We thank the referee for his very careful review, and his constructive suggestions. In the following, we answer his specific questions. In oder to facilitate the reference to the questions and proposed changes, we use the following color coding:

Color coding:

Reviewer comment
Our comment
Suggested changes in the manuscript

---

1) Regarding the dSCD retrieval and Section 3:
a) The authors use a zenith measurement prior to the scan to analyze the scan. If the upper atmospheric contribution to the dSCDs changes during the scan this can lead to signal in the measurements which is from the upper atmosphere being falsely attributed to lower altitudes especially in low eastward elevation angles at the end of the scan. The effect would be expected to lead to lower VCDs in the morning and higher VCDs in the evening, especially in winter. Because of the short 7 minute scan time this effect would likely only be significant at twilight. Does the instrument acquire data at twilight which are included in the analysis? Can the authors bound the impact of such an effect and compare it to the magnitude of their error budget?

Currently we are not considering any data measured at twilight in the analysis and hence the effect of geometry should be small within a 7 min time window. The difference for the zenith measurements is likely to reflect the difference in atmospheric conditions over 7 minutes. If these differences are huge, one of the assumptions made in the concept of MAXDOAS, same conditions for measurements in all off-axis directions, is violated and the scan in question most likely not suitable for profile retrieval.
A way of estimating the error from using the same reference would be to look at the difference between two zenith dscds of two consecutive scans using the noon zenith sky as reference. We made such a test and found that typically, these differences are of the order of several $10^{41.}$ $molec^2/cm5$. This is smaller/ of the order of the typical fitting errors for O4 of typically a couple 10^42.

b) For the fitting setting in the retrieval the authors use older cross-sections where newer cross-sections for the same gases are increasingly standard in the community e.g. (Damadeo et al., 2013; Peters et al., 2017). For O3 they use (Burrows et al., 1999) rather than (Bogumil et al., 2003) or (Serdyuchenko et al., 2014) and for O4 they use (Hermans et al., 1999) rather than (Thalman & Volkamer, 2013). Was there any particular reason for these choices? Is it based on Orphal, 2002 cited later?

There was no particular reason for using the chosen cross-sections, they were chosen because they are quite standardly used for retrievals. At the moment, we are re-running all our dataset using fitting settings as in Peters et al. (2017). A comparison among our chosen settings and the Peters et al., (2017) settings is shown in Fig.1 (data for one scan sequence).
a) Our retrieval settings.
b) Peters et al., 2017 retrieval settings.
c) a and b plotted together

d) Difference (a-b) with the error bars: $\sqrt{\varepsilon_a^2+\varepsilon_b^2}$

From this graphs it is evident that the difference arising from the use of different cross-sections is small compared to the typical fitting error in the DOAS retrieval.

