This markup file is organized as follows:
First, the answers (including original comments and changes in the manuscript) to reviewer 1 are presented followed by the answers to reviewer 2 and reviewer 3. After this, we include the marked up version of the manuscript.

Answer to Reviewer 1

We would like to thank the reviewer for his comments. We agree with the referee that the text was a bit hard to follow in places and we hope that we could improve this by following both his and the other reviewers advices and suggestions for improvement. Since we received a comment from reviewer two before we received a comment from reviewer one, we are sometimes referring in this answer to the answer to reviewer 2.
Below, we comment on the first reviewers specific comments. For easier reference, we added a number to each comment. We use the following color coding:

Color coding:

reviewer comment
our answer
proposed change in manuscript

Review of ”NO2 vertical profiles and column densities from MAX-DOAS measurements in Mexico City” by Friedrich et al.
This manuscript discusses a newly developed profile retrieval code - the Mexican Max-doas Fit (MMF). Note: The first author’s initials match the acronym, nicely done! The retrieval code consists of 2 parts, 1) an aerosol retrieval and 2) a trace gas retrieval using the previously retrieved aerosol profiles. This code is then used on 19 months of MAX-DOAS data measured at a location in Mexico City and the results are discussed. A comprehensive error analysis (which is great to see!) is also included in the manuscript.
It certainly is interesting to look at the complete 19 months NO2 data set (e.g. see the discussed averages of the diurnal variation) but my guess is the more interesting studies (specially from an environmental view point) can be done by looking at individual days and using the right ancillary data to understand the NO2 variability and what causes the observed peaks.
Overall, the manuscript is well structured and the figures and table are clear and straight forward to understand. However, in some places (e.g. in Section 5, Error analysis) the text can be somewhat difficult to follow, and the manuscript could gain from having another go at streamlining the text a bit more and simplifying the structure of some of the more complicated sentences.

1) Page 2, line 16: Replace ‘giving’ with ‘with’
corrected
2) Page 2, line 17: Section 2 should be Sect. 2 for consistency, check whole manuscript
checked in whole manuscript, thank you!
3) Page 2, line 19: ‘(constituting the forward model)’ - What exactly does this mean?
This means that the forward model in our case is a radiative transfer code and that we talk about radiative transfer codes here, because our forward model is a radiative transfer model, more specifically, VLIDORT.
4) Page 2, line 30: UNAM – can you please spell this out once
done at first appearance now, i.e. at the beginning of Sect. 2

5) Page 3, Figure 1, caption. Nice overview figure. For completeness, can you please also include a brief description of the yellow and red box in the caption.
Yes. As response to reviewer 2, we also made small adjustments to the figure, see answer to referee 2. The new caption is describing that version of the figure. We added to the caption:
"The yellow boxes represent the forward modelling steps. The red boxes are the inversion steps, using Thikonov regularization for aerosol retrieval and optical estimation (OE) for tracegas retrieval."
As a response to (23) below, we also now added "& rates of change" in the orange and green boxes before the yellow "VLIDORT box".
6) Page 3, line 3: Replace 'large' with 'long'
corrected
7) Page 4, line 8: Typo: 'receiving'
corrected
8) Page 5, line 4: Typo: 'an average'
corrected
9) Page 5, lines 16-25: Why was O4 not retrieved using the same wavelength interval as NO2? The much older O4 XS from Hermans et al. 1999 was used for the O4 retrieval, why not Thalman and Volkamer (2013)?
Thalman, R. and R. Volkamer, Temperature dependent absorption cross-sections of O2-O2 collision pairs between 340 and 630 nm and at atmospherically relevant pressure., Phys. Chem. Chem. Phys., 15(37), 15371–81, doi:10.1039/c3cp50968k, 2013.
The windows for NO2 and O4 dscd fitting are chosen to enclose pronounced absorption lines for the species in question and are widely used windows. We agree with the referee that there are O4 windows that are closer to the chosen NO2 window (405 -- 465 nm) than our choice (336 -- 390 nm) and that it had been a better choice to use one of those (e.g. 450 -- 520 nm). However, the difference in middle wavelength had only been 22 nm less (i.e. +50 instead of -72). We would also like to refer to our answer to reviewer 3 question 2 regarding correction for aod. Regarding the choice of cross-sections, there was no specific reason for the choice of cross-section. We would like to refer the reviewer 1 to answer 1b to the mayor comments from reviewer 2, where we also include a test for changing the retrieval settings. Our main finding is that the effect of changing the cross-sections is small.
10) Page 5, line 26: Would it be possible to say something briefly here about how the errors were determined?
The dscd error is calculated directly within qdoas. We use this error as dscd error without any modification or addition. We refer the reviewer to pages 28 -- 29 ("Errors on Slant Column Densities) of the qdoas manual (http://uv-vis.aeronomie.be/software/QDOAS/QDOAS_manual.pdf) for details on the dscd error calculation within qdoas.
11) Page 7: The authors explain that the retrieval code was recently updated from using the Gauss-Newton scheme to the more stable Levenberg Marquardt iteration scheme. However, this is not really relevant for the work presented here and seems to unnecessarily complicate the discussion. Unless there is a compelling reason to keep this information, I suggest to drop the relevant equations and just briefly mention in a couple of sentences (or one paragraph) that the retrieval code has been updated and how. It would also be better to have all the variables explained straight after Equation (1) and not further down the page.
We fully agree with the referee. Referee 2 (first report received) had a similar comment. We moved all explanations regarding changes to the code into an appendix. This also leads to all symbols in Eq. 1 being defined right after its appearance. For details on the changes, please see the answer 2b to the major comments section from referee report from reviewer 2.
12) Page 7, line 6: Change to 'non-linear'
corrected
13) Page 7, line 12: Change to 'dimension which is the number of telescope'

