Responses to reviewers

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*

Anonymous Referee #2 Received and published: 4 August 2020

This paper presents a very interesting comparison of HCHO between two different ground-based measurement techniques: MAX-DOAS and FTIR as well as their comparison with satellite-based (TROPOMI and OMI) data. The research presented in this paper is relevant to the scientific community since in several occasions questions have been raised regarding the differences between both MAX-DOAS and FTIR measurement techniques as well as their comparison with satellite-based measurements. Moreover, this research presents novel results in the Australasian region. For these reasons I recommend the publication of this paper in AMT. A few specific and minor/technical comments are suggested below.

Specific comments/suggestions:

a) Page 8, line 173 and Page 9, lines 182-183: When explaining the criteria to filter satellite-based datasets it is mentioned that cloud-free data was used (cloudy scenes were avoided or excluded from the analysis), was there a threshold value used? *– Cloudy scenes were determined by filtering for the recommended quality control parameter < 0.5, which incorporates cloud radiance fraction (at 340 nm) < 0.5. The relevant text on page 8 has been updated:*

“The recommended quality control (QC) filtering was applied, excluding retrieved values where the QC flag was less than 0.5 (on a scale of 0-1), which ensures scenes with cloud radiance fraction (at 340 nm) < 0.5 are excluded from the comparisons.”

b) Page 21, Figure A4: When comparing ground-based measurements with OMI, which data field was used?, the “ColumnAmount” or the “ ReferenceSectorCorrectedVerticalColumn”? *– The “ColumnAmount” field was used. This has now been noted in the figure caption of A4.*

c) Page 18, line 376: Would it be possible to please clarify the phrase " This would begin to address the current Northern-Hemispheric bias in satellite validation studies.", Is this in reference to the representativeness or the amount of monitoring stations used for satellite validations in both hemispheres? – *Yes, a sampling bias of SH vs NH validations studies – see response to the same question from Reviewer 1.*

Minor/technical comments:

Page 2, line 37: please review the phrase "Photolysis and reaction with OH and photolysis" – *amended to "Photolysis and reaction with OH limit…"*

Page 6, Figure 3a: a parenthesis is missing on the "y" axis of Figure 3a. – *fixed*

Page 7, line 144: a parenthesis is missing after 2018) – *fixed*

Page 9, line 187: cycle of formaldehyde? – *fixed*

Page 10, line 216: the forested west coast – *fixed*

Page 11, Figure 5: panels (b) and (c) have as insets a different panel letter (a) and (b) – *fixed*

Page 14, line 298: methods use different... – *fixed*

Page 15, line 312: well with the... – *fixed*

Page 17, line 327: please review the phrase "points lie are within" – *fixed, "points lie within"*

Page 18, line 371: please review the phrase "will be critical tool" – *fixed, "will be a critical tool"*

Anonymous Referee #3 Received and published: 5 August 2020

This paper presents a comparison of HCHO columns from MAX-DOAS measurements in Australia and New Zealand against ground-based FTIR and satellite measurements. The ground-based measurements consist of two novel MAX-DOAS stations located at Broadmeadows and Lauder. The MAX-DOAS measurement period consists of more than 2 years of observations, which provides information of HCHO amounts on these regions. The comparison of MAX-DOAS and FTIR at Lauder station result in a linear correlation of 0.81, while the comparison of MAX-DOAS and satellite observations at Broadmeadows and Lauder result in correlations of 0.95 and 0.61, respectively. This study provides results of HCHO at southern hemisphere where long-term observations of VOCs are missing. In addition, this study demonstrated the improvement offered by high spatial resolution of TROPOMI measurements in comparison to previous satellite instruments. The topic of this work fits well within the scope of AMT, the main findings are well described, the paper is well structured. I recommend acceptance to AMT after addressing few specific comments below.

Page 4, line 93, it would be nice to have a summary table of the retrieval settings used in the HCHO retrieval - *We have decided to list retrieval settings for all instruments and processes*

(MAX-DOAS spectral settings, MAX-DOAS profile retrieval, FTIR and TROPOMI retrieval) rather than employ tables. With respect, we feel that the DOAS spectral settings are summarised clearly and succinctly enough in the text as is.

Page 4, line 100, why only these two fitting windows were selected for comparison? What is the impact of O_3 absorption on the selection of HCHO fit window? *– As noted in lines 98-99 and in more detail in lines 102-105, the two bands were chosen because they are the two most commonly used in the literature. This is because of a trade-off between incorporating five HCHO absorption bands in the fit (324.5-359 nm) with more chance of spectral interference from ozone, or only three HCHO bands in the fit (336-359 nm). The aim of the sensitivity study was to determine whether there was any improvement in the fit possible from including more spectral information.*

While O_3 absorption is undoubtedly more significant going towards the UV end of the wavelength ranges, any concern that extra O_3 absorption is problematic in the 324.5-359 range is alleviated by the fact that neither the residual RMS nor HCHO dSCD (figure 2)) are impacted in this sensitivity test.

Page 5, figure 2, how is the fit error estimated? *– Thank you for this question because it allows a more insightful analysis of why the 324.5-359 nm fitting range gave a lower fit error. The fit error quoted in Fig. 2 is the fit error calculated in QDOAS, resulting from the covariance matrix E_x of the optical density fit result. This is estimated by:*

$$E_x = \chi^2 (A^T A)^{-1}$$

where A contains the linear components of the fit and

$$\chi^2 = \frac{\sum_{i=1}^{M} (f^i)^2}{M - N}$$

with $(f^i)^2$ being the sum of squares of the residuals, M being the number of wavelengths included in the fit and N being the number of fitted parameters. Given that we know neither the result of the fit (Fig. 2(a)) nor residuals (Fig. 2(c)) were substantially impacted by the choice of wavelength range we can conclude that the improvement in fit error comes from the M-N term, i.e. the information content of the retrieval. In other words, more information is supplied to the retrieval using the longer wavelength range with no accompanying deterioration in the residual RMS. The text has been modified to make this point.

Original text: "As shown in figure 2, the change of wavelength range had minimal impact on the residual RMS throughout the day, and minimal impact on the magnitude of the HCHO dSCD. However, despite minimal change in dSCD, the fit error was greatly reduced by using the longer wavelength range, therefore the 324.5-359 nm range was adopted as the standard employed for fitting formaldehyde"

Modified text: "The calculation of fit error in QDOAS depends on the linear fit parameters, the residuals and the information content of the retrieval, which depends in turn on the number of wavelengths in the fit. Neither the residual RMS (Fig. 2(c)) nor magnitude of the dSCD (Fig. 2(a)) were substantially impacted by the choice of wavelength range, suggesting that the improvement in fit error for the 324.5-359 nm range (Fig. 2(b)) results from increasing the information content of the retrieval. As a result of the increased information content and resulting lower fit errors, the 324.5-359 nm range was adopted in this paper for formaldehyde."

Page 7, line 144, a parenthesis is missing after "2018)" *– fixed*

Page 8, line 171, what is the impact of the selection criterion o colocation distance around the station (~10 km) in the comparison? – *The comparison statistics do change for different TROPOMI resolutions sampled around the measurement location. We have re-run the comparisons at different sampling resolutions and added the following sentences in Section 3.3:*

"The resolution selection criterion did not have a large effect on the comparison, with a regression slope of 0.68 (monthly averages) for averaging TROPOMI 50 km either side of Lauder as opposed to 20 km."

"Considering TROPOMI sampled 10km and 50 km either side of Broadmeadows, regression slopes were 0.56 and 0.65 respectively, with the low bias of TROPOMI compared to MAX-DOAS at high HCHO consistent across sampling resolution"

Page 9, line 181, why OMI use a distance of about 25 km and not 10km as TROPOMI? Would it explain the poor agreement between OMI and MAX-DOAS measurements in Figure A4, more background is captured for OMI than TROPOMI? – *As noted in section 2.5, the quoted spatial resolution of OMI is 13x24 km, compared to TROPOMI 3.6×7.2 km (before 6 August 2019) and 3.6×5.6 km (after 6 August 2019). This is the reason why the OMI sampling for analysis was 25 km. However, an error was made in the manuscript with TROPOMI columns actually compared at 0.2 degree resolution rather than 0.1 degree owing to the small number of pixels per cell at Lauder. This has been amended in the manuscript. For the OMI comparison, OMI will sample a little more of the background but we do not consider that this effect should remove the ability to see the HCHO seasonality – it doesn't for TROPOMI at 25 km either side of the measurement location. We have added the following in the discussion of Fig. A4 at the end of Section 3.3: "…the fact that OMI is sampled 25 km either side of the measurement location compared to approximately 20 km for MAX-DOAS, thereby taking in more of the background. However, this could not explain why no seasonality is evident in the OMI results."*

Page 10, line 235, although the temporal agreement is good in overall, figure 6 shown large differences from 2018-10 to 2019-01 between both dataset. Would you please discuss the possible reason of differences on the temporal variability between both datasets? – *This is a good question. We cannot provide a definitive answer for the difference in terms of HCHO sources because there is limited information on formaldehyde in NZ. However, given the sensitivities of the instruments, we suspect that the difference between the agreement in summer between the instruments lies in lower HCHO plumes in summers 2016-17 and 2018-19 which can be better detected by the MAX-DOAS than FTIR. We have updated the text describing this in Section 3.2:*

Previous: "The month-to-month variation in formaldehyde is in especially good temporal agreement for summer (DJF) 2017-18, while the summer increase is less clearly captured by the FTIR in summer 2016-17 or 2018-19."

Revised: "The month-to-month variation in formaldehyde is in especially good temporal agreement for summer (DJF) 2017-18, while both the timing and magnitude of HCHO in summer 2016-17 and 2018-19 are poorly replicated by the FTIR. Due to the higher

sensitivity of the MAX-DOAS to the lower troposphere this suggests HCHO plumes were lower in 2016-17 and 2018-19, and therefore not captured as well by the FTIR, than summer 2017-18."

Page 13, figure 7, it would be nice to add the latitude and longitude coordinates in the figure. *– this has been done.*

Page 14, figure 8, what is the impact of HCHO amounts from fire emissions in Broadmeadows? Although, the station is located on the city and HCHO, however in last years many fires have been observed in southeast of Australia, which could emit and transport HCHO from far location similar to observed in recent study for Canada wildfires. – *good point, indeed we expect biomass burning should be a source of HCHO in Melbourne. We have discussed biomass burning above in the responses to Reviewer 1.*

Page 14, line 298, please replace "methods usec" by "methods use" *– fixed*

Page 21, figure A4, the low spatial resolution of OMI in comparison to TROPOMI could be a reason for the not variability of HCHO. Would you expect a similar result for instruments with morning overpass like GOME2 and lower spatial resolution than OMI? – this follows on from the previous point about Fig. A4 *– yes, the lower spatial resolution of OMI may help blur the seasonal variation somewhat. However, TROPOMI and the MAX-DOAS show a strong seasonality at this location and also a strong dependence of HCHO on wind direction (indicating there are clearly strong HCHO sources in the region). Therefore I am surprised at the lack of seasonality in OMI results. Regarding GOME2, the diurnal variation of HCHO at this location peaks in the middle of the day and afternoon, so it would be interesting to see whether the morning overpass would also make it difficult to observe strong seasonality. This is outside the scope of the present study however, which is primarily to compare MAX-DOAS with FTIR and TROPOMI.*

Comparison of formaldehyde tropospheric columns in Australia and New Zealand using MAX-DOAS, FTIR and TROPOMI

Robert G. Ryan[1,2], Jeremy D. Silver[1], Richard Querel[3], Dan Smale[3], Steve Rhodes[4], Matt Tully[4], Nicholas Jones[5], and Robyn Schofield[1,2,6]

[1]School of Earth Sciences, The University of Melbourne, Melbourne, Australia
[2]ARC Centre of Excellence for Climate System Science, Sydney, Australia
[3]National Institute for Water & Atmosphere Research, Lauder, New Zealand
[4]Bureau of Meteorology, Melbourne, Australia
[5]School of Earth, Atmospheric & Life Sciences, University of Wollongong, Australia
[6]ARC Centre of Excellence for Climate Extremes, Sydney, Australia

Correspondence: Robert Ryan
(rgryan92@gmail.com)

Abstract. South-eastern Australia has been identified by modelling studies as a hotspot of biogenic volatile organic compound (VOC) emissions, however long term observational VOC studies are lacking in this region. Here, two and a half years of MAX-DOAS formaldehyde (HCHO) measurements in Australasia are presented, from Broadmeadows in northern Melbourne, Australia and from Lauder, a rural site in the South Island of New Zealand. Across the measurement period from December 2016 to November 2019, the mean formaldehyde column measured by the MAX-DOAS at Lauder was $2.50 \pm 0.61 \times 10^{15}$ molec.cm^{-2} and at Broadmeadows was $5.40 \pm 1.59 \times 10^{15}$ molec.cm^{-2}. In both locations the seasonal cycle showed a pronounced peak in Austral summer (DJF) consistent with temperature-dependent formaldehyde production from biogenic precursor gases. The amplitude of the seasonal cycle at Lauder was 0.7×10^{15} molec.cm^{-2} while it was 2.0×10^{15} molec.cm^{-2} at Broadmeadows. The Lauder MAX-DOAS HCHO measurements are compared with 27 months of co-located fourier-transform infrared (FTIR) observations. The seasonal variation of Lauder MAX-DOAS HCHO, smoothed by the FTIR averaging kernels, showed good agreement with the FTIR measurements, with linear regression slope of 1.03 and R^2 of 0.66 for monthly averaged formaldehyde partial columns (0-4 km). In addition to ground-based observations, a clear way to address the VOC measurement gap in areas such as Australasia is with satellite measurements. Here we demonstrate that the Tropospheric Monitoring Instrument (TROPOMI) can be used to distinguish formaldehyde hotspots in forested and agricultural regions of south-eastern Australia. The MAX-DOAS measurements are also compared to TROPOMI HCHO vertical columns at Lauder and Melbourne; very strong monthly average agreement is found for Melbourne (regression slope of 0.61, R^2 of 0.95) and a strong agreement is found at Lauder (regression slope of 0.73, R^2 of 0.61) for MAX-DOAS vs TROPOMI between May 2018 and November 2019. This study, the first long term satellite comparison study using MAX-DOAS in the southern hemisphere, highlights the improvement offered by TROPOMI's high resolution over previous satellite products and provides the groundwork for future studies using ground based and satellite DOAS for studying VOCs in Australasia.

