shorter periods are regressed. Perhaps in these cases the non-trend terms were regressed over the entire time period?

4 Minor comments

p.1, l.11: Be specific about what you mean with "remarkable variability".

p.2: Very nice and concise overview of ozone-related processes.

p.3, l.19: Identify "MLS" as "Aura MLS" here and throughout the rest of the paper. You don't want to confuse with the first MLS instrument which was flown in the 1990s-2000s on the UARS satellite.

p.3, l.19-21: You should introduce SAGE II over here, instead of two instruments (ACE-FTS and SAGE III/ISS) that are not mentioned in the rest of the manuscript.

p.3, l.25: Please rephrase. Harris et al (2015) did not merge these data sets, but use them to derive trends.

p.3, l.34: Remove "applying a multilinear regression analysis". This information is evident and not different from the other analyses you refer to.

p.4, l.1: Vague statement "Ball et al. (2018) applied a method independent from the ozone turnaround point". The subsequent clause "showed for the first time some evidence of a negative trend in lower stratospheric ozone" seems to imply that the different regression method is leading to this discovery. I am not sure that is what Ball et al. claimed.

p.4, l.4-5: Hanging statement: "This analysis has recently been challenged by Chipperfield et al. (2018)". In what way?

p.4, l.6: Clarify what a "pointing drift" is. A general reader will not have a clue what pointing means in this context. Consider vertical pointing, altitude registration, ...

p.4, l.15: I am not sure LOTUS is "homogenizing" the merging procedures, please double check this with one of the LOTUS participants.

p.5, l.12: "separated by 4.25deg at the tangent point." I assume you mean the viewing angles differ by 4.25deg.

p.5, Tab. 1: Extend this table to SAGE II and Aura MLS.

p.5, Tab. 1: Add level 2 versions in this table as well. This will make life easier for readers in 5 years from now.

p.5, Tab. 1: I advise to show the analysis time period for both instruments. Right now, different information is conveyed in "data time series": SCIA (full mission period) and OMPS (analysis time period). Please use one or the other, but do not mix up.

p.5, Tab. 1: Align values of spectral coverage and spectral resolution with what is in the main text.

p.5, l.20: Add the version of SCIATRAN.

p.6, l.1: Add the version of the SCIAMACHY L1 data, as was done for OMPS-LP.

p.6, l.9: Clarify "pointing knowledge issues", see also my earlier comment. E.g. "[...] when the issues related to the vertical pointing of the instrument, currently under [...]"

p.6, l.9: Refer to Moy, AMT 2017.

p.6, l.13: Replace "scientific measurements" by a better description or simply drop "scientific".

p.6, l.18: You mention Hubert et al. (2016) for Aura MLS drift. But any trend paper should refer to published drift results for all instruments involved. I.e. add those for SCIAMACHY (Rahpoe-2015, Hubert-2016, LOTUS-2018, ...?) and OMPS-LP (Kramarova-2018) as well, unless the L1-L2 versions have changed sufficiently to question the validity of those values.

p.6, l.18: Find a better phrasing for "For technical reasons" as it suggests an instrument malfunction. The observations by SAGE II are sparse due to the chosen measurement geometry and is unrelated to the instrument itself.

p.6, l.31: "Aura MLS", see earlier comment.

p.6, l.33: "taking only the latitude covered daily by OMPS-LP". You lost me here, what latitudes are not covered by OMPS? Please clarify what you mean in the main text. And why does this resolve an inconsistency? SCIAMACHY measurements are also made during daytime.

p.7, l.2: "Aura MLS", see earlier comment. Please incorporate this comment in the rest of the manuscript.

p.7, l.3: Add a motivation for not using the 2002 SCIAMACHY data.

p.7, l.11-12: Figure 1 does not confirm the statement "In both cases we find about 100 profiles on average in each bin." for SCIAMACHY. Each monthly 5deg zonal bin has 1000 profiles, which translates to 56 (=1000/18) profiles per bin, not 100. Did I misunderstand? If not, please change the misleading statement.

p.7, l.14: Clarify what interpolation method was used.

p.7, Fig.1: Add 5° at the end of the caption: "[...] in each 5° zonal monthly bin [...]".

p.8, Fig.2: Is the SCIAMACHY time period identical to that of Aura MLS (2005-2012)? Please add the time period for all four lines in the legend, not just for Aura MLS.

p.8, l.5-6: The phrasing is not clear whether the time period was only adapted for MLS. In other words, did you compare 2005-2012 for both SCIAMACHY and MLS, and

2012-2016 for both OMPS-LP and MLS? See previous comment.

p.9, l.14: Add a short phrase that the unit of the plain-debiasing data set is ozone number density.

p.9, l.16: "[...] differences between the merged data set [...]". What merged data set? The zonal one? The longitudinally resolved one?

p.9, l.16-18: What is the sign of the relative difference? (SCIAOMPS - MLS) / MLS or the other way around?

p.9, l.28-29: Are the larger relative difference values at 15 km truly due to lower data quality or due to the smaller number densities in the UTLS region?