done
14) Page 7, line 17: Change to: ‘equal to 1’
corrected
15) Page 7, line 23: Change to: ‘for the trace gas’
included “the”
16) Page 8, line 8: Change to: ‘with the LM iteration scheme.’
included “the”
17) Page 8, line 10: Change to: ‘algorithms. For example, there are’
Added “For example”
18) Page 8, line 20: Typo: ‘high speed’
corrected
19) Page 8, line 21: ‘instead of the 2x the number of layers calls’
corrected
20) Page 8, line 23: Jacobians always with capital J, also on Page 12 & 13
now capitalized everywhere.
21) Page 8, line 23/32 and footnote: Why not refer straight to LIDORT if only that part if used anyway?
VLIDORT and LIDORT are actually different code packages with different version numbers. Since it might be that they, of course by accident, include different “features” (i.e. bugs), we think that it is more accurate to state exactly which code and which version was used.
22) Page 8, line 32: Maybe replace with ‘For each simulated atmospheric layer, ’
corrected
23) Page 9, lines 9-14: It is not quite clear to me how the rate of change is represented in Figure 1, can you please explain . . . or I might have misunderstood?
This was perhaps not clear. We did not refer specifically to the rate of change. We removed the sentence. However, we also realized that the rate of change as layer input was indeed missing in the diagram. We added “& rates of change” in the orange and green box before each of the yellow VLIDORT boxes.
24) Page 9, line 13: Should be either ‘enclosed’ or ‘included’ ?
enclosed
25) Page 9, line 21: Replace ‘is’ with ‘are’
corrected
26) Page 11, line 2: Comma needed after Qray
corrected
27) Page 11, line 14: ‘are assumed to be constant in all layers.’
corrected
28) Page 11, line 15: Replace ‘are’ with ‘is’
corrected
29) Page 11, line 16: Change to ‘density profiles in arbitrary units from . . .’
corrected
30) Page 11, line 17: Change to ‘heights h to provide
corrected
31) Page 11, line 19: I am not sure if all readers will know what is meant with an ‘intensive quantity’, maybe explain briefly in a footnote?
Ok, we add as a footnote: “bulk property which does not change when changing the size of the system”
We also noted that we did not explicitly mention the first step where we convert the relative intensive profile to an extensive one in the frist place (for scaling) before we convert it back to an intensive one. Which is very confusing. We would like to change this by changing line 18 to “This profile, turned into a partial optical depth per layer by multiplying with the layer thickness, is scaled to match the total aerosol extinction from AERONET $\tau_{\rm aer}$. The profile is then converted back into an...”

32) Page 13, Equation 16: Rogue bracket or is something missing?
Opening bracket removed
33) Page 13, line 18: Gain is written in a strange font, on purpose (why?)? If not, please fix.
Changed to same font as AK everywhere
34) Page 13, line 21: ‘produces’
corrected
35) Page 14, line 9: Add comma after fitting
corrected
36) Page 14, line 23: Change to ‘AK matrices from the other errors.’
corrected
37) Page 15, line 1: ‘the VMR(VMR)’ – is that correct?
Yes, it describes the difference between subscript “VMR” and “pcol”
38) Page 15, lines 16/17: Why would the vertical aerosol axtinction profile not be available?
Because the aerosol retrieval failed, or because it was judged to be a bad retrieval due to a large rms w.r.t. measured and simulated dscd
39) Page 16, Figure 4 caption, last sentence: ‘an ideal’
corrected
40) Page 16, line 6: Add comma after ‘operator’ – makes this sentence a bit easier to read.
corrected
41) Page 17, Equation 25: Should that be 3% instead of 0.3%?
yes, corrected
42) Page 17, line 8: Should either be ‘error . . . is’ or ‘errors . . . are’
corrected
43) Page 17, line 14: Comma after retrieval
corrected
44) Page 17, line 15: ‘contributions: a) smoothing error and b) .. error.
corrected
45) Page 18, line 1: Comma after (2017)
corrected
46) Page 18, line 3: Could use ‘dependent’ instead of ‘not independent’.
corrected
47) Page 18, line 5: Delete ‘it’. Comma after ‘However’
corrected
48) Page 18, line 9: Delete ‘the’ before ‘VLIDORT’.
corrected
49) Page 18, lines 10/11: Add commas after ‘(2017)’ and ‘the residual’
First coma added, second would be incorrect, we believe
50) Page 19, Figure 6: The two solid orange lines are hard to distinguish, could use dash or dash/dot for one of them.
We agree with the reviewer and will change one of the orange lines to a dashed orange line.
51) Page 19, Figure 6 caption: Change to ‘a) The square . . .’ and delete full stop after ‘total’
corrected
52) Page 19, line 3: Better: ‘errors for No2 and O4 calculated ‘
corrected
53) Page 19, line 5: Change to ‘errors’ and delete ‘fairly’
corrected
54) Page 19, line 7: Delete ‘relatively’
corrected
55) Page 19, line 8: Something is not right with this sentence & it doesn’t make sense as it is written. Maybe delete ‘to’ or rephrase altogether.

Changed to two sentences: "The error in the vertical column is smaller than the errors in the VMR profile for almost all layers (Fig.6). This can be explained by an anti-correlation in different partial column errors indicated by the full error covariance matrix."

56) Page 20, Table 1, caption: The last sentence is a bit hard to read; would help to add a comma after included and it needs a 'with' after better.