1 Introduction

Volatile organic compounds (VOCs) influence tropospheric oxidation chemistry in three key ways: oxidation by OH (thereby decreasing oxidation capacity), formation of OH (e.g. by photolysis of formaldehyde) and contributing to tropospheric ozone formation. Remote sensing of oxidised VOCs such as formaldehyde has been demonstrated as an effective method for constraining VOC emissions and for studying the role of VOCs in atmospheric reactivity (see Kefauver et al. (2014) and references therein).

Formaldehyde has atmospheric mixing ratios ranging from several hundred parts per trillion (ppt) in unpolluted marine air (Mahajan et al., 2010; Peters et al., 2012) to tens of parts per billion (ppb) in polluted urban air (e.g. Zhu et al., 2017). Primary sources of formaldehyde include direct emission from fossil fuel combustion and wild fires. The main secondary sources of HCHO are oxidation of methane, isoprene and monoterpenes. Methane is considered to be the primary background HCHO source globally (Pfister et al., 2008), and because it is a potent greenhouse gas, studying background formaldehyde levels has important climate change implications. Isoprene and monoterpenes emitted from vegetation constitute the main source of biogenic carbon to the atmosphere (Guenther et al., 2012). While methane is considered the most important OH sink in background oceanic air, over land isoprene and monoterpenes constitute the largest OH reactivity and hence these biogenic VOCs play a crucial role in determining oxidative capacity (Fuentes et al., 2000; Lelieveld et al., 2008). Isoprene and monoterpenes are also thought to play a strong role in the climate system through radiative forcing by secondary formation of organic aerosols (Henze et al., 2008). Photolysis and reaction with OH ~~and photolysis~~ limit the lifetime of formaldehyde to several hours during the daytime which facilitates the comparison of co-located measurements and also means that spatially resolved HCHO measurements closely resemble the distribution of its VOC sources (Zhu et al., 2016).

Biogenic VOC emissions in Australasia are among the highest in the world due to the abundance of Australian endemic eucalyptus trees, known to be high isoprene and monoterpene emitters (Winters et al., 2009; Guenther et al., 2012). Global scale modelling has suggested Australia has the highest isoprene-derived formaldehyde levels of any other continent (Pfister et al., 2008), however constraining biogenic VOC emissions has so far proven challenging in Australia. Formaldehyde measurements, for example from satellites, are common proxies for biogenic VOC emissions but the accuracy of this under low NO_x conditions has not been observationally verified (Zhu et al., 2016; Wolfe et al., 2016), likely due to uncertainties in differentiating HCHO from different anthropogenic, isoprene and monoterpene sources. Emmerson et al. (2016, 2018) highlighted this by demonstrating that the MEGAN biogenic emissions scheme, used in numerous global and regional-scale chemistry/climate models, overestimates isoprene and underestimates monoterpenes in the thickly eucalyptus-forested south-east of Australia. Therefore, reliable, long term biogenic VOC measurements are needed in the Australasian region.

The Multi-axis differential optical absorption spectroscopy (MAX-DOAS) technique, a passive spectroscopic method which uses scattered solar radiation, can facilitate this through measurement of formaldehyde. In the last decade HCHO MAX-DOAS measurements have been reported from many locations in the world (Hoque et al., 2018a, b; Heckel et al., 2005; Pinardi et al., 2013; Peters et al., 2012; Vigouroux et al., 2009), however none so far have been reported in Australasia.

Figure 1. Maps of (left) south-eastern Australia showing the location of Melbourne and the Broadmeadows measurement site and (right) the South Island of New Zealand showing the locations of Lauder and the largest South Island cities, Dunedin and Christchurch.

Developments in satellite sensors and retrievals of atmospheric trace gases over the past two decades can offer new insights into air quality and composition (Martin, 2008). Validation by ground-based instrumentation is an important step in understanding the utility of such satellite data products. Because satellite instruments and MAX-DOAS share the same spectroscopic technique for retrieving UV and visible absorbing trace gases, MAX-DOAS is an ideal validation tool as demonstrated for HCHO in several previous papers (e.g. Chance et al., 2000; Thomas et al., 1998; Hoque et al., 2018b; De Smedt et al., 2015; Vigouroux et al., 2009; Lee et al., 2015; Kurosu et al., 2007). However no such validation studies have so far been published for the Australasian region.

Measurements in two locations are discussed in this paper: Broadmeadows, on the northern fringe of Melbourne in south-eastern Australia and Lauder, a remote locality in the South Island of New Zealand, as shown in the map in figure 1. Australia's Bureau of Meteorology has operated an EnviMeS MAX-DOAS instrument (similar instruments demonstrated good performance at the CINDI-2 MAX-DOAS intercomparison, see Kreher et al. (2019)) on a laboratory roof at its training facility at Broadmeadows (37.690S, 144.947E, 110 m amsl) since December 2016. This location is close to some significant pollution sources including factories and major roadways. MAX-DOAS measurements of nitrogen dioxide and nitrous acid at the Broadmeadows site have been reported in Ryan et al. (2018).

Lauder is located in Central Otago, New Zealand (45.038S, 169.684E, 370 amsl), surrounded by irrigated farmland, ringed by distant mountain ranges and lying approximately 30 km north-east of the nearest large town, Alexandra. An EnviMeS MAX-DOAS has been operational at Lauder since November 2016 allowing a significant period of overlap between the Lauder and Melbourne timeseries'. The NIWA EnviMeS MAX-DOAS demonstrated good performance at the CINDI-2 international comparison campaign held in the Netherlands in 2016 (Kreher et al., 2019).

Both Broadmeadows and Lauder have regular co-located meteorological, aerosol, radiation and and trace gas measurements; the Lauder site is part of numerous international atmospheric monitoring networks (De Mazière et al., 2018; Pollard et al.,

2017; Tradowsky et al., 2018). In addition, formaldehyde vertical columns measured at Lauder using fourier-transform infrared (FTIR) spectroscopy (Vigouroux et al., 2018) are available for comparison with the MAX-DOAS measurements.

The paper is structured as follows: the methodology section presents the MAX-DOAS and FTIR HCHO retrieval approach used in this work. Section 3.1 presents the MAX-DOAS HCHO timeseries from Lauder and Broadmeadows, briefly discussing potential HCHO sources at each location. Section 3.2 compares the FTIR and MAX-DOAS HCHO results from Lauder and finally section 3.3 presents a comparison study for TROPOMI HCHO vs MAX-DOAS at both Lauder and Broadmeadows.

2 Methodology

2.1 MAX-DOAS measurements

MAX-DOAS measurements at Broadmeadows were made with a 2D EnviMeS instrument pointing to a fixed azimuth direction of 208°. The measurement, completed over 12 minutes, consisted of the elevation angles 90°, 30°, 20°, 10°, 5°, 3°, 2° and 1° as described in Ryan et al. (2018). At Lauder, a 1D EnviMeS instrument was used pointed at a fixed azimuth of 30° and the elevation angles used were 90°, 40°, 20°, 10°, 5°, 3° and 2°~~, and 1°~~. Dark current and offset corrections were made for each dataset using calibration spectra collected nightly, while initial wavelength and lineshape calibrations were facilitated by laboratory-measured mercury emission lamp spectra.

2.2 MAX-DOAS spectral analysis

The MAX-DOAS data analysis process consists of two parts: calculation of differential slant column densities (dSCDs) from the raw spectra and an inversion algorithm to retrieve vertical trace gas profiles from the dSCD information. The spectral retrieval was done in QDOAS (http://uv-vis.aeronomie.be/software/QDOAS/). Cross sections used in the analysis were NO_2 at 220 and 298 K (Vandaele et al., 1998), O_4 at 298 K (Thalman and Volkamer, 2013), O_3 at 223 and 243 K (Serdyuchenko et al., 2014), HCHO at 297 K (Meller and Moortgat, 2000), BrO at 223 K (Fleischmann et al., 2004), HONO at 298 K (Stutz et al., 2000) and a Ring cross section at 250 K (Grainger and Ring, 1962). All cross sections were pre-convolved with the line shape of the instrument and 5th order polynomial and second order offset terms were also included in QDOAS.

Differential slant columns (dSCDs) of O_4, used in MAX-DOAS aerosol retrievals, were determined using the wavelength range 338-370 as in Ryan et al. (e.g. 2018); Kreher et al. (e.g. 2019). A simple sensitivity study was run to determine the appropriate wavelength range for formaldehyde retrieval given that two wavelength ranges are common in previous papers: 324.5-359 nm and 336-359 nm. Formaldehyde absorption bands for formaldehyde are in theory measurable by the MAX-DOAS UV spectrometers used in this work down to 300 nm. Published research to date, however, tends to avoid fitting below 320 nm ~~however~~ due to strong ozone absorption. Retrieval strategies in other work use a fitting range of 336-359 nm (e.g. Kreher et al., 2019; Heckel et al., 2005; Pinardi et al., 2013; Vigouroux et al., 2009) encompassing the three highest UV HCHO absorption features. Here a simple sensitivity study was run to determine if any benefit can be derived from additional absorption bands in the extended range (e.g. Chan et al., 2019; Johansson et al., 2009; Wang et al., 2017b; Franco et al., 2015).

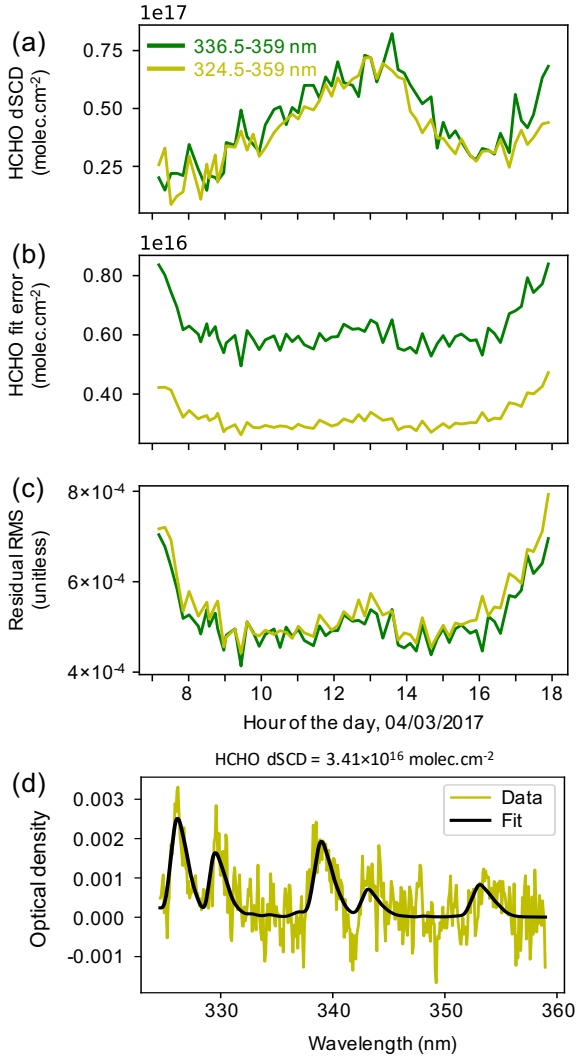

Figure 2. (a)-(c) Results of sensitivity tests to determine the appropriate fitting range for formaldehyde, using 3° elevation angle data from 04/03/2017 at Broadmeadows, showing dSCD, dSCD fit error and residual RMS respectively. (d) Example 3° elevation angle HCHO DOAS fit from 11 am local time on the same day.

