

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 O₃ 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 O₃ absorption is undoubtedly more significant going towards the UV end of the wavelength ranges, any concern that extra O₃ absorption is problematic in the 324.5-359 nm 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 of collocation 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 3.6x7.2 km (before 6 August 2019) and 3.6x5.6 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.