We reformulated to:

"However, if the algorithm error according to Wang et al. (2017) is included, the remaining error due to the uncertainty in the aerosol profile is slightly better: 9.3\% instead of the 9.8% without O_4 retrieval."

57) Page 20, line 2: Add comma after 'In this section'

corrected

58) Page 20, line 4: Typo: 'approx.'

corrected

59) Page 21, line 10: Typo: 'Currently'

corrected

60) Page 23, line 1: I would rather say: 'Generally, a better . . .'

corrected

61) Page 24, lines 8-11 and Figure 10: Would be really interesting to get higher resolved surface measurements as well, otherwise a small increase might be hidden in the surface data set as well. The peak only shows up clearly in the individual measurements with sufficiently high temporal resolution. Similar peak also shows up on Aug 15 and one could argue to some degree even on 9 Sep and 22 Dec with a bit of a time shift. Any idea what causes it?

We think this NO2 enhancements might be transported from somewhere within the basis. Definitely, surface data with higher temporal resolution would allow us to do a more in-depth analysis on a day-to-day basis, as mentioned to referee 2. We think, however, that such a detailed study would divert from the main objectives of the paper, which is to describe the methods and quality of these data.

62) Page 24, line 28: Change to: 'This might have to do with the fact that during . . .' otherwise something seems to be missing from this sentence.

Yes, indeed. corrected

63) Page 24, lines 20-23 & Page 26, Figure 11: Could you add a brief discussion here on the nicely (amazingly?) constant offset between surface and MAX-DOAS data, also including the uncertainties of both data sets in that discussion. Would you say that this is predominantly caused by NO2 having strong emissions on the surface which are then just diluted over the vertical range which the MAX-DOAS measurements are covering?

Thank you for this comment. We added the following sentence at the end of the paragraph (Page 24, line 23). "Despite the fact that the offset in the curves for surface- and MAX-DOAS measurements appears to be nearly constant throughout the day, it would be interesting to investigate further how this offset varies in different seasons particularly when vertical mixing is not favoured".

64) Page 26, line 1: Change to: 'and certain trace gases. We . . . NO2 at one . . .'

corrected

65) Page 27, line 1: delete ' 's'

corrected

66) Page 27, line 4: Add something like 'Sincs this study, it has been . . ..'

We take this suggestion

67) Page 27, line 10: Add 'the' before 'NO2'.

corrected

68) Page 31-34, References: There seems to be some doubling up of information, please check through all the references for correct formatting.

Thank you for the note on this, we looked over the reference formatting, see also reply to referee 2 on this subject. Unfortunately, there were problems with the mark-up in the reference section. There is no mark-up in the reference.

Answer to Reviewer 2

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

W.r.t the answer given in the comments, we slightly enhanced the new Appendix A and updated this here in this answer. W.r.t. the difference PDF: The updated references are not marked-up by color.

---

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.

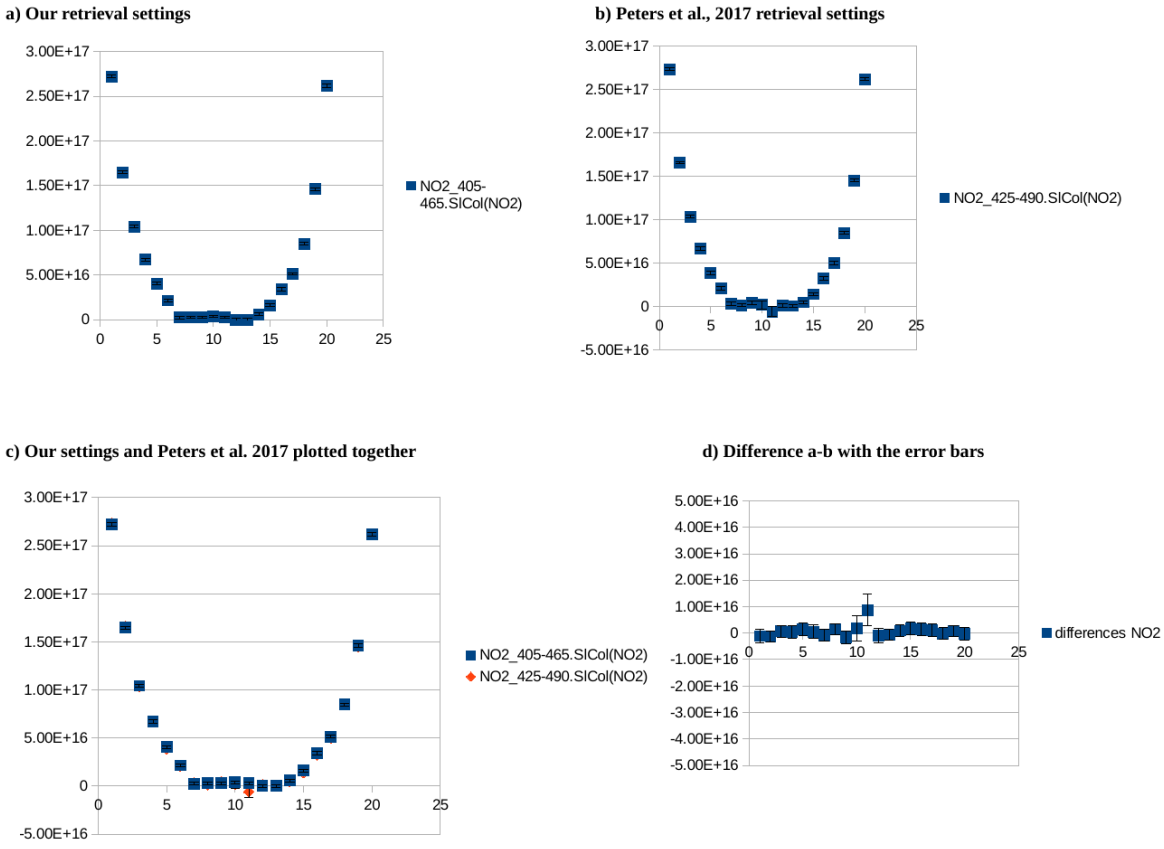

Fig1: Comparison of results for one scan using retrieval settings as used in the manuscript with results using retrieval settings as in Peters et al. (2017)

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.