Data for this test was chosen from a clear sky Autumn day at Broadmeadows with maximum HCHO dSCDs $\approx 7.5 \times 10^{16}$ molec.cm^{-2} at 3° elevation angle. ~~As shown in figure 2, the change of wavelength range had minimal impact on the residual RMS throughout the day, and minimal impact on the magnitude of the HCHO dSCD. However despite minimal change in dSCD, the fit error was greatly reduced by using the longer wavelength range, therefore the 324.5-359 nm range was adopted as the standard employed for fitting formaldehyde.~~ The calculation of fit error in QDOAS depends on the linear fit parameters,

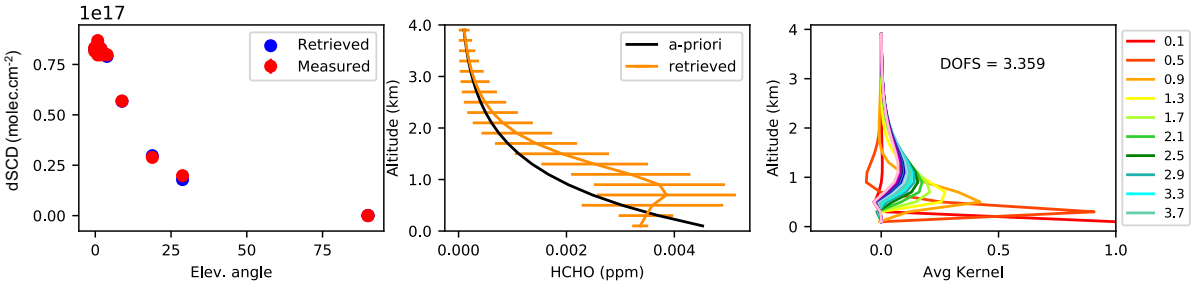

Figure 3. Example HEIPRO retrieval from Broadmeadows on 13/12/2017 at 14:00 local time. (a) shows ~~model-measurement~~retrieved-measured dSCD comparison, (b) shows the retrieved and a priori profiles, (c) the averaging kernels for this retrieval.

the residuals and the information content of the retrieval, which depends in turn on the number of wavelengths in the fit. Neither the residual RMS (figure 2(c)) nor magnitude of the dSCD (figure 2(a)) were substantially impacted by the choice of wavelength range, suggesting that the improvement in fit error for the 324.5-359 nm range (figure 2(b)) results from increasing the information content of the retrieval. As a result of the increased information content and resulting lower fit errors, the 324.5-359 nm range was adopted in this paper for formaldehyde. An example HCHO DOAS fit is shown figure 2(d) demonstrating the convincing retrieval of formaldehyde dSCDs using the extended range.

2.3 MAX-DOAS profile retrievals

Formaldehyde vertical columns and profiles from Broadmeadows and Lauder were retrieved from dSCDs using the HEIPRO inversion algorithm (Frieß et al., 2006). HEIPRO has previously been used for NO_2 and HONO gas profile retrievals at Broadmeadows (Ryan et al., 2018). In an initial step, aerosol profiles were determined from dSCDs of the O_4 dimer. These were used as input information on the light path for calculating air mass factors and HCHO vertical column density (VCD) in the second retrieval step. Vertical profiles were retrieved on a 20 layered grid with 200 m resolution from 0-4 km, aerosol retrievals were calculated at 360.8 nm and HCHO retrievals at 338.9 nm. A priori profiles used in the inversion were chosen to be exponentially decreasing functions of altitude, characterised by a set surface mixing ratio and scale height, which were 0.5 ppb and 1 km respectively for formaldehyde. HEIPRO was run in 15 minute intervals ensuring that each measurement set contained a full set of elevation angles. MAX-DOAS retrievals were filtered for results with less than one independent piece of information and for the presence of clouds. At Broadmeadows this was determined using an empirical algorithm based on colour indices (e.g. Gielen et al., 2014; Wagner et al., 2014, 2016), also described in Ryan et al. (2018), and at Lauder using the SkyNet AOD flag which is calculated using the method outlined in Khatri and Takamura (2009).

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.

An example MAX-DOAS HCHO retrieval from HEIPRO is shown in figure 3, including model-measurement comparison, retrieved and a priori profile and averaging kernels. These example averaging kernels at Broadmeadows show highest sensitivity at the surface, and 3.4 degrees of freedom for signal (DoFs). The Lauder retrievals consistently ~~had~~have reduced surface sensitivity and lower DoFs compared to Melbourne, which is likely related to lower amounts of formaldehyde at Lauder and the fact that 2° is the lowest possible elevation angle for MAX-DOAS at Lauder due to proximate mountain ranges. Across the whole measurement period, the average DoFs was 2.25 ± 0.34 (1σ) at Broadmeadows and 1.27 ± 0.11 (1σ) at Lauder. Detection limits for the MAX-DOAS vertical column densities at Lauder and Broadmeadows have been estimated using the method outlined in Peters et al. (2012):

$$DL_{VCD} = \frac{2 \times R_{avg}}{XS_{max} \times A - 1} \quad (1)$$

where R_{avg} is the average residual RMS, XS_{max} is the maximum value of the cross section (1.32×10^{-19} for HCHO) and A is the airmass factor taken here as 15 for low elevation angles. R_{avg} was 4.5×10^{-4} at Broadmeadows giving $DL_{VCD}(HCHO)$ as 4.9×10^{14} molecules.cm^{-2}. The average residual RMS was lower at Lauder, 2.9×10^{-4}, giving a calculated detection limit of 3.2×10^{14} molecules.cm^{-2}. Over the whole measurement period, the average vertical column was $2.50 \pm 0.61 \times 10^{15}$ molec.cm^{-2} at Lauder and $5.40 \pm 1.59 \times 10^{15}$ molec.cm^{-2} at Broadmeadows, meaning that HCHO VCDs were generally above the detection limit but measurements at Lauder were closer to the detection limit than at Broadmeadows.

2.4 FTIR retrieval

Solar FTIR measurements have been made since the early 1990s at Lauder as part of the Network for Detection of Atmospheric Composition Change (NDACC, (Jones et al., 1994; De Mazière et al., 2018)). Measurements are made on all possible clear sky days, throughout the day, using Bruker high resolution (0.0035cm^{-1}) spectrometers (www.bruker.com).

Initial retrievals of HCHO from the Lauder 1992-2005 FTIR dataset are described in detail in Jones et al. (2009). The HCHO retrieval strategy (under the auspices of the NDACC infrared working group) was harmonized across the network as detailed in Vigouroux et al. (2018). Lauder spectra HCHO reprocessing was part of this harmonization activity and is the retrieval strategy used to provide HCHO data in this study. The same HCHO dataset is also used in a TROPOMI comparison study comprising of globally distributed ground based FTIR measurements (Vigouroux et al., 2020). These studies show that HCHO abundances over Lauder exhibit a seasonal cycle peaking in the summer (DJF).

Pertinent to this study, and paraphrasing details in Vigouroux et al. (2018), the Lauder FTIR retrievals are performed on a 48-layer atmosphere (0.37-100km) of which 15 layers are between 0.37 and 10km. The retrievals use a static a priori originating from WACCM_v4 climate-chemistry model simulations (Garcia et al., 2007) and the retrievals are constrained using Tikhonov regularization (L1, alpha = 100). Combined with a measurement signal to noise ratio of 400, the retrieval strategy has sensitivity over the altitude range 0.37-26km with an average total column DoFs of 1.4±0.2 (1σ). The highest sensitivity is in the upper

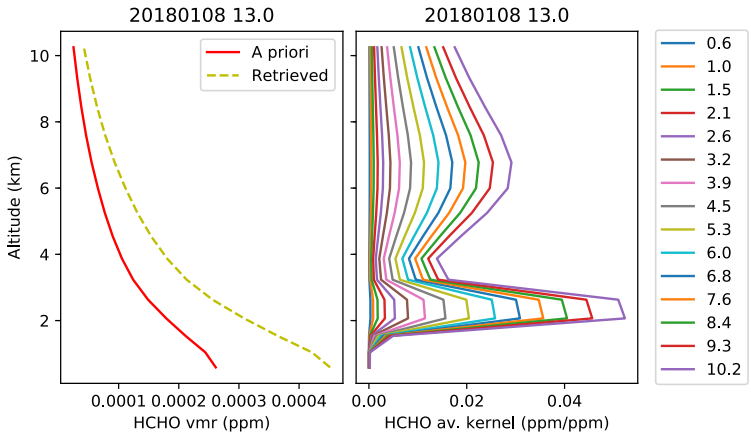

Figure 4. Example FTIR profile retrieval from 1 pm local time on 8th January 2018 at Lauder. The left hand panel shows a priori and retrieved HCHO profiles in the lowest 10 km, the right hand panel shows the corresponding averaging kernels.

troposphere peaking at 8km with a full width maximum of 16-18km. This differs from the MAX-DOAS measurements which has maximum sensitivity in the boundary layer. An example Lauder FTIR formaldehyde retrieval from 8th January 2018 is shown in figure 4. Attributed uncertainty analysis of the total column measurement gives an estimate of $\approx 2\%$ and $\approx 12\%$ for random and systematic error respectively. The systematic error is dominated by spectroscopic line strength uncertainty whilst the major component of the random error is measurement noise.

2.5 Satellite details

The TROPOspheric Monitoring Instrument (TROPOMI) is a nadir viewing imaging spectrometer aboard the European Space Agency's Copernicus Sentinel 5 Precursor (S5P) satellite. S5P, launched in October 2017 and is a low (afternoon) polar orbit ($\approx$ 824 km) mission providing daily global coverage for a range of UV, visible and infrared absorbing trace gases (Veefkind et al., 2012). The S5P overpass time is 13:30 (local time) and the spatial resolution of TROPOMI is 3.6 $\times$ 7.2 km (before 6 August 2019) and 3.6 $\times$ 5.6 km (after 6 August 2019).

Formaldehyde slant column densities are retrieved from analysis of absorption features over the wavelength range 328.5 - 359 nm. The SCDs are converted to vertical columns using airmass factors calculated at 340 nm using HCHO a priori vertical profiles simulated by the TM5-MP global chemistry-transport model as described in De Smedt et al. (2018).

For this study, TROPOMI data was regridded to $0.1 \times 0.1°$, (approximately 10 km). The recommended quality control (QC) filtering was applied, excluding retrieved values where the QC flag was less than 0.5 (on a scale of 0-1), ~~and cloudy scenes as determined by the satellite results were also excluded from the comparisons~~which ensures scenes with cloud radiance fraction (at 340 nm) < 0.5 are excluded from the comparisons. Given the satellite overpass was around 13:30 local time, MAX-DOAS results between 13:00 and 14:00 were averaged for the comparisons.

The Ozone Monitoring Instrument (OMI) is also a UV/Vis nadir viewing spectrometer providing near global daily coverage, housed on the National Aeronautics and Space Adminsitration's Earth Observing System Aura satellite (Levelt et al., 2006). The spatial resolution of OMI is 13 × 24 km and the overpass time is also around 13:30 local time. Formaldehyde slant columns ~~are~~ retrieved from OMI using a wavelength range of 327.5 - 356.5 nm (González Abad et al., 2015) are used along with GEOS-Chem simulated a priori profiles to calculate HCHO vertical columns (Bey et al., 2001). For comparison with the Broadmeadows MAX-DOAS dataset, OMI HCHO columns were regridded to 0.25×0.25°, meaning that columns approximately 25 km either side of the measurement site were used, and as with TROPOMI, cloudy scenes were excluded from the comparison.

3 Results and Discussion

3.1 Lauder vs Melbourne HCHO

The timeseries' of monthly formaldehyde vertical columns from Broadmeadows and Lauder MAX-DOAS measurements are presented in figure 5(a). Following the example of Jones et al. (2009), the seasonal cycle of formaldehyde was fitted with a cosine function described by the following equation:

$$C(t) = a_0 + a_1 t + a_2 \cos K(t - \phi) \tag{2}$$

where $C(t)$ is the formaldehyde vertical column as a function of time (in units of days since 1st January 2016), ϕ is the phase term with units of day of the year and $K = 2\pi/365$. Also fitted in the linear regression are a_2 (amplitude of the seasonal cycle), a_0 (the initial mean column amount) and a_1 (the magnitude of the linear trend in HCHO over time). At Lauder, the mean HCHO VCD was 2.5×10^{15} molecules.cm^{-2} and the amplitude of the fitted seasonal cycle was 6.9×10^{14} molecules.cm^{-2} while at Broadmeadows the average HCHO VCD was 5.4×10^{15} molecules.cm^{-2} with a seasonal cycle amplitude of 2.0×10^{15} molecules.cm^{-2}. The HCHO seasonal cycle from Lauder MAX-DOAS measurements is consistent with that found from FTIR measurements at Lauder from July 2002-July 2017 (Vigouroux et al., 2018). The fact that both the magnitude of the HCHO VCDs and amplitude of the seasonal cycle are much smaller at Lauder than Broadmeadows could be due to higher anthropogenic VOC precursors as Melbourne is a large city and/or ~~could be~~ due to higher biogenic VOC emissions from forests surrounding Melbourne.

The seasonal cycle of formaldehyde shows a distinct austral summer peak in both locations. This would be expected from biogenic production of formaldehyde, for example from isoprene, which depends strongly on temperature (Duncan et al., 2009; Palmer et al., 2006; Zhu et al., 2014). The phase of the cosine fit in each location is 31 days indicating that the HCHO seasonal cycle peaks at the end of January. This is also consistent with the results for Lauder in Vigouroux et al. (2018) and suggests that the same background mechanisms may be responsible for summer-time HCHO production at Lauder and Broadmeadows.

Polar bivariate plots showing the relationship between formaldehyde and wind direction and speed at Broadmeadows and Lauder are shown in figure 5(b) and (c) respectively. At Broadmeadows, HCHO concentrations are highest with wind from the

northern and eastern sectors, aligning with the direction of rural and densely forested regions, suggesting an important role for biogenic HCHO sources in this location. 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. At Lauder, maximum column amounts correspond with moderate wind speeds from the east. Over the course of the MAX-DOAS dataset, the wind came from this direction less than 10 % of the time, the same key source directions including the strong 'easterly maximum' are observed in polar bivariate plots of the 2001-2019 FTIR dataset (not shown). There is a large variation in vegetation types across New Zealand's South Island, including temperate rainforest in the west, dryland agricultural in the Central Otago region and intensive irrigated pasture in much of the east, south and south east, which might be expected to produce different volatile organic emissions and formaldehyde amounts. 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 kilometers 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.