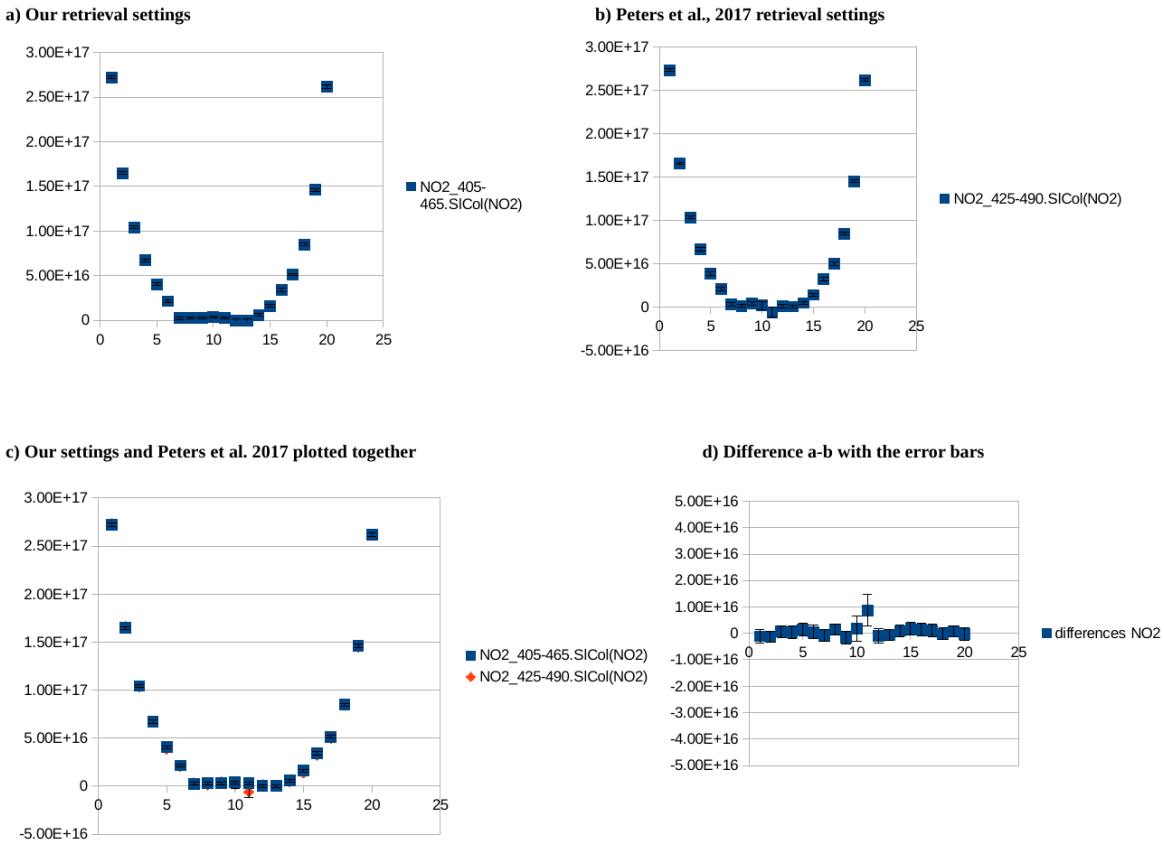

Fig1: Comparison of results for one scan using retrieval settings as used in the manuscript with results using retrieval settings as in Peters et al. (2017)

2) Regarding the profile retrieval method description and Section 4

a) At present the aerosol and trace gas inversions in Fig. 1 are presented as the same, whereas the former uses Tikhonov regularization and the latter optimal estimation. This should be reflected in the figure as it is in the text.

We agree with the reviewer and would like to make the following change in Figure 1. We include a suggested Figure to replace Figure 1 in the manuscript here as Fig. 2.

* Replace the first red box by: “Inversion using Thikonov Regularization”

* Add on the right hand side of that box a box with “L1 Scaling fator” and connect this box with a left-ward pointing arrow to the red box

* Replace the second red box by “Inversion using OE”

* Add on the right hand side of this box a box with “Sa from simulations”

Also, reviewer 1 has pointed out a missing “rates of change” as input quantities, we also add this to the orange and green box that leads to the yellow VLIDORT boxes.