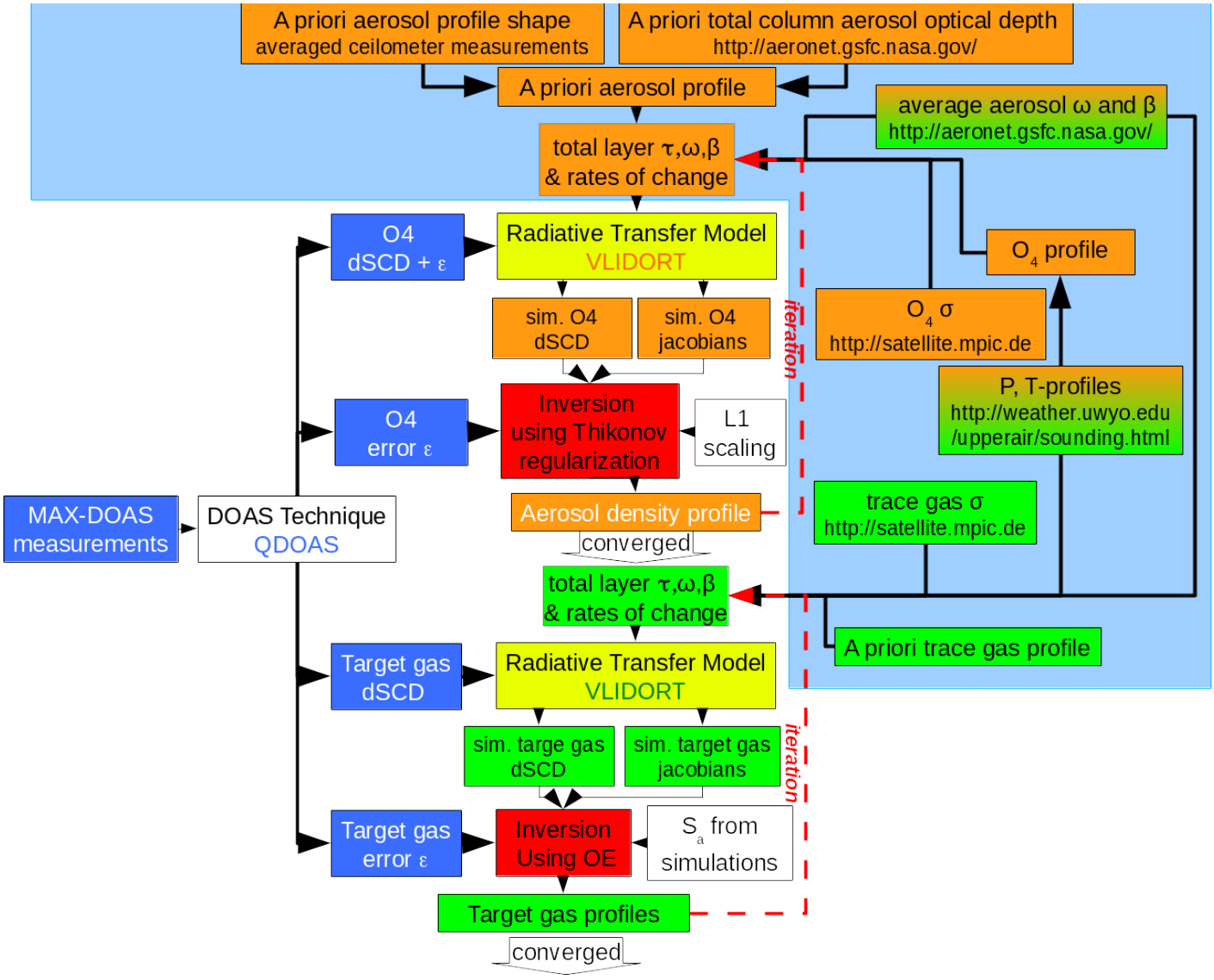

Fig2 : replaces Figure 1 in manuscript

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 ($R = L1T\,\alpha L1$) 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, i.e. Eq. 1 was replaced by Eq. A1 for more non-linear inversion problems (Rodgers, 2000)

Former Eq.2, now Eq. A1

The symbols mean the same as in Eq. 1: superscript T denotes transposed, superscript −1 denotes the inverse, subscript a indicates a priori values. S m is the measurement error covariance matrix, y denotes the vector of measured differential slant column densities, F (x i) are the simulated differential slant column densities at iteration i, calculated using the forward model with input profile x i , K i is the Jacobian matrix at the i-th iteration. The new x i+1 is only accepted if the cost function in Eq. A2 decreases w.r.t the previous cost-function

Former Eq.3, now Eq. A2

If this is the case and (1 + γ) is not yet equal to 1, the factor (1 + γ) is halved for the next iteration. If however, the cost-function increases, the newly calculated x i+1 is discarded and the i-th calculation repeated with a factor (1 + γ) increased by a factor of 16.

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 is around 5 seconds.