Table 1. Summary of MAX-DOAS formaldehyde results and error budget from Lauder and Broadmeadows

Site	Mean column (molecules.cm^{-2})	Seasonal Amplitude (molecules.cm^{-2})	Smoothing error %	Noise error %	Model params. error %	Systematic error %	Total error %
Lauder	2.5×10^{15}	6.9×10^{14}	13	9	10	9	21
Broadmeadows	5.4×10^{15}	2.0×10^{15}	12	11	10	9	21

3.2 MAX-DOAS vs FTIR at Lauder

One previous study highlights a comparison between MAX-DOAS and FTIR formaldehyde columns, from the tropical Reunion Island (Vigouroux et al., 2009). In that paper, the comparison period was four months. In this work, co-located measurements over a period of 27 months are compared, from November 2016 to January 2019, allowing comparison of HCHO over two annual cycles. The comparison method used here has been adapted from Vigouroux et al. (2009) and Rodgers and Connor (2003). Partial column amounts have been compared in the lowest 4 km of the atmosphere, the region of expected formaldehyde production and the region of highest sensitivity for MAX-DOAS measurements. Because the FTIR instrument is less sensitive to the HCHO partial column in the lowest 4 km as evident from the averaging kernels in figures 3(a) and 4), the MAX-DOAS partial columns have been smoothed by the FTIR total averaging kernel using the method outlined in Vigouroux et al. (2009).

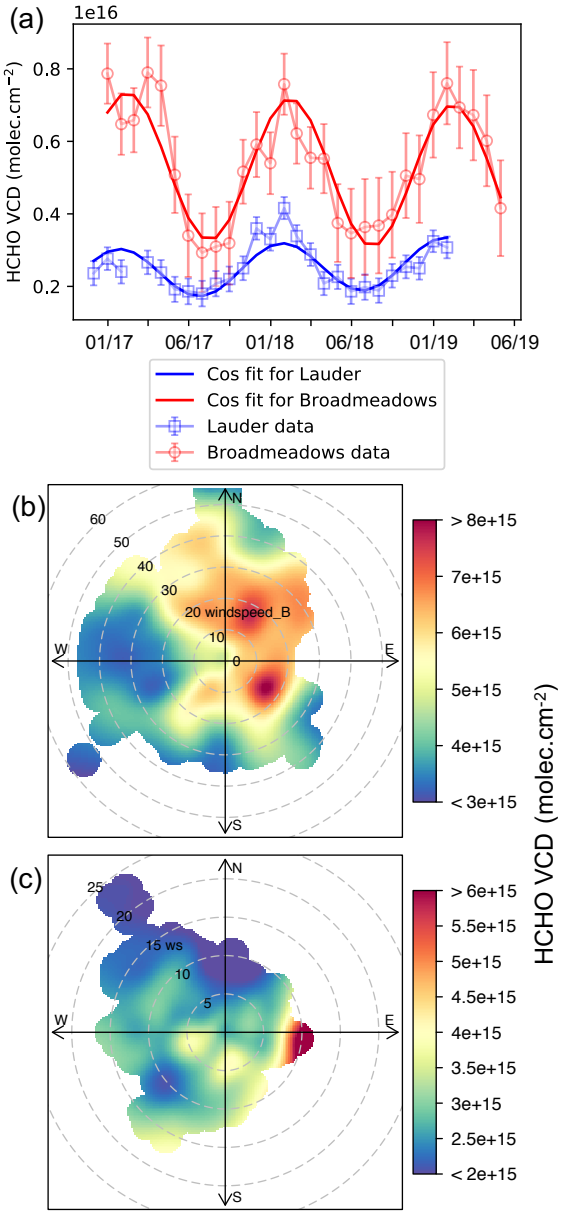

Figure 5. (a) Timeseries of monthly averaged MAX-DOAS formaldehyde VCDs at Broadmeadows (red points) and Lauder (blue squares). Cosine fits to each dataset are also plotted (solid lines), representing the seasonal cycle of HCHO at each location. (b) Polar bivariate plot showing the relationship between HCHO VCD and windspeed (in km/h) and direction at Broadmeadows across the measurement period. (c) Same as (b) but for Lauder.

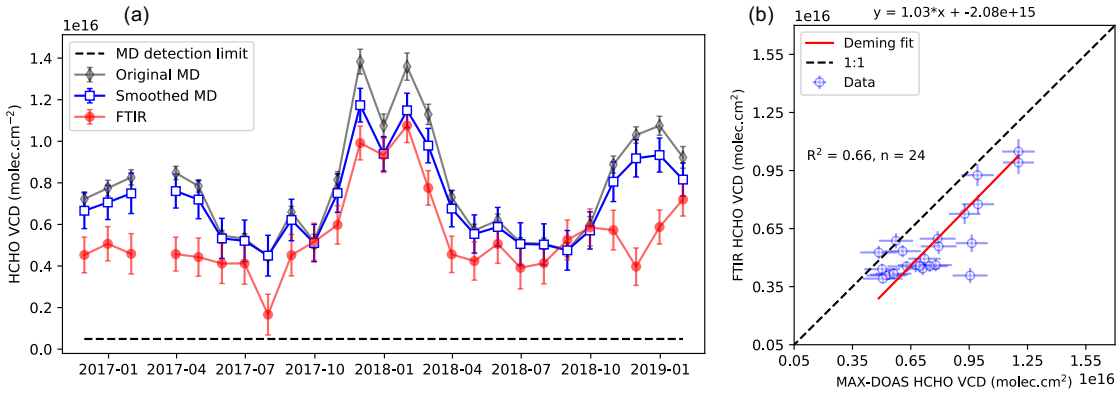

Figure 6. (a) Timeseries comparison of monthly averaged FTIR and MAX-DOAS (original and smoothed with FTIR averaging kernels) HCHO vertical columns from November 2016 to January 2019 at Lauder. (b) Scatter plot comparison of FTIR with MAX-DOAS smoothed HCHO columns.

As in Vigouroux et al. (2009), the equation for the smoothing is simplified by the fact that the same a priori profile was used to retrieve MAX-DOAS and FTIR profiles, allowing the smoothed DOAS column to be given by:

$$C_{DOAS,smooth} = C_a + \mathbf{A_F}(\mathbf{x_D} - \mathbf{x_a}) \tag{3}$$

where $\mathbf{A_F}$ is the FTIR total column averaging kernel matrix (from 0-4 km), which has units of VMR/VMR, C_a is the common a priori column amount, $\mathbf{x_D}$ is the original retrieved MAX-DOAS profile, $\mathbf{x_a}$ is the common a priori profile and $C_{DOAS,smooth}$ is the smoothed MAX-DOAS column amount. Only columns between 8 am and 6 pm local time contributed to the monthly averages examined here.

The timeseries of monthly averaged results is presented in figure 6(a), showing that both measurements capture the same broad seasonal cycle at Lauder and that monthly average columns for both measurements were clearly above the calculated MAX-DOAS detection limit. The month-to-month variation in formaldehyde is in especially good temporal agreement for summer (DJF) 2017-18, ~~while the summer increase is less clearly captured by the FTIR in summer 2016-17 or 2018-19~~ while both the timing and magnitude of HCHO in summer 2016-17 and 2018-19 are poorly replicated by the FTIR. Due to the higher sensitivity of the MAX-DOAS to the lower troposphere this suggests HCHO plumes were lower in 2016-17 and 2018-19, and therefore not captured as well by the FTIR, than summer 2017-18. There is a clear offset between the MAX-DOAS and FTIR columns with the FTIR consistently lower across the comparison period. Comparing the measurements by linear (Deming) regression, (incorporating errors in both the x and y ordinates), the offset is found to be 2.92×10^{15} molecules.cm^{-2} and nearly constant as indicated by the regression slope (1.17, see figure 6(b)). The timeseries also shows that smoothing the DOAS partial columns brought them more into line with the FTIR columns, especially in the peak months (November-March). The R^2 value

of 0.65 (n=27) for the regression in figure 6(b) highlights the moderate temporal agreement. Considering daily averages, a slope of 1.31 and R^2 of 0.42 (n=810) were found, while for weekly averages the slope of the Deming regression was 1.19 with $R^2 = 0.47$ (n=116). The weekly and daily average timeseries and scatter plots are shown in figure A1 in Appendix A.

The differences and errors on the differences between MAX-DOAS and FTIR columns were calculated for the smoothed and original MAX-DOAS columns following the method outlined in Vigouroux et al. (2009). For the raw MAX-DOAS columns the difference (MAX-DOAS - FTIR, $\pm 1\sigma$) was 15.1 $\pm$ 26.3 % while for the smoothed comparison it was 10.1 $\pm$ 26.1 %. These results and the breakdown of random and systematic errors on the differences are compiled in table 2. ~~The percentage differences between the MAX-DOAS and FTIR (mean $\pm$ standard deviation) were $63 \pm 36\%$ for the original MAX-DOAS columns and $53 \pm 36\%$ when comparing the smoothed MAX-DOAS columns. As expected given the offset between FTIR and MAX-DOAS column amounts, the differences were on average greater than the combined error on the differences of the columns, calculated according to $E_{TC} = \sqrt{E^2_{MAX-DOAS} + E^2_{FTIR}}$, where E_{FTIR} and $E_{MAX-DOAS}$ are the combined random and systematic errors of each instrument. For the whole timeseries the mean combined error on the differences was $16.7 \pm 4.9\%$ (mean $\pm 1\sigma$). The differences between FTIR and MAX-DOAS columns here are greater~~The differences and standard deviations of the column comparisons are slightly larger here than for the results found in the Reunion Island comparison (Vigouroux et al., 2009), where no significant offset between measurements was found. In contrast to their study, here the smoothing was found to improve the mean difference between the columns. The greater mean difference and standard deviations of the differences at Lauder compared to Vigouroux et al. (2009) likely reflect the much longer comparison period, incorporating variations across a much wider range of atmospheric conditions, and the fact that here only the altitude range 0-4 km is examined rather than 0-10 km as in Vigouroux et al. (2009). In addition, differences in site characteristics may play a role in the greater offset observed at Lauder. Reunion Island being a coastal site is likely to be measuring marine background formaldehyde, as indicated by the fact that the 2007 measurements in Vigouroux et al. (2009) rarely not exceeded 7.7×10^{15} molecules.cm^{-2}, with little local surface HCHO production. In comparison, the mean smoothed DOAS column across the 27 month comparison period was 7.7×10^{15} molecules.cm^{-2}, suggesting greater local production which will occur at the surface, where the MAX-DOAS sensitivity is greatest and the FTIR least sensitive.

3.3 MAX-DOAS vs TROPOMI

In this section, MAX-DOAS formaldehyde columns are compared with satellite results. Firstly, Lauder HCHO MAX-DOAS columns are examined alongside results from TROPOMI. Following the example of MAX-DOAS vs satellite formaldehyde comparisons in Hoque et al. (2018b) and De Smedt et al. (2015), vertical columns are compared rather than profiles.

TROPOMI reports an uncertainty on the column amount, however it was found that this uncertainty was highly correlated with the magnitude of the column amount. Therefore, we estimated the uncertainty on the satellite column retrievals from the number of retrievals contributing to the averaged column in the $0.1 \times 0.1^\circ$ grid box (number per cell, N_{pc}) and the standard deviation of those retrievals (S_T): $E_{sat} = \frac{S_T}{\sqrt{N_{pc}}}$. More measurements were available from TROPOMI over Broadmeadows than at Lauder, with the average N_{pc} across the comparison period and considering TROPOMI pixels 0.1° either side of the ground-based station, of 1.18 in New Zealand and 2.76 in Melbourne. Because N_{pc} was often below one for 0.1° resolution, comparison

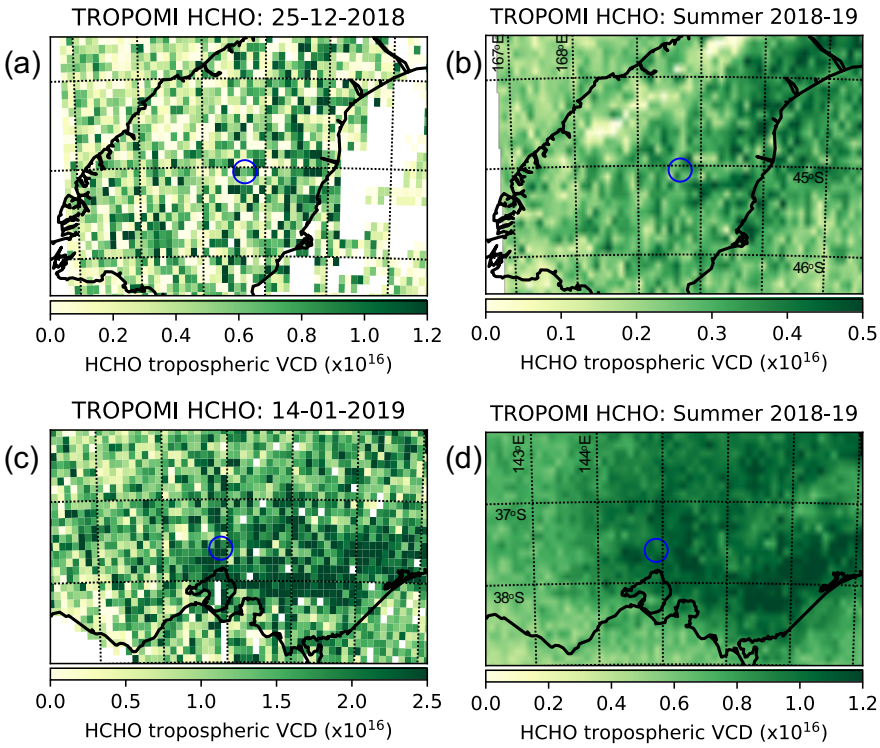

Figure 7. Examples of TROPOMI $0.1 \times 0.1^{\circ}$ gridded output. The blue circles indicate the MAX-DOAS measurement locations. The grid lines drawn are at 1° latitude and longitude intervals. (a) Example retrieval from the South Island of New Zealand, 24th of December 2018. (b) All New Zealand retrievals from December 2018-February 2019 averaged. (c) Example from Victoria, 14th January 2019, and (d) all Victorian retrievals from December 2018-February 2019 averaged. Note that the colour bar scale is different for each plot to emphasise the spatial patterns.

with MAX-DOAS results was carried out at 0.2° resolution. Final compared results filtered out pixels with $N_{pc} < 1$, giving a used average N_{pc} of 1.84 for Lauder and 2.94 for Broadmeadows. The discrepancy in N_{pc} could be due to more cloud over New Zealand than Victoria, or because HCHO columns over Lauder are low enough to be approaching the detection limit. TROPOMI results showed greater spatial variation over New Zealand than Victoria, as illustrated in the example map in figure 7(a). This is reflected in the standard deviation (S_T) of HCHO retrievals contributing to the Lauder and Broadmeadows average TROPOMI columns; the mean $\pm S_T$ was $1.66 \times 10^{15} \pm 1.50 \times 10^{15}$ molec.cm^{-2} and $7.53 \times 10^{15} \pm 1.10 \times 10^{15}$ molec.cm^{-2} for Lauder and Broadmeadows respectively. Overall, these factors combined to give a high mean percentage variancce for Lauder TROPOMI columns of 129 $\%$, while for Broadmeadows it was only 9.7%.

