Reviewer comments in black text
Author responses in orange italic text
Additions to the manuscript in blue italic text

Authors' note:
An error was found in the time-matching of FTIR and MAX-DOAS columns, so the comparison plots (figures 6 and A1) have been updated accordingly.

Anonymous Referee #1 Received and published: 30 July 2020

The study of Ryan et al. (2020) presents time-series of formaldehyde tropospheric columns from two ground-based MAX-DOAS stations, located in Australia and New Zealand, and their comparisons with ground-based FTIR and satellite (mainly TROPOMI) data. This paper is relevant for a publication in AMT, because the monitoring of biogenic VOCs, especially in such a region that was lacking of such measurements, is of great importance for our understanding of the complex tropospheric chemistry. The paper provides information on the observed amounts of HCHO in this region, and on the seasonal cycles, both important for modelers. The MAX-DOAS data are compared carefully with FTIR and TROPOMI data, and the paper demonstrates the improvement of the TROPOMI HCHO data compared to the OMI satellite. The paper is also very well-structured and clear. Therefore, I recommend the publication of this paper in AMT, with a few minor comments, suggestions and questions (listed below) that should / could be addressed for some clarification.

I) Specific comments/suggestions:

- It could be nice to add a Table summarizing the main results discussed in the paper: the 2 ground-based MAX-DOAS data, with mean values, and seasonal amplitudes, and their error budget. I don't see any random and systematic uncertainty numbers given in the paper for MAX-DOAS, while I see error bars in e.g. the scatter plots (with Deming fit that is using the MAX-DOAS uncertainty). It would be clearer in a Table. Then, possibly in the same Table (or another one), the different comparisons results (slope, R2, Diff+/-Std, ErrTC,. . .), as in Table A1 first 3 lines: I think these lines deserve to be in the main text, not in the Appendix.
– The following text has been added in Section 2.3 by way of overview of random and systematic errors in the MAX-DOAS process:

"The errors associated with the MAX-DOAS retrieval include systematic errors, which derive primarily from HCHO cross section uncertainty which is around 9% (Vigouroux et al., 2009). Random errors include model parameter uncertainty (such as uncertainty in a priori parameters), estimated to be 10% following the methodology outlined in Ryan et al. (2018), along with retrieval noise and smoothing errors which were calculated in HEIPRO."

We have added the first table requested in Section 3.1, summarising the random, systematic and overall errors along with the mean HCHO column and seasonal amplitude at each MAX-DOAS measurement site. We have also added the requested comparison summary table incorporating all relevant statistics and error breakdown.

- Sect. 3.1, discussion Lauder vs Melbourne: I understand that the authors want to focus on the "background" explanation (normal conditions) of HCHO. However (suggestion), a discussion on the added possible impact of biomass burning could be interesting as well (e.g. the peak in January 2018 seen at both stations might be related to fires?). Can plumes be observed from TROPOMI this month passing over Melbourne and Lauder as well on specific days? Also, maybe a Google Earth-type map (or a vegetation map found elsewhere) could help for the discussion on the different type of vegetation (forest / agriculture,. . .), e.g. in replacement of Fig. 1? *– This is a good point; biomass burning is also likely to be a large source of oxidised VOCs especially in south-eastern Australia. We have included at line 215 the following (following from noting that the source directions of HCHO in Melbourne were from the north and east, the directions of the the most proximate forests): "The dominant source directions from forested and rural regions, along with summer time peak, are also consistent with biomass burning being a source of formaldehyde in Melbourne."*

We have updated Figure 1 to Google Earth-derived maps which allow vegetation to be seen, and aids in the interpretation of source directions.

The idea of using TROPOMI to study smoke plumes extending from south-eastern Australia, including over New Zealand, is a good one especially given the recent (2019-20) large fire events. We believe this work could comprise a paper all of its own and is outside the scope of this comparison study.