p.9, l.30-31: Replace by "[...] deseasonalized relative anomalies from [...]" to clarify that you are not working with absolute anomalies.

p.9, l.30-31: What is the motivation behind debiasing the deseasonalized relative anomalies? By computing the anomaly any multiplicative biases are removed by definition.

p.9, l.31: Replace by "[...] month of the year, m, the (relative) anomalies, [...]".

p.10, Fig. 3 and p.11, Fig. 4: Add in the colour scale the exact sign of the difference: (merged - MLS) / MLS or (MLS - merged) / merged ?

p.10, Fig. 3 and p.11, Fig. 4: Add in the caption that MLS data has been offset to SCIA prior to the comparison.

p.10, l.4: Eq. 5 is not really used in the rest of the paper (p.12, l.7-8). I would therefore suggest to drop it, also because (a) you do not explain how the uncertainty for the plain-debiasing data is computed and (b) there is no term for the uncertainty in the seasonal cycle.

p.10, l.9: Replace by "Figure 4 shows the absolute differences [...]" to clarify that these are absolute differences of relative anomalies.

p.11, l.1: "whereas below 20 km the pattern becomes rather chaotic".

p.11, Fig.4: Larger differences are found around 35 km in 10S-10N during the OMPS-LP period. What is the cause of this? Does the different MLS SC in the two periods play a role?

p.11, l.4-5: "The drift is computed as the linear change of the differences between the merged time series and MLS data [...]". Are these relative differences for plain-biased merged data and absolute differences of anomaly-merged data? Clarify this in the text. And add the unit of the drift: % per decade/year/... .

p.12, l.1-2: "Very similar results for the drift are obtained using anomalies time series". The timeseries in Figs. 3-4 look fairly different in places, and I am surprised the drift results are very similar for the anomaly time series. This plot has to be in the main paper, also since it may be the basis of an interesting discussion on what merging technique led to most stable results for this particular case. (See also one of my major comments.)

p.12, Fig. 5: Add in the colour scale the full unit (% per decade/year/...) and the exact sign of the difference: (merged - MLS) / MLS or (MLS - merged) / merged ?

p.12, l.6: What are the units of O3 in Eq. 6? The plain-debiased time series are in molec cm^{-3}, the anomaly-merged time series are in %?

p.12, l.14-15: The phrase "The t-th row of the X matrix contains the values of the fit terms for the selected t." does not add information. It could easily be dropped.

p.13, l.1: The equivalence of the 2σ rule to 95% confidence level is introductory statistics, hence the reference to (Tiao et al., 1990) is not needed.

p.13, l.3: How are the plain-debiased time series in molec cm^{-3} regressed to obtain % per decade?

p.13, l.6-9: Do I understand you correctly that the EHF term is used instead of the harmonic terms, below 25 km and only for the 50-60°N band? Why not for 50-60°S as well, or at other latitudes? Please add that this modified regression model is not applied to the analysis of anomaly-merged time series.

p.13, l.23: Please cite more recent work, at least Maycock et al. (2016), perhaps others as well (Ball et al., 2016; Damadeo et al., 2018; ...).

p.14, l.1: Add the source of the El Nino 3.4 index data, as you did for the other proxy data sets.

p.14, l.6: Add that N_34 represents the El Nino 3.4 index anomaly data.

p.14, l.11-12 and Fig. 6: Are the trends in Fig. 6 regressed directly from the zonally averaged merged time series, or are these trend results regressed from lat-lon resolved merged time series then averaged over the latitude bands? In the first case, this contradicts an earlier statement that only analysis of lon-resolved data would be described (p.7, l.15). In the latter case, how do you propagate the uncertainties?

p.15, l.2: You may consider adding the LOTUS report and replace the WMO reference in this phrase (WMO, 2018).

p.15, l.9: Perhaps the cause is not instrumental, but related to the interference of the solar and trend term? See my major comment above.

p.15, l.35: Are all terms (seasonal, QBO, solar, ENSO, ...) constrained by the 2003-2011 period or just the linear trend term? This shorter period potentially makes the interference between solar and trend terms even larger. Have you looked into this? The trend results may be more stable/robust when you constrain all non-trend terms (especially solar) to the larger 2003-2018 period.

p.16, Fig. 7: Add complete unit to y-axis label: molec cm^{-3}.

p.16, Fig. 7: Adding the anomaly-merged time series and fits would make a fine illustration of how the merging strategy can overcome some of the issues in the data sets as mentioned e.g. in p.16, l.4-7.

p.16, l.1: There is the switch to OMPS-LP in 2012. Can this be a viable alternative explanation to the "discontinuity"? The fits themselves will, in addition, likely be impacted by the solar-trend interference as well.

p.16, l.11-12: Please substantiate why the longitudinal trend results are reliable? A figure like Fig. 8 for neighbouring levels z=41 and 44 km will help to demonstrate the stability of the results in the vertical domain, especially in the US where trends are mostly significant (also in other studies).