Nevertheless, the average summer (DJF) 2018-19 TROPOMI retrieval map for the central New Zealand South Island shown in figure 7(b), supports the conclusion (from the MAX-DOAS measurements) that highest formaldehyde amounts are in the

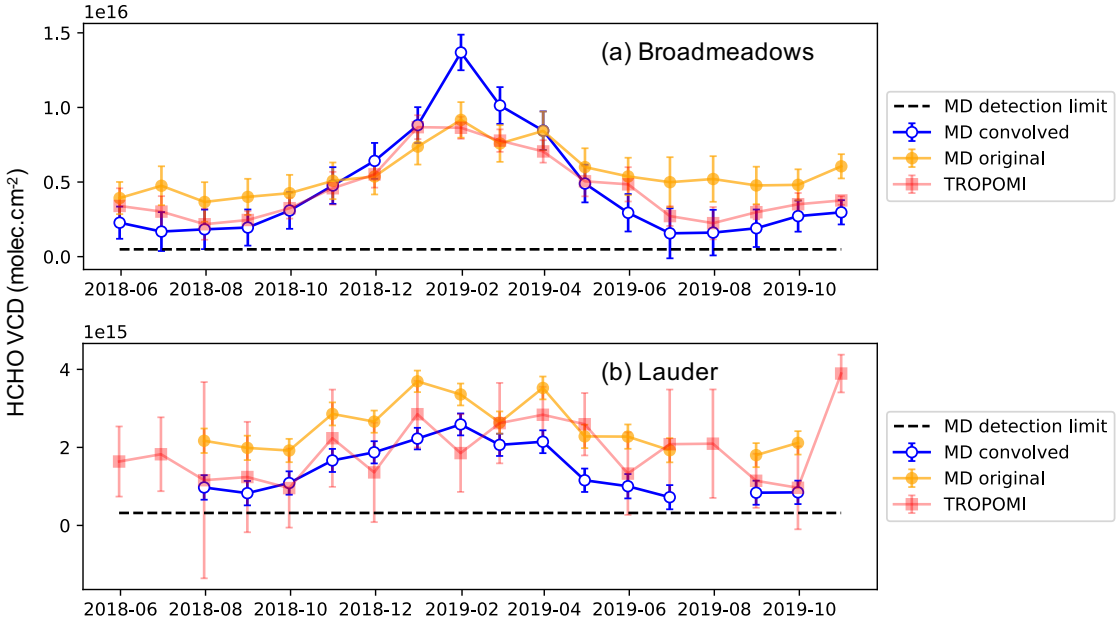

Figure 8. Timeseries comparison of monthly mean formaldehyde columns from TROPOMI and the MAX-DOAS between May 2018 and November 2019 at (a) Broadmeadows and (b) Lauder. Both the original (orange) and convolved (blue and white, i.e. accounting for different retrieval a priori and sensitivities) MAX-DOAS columns are shown.

agricultural and more densely populated eastern parts of the island. There are no standout HCHO hot spots in the thickly forested west coast or south-western Fiordland regions. The New Zealand Alps are highlighted in this figure by the lack of formaldehyde, possibly due to minimal vegetation in this region and because the satellite retrieval will not work over areas of high albedo (i.e. snow). The inference that formaldehyde is close to background levels is supported by the fact that average summer column amounts over the Tasman Sea and Pacific Ocean off the coast of the South Island appear similar to those over land. In comparison, the average summer 2018-19 map from Victoria highlights some clear features especially high formaldehyde levels over the densely forested regions in the east of the state. The irrigated agricultural land north of Melbourne stands out compared to the drier grazing country in the west and north-west; these areas highlighted by TROPOMI correspond to the directions of highest measured HCHO at Broadmeadows in figure 5(b).

Formaldehyde columns from TROPOMI and MAX-DOAS at Broadmeadows and Lauder were compared over the course of 18 months (May 2018-November 2019). For the comparison, TROPOMI results (columns and associated a priori profiles and averaging kernels) were averaged 0.1° either side of the Broadmeadows and Lauder MAX-DOAS locations. MAX-DOAS columns (along with averaging kernels) were averaged between 13:00-14:00 local time, around the time of the TROPOMI overpass. TROPOMI vertical profiles are not available for download and hence, to accurately compare tropospheric columns

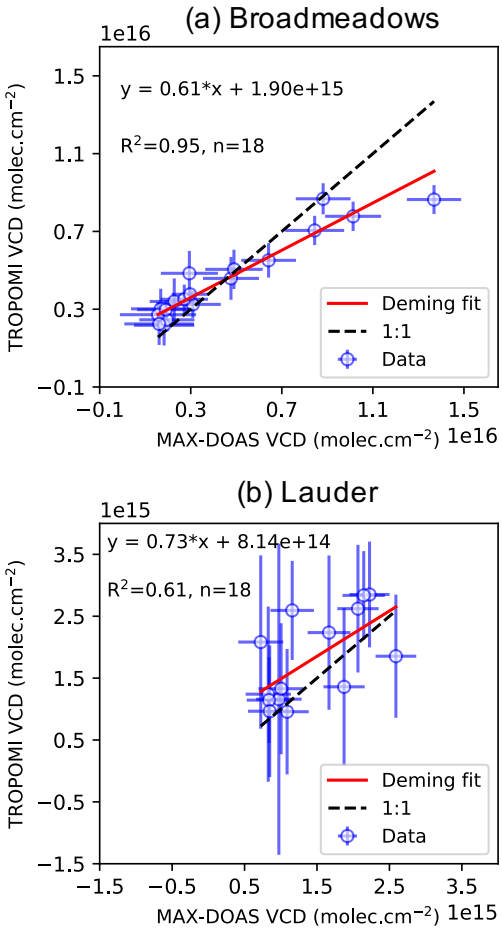

Figure 9. Scatter plot and Deming regression results for the comparison of monthly mean formaldehyde columns from TROPOMI and the MAX-DOAS (convolved with TROPOMI averaging kernels) between May 2018 and November 2019 at (a) Broadmeadows and (b) Lauder.

across the same altitude range, the MAX-DOAS retrievals for this comparison were run to 10 km rather than 4 km as in the FTIR-MAX-DOAS comparison in section 3.2.

For direct comparison of TROPOMI and MAX-DOAS formaldehyde vertical columns, accounting for the different instrumental a priori profiles and vertical sensitivities, the method outlined in Vigouroux et al. (2020) for comparing TROPOMI with FTIR was adapted. Firstly, to account for the fact that the two retrieval methods use different a priori profiles, the following equation was used to produce an adjusted MAX-DOAS profile $\mathbf{x'_D}$:

$$\mathbf{x'_D} = \mathbf{x_D} + (\mathbf{A_M} - \mathbf{I})(\mathbf{x_{D,a}} - \mathbf{x_{T,a}}) \quad (4)$$

Table 2. Summary of MAX-DOAS formaldehyde results and error budget from Lauder and Broadmeadows. The slope and R^2 are from the linear regression fit to y = Slope×x + c where y = FTIR or TROPOMI HCHO columns and x = MAX-DOAS HCHO columns. The random and systematic error columns represent the breakdown of errors on the column differences. Percentages are with respect to the MAX-DOAS columns.

Site	MAX-DOAS comparison	Slope (monthly avg)	R^2 (monthly avg)	Mean diff ± Std (%)	Systematic error (%)	Random error (%)
Lauder	raw vs FTIR	0.87	0.66	15.1 ± 26.3	10.9	18.4
Lauder	smoothed vs FTIR	1.17	0.65	10.1 ± 26.1	10.9	3.8
Lauder	raw vs TROPOMI	1.21	0.69	28.7 ± 61.3	31.4	83.8
Lauder	smoothed vs TROPOMI	0.73	0.61	-22.4 ± 132.3	26.3	79.7
Broadm.	raw vs TROPOMI	0.68	0.90	21.4 ± 45.2	20.2	62.6
Broadm.	smoothed vs TROPOMI	0.61	0.95	-5.3 ± 93.7	12.6	37.3

where $\mathbf{x_D}$ is the original MAX-DOAS profile, $\mathbf{A_M}$ is the MAX-DOAS averaging kernel matrix, $\mathbf{I}$ is the identity matrix, $\mathbf{x_{D,a}}$ is the MAX-DOAS a priori profile and $\mathbf{x_{T,a}}$ is the TROPOMI a priori profile expressed on the MAX-DOAS altitude grid. The integrated adjusted column gave an adjusted MAX-DOAS HCHO tropospheric column, which was then smoothed using the TROPOMI averaging kernels (expressed on the MAX-DOAS altitude grid) using the same method as for smoothing the FTIR columns in section 3.2 (Rodgers and Connor, 2003):

$$C_{D,smooth} = C_{T,a} + \mathbf{a_T}(\mathbf{x'_D} - \mathbf{x_{T,a}}) \tag{5}$$

where $C_{D,smooth}$ is the smoothed MAX-DOAS tropospheric column, $C_{T,a}$ is the TROPOMI a priori tropospheric column and $\mathbf{a_T}$ is the TROPOMI column total averaging kernel.

The monthly average timeseries of HCHO tropospheric columns at Broadmeadows measured by MAX-DOAS and TROPOMI is shown in figure 8(a). The seasonal variation of formaldehyde with its strong summer peak is clearly captured by TROPOMI, with all MAX-DOAS and TROPOMI data points above the calculated MAX-DOAS detection limit. The original MAX-DOAS retrieved columns agree well with the magnitude of the TROPOMI observations between October 2018 and June 2019, including over the summer peak, but are greater than TROPOMI outside these months. The MAX-DOAS columns adjusted for a priori differences and convolved with TROPOMI averaging kernels agree well with TROPOMI, within uncertainty, for all months except the height of the summer peak in January-February 2019. This discrepancy during times of peak HCHO production in the boundary layer highlights the much greater sensitivity of the MAX-DOAS to the lower atmosphere than TROPOMI. ~~On average the difference between TROPOMI and the smoothed MAX-DOAS HCHO tropospheric columns at Broadmeadows was 21% ± 34%~~ The average difference between TROPOMI and the smoothed and raw MAX-DOAS columns, along with the

breakdown of random and systematic errors on the differences (calculated following the methodology outlined in Vigouroux et al. (2009)) is presented in table 2. Smoothed MAX-DOAS columns were on average 5 % higher than TROPOMI but for individual measurements, the difference was highly variable (standard deviation 94 %). This small average bias towards MAX-DOAS is consistent with the bias found between ground-based FTIR stations and TROPOMI for locations with comparable average HCHO column amounts in Vigouroux et al. (2020).

Figure 8(b) shows the same as figure 8(a) but for Lauder. As with Broadmeadows, the broad seasonal variation is captured by TROPOMI and all data points are above the calculated MAX-DOAS detection limit although TROPOMI errorbars are greater than at Broadmeadows, often extending below the MAX-DOAS detection limit, due to the lower number of available TROPOMI retrievals over Lauder. The convolved MAX-DOAS HCHO columns compare well within error for a majority of months. ~~As with Broadmeadows, TROPOMI columns were greater on average than MAX-DOAS, by an average of $38 \pm 59\%$, although the difference was much less up to April 2019 ($12 \pm 30\%$)~~On average, TROPOMI was 29 % lower than MAX-DOAS raw columns and 22 % higher than smoothed MAX-DOAS columns, however the smoothing process accentuated the largest differences resulting in a standard deviation for the smoothed comparison greater than 100 %. The average bias found for Lauder MAX-DOAS vs TROPOMI is consistent within the uncertainty with the negative bias for TROPOMI vs FTIR for Lauder in Vigouroux et al. (2020).