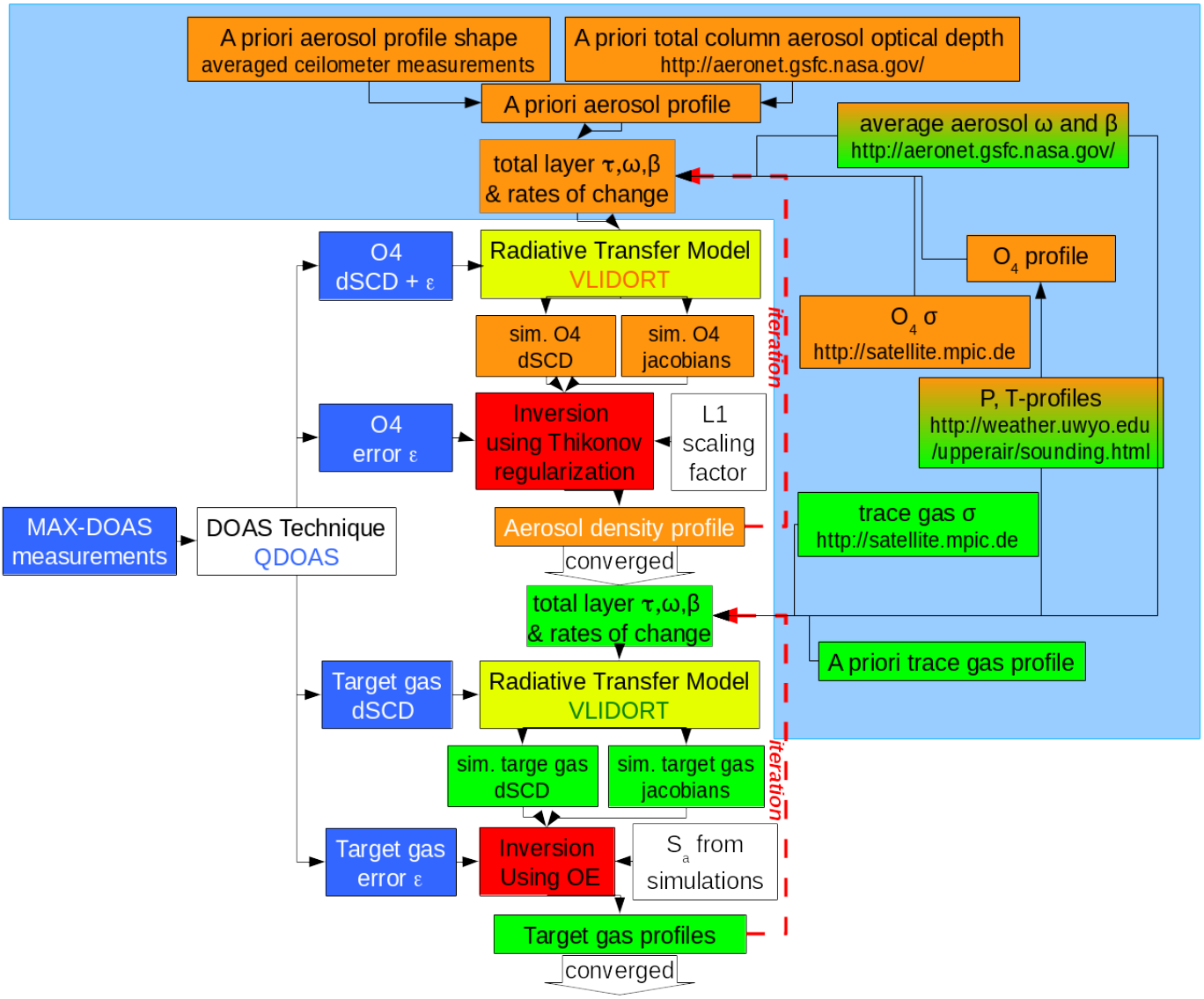

Fig2 : replaces Figure 1 in manuscript

b) **T**he code used to analyze the 18 month data set presented in the work utilized a Gauss-Newton (GN) iteration scheme for inversion, however, MMF has since been updated to utilize a Levenberg Marquardt (LM) iteration scheme, as well as other more minor updates. At present both schemes are described somewhat in parallel, and the authors are diligent in describing which scheme they are discussing. Nonetheless, equations for the GN scheme, which was used, are sometimes left out in favor of the more current LM scheme equivalents, leaving the methods applied not fully transparent to the reader. I would

recommend describing the GN scheme as default as it is most relevant to the titular topic of the work and collecting and describing the changes for the LM scheme either all together in a dedicated section or within the relevant subsections.

We agree with the reviewer that it is not very clear and we intent to improve this by moving the changes to an appendix and concentrate on the description of the actual used code in the main text. Specifically, this means:

* moving the mentioning of the change to LM to a footnote, which then refers to a new appendix A

* Creating an Appendix A (Current appendix A will then change to appendix B). The current footnote 1 will then be part of that appendix A. Eq. 2 and the cost-function (Eq. 3) Is then moved to that appendix.

* The 3rd paragraph of Sect. 4 lines 8--10 ["MMF has been participating...] is moved to the Appendix as well.

* Last paragraph of Sect. 4.1 (page 8, lines 4–8) is moved (and slightly reformulated) to the new Appendix A

However, we are not sure which equations the author sees missing.

Page 6, line 8 –11 replaced by:

"The retrieval time per aerosol and trace gas retrieval with the Mexico City set-up is roughly half a minute for each scan, but highly dependent on the conditions."

Sect. 4.1 Inversion theory replaced by (This also includes a change addressing point c):

"The inversion strategy relies on the fact that the problem is sufficiently linear so that in the iteration procedure, the new value for the quantity vector in question x (either the aerosol total extinction per layer or the trace gas optical depth per layer) can be calculated using a Gauss-Newton (GN) scheme[1] according to Eq.1 (Rodgers, 2000). This step corresponds to the red box and arrows in Fig. 1

[Eq. 1]

Here, superscript T denotes transposed, superscript −1 denotes the inverse. The index i is the iteration index, the subscript a indicates a-priori values. Sm is the measurement error covariance matrix, y denotes the vector of measured differential slant column densities. F (xi) are the simulated differential slant column densities, calculated using the forward model with input profile xi . Both y and F (xi) are vectors of dimension (# telescope viewing angles). Ki = ∂F (x)I /∂xn is the jacobian matrix at the i-th iteration describing the change of simulated dSCD for viewing angle I when the profile x in layeris varied.

In the case of optimal estimation (OE), the regularization matrix R is equal to the inverse of the a-priori covariance matrix, $R = Sa^{-1}$. OE regularization is used for trace gas retrieval. Other regularization matrices are possible, see e.g. Steck (2002).

For the aerosol retrieval used in this study, we use the L1 operator (R = L1T αL1) where the scaling parameter α is set to a constant value of 20 and is supplied via an input script to limit the degrees of

freedom (DOF) to just slightly above 1. Different scalings for the upper layers and lower layers could be supplied, as well as a complete regularization matrix R.

New footnote (1):

In a recent update, the GN scheme was replaced by the more stable Levenberg Marquardt (LM) iteration scheme, more details on recent changes can be found in Appendix A.