- Sect.3.2, scatter plots: I do not understand the correspondence between the slope provided in the scatter plots and the red line: in Fig. A1, the slope of the red lines look larger than 1 to me, while the number given are 0.64 and .0.71. Could you explain? *– thanks for alerting me to this error. The linear least squares regression equations have been printed by mistake, rather than the Deming regression equation (which is plotted). They have been updated now.*

- p.12, l.248: calculation of the error on the differences: in the EMAX_DOAS and EFTIR total uncertainty, do you include the smoothing error? It should be included for your comparisons of FTIR and original DOAS data (Rodgers and Connor, 2003, Eq. 13). And for comparisons of FTIR and smoothed DOAS data, then it becomes Eq.30 of C2 Rodgers and Connor (2003). It would be nice to have separate information on random and systematic part of the uncertainties. (of DOAS, and of the comparisons). However, the systematic part of the smoothing might be negligible in your case compared to the other sources, but it should be better to check and say it. *– yes, the smoothing error is incorporated in the errors used to calculate the error on the differences. Following from this reviewer comment, the calculation of the differences of the raw and smoothed columns, along with breakdown of random and systematic errors on the differences, have been carried out using the methodology in Vigouroux 2009 and inserted into the tables requested above.*

- p.14, l.278-298, discussion on Fig.7: may the enhanced TROPOMI HCHO columns on the east be (also) due to transport pathways from the polluted city?
The largest cities in the South Island lie to the south east and north east of Lauder, rather than directly to the east. Nevertheless, the point is in general terms a good one because the population density is highest along the east coast of the South Island. This has now been noted in the discussion of Fig. 5 on page 10:

"The highest population density in the South Island, including the cities of Dunedin and Christchurch, lies along the east coast. Given that the lifetime of formaldehyde is on the order of hours, transport on the order of a hundred kilometres is possible meaning the different source directions can reasonably be compared. On the available evidence it could be hypothesised that the agricultural and more densely populated eastern sector is a stronger source of formaldehyde to Lauder than the forested the west coast".

Also noted on page 10:

"...figure 7(b) supports the conclusion (from the MAX-DOAS measurements) that the highest formaldehyde amounts are in the agricultural and more densely populated eastern parts of the island."

- p.14, 290-292, averaging of averaging kernels: it is usually preferred to make all smoothing part before with individual data and ak, and then average the smoothed results. (cf von Clarmann, T. and Glatthor, N. , AMT, 2019). Maybe low impact in your case, did you check? *– thank you for alerting me to this interesting point. I had not checked initially, but now have, and find that the correct smoothing approach does not change the regression or bias results of the comparison.*

- p.16, l.318 - p.17, l.333: discussion comparisons DOAS-TROPOMI: maybe these results (slope, diff, seasonal cycle comparisons Fig.8, . . .) should be put in perspective with previous TROPOMI HCHO validation (Vigouroux et al. 2020), especially for similar stations (in HCHO amounts). Do these studies agree? Maybe the public routine validation report (not peer-reviewed) that can be found here: http://mpcvdaf.tropomi.eu/index.php/formaldehyde can also help with MAX-DOAS / TROPOMI comparisons made at a few stations, but the sites used there are more polluted sites. *– yes, good point, following from the revised calculation of column differences, comparisons have been made with Vigouroux 2020 (which is also the most pertinent reference for validation in the TROPOMI public routine validation report!).*

- p.17, l.340-342: the TROPOMI a priori profiles are also from chemical transport model. Did you compare the ones used for TROPOMI and for OMI retrievals? Are they so much different? A possibility to learn about the improvement made from OMI to TROPOMI would be to ask the satellite data providers (e.g. Isabelle De Smedt, BIRA-IASB). *– Good point, in fact according to De Smedt et al., 2018, the HCHO a priori profiles are calculated using the same chemical transport model in both OMI and TROPOMI retrievals. The key difference between the instruments lies in the resolution improvement. The text has been revised as follows:*

Previous text: "Another possibility (for discrepancy between OMI and TROPOMI comparisons) could be inappropriate a priori profiles, which are calculated in OMI using chemical transport models. Knowing that formaldehyde production from isoprene and monoterpene emissions is poorly simulated in south-eastern Australia using standard VOC mechanisms (Emmerson et al., 2016, 2018), relying on these mechanisms to produce useful formaldehyde a priori profiles in this area may be problematic. Previous studies (e.g. De Smedt et al., 2015; Wang et al.,2017a) found that agreement between OMI and MAX-DOAS measurements was found when using the MAX-DOAS a priori profiles to retrieve satellite columns; it would be interesting in future work to do the same for HCHO satellite-based retrievals over Australasia."