The agreement between TROPOMI and MAX-DOAS is further examined using linear Deming regression analysis in figure 9. For Lauder, figure 9(b) shows the monthly average scatter plot with overall regression slope of 0.73 and $R^2 = 0.61$ (n=18). The majority of data points lie ~~are~~ within error of the 1:1 line. The regression values for the daily measurements at Lauder were slope = 0.40 and $R^2 = 0.22$ (n=510), while weekly averages gave a slope of 0.66 and R^2 of 0.45 (n=73). The resolution selection criterion did not have a large effect on the comparison, with a regression slope of 0.68 (monthly averages) for averaging TROPOMI 50 km either side of Lauder as opposed to 20 km. At Broadmeadows, data points lie along the 1:1 line within error except for the highest two values, which are January and February 2019 as highlighted in the timeseries, giving a regression slope of 0.61. This further highlights the finding, in line with Vigouroux et al. (2020), that low bias of TROPOMI compared to ground-based measurements is accentuated at high HCHO levels. The very strong temporal consistency is highlighted by an R^2 of 0.95 (n=18). Considering the individual daily measurements at Broadmeadows, the slope of the regression was 0.77 with $R^2 = 0.69$ (n=506) and for weekly averages the slope was 0.66 with $R^2 = 0.89$ (n=73) (plots for Lauder and Broadmeadows daily measurements and weekly averages are shown in figures A2 and A3 in Appendix A. Considering TROPOMI sampled 10 km and 50 km either side of Broadmeadows, regression slopes were 0.56 and 0.65 respectively, with the low bias of TROPOMI compared to MAX-DOAS at high HCHO consistent across sampling resolution.

The success of this comparison study for formaldehyde with TROPOMI, especially at Broadmeadows, is highlighted by a comparison (2017-2019) at the same Broadmeadows location between OMI and the MAX-DOAS. As shown in figure A4, OMI does not clearly capture any of the seasonal formaldehyde variation in Melbourne, and as such fails to replicate the MAX-DOAS values. The errorbars shown in this figure are the quoted uncertainty on the OMI columns, and represent 67 % of the total column on average, perhaps due to the poorer resolution of OMI compared to TROPOMI, making observation of the seasonal cycle difficult in this data. Monthly OMI HCHO columns are on average 200 % higher than the MAX-

DOAS (see table A1 in Appendix A), far greater than any discrepancy reported in the literature for a MAX-DOAS vs satellite retrieval. One possibility for the disparity is the fact that OMI is sampled 25 km either side of the measurement location compared to approximately 20 km for MAX-DOAS, thereby taking in more of the background. However, this could not explain why no seasonality is evident in the OMI results. ~~Another possibility 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; , 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~~ Given that both OMI and TROPOMI retrievals rely on a priori formaldehyde profiles calculated using the same chemical transport model (TM5, De Smedt et al. (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).

4 Conclusions

This paper presents comparison studies of MAX-DOAS formaldehyde measurements in two distinctly different environments: the remote Central Otago region in New Zealand and the suburban fringe area of Broadmeadows in Victoria. This work is the first long-term comparison and validation study undertaken using MAX-DOAS measurements in the Southern Hemisphere.

For MAX-DOAS measurements between December 2016 and November 2019, the mean formaldehyde column measured by the MAX-DOAS at Broadmeadows was $5.40 \pm 1.59 \times 10^{15}$ molec.cm^{-2} compared to $2.50 \pm 0.61 \times 10^{15}$ molec.cm^{-2} at Lauder. The amplitude of the seasonal cycle was also greater at Broadmeadows than Lauder, 2.0×10^{15} molec.cm^{-2} compared to 0.7×10^{15} molec.cm^{-2}. The seasonal cycles at Lauder and Broadmeadows could be described by a periodic function peaking at the end of January, i.e. at the height of the austral summer, consistent with biogenic temperature-dependent formaldehyde production.

At Lauder, 27 months of MAX-DOAS measurements were compared with FTIR formaldehyde partial columns between 0-4 km. Smoothing of the FTIR columns using the MAX-DOAS averaging kernels to resolve for the different vertical sensitivities was carried according to the methodology outlined in Rodgers and Connor (2003) and Vigouroux et al. (2009). The seasonal cycle of formaldehyde at Lauder, with pronounced summer peak, was clearly replicated by both sets of observations and the smoothed FTIR columns correlated more strongly than the original with the MAX-DOAS results. The timing of the HCHO

seasonal cycle peak were very similar between Broadmeadows and Lauder suggesting similar HCHO sources, however the source strength at Lauder seems to be weaker with a lower seasonal cycle amplitude.

In the first TROPOMI - MAX-DOAS Southern Hemispheric comparison study, TROPOMI performed especially well compared to the Broadmeadows monthly average columns in terms of temporal variation and magnitude ($R^2 = 0.95$, slope = ~~0.81~~0.61). This result is a significant improvement in both the comparison with OMI both at this location and in previous literature reports. Higher spatial variability and lower absolute amounts of HCHO made the comparison more difficult at Lauder, however, the linear regression analysis also indicated moderate temporal agreement in most months of the comparison (R^2 = 0.61, slope = 0.73).

Using maps of average TROPOMI HCHO retrievals this study also demonstrates the utility of the satellite product to identify hotspot regions of biogenic VOCs, which will be a critical tool in addressing the current gap in understanding of isoprene and monoterpene chemistry in south-eastern Australia.

This TROPOMI comparison study, especially over Melbourne, raises many exciting possibilities for future work. This study shows the importance of long-term timeseries MAX-DOAS measurements for satellite validation, and could ~~be contributed~~contribute to international validation efforts. This study could also be extended to consider not only formaldehyde validation but also NO_2, HONO and glyoxal. ~~This would begin to address the current Northern-Hemispheric bias in satellite validation studies~~This would continue to address the lack of Southern-Hemispheric satellite validation studies using ground-based remote sensing. This work also shows the utility of the MAX-DOAS technique for studying formaldehyde in the VOC-hotspot of south-eastern Australia, and it would be interesting in future studies to deploy MAX-DOAS instruments into the forested areas highlighted in TROPOMI as large formaldehyde source regions. Moreover, this work has shown that improvements in satellite technology, culminating (at this point in time) in TROPOMI, mean that space-based HCHO measurements will also be of great benefit in constraining the temporal and spatial distribution of VOC emissions in this region. With such assurance, related tropospheric oxidation and ozone chemistry, with their associated air quality and climate implications, can be studied on a much grander scale.

Data availability. Data available on request from the authors

Appendix A

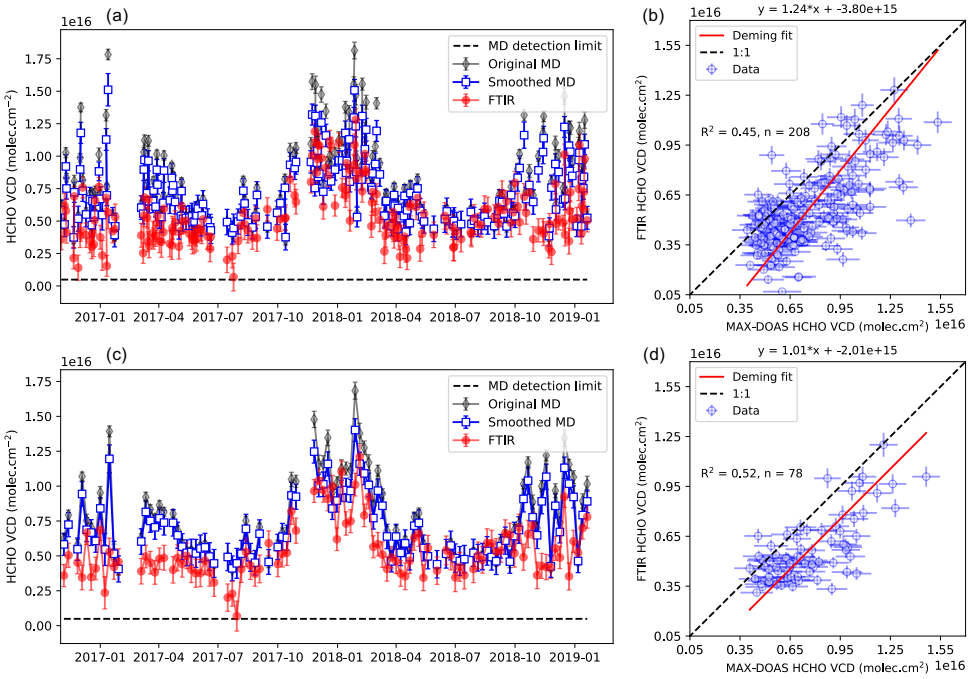

Figure A1. Results for the comparison of formaldehyde columns from FTIR and the MAX-DOAS (convolved with FTIR averaging kernels) at Lauder: (a) Daily average timeseries, (b) daily average scatter plot weekly averages, (c) weekly average timeseries and (d) weekly average scatter plot.

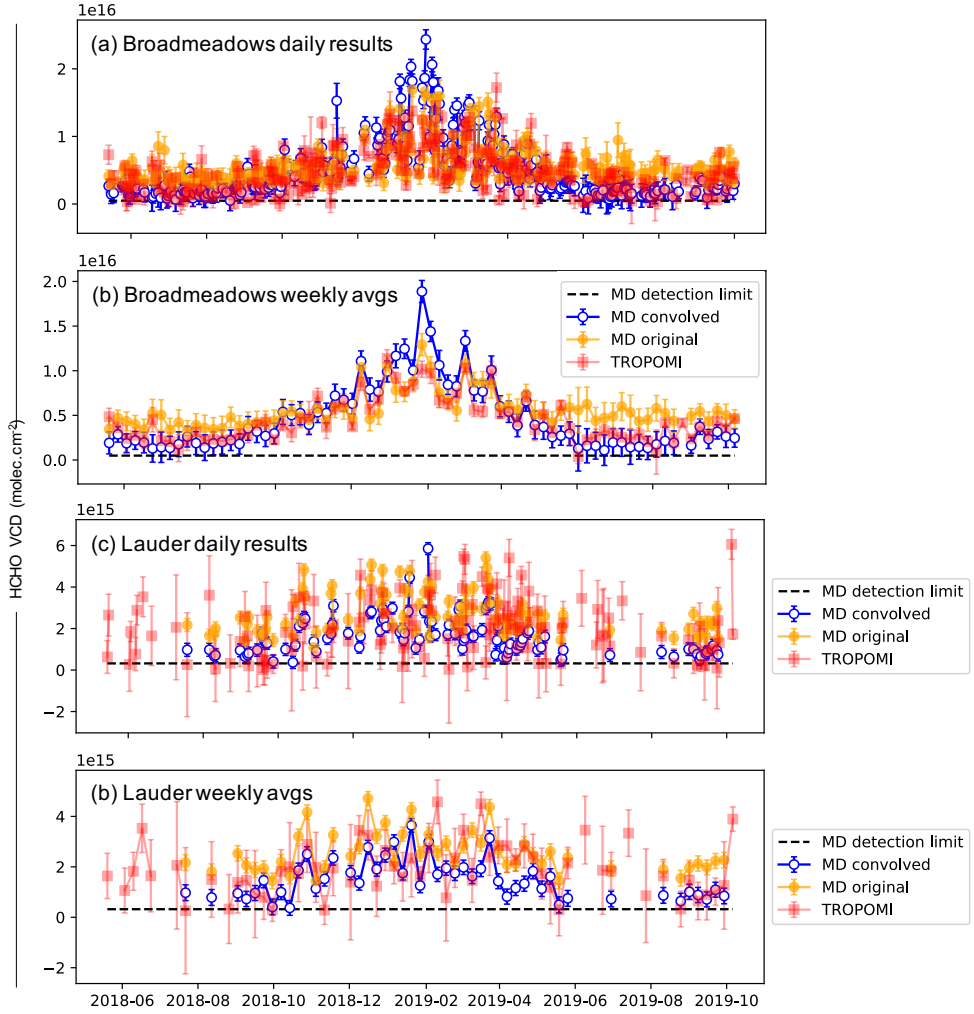

Figure A2. Timeseries results for the comparison of formaldehyde columns from TROPOMI and the MAX-DOAS (convolved with TROPOMI averaging kernels) between May 2018 and November 2019: (a) Broadmeadows daily measurements, (b) Broadmeadows weekly averages, (c) Lauder daily measurements and (d) Lauder weekly averages.