New appendix A: Recent updates of the code

In a recent update of the code, implemented after the analysis presented here (i.e. not used for obtaining the results here) the retrieval space was changed from linear space to logarithmic retrieval space. This means that the retrieval works in a linear (dscd measurement)-logarithmic(profile retrieval) space now. This enhances the nonlinearity of the problem and required a change of iteration scheme.

The GN iteration scheme (Eq. 1), was replaced by a slightly slower but more stable Levenberg Marquardt (LM) iteration scheme (Rodgers, 2000):

[Equation A1, Former Eq. 2]

In order to counteract the slowdown of the retrieval, more restrictions were placed on the observation geometry for a single scan: a single relative azimuth angle and a single solar zenith angle per scan. This means in particular that two different viewing directions cannot be treated as a single scan any longer. Although this means a significant cut in flexibility, it results in a retrieval time speed up of a factor of 4 and a more typical retrieval time per scan (for each component) is around 5 seconds.

Tests using the logarithm of the partial layer vertical column density (for NO2 retrieval) or layer extinction profile (for aerosol retrieval) motivated the change to the LM iteration scheme due to the increased non-linearity when working in a semi-log space as state-measurement space.

With this new configuration, MMF has been participating in the Round-Robin comparison of different retrieval codes for the FRM4DOAS project (Frieß et al., in preparation). It has also participated in the profile retrieval from dSCD from the CINDI2 campaign, both for NO2 and HCHO (Tirpitz et al., in preparation) as well as for HONO (Wang et al., in preparation).

The LM scheme of Eq. 2 has currently only been tested with OE and not with Thikonov regularization, i.e. the aerosol retrieval was also performed using OE.

c) At the top of page 8 is the following paragraph: “For the aerosol retrieval used in this study, we use the L1 operator (R = L1 T α L1) where the scaling parameter α is supplied via an input script to limit the degrees of freedom (DOF) to just slightly above 1. Different scalings for the upper layers and lower layers can be supplied, as well as a complete regularization matrix R.” I understand the latter sentence to describe a capability of MMF, but how was the regularization conducted for the analysis presented later? Was a constant α determined such that the DOF was just over 1 or was something else done?

We used a constant scaling of 20 for all layers. This ensured an average dof slightly larger than 1. We would

like to add this in the text, see change suggestion to (b), or copied here below:

"For the aerosol retrieval used in this study, we use the L1 operator (R = L1T αL1) where the scaling parameter α is set to a constant value of 20 and is supplied via an input script to limit the degrees of freedom (DOF) to just slightly above 1. Different scalings for the upper layers and lower layers could be supplied, as well as a complete regularization matrix R."

d) Discussing the advantages of MC RTM codes, the ability to model statistically rare photons and output information of the distribution of photons is also useful. In particular the statistics are worth mentioning as they quite intuitively play into the time trade-off.

We agree with the reviewer, that the discussion of MC is very interesting, however we prefer not to include that discussion here. It is not really the topic of this paper to give an extended review on RT codes not used in the retrieval technique presented here. We do mention the greater accuracy of MC codes though. We are no experts in MC codes and hence would also not feel very comfortable to discuss them in great detail.

e) For the aerosol retrieval on page 11, line 13-14 "The average sing scattering albedo ω and asymmetryparameter g are not subject to retrieval and are constant in all layers". What values are used?

We use values from Aeronet as is mentioned on the same page in lines 9–11. This was not clear, we will move the definition of g and ω to that paragraph and mention more explicitly that we used extra-/ interpolations:

Replace lines 9 – 12 on page 11 (first paragraph of Sect. 4.2.2) by:

"The (a-priori) aerosol data for total optical depth, average single scattering albedo ω and asymmetry parameter g (used to calculate the phase function moments) are time interpolated values from the co-located AERONET (Aerosol Robotic Network) station in Mexico City (V2, level 1.5 at http://aeronet.gsfc.nasa.gov). They are also extra-/interpolated at the retrieval wavelength. The a-priori shape of the profile is taken from hourly averaged ceilometer data (García-Franco et al., 2018), interpolated at the middle layer height h of each layer."

f) The necessary inputs for VLIDORT are normalized as the authors state e.g. page 9 line 4-5, but this should be made clear more consistently. For instance the listed elements 4-6 on page 9 lines 9-11 should be "normalized rate of change ….". Similarly on page 12, line 10 "… what needs to be done is to calculate the normalized derivatives …" as this is what is presented in Eq. 11,12

We agree with the reviewer that this is not consistently written. We will implement the changes suggested by the reviewer and also make the following change:

Page 12, line 18: "For the trace gas jacobian calculation, the corresponding normalized derivatives are:"

3) Regarding Section 5 and error analysis:

a) In the description of the averaging kernels and degrees of freedom it should be noted that both are

relative to the a priori information. This is especially important for the aerosol retrieval which uses Tikhonov regularization which yields an unbiased estimator contingent upon the a priori. E.g. on page 14 line 21 language similar to should be used.