p.17, Fig. 8: Add to the caption what merged data set was used: plain-debiased or anomaly-based?

p.17, Fig. 8: Remove the results in the latitude-range that you mentioned earlier on was not reliable (60° for plain-debiased and 70° for anomaly-based).

p.17, l.10-11: Move this discussion to previous paragraph and elaborate on how stable results are in vertical domain.

p.17, l.15-16: Motivate why you use the anomaly approach.

p.18, l.6-7: Add brief explanation why the harmonic terms are not included (deseasonalized anomaly time series).

p.18, l.6-7: Slightly confusing, since the trend model is very different from that in previous section. Please clarify whether it is an independent trend (ILT) or a piece-wise trend (PWLT).

p.18, l.13: Add correct unit : "[...] about -2% per decade is detected [...]".

p.18, l.13: As asked before (p.15, l.35), what time period was used to constrain the non-trend terms? And how robust are -especially- the 2012-2018 trend results given the low frequency of the 11 year solar cycle proxy?

p.18, l.19: You mention 2010-2018 here, while the figure caption says 2012-2018. Which one is correct?

p.18, l.32: What do you mean with "up to" the polar regions?

p.19, Fig. 9: You mention 2012-2018 in the caption, while the main body text says 2010-2018. Which one is correct?

p.19, Fig. 9: Add the time period to each panel, in addition to (a), (b), ...

p.19, l.6-7: Strong claim that needs demonstration: how reliable is the observed lon-resolved structure?

Supplement, Fig. S1: Add in the caption which merged time series are shown: plain-debiased or anomaly?

Supplement, Figs. S1 and S2: Add sign and correct unit (% per decade?) to colour scale or in the caption.

Supplement, Figs. S1 and S2: Each subpanel represents one longitude-bin, all together they convey information about longitude structure of drift of SCIAOMPS wrt MLS. However, the longitude structure would be much more obvious if you would have shown one latitude-bin per subpanel, and then plot drift vs altitude vs longitude. Can you add this to the supplement?

Supplement, Figs. S3 and S4: Add to the caption what merged data set was used: plain-debiased or anomaly-based?

Supplement, Figs. S3 and S4: Remove lat-range with data that you claimed earlier in the paper are unreliable (poleward of 60° or 70° latitude).

5 Technical corrections

p.1, l.4-5: Replace by "[...] is performed by the processor of the University [...]"

p.1, l.10: Replace "high" horizontal sampling by "dense" horizontal sampling.

p.1, l.24: Remove either "important" or "key". Important implies key and vice versa.

p.2, l.7: Replace by "during the 1990s".

p.2, l.10: As non-native speaker I expected "GHGs such as", but perhaps "such" is not needed.

p.3, l.27: Add a "," in "[...] before 1998, and a positive trend of [...]".

p.3, l.29: Replace by "[...] with uncertainty estimates. [...]".

p.3, l.30: Add "-", replace by "[...] satellite and ground-based data sets [...]".

p.3, l.35: Replace by "[...] significant trends [...]".

p.3, l.35: Replace by "[...] in the upper stratosphere at mid-latitudes [...]".

p.5, l.4: Replace "[...] in-flight direction [...]" by "in flight direction" or "in the direction of flight".

p.5, l.13: Remove the first "charged" in "charged charged-coupled device".

p.5, l.22: Replace "application of SCIAMACHY retrieval scheme" by "application of SCIAMACHY's retrieval scheme".

p.6, l.3: Replace "we take into account in addition" by "we also take into account".

p.7, l.23: Replace by "[...] the SC of all single instrument data sets [...]".

p.8, l.3: Replace by "[...] the three ozone profile data records in number density [...]".

p.13, l.16: Replace by "[...] and the in-phase at mid-latitudes [...]".

p.13, l.32: Drop "a" in "[...] leading to longitudinally dependent modifications of ozone [...]".

p.14, l.17: Replace by "Bourassa et al. (2018)".

p.15, l.11: Replace "detected" by "found" or "observed". In my view, "detected" implies that the result is significant which is not the case.

p.15, l.33-34: Remove newline after "[...] panel (a).".

p.17, l.13: Add "s" to "[...] SAGE II occultation observations [...]".

p.18, l.4-5: Replace by "[...] the mean SAGE II latitude plus or minus its standard deviation [...]".

p.18, l.7: Replace by "[...] 60° latitude [...]".

p.18, l.29: Replace by "[...] is first removed [...]".

p.18, l.30: Remove "one" from "[...] MLS one [...]".

p.18, l.32: Replace by "[...] with respect to the MLS time series [...]".

p.19, caption: Replace by "[...] and in panel (d) over 2012–2018 [...]".

p.19, l.4: Replace by "[...] from 2003 until early 2018 [...]".

p.19, l.5: Replace "detected" by "found" or "observed". In my view, "detected" implies that the result is significant which is not the case.

p.19, l.10: Replace by "[...] has vanished when adding [...]".

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2018-275, 2018.