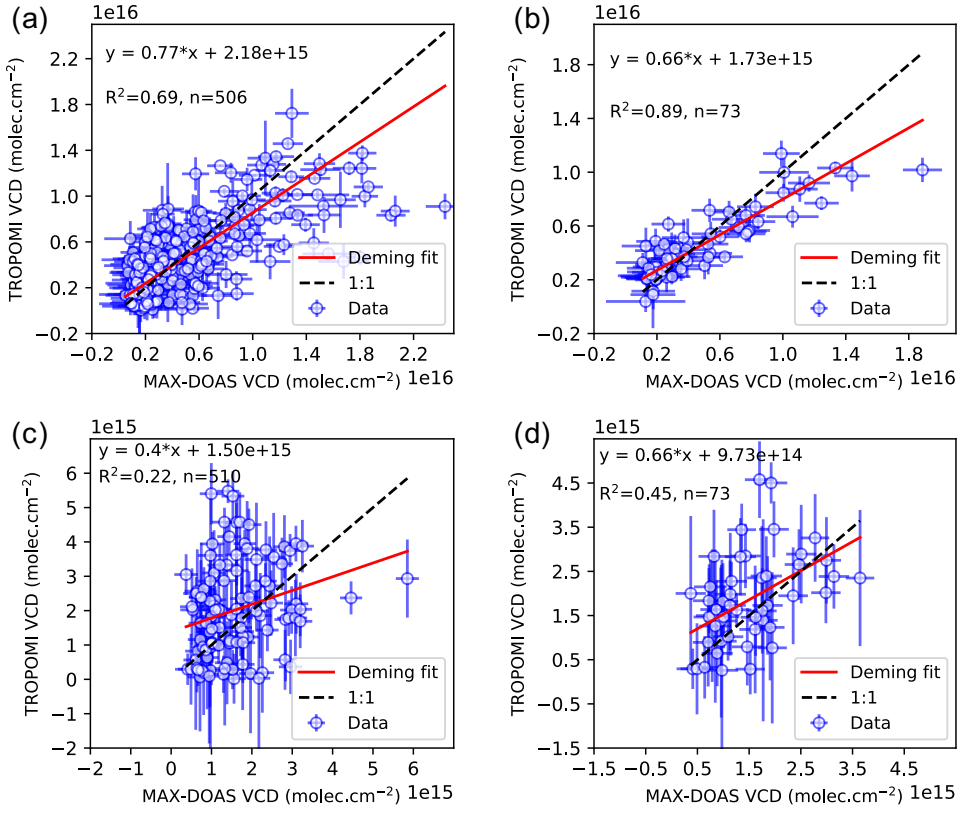

Figure A3. Scatter plot results for the comparison of formaldehyde columns from TROPOMI and the MAX-DOAS (convolved with TROPOMI averaging kernels) between May 2018 and November 2019: (a) Broadmeadows daily measurements, (b) Broadmeadows weekly averages, (c) Lauder daily measurements and (d) Lauder weekly averages.

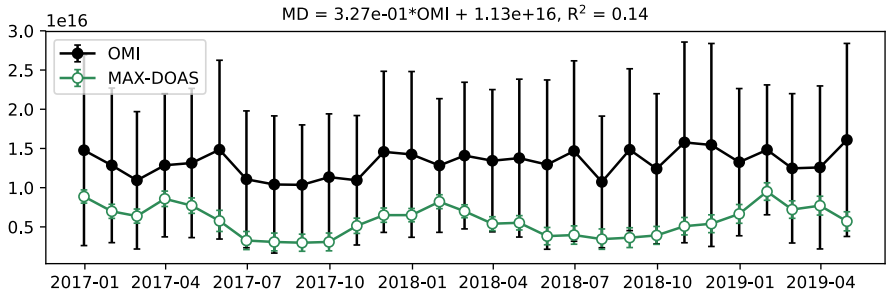

Figure A4. Timeseries of monthly averaged MAX-DOAS and OMI HCHO vertical columns ("ColumnAmount" field) at Broadmeadows from January 2017 to May 2019.

Table A1. Results from this and previous literature studies comparing formaldehyde vertical columns from MAX-DOAS and satellite retrievals. Note that 'Diff' is for MAX-DOAS - satellite. Slope is the gradient (m) of the linear regression for Satellite = $m\times$MAX-DOAS + C. [1]This study, [2]Tian et al. (2018), [3]Chan et al. (2019), [4]Wang et al. (2017a), [5]Hoque et al. (2018b), [6]Vigouroux et al. (2009). M = Monthly, D = Daily

Satellite	Location	Time period	Avgs	Max. dist. (km)	Diff. $\pm$ std (%)	R^2	Slope
TROPOMI[1]	Lauder, NZ	May 2018-Jan 2019	M	10	-38$\pm$59	0.61	0.73
TROPOMI[1]	Melbourne, Australia	May 2018-May 2019	M	10	-21$\pm$34	0.95	0.61
OMI[1]	Melbourne, Australia	Jan2017-May 2019	M	50	-187$\pm$103	0.33	7.14
OMI[2]	Yangtze R. Delta, China	Dec 2013-Dec 2016	M	25	-	0.56	0.84
OMI[3]	Nanjing, China	Apr 2013-Apr 2017	M	20	-	0.56	1.01
OMI[4]	Wuxi, China	May 2011-Nov 2014	D	50	-	0.17	2.08
GOME-2[4]	Wuxi, China	May 2011-Nov 2014	D	50	-	0.18	1.64
GOME-2[5]	Pantnagar, India	Feb-Nov 2017	M	50	-	0.50	-
SCIAMACHY[6]	Reunion Island	Jul 2004-Jul 2005	D	500	-11.2$\pm$30.5	-	-

Author contributions. RGR processed the MAX-DOAS data, performed the comparison calculations, made the figures and wrote the paper. The study was designed jointly by RGR, JDS, RS, RQ and DS. JDS downloaded TROPOMI data and assisted both in comparison calculations and drafting the paper. RQ and DS were responsible for MAX-DOAS and FTIR data collection and the Lauder site and contributed to writing the manuscript. SR and MT were responsible for MAX-DOAS data collection and Broadmeadows. NJ contributed to vertical column HCHO calculations for FTIR and MAX-DOAS results and assisted in writing the paper. RS directed the project and contributed to drafting the manuscript.

Competing interests. The authors declare no competing interests exist.

Acknowledgements. RS and RGR acknowledge support from the Australian Research Council's Centre of Excellence for Climate System Science (CE110001028), the Centre of Excellence for Climate Extremes (CE170100023) and Australian Research Council's Discovery project: Tackling Atmospheric Chemistry Grand Challenges in the Southern Hemisphere (DP160101598). RR wishes to thank Prof. Peter Rayner for helpful discussions on inverse methods, Dr Udo Frieß for access to and assistance with the HEIPRO algorithm, and the Dr Albert Shimmins Memorial Fund (through the University of Melbourne) for financial support contributing to the writing of this manuscript. The authors acknowledge Thomas Danckeart, Caroline Fayt and Michel van Roozendael, authors of the QDOAS software used in DOAS analysis in this work and Blair Trewin at the Bureau of Meteorology for providing meteorological data. MAX-DOAS and FTIR measurements at Lauder are core-funded by NIWA through New Zealand's Ministry of Business, Innovation and Employment Strategic Science Investment Fund.

References

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q., Liu, H. Y., Mickley, L. J., and Schultz, M. G.: Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, Journal of Geophysical Research: Atmospheres, 106, 23 073–23 095, 2001.

Chan, K. L., Wang, Z., Ding, A., Heue, K.-P., Shen, Y., Wang, J., Zhang, F., Shi, Y., Hao, N., and Wenig, M.: MAX-DOAS measurements of tropospheric NO_2 and HCHO in Nanjing and a comparison to ozone monitoring instrument observations, Atmospheric Chemistry and Physics, 19, 10 051–10 071, 2019.

Chance, K., Palmer, P. I., Spurr, R. J., Martin, R. V., Kurosu, T. P., and Jacob, D. J.: Satellite observations of formaldehyde over North America from GOME, Geophysical Research Letters, 27, 3461–3464, 2000.

De Mazière, M., Thompson, A. M., Kurylo, M. J., Wild, J. D., Bernhard, G., Blumenstock, T., Braathen, G. O., Hannigan, J. W., Lambert, J.-C., Leblanc, T., McGee, T. J., Nedoluha, G., Petropavlovskikh, I., Seckmeyer, G., Simon, P. C., Steinbrecht, W., and Strahan, S. E.: The Network for the Detection of Atmospheric Composition Change (NDACC): history, status and perspectives, Atmospheric Chemistry and Physics, 18, 4935–4964, 2018.

De Smedt, I., Stavrakou, T., Hendrick, F., Danckaert, T., Vlemmix, T., Pinardi, G., Theys, N., Lerot, C., Gielen, C., Vigouroux, C., et al.: Diurnal, seasonal and long-term variations of global formaldehyde columns inferred from combined OMI and GOME-2 observations, Atmospheric Chemistry & Physics Discussions, 15, 2015.

De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compernolle, S., Van Roozendael, M., Richter, A., Hilboll, A., Peters, E., Pedergnana, M., Loyola, D., Beirle, S., Wagner, T., Eskes, H., van Geffen, J., Boersma, K. F., and Veefkind, P.: Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and from the QA4ECV project, Atmospheric Measurement Techniques, 11, 2395–2426, 2018.

Duncan, B. N., Yoshida, Y., Damon, M. R., Douglass, A. R., and Witte, J. C.: Temperature dependence of factors controlling isoprene emissions, Geophysical Research Letters, 36, 2009.

Emmerson, K. M., Galbally, I. E., Guenther, A. B., Paton-Walsh, C., Guerette, E.-A., Cope, M. E., Keywood, M. D., Lawson, S. J., Molloy, S. B., Dunne, E., et al.: Current estimates of biogenic emissions from eucalypts uncertain for southeast Australia, Atmospheric Chemistry and Physics, 16, 6997–7011, 2016.

Emmerson, K. M., Cope, M. E., Galbally, I. E., Lee, S., and Nelson, P. F.: Isoprene and monoterpene emissions in south-east Australia: comparison of a multi-layer canopy model with MEGAN and with atmospheric observations, Atmospheric Chemistry and Physics, 18, 7539–7556, 2018.

Fleischmann, O. C., Hartmann, M., Burrows, J. P., and Orphal, J.: New ultraviolet absorption cross-sections of BrO at atmospheric temperatures measured by time-windowing Fourier transform spectroscopy, Journal of Photochemistry and Photobiology A: Chemistry, 168, 117–132 2004.

Franco, B., Hendrick, F., Van Roozendael, M., Müller, J.-F., Stavrakou, T., Marais, E., Bovy, B., Bader, W., Fayt, C., Hermans, C., et al.: Retrievals of formaldehyde from ground-based FTIR and MAX-DOAS observations at the Jungfraujoch station and comparisons with GEOS-Chem and IMAGES model simulations, Atmospheric Measurement Techniques, 8, 1733–1756, 2015.

Frieß, U., Monks, P. S., Remedios, J. J., Rozanov, A., Sinreich, R., Wagner, T., and Platt, U.: MAX-DOAS O_4 measurements: A new technique to derive information on atmospheric aerosols: 2. Modeling studies, Journal of Geophysical Research: Atmospheres, 111, 2006.

Fuentes, J. D., Lerdau, M., Atkinson, R., Baldocchi, D., Bottenheim, J. W., Ciccioli, P., Lamb, B., Geron, C., Gu, L., Guenther, A., Sharkey, T. D., and Stockwell, W.: Biogenic Hydrocarbons in the Atmospheric Boundary Layer: A Review, Bulletin of the American Meteorological Society, 81, 1537–1576, 2000.

Garcia, R., Marsh, D., Kinnison, D., Boville, B., and Sassi, F.: Simulation of secular trends in the middle atmosphere, 1950–2003, Journal of Geophysical Research: Atmospheres, 112, 2007.

Gielen, C., Van Roozendael, M., Hendrick, F., Pinardi, G., Vlemmix, T., De Bock, V., De Backer, H., Fayt, C., Hermans, C., and Gillotay, D.: A simple and versatile cloud-screening method for MAX-DOAS retrievals, Atmospheric Measurement Techniques, 7, 3509

González Abad, G., Liu, X., Chance, K., Wang, H., Kurosu, T. P., and Suleiman, R.: Updated Smithsonian Astrophysical Observatory Ozone Monitoring Instrument (SAO OMI) formaldehyde retrieval, Atmospheric Measurement Techniques, 8, 19–32, 2015.

Grainger, J. F. and Ring, J.: Anomalous Fraunhofer Line Profiles, Nature, 193, 762–762, 1962.

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.: The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions, Geoscientific Model Development, 5, 1471–1492, 2012.

Heckel, A., Richter, A., Tarsu, T., Wittrock, F., Hak, C., Pundt, I., Junkermann, W., and Burrows, J.: MAX-DOAS measurements of formaldehyde in the Po-Valley, Atmospheric Chemistry and Physics, 5, 909–918 2005.

Henze, D., Seinfeld, J., Ng, N., Kroll, J., Fu, T.-M., Jacob, D. J., and Heald, C.: Global modeling of secondary organic aerosol formation from aromatic hydrocarbons: high-vs. low-yield pathways, Atmospheric Chemistry and Physics, 8, 2405–2420, 2008.

Hoque, H., Irie, H., and Damiani, A.: First MAX-DOAS Observations of Formaldehyde and Glyoxal in Phimai, Thailand, Journal of Geophysical Research: Atmospheres, 123, 9957–9975, 2018a.

Hoque, H. M. S., Irie, H., Damiani, A., Rawat, P., and Naja, M.: First simultaneous observations of formaldehyde and glyoxal by MAX-DOAS in the Indo-Gangetic Plain region, SOLA, 2018b.

Johansson, M., Rivera, C., Foy, B. d., Lei, W., Song, J., Zhang, Y., Galle, B., and Molina, L.: Mobile mini-DOAS measurement of the outflow of NO_2 and HCHO from Mexico City, Atmospheric Chemistry and Physics, 9, 5647–5653, 2009.

Jones, N., Koike, M., Matthews, W., and McNamara, B.: Southern hemisphere mid-latitude seasonal cycle in total column nitric acid, Geophysical Research Letters, 21, 593–596, 1994.

Jones, N., Riedel, K., Allan, W., Wood, S., Palmer, P., Chance, K., and Notholt, J.: Long-term tropospheric formaldehyde concentrations deduced from ground-based fourier transform solar infrared measurements, Atmospheric Chemistry and Physics, 9, 7131–7142, 2009.