We agree with the reviewer and thank him for his suggestion, which we will use in the revised manuscript:

"DOF, the number of pieces of information independent of the a priori in the profile retrieval, …"

b) Section 5.3.1 is difficult to parse, particularly the first sentence: "The error originating from the cross-section is estimated by assuming that the column amount regarding to the used cross-section has a uncertainty of 3% (Wang et al., 2017)". I assume Eq. 25 has an error and should have 3% or 3.0% rather than 0.3%, otherwise I am misunderstanding. A clearer distinction in the language regarding errors in the measurements (y) as opposed to in the column or partial columns (x).

The reviewer is correct: it should be 3.0% in equation 25.

We agree also that the first sentence is confusing and therefore we change it to : "The error originating from an uncertainty in the cross-section of 3% (Wang et al., 2017) is also around 3.0% in the vertical column" and similar in the lower profile.

The reviewer well understood, that the profile shows smaller spectroscopic errors where it is dominated by the a priori information.

c) The error budget is composed in a number of different ways with some common terminology describing similar errors in the aerosol and NO2 retrievals. This is relatively clear and transparent in Table 1, but can be difficult to follow in the text.

For instance the measurement of error in NO2 is 2.4% first quoted on page 16 line 6. Later on page 17 line 27 "measurement of noise" of 2.2% is quoted, this latter number is measurement noise in O4 propagated to the NO2 retrieval, a different quantity, nonetheless it can seem inconsistent.

Earlier and more frequent reference to Table 1 would be useful I offer a key example:

The language at the end of Section 5.3.2 should be revised, it is difficult to understand precisely. Starting at page 17 line 27:

"The propagation of the smoothing (4.6%) and measurement noise (2.2%) errors of the O4 retrieval into the NO2-retrieval results in a 5.1% error in the NO2 VCD" this appears to refer to Table 1 line 9 and is reasonably clear perhaps end the sentence here. Continuing, "while if no O4- retrieval is performed successfully the error would be in our example 9.8%", here as I understand it line 7 of Table 1 is now substituted without reference to other errors, this should be stated explicitly.

Finally, "In case we would include the algorithm error (7.8%) introduced by Wang et al. (2017) the error when a O4-retrieval is performed successfully would be 9.4%." This is reasonably clear but there appears to be a discrepancy with line 10 of Table 1.

The referee is correct, there is a mistake: the error is 9.318% -> 9.3 %

The error when a O4-retrieval is performed successfully would be **9.3%**. However the algorithm-error is calculated from the resulting residual of the fit and is not independent on the other error sources as mentioned earlier.

In the revised manuscript we also will refer to table 1 as son as possible as the reviewer suggested.

In Fig. 2 the NO2 dSCD errors are shown, is the variability largely a reflection of the relative magnitude of the underlying dSCDs? Are the proportional errors reasonably constant areound the 2.4% value quoted in Table 1, or do they vary with viewing angle also?

The referee makes a very good point: The 2.4% is just the average, and the error is not constant in percentage. For high elevation angles, the dSCD can take values of 0 or even below 0. Hence, to express the error in terms of percentage for low elevation angles is a bit tricky. In order to give an idea of the dependence of the error in terms of percentage for different elevation angles (i.e. like Fig. 2 in the manuscript but in percentage instead of absolute errors), we calculate the percentage, but w.r.t the average dscd at that elevation angle instead of the (sometime negative) actual dscd:

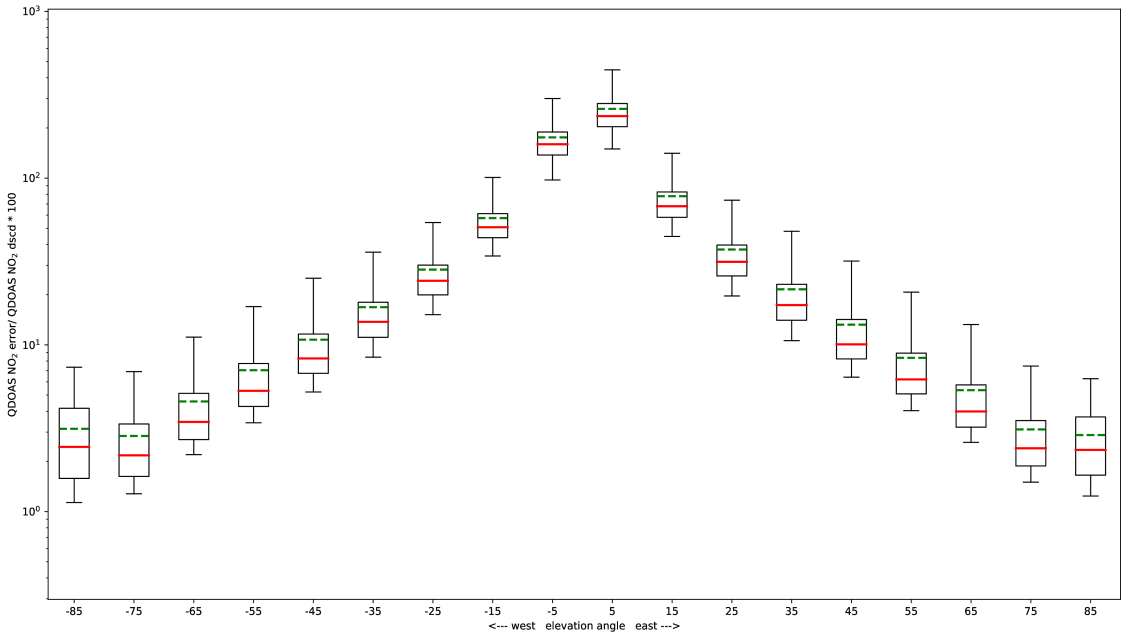

Fig. 3: As Figure 2 in manuscript but in terms of percentage w.r.t average dscd.

4) Regarding the Results and Conclusions

a) For the limited degrees of the aerosol retrieval, the authors state (page 21 lines 2-3) that “Currently, the integration times in the spectra from which the O4 dSCDs are calculated, are not long enough to ensure an O4 dSCD error resulting in DOF larger than 1 for the aerosol retrieval.” However, based on the error budget presented in Table 1, the measurement noise in O4 is the smallest component. Should increased integration times be expected to yield significant improvement? In the next sentence: “Since we use a Tihkonov regularization for aerosol retrieval, this means that we can basically retrieve the total aerosol extinction.” Based on Fig. 4(b) the retrieved DOF is approximately a total column below ~5.5km, very likely similar to AOD under most circumstances, but not necessarily the same.

The statement "Currently, the integration times in the spectra from which the O4 dSCDs are calculated, are not long enough to ensure an O4 dSCD error resulting in DOF larger than 1 for the aerosol retrieval” refers to the limited DOF that could be achieved from a profile retrieval due to the large (typically of the order of a couple of 10^{42} molec^2/ cm^5) dscd error. The reviewer also cites our sentence “Since we use a Tihkonov regularization for aerosol retrieval, this means that we can basically retrieve the total aerosol extinction”. This means that we retrieve the total column but likely not the correct profile if this is hugely different from the a priori profile. We would like to draw the attention of the reviewer to page 15 line 20: "The a priori information about the optical properties described by the aerosol extinction profile is designed for cloud free days and therefore the error analysis is just valid for such cloud free days". This means that the errors estimated only hold for those days were the true profile shape is close to the assumed a-priori shape. A 100% error variance was assumed in the error estimation.

b) Regarding the comparison with in situ NO2 measurements, the authors highlight the impact of clouds on the comparison in Figs. 8 and 9 and examine the diurnal and seasonal components of the comparison in Figs. 11 and 12 respectively. Figure 10 to some degree combines all these aspects in the context of case studies. I wonder whether it is possible to build on this further. For instance, the slope of a MAX-DOAS – in situ comparison can be to some degree inferred from the information presented in Figs. 11 and 12, are there sufficient statistics to present Pearson’s R on these graphs also? If so it might bring greater precision to some of the discussion. Similarly, the results in Figure 8 should have some diurnal and seasonal variation which would help point to the representativeness of the effects highlighted in the Fig. 10 case studies and accompanying discussion.

We agree completely that looking at the correlation coefficients of individual days will give further insight into understanding in more depth the effects of the local dynamics and vertical and horizontal inhomogeneities. This will be interesting even more so when the surface in situ data is analyzed with a higher temporal resolution (currently we had only the hourly mean data available from the monitoring station which does limit the calculation of a reliable Pearson’s coefficient). However, we believe that these comparisons succeed in the general objective sought of this study which is to show that the MAX-DOAS results for the lower layers follow reasonably well what is being measured at the surface with a more conventional methodology. We do have the intention of using our data in future investigations to study specific events and understand the individual characteristic that each instrument is capturing depending on their location within the city.

The caption to Fig. 8 says the slopes where forced to zero, while in Fig. 9 the fits have non-zero intercepts. Why the inconsistency? Does this have any significant impact? At present it is difficult to make much of the point cloud in Fig. 9, are the correlations reasonably linear across the space? Binning data and presenting statistics might provide better insight than the present graph.