Tests using the logarithm of the partial layer vertical column density (for NO 2 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., 2018). It has also participated in the profile retrieval from dSCD from the CINDI-2 campaign, both for NO 2 and HCHO (Tirpitz et al., in preparation) as well as for HONO (Wang et al., in preparation). The LM scheme of Eq. A1 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:

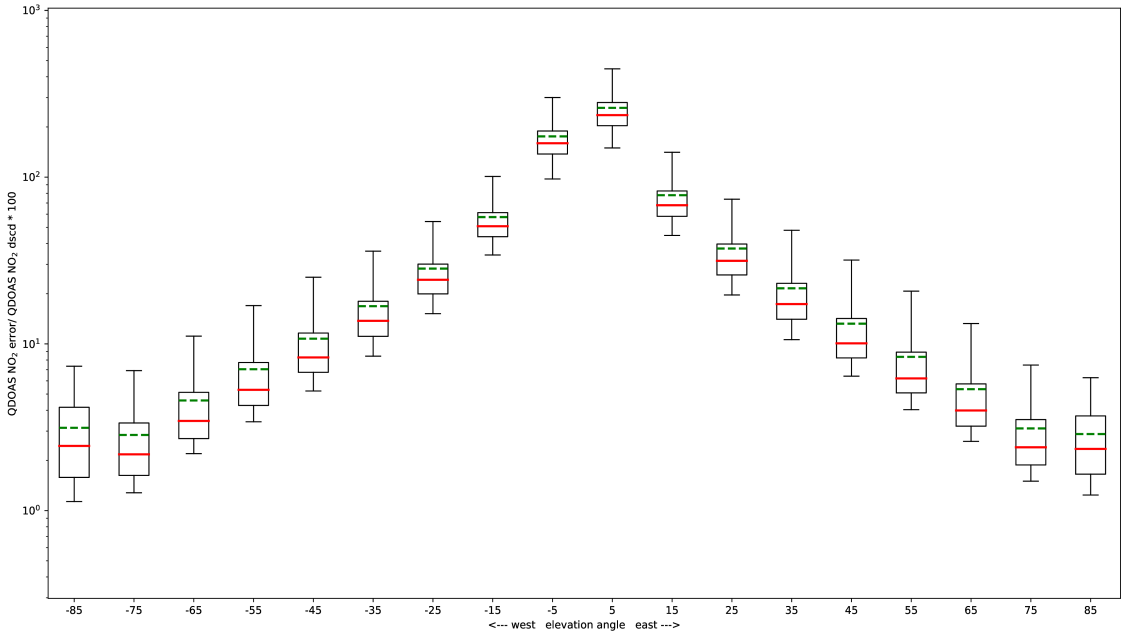

Fig. 3: As Figure 2 in manuscript but in terms of percentage w.r.t average dscd.

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 are replacing Fig. 9 with one where the intercept is also forced to 0 in accordance to Fig.8. This would be Fig.4 below:

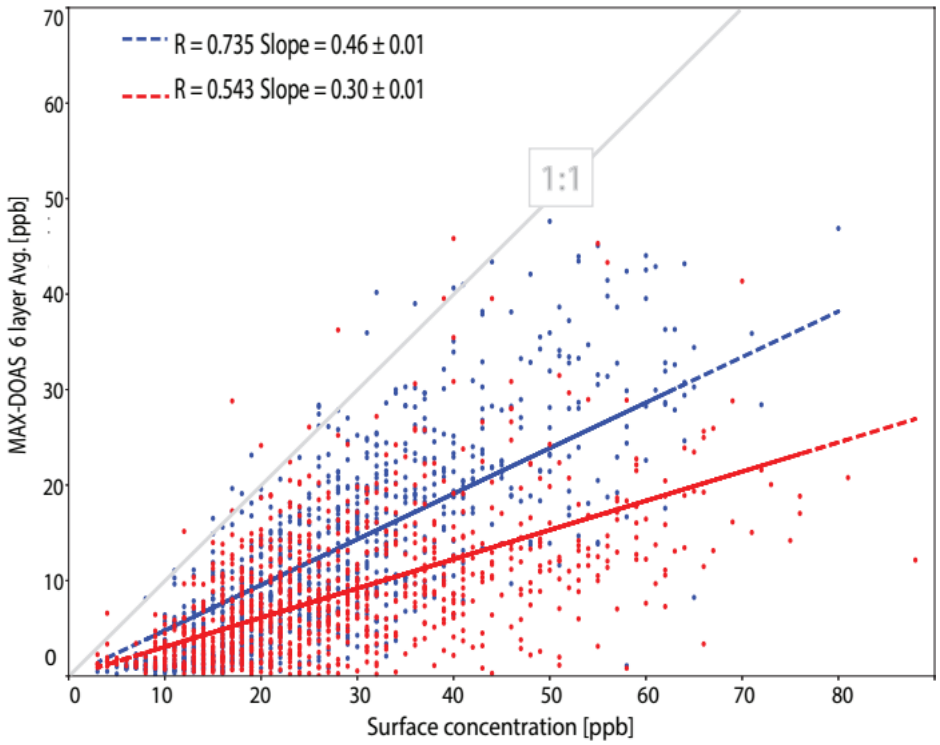

Fig. 4: As Figure 9 in manuscript but with forced offset 0. Replaces Fig. 9

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.

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 keep the url. However, there was an issue with the doi of Arellano et.al. For this one, we report the url instead. (The editors were informed and the issue is being solved).