Kefauver, S. C., Filella, I., and Peñuelas, J.: Remote sensing of atmospheric biogenic volatile organic compounds (BVOCs) via satellite-based formaldehyde vertical column assessments, International journal of remote sensing, 35, 7519–7542, 2014.

Khatri, P. and Takamura, T.: An Algorithm to Screen Cloud-A? ected Data for Sky Radiometer Data Analysis, Journal of the Meteorological Society of Japan. Ser. II, 87, 189–204, 2009.

Kreher, K., Van Roozendael, M., Hendrick, F., Apituley, A., Dimitropoulou, E., Frieß, U., Richter, A., Wagner, T., Abuhassan, N., Ang, L., Anguas, M., Bais, A., Benavent, N., Bösch, T., Bognar, K., Borovski, A., Bruchkouski, I., Cede, A., Chan, K. L., Donner, S., Drosoglou, T., Fayt, C., Finkenzeller, H., Garcia-Nieto, D., Gielen, C., Gómez-Martín, L., Hao, N., Herman, J. R., Hermans, C., Hoque, S., Irie, H., Jin, J., Johnston, P., Khayyam Butt, J., Khokhar, F., Koenig, T. K., Kuhn, J., Kumar, V., Lampel, J., Liu, C., Ma, J., Merlaud, A., Mishra, A. K., Müller, M., Navarro-Comas, M., Ostendorf, M., Pazmino, A., Peters, E., Pinardi, G., Pinharanda, M., Piters, A., Platt, U., Postylyakov, O., Prados-Roman, C., Puentedura, O., Querel, R., Saiz-Lopez, A., Schönhardt, A., Schreier, S. F., Seyler, A., Sinha, V., Spinei, E., Strong, K., Tack, F., Tian, X., Tiefengraber, M., Tirpitz, J.-L., van Gent, J., Volkamer, R., Vrekoussis, M., Wang, S., Wang, Z., Wenig, M., Wittrock, F.,

Xie, P. H., Xu, J., Yela, M., Zhang, C., and Zhao, X.: Intercomparison of NO_2, O_4, O_3 and HCHO slant column measurements by MAX-DOAS and zenith-sky UV-Visible spectrometers during the CINDI-2 campaign, Atmospheric Measurement Techniques Discussions, 2019, 1–58, 2019.

Kurosu, T. P., Chance, K., Liu, X., Volkamer, R., Fu, T.-M., Millet, D., Jacob, D. J., and Levelt, P.: Seasonally resolved global distributions of glyoxal and formaldehyde observed from the Ozone Monitoring Instrument on EOS Aura, Proceeding of Anais XIII Simpósio Brasileiro de Sensoriamento Remoto, 21, 6461, 2007.

Lee, H., Ryu, J., Irie, H., Jang, S.-H., Park, J., Choi, W., and Hong, H.: Investigations of the diurnal variation of vertical HCHO profiles based on MAX-DOAS measurements in Beijing: Comparisons with OMI vertical column data, Atmosphere, 6, 1816–1832, 2015.

Lelieveld, J., Butler, T., Crowley, J., Dillon, T., Fischer, H., Ganzeveld, L., Harder, H., Lawrence, M., Martinez, M., and Taraborrelli, D.: Atmospheric oxidation capacity sustained by a tropical forest, Nature, 452, 737–740 0028–0836, 2008.

Levelt, P. F., van den Oord, G. H. J., Dobber, M. R., Malkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J. O. V., and Saari, H.: The ozone monitoring instrument, IEEE Transactions on Geoscience and Remote Sensing, 44, 1093–1101, 2006.

Mahajan, A. S., Whalley, L. K., Kozlova, E., Oetjen, H., Mendez, L., Furneaux, K. L., Goddard, A., Heard, D. E., Plane, J. M., and Saiz-Lopez, A.: DOAS observations of formaldehyde and its impact on the HO x balance in the tropical Atlantic marine boundary layer, Journal of atmospheric chemistry, 66, 167, 2010.

Martin, R. V.: Satellite remote sensing of surface air quality, Atmospheric environment, 42, 7823–7843, 2008.

Meller, R. and Moortgat, G. K.: Temperature dependence of the absorption cross sections of formaldehyde between 223 and 323 K in the wavelength range 225–375 nm, Journal of Geophysical Research: Atmospheres, 105, 7089–7101 2156–2202, 2000.

Palmer, P. I., Abbot, D. S., Fu, T.-M., Jacob, D. J., Chance, K., Kurosu, T. P., Guenther, A., Wiedinmyer, C., Stanton, J. C., Pilling, M. J., et al.: Quantifying the seasonal and interannual variability of North American isoprene emissions using satellite observations of the formaldehyde column, Journal of Geophysical Research: Atmospheres, 111, 2006.

Peters, E., Wittrock, F., Großmann, K., Frieß, U., Richter, A., and Burrows, J.: Formaldehyde and nitrogen dioxide over the remote western Pacific Ocean: SCIAMACHY and GOME-2 validation using ship-based MAX-DOAS observations, Atmospheric Chemistry and Physics, 12, 11 179–11 197, 2012.

Pfister, G., Emmons, L., Hess, P., Lamarque, J., Orlando, J., Walters, S., Guenther, A., Palmer, P., and Lawrence, P.: Contribution of isoprene to chemical budgets: A model tracer study with the NCAR CTM MOZART-4, Journal of Geophysical Research: Atmospheres, 113, 2008.

Pinardi, G., Van Roozendael, M., Abuhassan, N., Adams, C., Cede, A., Clémer, K., Fayt, C., Frieß, U., Gil, M., and Herman, J.: MAX-DOAS formaldehyde slant column measurements during CINDI: intercomparison and analysis improvement, Atmospheric Measurement Techniques, 6, 167–185, 2013.

Pollard, D. F., Sherlock, V., Robinson, J., Deutscher, N. M., Connor, B., and Shiona, H.: The Total Carbon Column Observing Network site description for Lauder, New Zealand, Earth System Science Data, 9, 977, 2017.

Rodgers, C. D. and Connor, B. J.: Intercomparison of remote sounding instruments, Journal of Geophysical Research: Atmospheres, 108, 2003.

Ryan, R. G., Rhodes, S., Tully, M., Wilson, S., Jones, N., Frieß, U., and Schofield, R.: Daytime HONO, NO_2 and aerosol distributions from MAX-DOAS observations in Melbourne, Atmospheric Chemistry and Physics, 18, 13 969–13 985, 2018.

Serdyuchenko, A., Gorshelev, V., Weber, M., Chehade, W., and Burrows, J.: High spectral resolution ozone absorption cross-sections, Atmospheric Measurement Techniques, 7, 1867–8548, 2014.

Stutz, J., Kim, E., Platt, U., Bruno, P., Perrino, C., and Febo, A.: UV visible absorption cross sections of nitrous acid, Journal of Geophysical Research: Atmospheres, 105, 14 585–14 592, 2000.

Thalman, R. and Volkamer, R.: Temperature dependent absorption cross-sections of O_2–O_2 collision pairs between 340 and 630 nm and at atmospherically relevant pressure, Physical chemistry chemical physics, 15, 15 371–15 381, 2013.

Thomas, W., Hegels, E., Slijkhuis, S., Spurr, R., and Chance, K.: Detection of biomass burning combustion products in Southeast Asia from backscatter data taken by the GOME spectrometer, Geophysical Research Letters, 25, 1317–1320, 1998.

Tian, X., Xie, P., Xu, J., Li, A., Wang, Y., Qin, M., and Hu, Z.: Long-term observations of tropospheric NO_2, SO_2 and HCHO by MAX-DOAS in Yangtze River Delta area, China, Journal of Environmental Sciences, 71, 207–221, 2018.

Tradowsky, J. S., Bodeker, G. E., Querel, R. R., Builtjes, P. J., and Fischer, J.: Combining data from the distributed GRUAN site Lauder-Invercargill, New Zealand, to provide a site atmospheric state best estimate of temperature, Earth System Science Data, 10, 2195–2211, 2018.

Vandaele, A. C., Hermans, C., Simon, P. C., Carleer, M., Colin, R., Fally, S., Merienne, M.-F., Jenouvrier, A., and Coquart, B.: Measurements of the NO_2 absorption cross section from 42 000 cm-1 to 10 000 cm-1 (238–1000 nm) at 220 K and 294 K, Journal of Quantitative Spectroscopy and Radiative Transfer, 59, 171–184, 1998.

Veefkind, J., Aben, I., McMullan, K., Förster, H., De Vries, J., Otter, G., Claas, J., Eskes, H., De Haan, J., Kleipool, Q., vanWeele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sensing of Environment, 120, 70–83, 2012.

Vigouroux, C., Hendrick, F., Stavrakou, T., Dils, B., Smedt, I. D., Hermans, C., Merlaud, A., Scolas, F., Senten, C., Vanhaelewyn, G., et al.: Ground-based FTIR and MAX-DOAS observations of formaldehyde at Réunion Island and comparisons with satellite and model data, Atmospheric Chemistry and Physics, 9, 9523–9544, 2009.

Vigouroux, C., Bauer Aquino, C. A., Bauwens, M., Becker, C., Blumenstock, T., De Mazière, M., García, O., Grutter, M., Guarin, C., Hannigan, J., et al.: NDACC harmonized formaldehyde time series from 21 FTIR stations covering a wide range of column abundances, Atmospheric Measurement Techniques, 11, 5049–5073, 2018.

Vigouroux, C., Langerock, B., Bauer Aquino, C. A., Blumenstock, T., De Mazière, M., De Smedt, I., Grutter, M., Hannigan, J., Jones, N., Kivi, R., Lutsch, E., Mahieu, E., Makarova, M., Metzger, J.-M., Morino, I., Murata, I., Nagahama, T., Notholt, J., Ortega, I., Palm, M., Pinardi, G., Röhling, A., Smale, D., Stremme, W., Strong, K., Sussmann, R., Té, Y., van Roozendael, M., Wang, P., and Winkler, H.: TROPOMI/S5P formaldehyde validation using an extensive network of ground-based FTIR stations, Atmospheric Measurement Techniques Discussions, 2020, 1–24, 2020.

Wagner, T., Apituley, A., Beirle, S., Dörner, S., Friess, U., Remmers, J., and Shaiganfar, R.: Cloud detection and classification based on MAX-DOAS observations, Atmos. Meas. Tech., 7, 1289–1320, 2014.

Wagner, T., Beirle, S., Remmers, J., Shaiganfar, R., and Wang, Y.: Absolute calibration of the colour index and O_4 absorption derived from Multi-AXis (MAX-) DOAS measurements and their application to a standardised cloud classification algorithm, Atmos. Meas. Tech. Discuss., 2016, 1–34, 2016.

Wang, Y., Beirle, S., Lampel, J., Koukouli, M., Smedt, I. D., Theys, N., Li, A., Wu, D., Xie, P., Liu, C., et al.: Validation of OMI, GOME-2A and GOME-2B tropospheric NO_2, SO_2 and HCHO products using MAX-DOAS observations from 2011 to 2014 in Wuxi, China: investigation of the effects of priori profiles and aerosols on the satellite products, Atmospheric Chemistry and Physics, 17, 5007–5033, 2017a.

Wang, Y., Lampel, J., Xie, P., Beirle, S., Li, A., Wu, D., and Wagner, T.: Ground-based MAX-DOAS observations of tropospheric aerosols, NO_2, SO_2 and HCHO in Wuxi, China, from 2011 to 2014, Atmospheric Chemistry and Physics, 17, 2189–2215, 2017b.

Winters, A. J., Adams, M. A., Bleby, T. M., Rennenberg, H., Steigner, D., Steinbrecher, R., and Kreuzwieser, J.: Emissions of isoprene, monoterpene and short-chained carbonyl compounds from Eucalyptus spp. in southern Australia, Atmospheric Environment, 43, 3035–3043

Wolfe, G., Kaiser, J., Hanisco, T., Keutsch, F., De Gouw, J., Gilman, J., Graus, M., Hatch, C., Holloway, J., Horowitz, L., et al.: Formaldehyde production from isoprene oxidation across NOx regimes, Atmospheric chemistry and physics, 16, 2597–2610, 2016.

Zhu, L., Jacob, D. J., Mickley, L. J., Marais, E. A., Cohan, D. S., Yoshida, Y., Duncan, B. N., Abad, G. G., and Chance, K. V.: Anthropogenic emissions of highly reactive volatile organic compounds in eastern Texas inferred from oversampling of satellite (OMI) measurements of HCHO columns, Environmental Research Letters, 9, 114 004, 2014.

Zhu, L., Jacob, D. J., Kim, P. S., Fisher, J. A., Yu, K., Travis, K. R., Mickley, L. J., Yantosca, R. M., Sulprizio, M. P., and Smedt, I. D.: Observing atmospheric formaldehyde (HCHO) from space: validation and intercomparison of six retrievals from four satellites (OMI, GOME2A, GOME2B, OMPS) with SEAC 4 RS aircraft observations over the southeast US, Atmospheric Chemistry and Physics, 16, 13 477–13 490

Zhu, L., Jacob, D. J., Keutsch, F. N., Mickley, L. J., Scheffe, R., Strum, M., Gonzalez Abad, G., Chance, K., Yang, K., Rappengluck, B., et al.: Formaldehyde (HCHO) as a hazardous air pollutant: Mapping surface air concentrations from satellite and inferring cancer risks in the United States, Environmental science & technology, 51, 5650–5657, 2017.