The slopes in the fits for Fig. 8 were forced to zero deliberately in order to have a robust way to capture the changes in the slope as the number of layers was increased. As can be seen in the offsets reported in Fig. 9, -3.4 and 0.1 ppb, y-intercepts are small for both data sets and will produce insignificantly small changes in the slopes in case they would also be forced to zero. The purpose of this figure was to highlight how the correlation is affected in cloud (R=54) vs. clear-sky (R=74) conditions, and the red and blue solid straight lines are clearly depicting the change in both data sets. We consider replacing Fig. 9 with one where the intercept is also forced to 0 in accordance to Fig.8. This would be Fig.4 below:

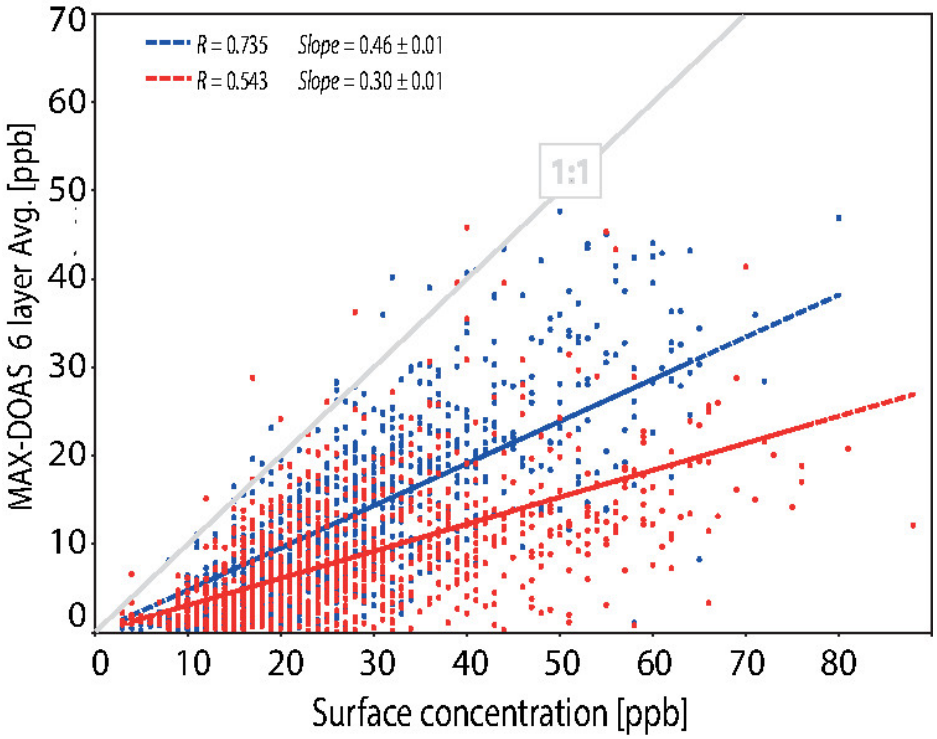

Fig. 4: As Figure 9 in manuscript but with forced offset 0.

Regarding the binning, we don't think binning the data will help making this distinction clearer.

c) The authors conclude that the MAX-DOAS "systematically underestimates the ground level concentrations…", however, this is relative to a single in situ sensor and could in part reflect systemic persistent horizontal inhomogeneity. Such effects have been observed before e.g. (Dunlea et al., 2007; Oetjen et al., 2013; Ortega et al., 2015; Shaiganfar et al., 2011). Particularly at UNAM and in Mexico City, (Rivera et al., 2013) highlights that the MAX-DOAS at UNAM is likely to sample across a significant horizontal gradient. This is relevant to the later discussion of future plans to compare with more sites and with satellites, especially as the Acatlán and Vallejo sites should have overlap in their sampling (Arellano et al., 2016).

We fully agree with these statements, and as mentioned before, this will be investigated further in future work. For example, we are exploring the differences in VCD's obtained when MMF analyses distinctively data measured from easterly or westerly scans. The horizontal gradients are evident already from satellite data as has been accurately pointed out by the reviewer.

Minor comments:

1. Page 1, line 14: "… the total error is considerably large …" large relative to what? The errors do not seem atypical, further they are quantified immediately thereafter.

Change to "The total error, depending on the exact counting is 14 –20 % and this work provides new and relevant information about NO2 in the boundary layer of Mexico City."

2. Page 1, line 19-20: "it is indispensable to have the proper tools to measure them not only at ground level but also throughout the boundary layer." Consider including a citation to support that contention that boundary layer measurements are indispensable as this is quite a strong statement.

Suggested reference is Franco-García et al., 2018.

3. Page 2, line 5: insert "been" to have: "applications of this technique have been demonstrated to"

corrected

4. Page 2, line 11: change "in" to "on" for "restrictions on the usage"

corrected

5. Page 2, lines 12-14: this is a long sentence, consider breaking. Also consider changing "and" to "which" at the end otherwise to clarify relation of clauses, i.e. "… (MCMA) which has been …"

We will follow the suggestions

6. Page 5 line 4: "and average" here should be "an average"

corrected

7. Page 5 line 5: Multiple errors, homogeneity vs inhomogeneity, something can be true or untrue, consider rephrasing e.g. "… since the assumption of horizontal homogeneity likely holds less well." or "… since this likely deviates further from the assumption of horizontal homogeneity."