ANSWER TO REVIEWER 3

We would like to thank reviewer 3 for their comments. In the following, we respond to each of the comments individually. In order to facilitate the tracking, we use the following color coding:

Color coding:

reviewer comment

our answer

proposed change in manuscript

General comments:

1/I think it would help the reader to include a map (that could be Figure no 1 of the manuscript) of the Mexico City area showing the locations of the different instruments (MAX-DOAS, AERONET, in-situ). Moreover, indicating the pointing directions of the MAX-DOAS instrument involved in this study could maybe also give some insights on the interpretation of the discrepancy between MAX-DOAS and in-situ NO2 surface concentration values. For instance, part of the underestimation of the in-situ values by the MAX-DOAS could be related to the fact that the MAX-DOAS instrument points towards a part of Mexico City which is less polluted than the location of the in-situ instrument.

In addition to the map, the location (latitude, longitude, altitude) of all the instruments should appear in the text.

The AERONET site and the in-situ measurement site are at the same position (give a nd take a few tens of meters) as the MAXDOAS station. The site for the ballon launch for the sounding is at the airport (a few kilometers to the north). However, since the position of the ballon quickly changes, we do not think that the actual launching position is very important. We will however state the coordinates in brackets below the first mentioning of the measurements. As for a map for the MAXDOAS instrument and its orientation, such a map is included in Arellano et al. 2016. We will refer to that in the manuscript specifically.

2/AERONET data are used as input in the retrieval but also as ancillary data for the sky conditions screening. Was there any attempt to compare the retrieved AODs with those from AERONET ? It can be a good check for the aerosol retrieval part of the profiling. Also related: it seems that the availability of AERONET observations has been used as a quality control (QC) flagging for the MAXDOAS retrievals. Was there any attempt to apply a QC flagging which is more specific to the MAX-DOAS retrievals, e.g. using parameters like DOF and the RMS of the differences between measured and calculated dSCDs ?

We first clarify which aeronet data is used in which way (a) and then answer the question about the filtering (b):

(a) We use the AERONET data of omega and g as input for the forward model (time interpolated). We do not attempt to retrieve those values. Further, the extrapolated (at the aerosol retrieval wavelength, and time interpolated) aod value is used as a-priori. As input for the NO2 retrieval, an interpolation between the retrieved aod and the nearest AERONET wavelength (time interpolated) was performed, see also answer to question “specific 2” below. In case of failure of the aerosol retrieval (non-convergence, or a bad fit in terms of the average absolute value of the difference in measured and simulated dscd in units of dscd error [sum(abs(DSCD_sim-DSCD_meas)/DSCD_meas)/number_of_elevation_angles] but no filtering on DOF because we designed the scaling of the Thikonov constrained in such a way to have a DOF of just above 1), an extrapolation of the two nearest (both to the long-wavelength side) AERONET values (time interpolated) was performed. Hence we do not see too much sense to compare to AERONET.

(b) In order to ensure a small forward model error it is important to have good estimations on the aerosol parameters g and omega. Without AERONET data available, we use an interpolation of the nearest available data in time. Hence the forward model error is expected to be smaller if AERONET data is available close in time. We also use the presence of AERONET data as a proxy for cloud free conditions. See also the answer to question 4a from reviewer 2.
Regarding the NO2 retrieval we do currently not use any filtering, not on RMS not on DOF. We looked at the distribution of DOF. We show a histogram in Fig. 1

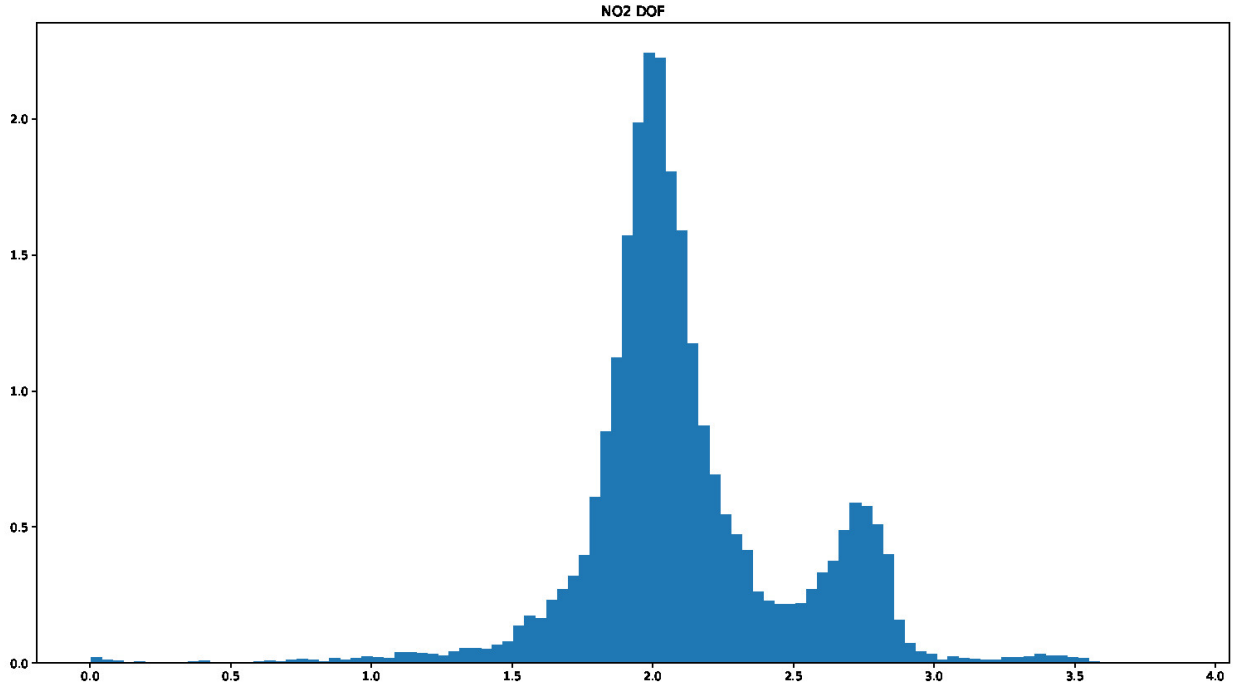

Fig.1 Distribution of DOF for NO2 retrievals.

