I found the section on the optimization of the spectral fitting window to be very good (incl. Fig 3). The homogenization of vertical columns by applying the reference sector method is an interesting result. The idea to have study regions is also meritorious. I was also pleased to see the use of orthogonalized temperature-dependent $O_3$ absorption cross-section spectra.

There are other superior spectral fitting techniques that have not been mentioned such as weighting-function-DOAS. One criticism is that the authors have some minor difficulties communicating in English. Also, there are several instances where the authors claim to be providing analysis of results, but they are simply speculating because they do not offer proof to support their statements or the proof they provide is inadequate. There are other occasions where the reader requires further elaboration by the authors to understand the points they are making.

I am also troubled with their use of statistics. For example, they look at the RMS of the residuals to decide whether the inclusion of a basis function is warranted. However, any basis function has the capability of reducing the RMS of the residuals, so an appropriate statistic is required. They have one available in terms of the standard deviation of the mean but they could explore even more appropriate statistics. I am not sure if QDOAS provides a $\chi^2$. The statistic used to judge whether the inclusion of an additional basis function is warranted must account for the additional degree of freedom used.

There are a couple of sections that can be omitted (see below).

**Scientific comments**

Eq. 1 – The quantity in the brackets in the numerator should be squared. This must simply be a typo otherwise the RMS values in Table 2 are not very good.

P7101L26 "…three step pre-fitting…" -> "…three step fitting…"

Section 3.1.1- The authors have tested the idea of pre-fitting but have not used the non-DOAS method of Chance et al. (2000), which has an advantage over the DOAS method used in this work in that a polynomial scales the absorption optical depth as a function of wavelength (for both pre-fitting and fitting cases). This should be noted. The issue with pre-fitting, particularly with a DOAS approach, is that the BrO and $O_3$ slant columns are fitting-window-dependent primarily because of the wavelength dependence of atmospheric scattering. For $O_3$, this will be a more significant problem since the mean wavelength for its pre-fitting window is significantly shorter than the mean wavelength of the HCHO fitting window. A comment here about using WF-DOAS or any more sophisticated fitting technique that does not simply fit cross-sections to (differential) optical depth spectra) in the future would be well-placed. Limiting to SZA<60° mitigates this issue.

Section 3.2.1 – This section was essentially pointless to me because an apples-to-apples comparison of Ring basis functions was not done because of the slit function issue.

Section 3.2.2 – The authors speculate that the larger fitting window (extending to >350 nm) are "substantially affected by unquantified fitting interference with $O_4$". The source of interfering spectral structure could be from the instrument or earth/atmosphere for example. To demonstrate that $O_4$ is the

likely source of the interference, the authors may choose to show the exponential decay of this residual feature as a function of cloud top height or simply remove the speculation. As a result, I disagree with the authors' conclusion regarding whether $O_4$ should be included in the optimized fit. I think that the authors should state their literature reference for the absorption spectroscopy of the $(O_2)_2$ collisional complex ($O_4$), and for OClO if they don't remove OClO from the paper (see below).

Section 3.2.3 – This section could be omitted since OClO is not expected to be found for SZA<60°. Including this absorber has no physical basis.

P7106L26 "concentration" -> "slant column"

Section 3.2.5 – The use of 'add back' is misleading in my opinion. I believe the authors are referring to the slant column that is used in generating the $I_0$-corrected cross-section following Appendix A of Aliwell et al. The SCD used to generate the $I_0$-corrected cross-section should be stated for each species. Test 3f is 'apples-to-apples' only if the $O_3$ slant column used to generate the $I_0$-corrected cross-section is $0.8x10^{19}$ molec/cm$^2$. A dynamic $I_0$-correction might be more appropriate, where the $O_3$ SC used to generate the $I_0$-corrected cross-section is obtained from the BrO pre-fit but this is probably not too important for HCHO since the data of the latter species is probably discarded in post-processing for highly perturbed $O_3$ conditions. Figure 5 should extend to ~$2.3x10^{19}$ molec/cm$^2$ to cover the $O_3$ SC shown in Figure 6.

Section 3.2.6 – To determine the suitable range of cross-section temperatures, an extreme atmosphere should be considered (tropical) rather than a ~mid-latitude one since temperatures at the tropopause near the equator can range as low at 195 K and are warm in the upper stratosphere. At these latitudes, 228 K may be too warm as the lower T of the cross-section pair.

P7109L7 "…mean tropospheric temperature…" -> "a single, effective temperature for tropospheric HCHO absorption…"

Section 3.3 – Remove "reduction in width" (L14) since reducing the width reduces the need for a higher order polynomial. The authors should not use residuals as a statistic to justify whether the polynomial order is sufficient (see introductory comments). The last sentence of this section should be removed. They can discuss 1σ instead or use more relevant statistics.

P7109L25 – "sufficient" is not quantified. Proper statistical tests demonstrate sufficiency.