Thank you for your suggestions. We will implement the latter.

8. Page 6 line 13: eliminate double negative, perhaps replace "… is not too non-linear so …" with "… is sufficiently linear such …"

We will adapt the suggestion.

9. Page 7 line 6: should “unlinear” here be “non-linear”?

corrected

10. Page 7 line 17: “equals” should be “equal”

corrected

11. Page 8 line 7: References to manuscripts in preparation do not appear in the reference section. Here there is a reference to a manuscript by Wang et al. on IO whereas previously on page 6 line 10 there is a reference to a manuscript by Wang et al. on HONO. Are these two difference manuscripts or is one instance a typo?

It is a typo, it should be HONO. Thank you! We will include the manuscripts in preparation to the reference section.

12. Page 9 line 4: Here “Jacobians” is capitalized whereas it was not previously, check consistency.

Jacobian everywhere now.

13. Page 9 line 13: “enclosed” here should replace “inclosed” which is no longer standard.

corrected

14. Page 10 line 15: As described above, the temperature dependence of the cross-section is not presently implemented, as such it should likely be eliminated from Eq. 4.

That is correct, temperature dependence from Eq.4 removed

15. Page 12 lines 14-16: There are a number of formatting errors in equations specifically, Eq. 2, 16, 22, 25. Here three equations appear but only two are numbered, specifically the normalized derivative of ω is not assigned an equation number. In Eq. 12 unlike the previous equations only the simplified expression is given, not an intermediate step in the derivation.

Eq.2: removed closing square bracket after superscript -1 (This Eq. will be moved to a new Appendix A as outlined above)

Eq. 11 – 12: This indeed should be 3 numbers for three equations, this will be fixed.

Eq. 12: That is correct, we skipped it because it was not very readable and we felt that it did not add to the understanding. However, we can of course add it if the reviewer thinks that it adds to the understanding.

Eq. 16: Removed last opening round bracket and changed log to ln.

Eq.22: VMR moved to subscript

Eq.25: Changed 0.3^2 -> 3.0^2

16. Page 13 lines 4,6: The logarithm in Eq. 16 is base e, since ln(x) specifically appears in the text below, these should probably match to avoid the potentially for an apparent difference.

This is a good point, this will be changed, see comment to 15 above.

17. Page 13 line 21: “produce” here should be “produces”

corrected

18. Page 14 line 9: “constraint” here should be “constrained”

corrected

19. Page 14 line 12: Should “not symmetrical” here be “asymmetric”?

corrected

20. Page 18 line 7: The word “most rigorous” is probably not the best choice. Depending on what the authors wish to communicate, most imposing, or least supported might be alternatives.

We take “least supported” then.

21. Page 18 line 9: eliminate the before VLIDORT

corrected

22. Page 19 line 7: eliminate “relatively” it is not needed.

corrected

23. Page 20 line 3: Here “in situ” appears as two word in italics which I believe is the Copernicus standard for such phrases derived from Latin; “a priori” should I believe appear the same way.

We changed all “a-priori” to italic “a priori”.

24. Page 21 line 12: eliminate “in” to get “about half of”, it is not necessary

we changed this to “for about half of the coincident measurements there was…” (i.e. removed “in” and added “for” before about”.

25. Page 23 lines 16-17: consider rephrasing sentence for clarity, perhaps “When all the coincident data is considered, regardless of if the retrieval had data available from the AERONET instrument on that day or not, the R and slope values are 0.62 and 0.39, respectively.”

Thank you for the suggestion which we are happy to use.

26. Page 24 line 16: “relatively” is not needed; change “despite that there are more” to “despite there being more”

corrected

27. Page 26 line 8: Based on Fig. 8 and the prior text, the MAX-DOAS results are on average 0.4 (or 40%) of the ground level in situ measurement. The underestimate then is the difference, namely 0.6 or (60%) is this not the case?

Agreed, changed to:

"However, the MAX-DOAS systematically underestimates the ground level concentrations by a factor of about 0.6. "

28. Page 27 line 1: I don't think the 's is needed after NO2

corrected

29. Appendix A equations: Some of the numbers in these equations given with decimal precision are numeric factors and I don't think require the decimal precision. Some instances are the leading 1's in A8, A9, and A11, and I think all whole numbers in A16 and A20.

Will be modified.

30. References: There are some formatting oddities in the references. Many but not all papers appear with both a DOI code and also a url which in many instances are redundant. The Bates citation includes a citation statistic.

The references were checked, we will only keep the url and modify this in the file.