If we were to include a filter on DOF, we would likely choose a limit of 1.5, hence the impact had been rather small.

Specific comments:

1/Page 4, beginning of Section 2.3: It would be good to list the exact elevation angle values of a typical scan.
We will include this in the last line on page 4, just before mentioning the likely change in measurement sequence.

2/According to Section 3, it seems that aerosol profile retrieval is done in the UV range and then retrieved profiles are used as input for the NO2 profile retrieval in the visible range. Has a correction been applied to the retrieved extinction/AOD for taking into account for the wavelength dependence of the AOD/extinction ? If not, then this approximation should be included as an additional error source in Table 1.
We perform a linear interpolation at the NO2 retrieval wavelength using the value retrieved value in the UV and the closest AERONET value. We plan to use an O4 retrieval window closer to our NO2 retrieval window in the future. The error arising from this would be one contribution to the estimated algorithm error which is already in the table (The error in NO2 from errors in the aerosol profile). The easiest way to get this error contribution would be perhaps to use PANDORA

instruments (calibrated-direct sun measurements) collocated to AERONET sites. But up to now there are not yet sufficient coincident measurements in Mexico City

3/Page 6, lines 8-11: Maybe you could add a couple of sentences about the performance of MMF in these profile comparison exercises. Please note that in the meantime, Friess et al. is now published in AMTD.
Thank you for the note that Friess et al. Is now published. This will be added to the references. The inclusion of the changes in MMF is moved to an appendix (see comment to question 2b from reviewer 2), a note on the time performance is included.

4/Page 8, line 7: Is it IO or HONO (cf page 6, end of Sect. 4) ?
It is HONO, we corrected this and also added the missing reference to the list of references.

5/Page 8, Sect. 4.2: You should add a paragraph on the SCIATRAN RTM, which has been also used in past MAX-DOAS profiling studies (see e.g. Friess et al., AMTD, 2018).
We will add SCIATRAN RTM in the list of examples for radiative transfer models used as forward models for profile retrieval with MAXDOAS.

6/Page 10, line 1: what type of interpolation is done for the pressure, temperature profiles ?
We use simple linear interpolation. We know that this can be improved upon for P. However we expect the effect to be rather small since the grid for the T and P profiles is of a similar resolution than the internal retrieval grid.

7/Page 10, line 14: a correlation length of 500m is used. Did you perform sensitivity tests on this parameter in order to estimate its impact on the retrieved profiles and on the level of agreement with in-situ measurements ?
The correlation length was only used for the error calculation, not for the retrieval. For the retrieval, no Sa matrix was constructed, but a Thikonov constrained used.

8/Page 19, Figure 6b: A priori profile should be also included in this Figure in order to see how far the retrieved profile differs from the a priori one.
We will include the a-priori profile in the plot. Also, as response to reviewer 1, we change one of the orange line for easier distinction. A proposed new Figure 6 (in the manuscript) is reproduced here as Fig. 2 (we will adjust the axis labels and tick labels to a more readable font size)

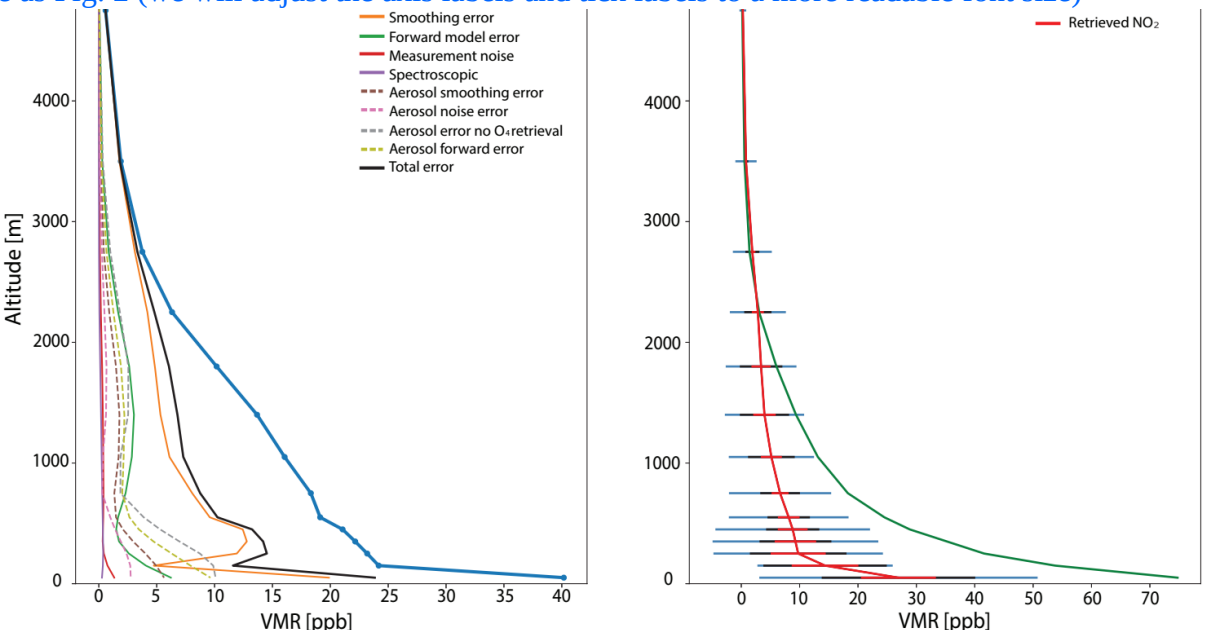

Fig. 2: as Figure 6 in the manuscript but including the a-priori (right) and a better distinction of the curves (left).