Revised text: "Given that both OMI and TROPOMI retrievals rely on a priori formaldehyde profiles calculated using the same chemical transport model (TM5, De Smedt 2018), a priori differences cannot explain the difference in the comparison. However, revious studies (e.g. De Smedt et al., 2015; Wang et al.,2017a) found that agreement between OMI and MAX-DOAS measurements improved when using the MAX-DOAS a priori profiles to retrieve satellite columns; it would be interesting in future work to do the same for HCHO satellite-based retrievals over Australasia. Examining the influence of a priori profiles calculated by chemical transport models on formaldehyde retrievals is also of particular interest in south-eastern Australia given that biogenic VOC emissions have been shown to be poorly simulated in this region (Emmerson et al., 2016, 2018)."

- P.18, l. 376: "This would begin to address the current Northern-Hemispheric bias in satellite validation studies". I do not understand this sentence, sorry. Which Northern Hemispheric bias? Which molecule? How the data in Australasia can help addressing a bias in Northern Hemisphere? Please, clarify. – *We refer here to a sampling bias rather than a numerical bias, that is, the lack of satellite validation results in the Southern Hemisphere. The sentence in question has been revised to "This would continue to address the lack of Southern-Hemispheric satellite validation studies using ground-based remote sensing."*

II) Minor or technical comments:

- abstract, p.1, l.5 + Sect.2.3, p7, l.139 + Sect.3.1,p.9,l.193 + Sect.conclusions,p.17,l.352: the numbers given for the mean columns at Lauder are not consistent in these sections. – *The correct value of molecules cm^{-2} has been updated.*

- abstract, p.1, l.12: I would define the partial columns. – *fixed, now reads "... partial columns (0-4 km)."*

- p.2, l.51: uses (not use) – *fixed*

- p.4, l.101-102: remove one of the two "however". – *fixed*

- p.6, legend box inserted in Fig.3(a): it is written Retrieved (blue) and Measured (Red). I guess "model" is the "Retrieved"? Maybe use the same terminology in legend of Fig.3 (and text) and the legend box in Fig.3a. – *fixed, Fig. 3(a) caption now reads "shows retrieved-measured dSCD comparison"*

- p.6, l.129: have, not had. – *fixed*

- p.6, l.131: "2◦ the lowest possible elevation angle at Lauder": it is written 1◦ in the Sect. 2.1. (p.4,l.86). One should be corrected. – *lowest is 2º, p.4 l.86 has been corrected*

- p.7, l.159: The 2% random error for FTIR measurement: is it the value for the specific day 8th January 2018 or for the mean of FTIR data? It is a low value compared to the 6% given in Vigouroux et al., 2018. If this is the value for the specific day, maybe give the total column value as well or the random error in absolute values. – *This value is the mean for the FTIR data.*

- p.8, l.164: “in October 2017, and is a low. . .” (add and ?) *- fixed*

- p.8, l.179: “(Gonzalez Abad et al. 2015) and are. . .” (add and ?) *– in fact I have removed the “are” from earlier in the sentence to make it read more clearly.*

- p.9,l.199: I would change for: “. . .as Melbourne is a large city, and/or to higher biogenic VOC. . .” (a combination of both being likely) *- fixed*

- p. 12, l.245: You should specify how the percentage is calculated (mean(relative diff) or mean(absolute diff)/mean(levels)), and specify what is levels: DOAS ?FTIR? mean of both

- p.17, l. 327 “The majority of. . .. line”: English? *- fixed (“are” removed).*

- p.18, l. 365: slope=0.61 (and not 0.81, right ?) *– yes, typo corrected*.

- p.18, l.374: “could contribute” (not could be contributed) *– fixed*

- p. 21, Figure A4, legend: Remove one of two repeated sentences *– fixed*