P7110L9 – "lower UV" is unconventional. Suggestions: "UV-B" (for ≤320 nm), "short UV wavelengths"

P7111L9 – Standard deviations are slightly higher globally with undersampling correction. The authors have not demonstrated using appropriate statistics that the use of undersampling spectra is justified in the spectral fitting. Again, they have failed to account for the fact that they are using an extra basis function. Until some revision is made, "strongly" should be removed from the last sentence in this sub-section.

Section 3.4.3 – It is not clear how a linear offset term in the DOAS fit is not already included by virtue of the 5$^{th}$ order polynomial. Again, the use of the RMS statistic (P7111L25) is not appropriate to make a judgement on the utility or necessity of a basis function. The standard deviation could be used, after the authors have clarified the above issue (re: 5$^{th}$ order polynomial) to the reader.

Section 4.4 – This section needs substantial revision. It is not really the resolution that enables different cloud fraction thresholds to be tested but rather the spatial sampling frequency. Suggested wording: "The high spatial sampling frequency … provides the opportunity to test a higher rejection…"
To which "statistic" are the authors referring: "…the statistic is affected…"

The authors claim significant "above cloud enhancements" are occurring in the AMA region. I doubt this since the majority of the clouds are at 600 hPa (4 km) or lower pressure (higher altitude) as shown in Figure 10 and as stated in the text. Then the authors contradict themselves by stating that the Amazon has dense, low cloud cover (on 9 Aug 2007). This should be removed. Figure 10 shows that the highest formaldehyde columns in the AMA region are occurring when the cloud tops are low for partly cloudy scenes. This suggests that the formaldehyde is below 600 hPa (near the surface… as expected).

The authors also state that the "AMF (takes) a greater fraction of cloud albedo into account". Why would the cloud fraction be higher in the AMF than in the observations? My guess is that there is an issue with the AMF simulation (oversimplification of cloud radiative transfer) for mostly-cloudy scenes that leads to the increase in vertical columns with increasing cloud cover. An increase in the vertical column due to 'above cloud enhancements' would also be seen in the slant columns, which is not the case (Figure 9) in either hotspot region.

P7116L9 The offset correction (tests 5c-d) appears to not "be a smaller effect" than the $I_0$ correction (test 3f), so I suggest that one of these concluding statements is revised.

P7116L11 "…equal importance…" is probably too strong. How about "…high importance…"?

**Technical**

P7096L20 "…occur to…" -> "…occur near…"

P7097L17 " …its'… " -> "…its…"

P7098L16 "…on an assumed/known…" -> "on assumed or known…"

P7099L22 "…example…" -> "…sample…"

P7099L26 "…statistic…" -> "…statistics…"

P7101L13 "…in Theys et al. (2011)'s original retrieval." -> "…in the retrieval of Theys et al. (2011)." [avoid the possessive form]

P7102L16 "…serve…" -> "…serves…"

P7102L27 "…removal…" -> "…exclusion…"

P7103L1 “…and…” -> “…an…”

P7105L2 “…minima…” -> “…minimum…”

P7105L9 “…finding…” -> “…where it is found that…”

P7105L15 “higher” -> “wider” [this unconventional usage appears in several instances]

P7106L19 “…0.31…” -> “…by 0.31…”

P7106L22 “high” -> “large”

P7108L22 “…0.11% per extra K.” -> “…0.11%/K.”

P7108L28 “…the the…” -> “…the…”

P7109L2 “… x $10^{-7}$.” -> “… $10^{-7}$.”

P7109L4 “…spectra recording. However… likelihood…” -> “observation. However…weak likelihood…”

P7109L18 define “SC”

P7110L11 “…vectors…” -> “…spectra…” [also at P7110L15, use “polarisation spectra”]

P7110L22 “…retrieval based…” -> “…retrieval-based…”

P7111L7 “…as a pseudo-absorber…” -> “…as pseudo-absorbers…” [there are two of them, right?]

P7112L17 “…accounted for…” -> “…taken into account…”

P7113L7 suggested rewording: “ SCs from the reference strip … are fitted with a 3$^{rd}$ order polynomial over all latitudes, and the latitude-dependent fitted SC from this strip is subtracted globally from each day’s measurements to serve as a daily correction.”

P7114L17 It is not what is meant. Suggested rewording: “…due to the much smaller relative differences when vertical columns are considered instead of SCs.”

P7114L22 Don’t start a new paragraph.

P7115L9 “…higher signal to noise ratio…” -> “…less variability…”

P7115L10 “…in slant columns whose…” -> “…in the number of observations for which the …”

P7115L26 “…the contrasting vertical column increase”. The vertical columns increase relative to what or as a function of what? I assume the authors mean: “… the contrasting behaviour of the vertical columns versus cloud fraction threshold between the two regions…”

Table 1: “$I_0$-corrected to…” -> “$I_0$-corrected using…”

Figure 9 caption: “Removing…” -> “Altering…”

Figure 10 caption: You should define “high CF” since it appears that not all high CF pixels are in the 300-600 hPa region. One point with a high CF, for example, is near 900 hPa and shows a HCHO SC of $1x10^{16}$ molec/$cm^2$.