9/Page 26, lines 5-10: According to the authors, a possible reason for the underestimation of the in-situ surface concentration by the MAX-DOAS is the fact that the MAX-DOAS instrument has a maximum sensitivity around 1km and less sensitivity close to the ground. This feature is quite unusual since normally lowest elevation angles have a higher weight in the retrieval due to higher AMFs, and therefore MAX-DOAS measurements close to the ground. Could the authors elaborate on that ? Another possible explanation for the discrepancy is that, related to the horizontal extent and the pointing direction of the MAX-DOAS measurements, both instruments probe different air masses. I think this point should be also added in the discussion.

The averaging kernel, for a typical AK see Fig.4(a), shows that we expect an underestimation at the surface. It peaks at around 1km. The main reason for the lower sensitivity is the measurement angle distribution (too few low elevation angles) and the rather huge dscd errors at these low elevation angles, see Fig.2. AK=(K^T Se-1 K+Sa-1)-1 (K^T Se-1 K). While Sa^-1 is constant, a full covariance matrix from model averages, Se-1, the measurement error and K, the Jacobian, mainly dependent on the aerosol content, are variable. Therefore, the AK is only an example.

We can actually estimate the slope and underestimation theoretically using a typical Averaging Kernel using the variability of NO2 in the Mixing layer described by the Sa. However, since Sa is only an estimation, we tried two Strategies a) either using the Sa calculated from profiles of the model run and another taken from the Literature Wang et al., 2017. The first one is used as constraint in the OET-retrieval and described in the Manuscript. The latter uses a 100% variability of the a priori an the diagonal and a 500m exponential correlation length for off-axis elements as in Wang et al. (2017).

The Slope between a retrieved quantity, either the total column or the average of some layers, is calculated, respectively, by applying an operator g=(1 1 1 1 1 1 ....) on the profile in units of partial columns or g6=(1/6, 1/6 ,1/6, 1/6, 1/6, 1/6,0,0,0....) in VMR. "g6" is the operator which calculates the average of the lowest 6 layers. To get the in situ value we apply g1=(1 0 0 0 0...) on the profile.

The linear relation between retrieved values (e.g. averages of 6 layers) and in situ values depend on the correlation between all layers and is theoretically described by the following expression:

$< g_6\, AKV_{VMR} | SA_{VMR} |\, g_1 >$ which assumes that the profile variability is described by a normal distribution P(x)=1/sqrt(Pi Det(SA)) Exp(-(x-xa)^t SA-1 (x-xa)) (Rodgers 2000) of the Variability.

The theoretically calculated slope can be compared to the experimental obtained slope (Fig. 8) and so the Sa- matrix can be tested for, how plausible the estimation was.

In Fig. 3 here, we show the slopes (y-axis of Maxdoas v.s. insitu for the average of a different numbers of the lowest layers indicated by the x-axis (just as in Fig. 8 in the paper).

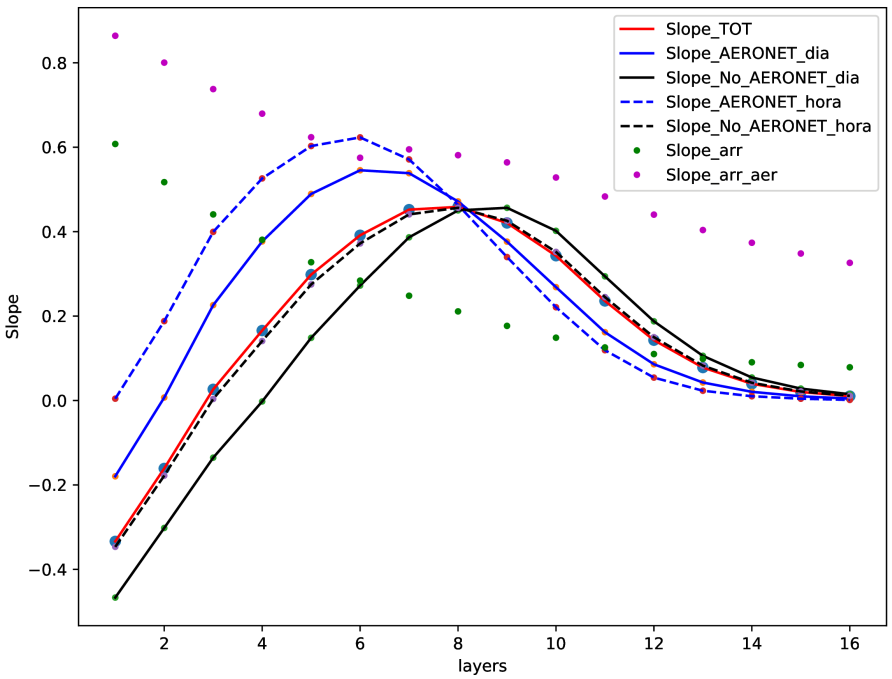

Fig3: As Figure 8 in the manuscript, just with a few more lines with respect to different filters related with the aeronet data and we added the two theoretical calculations shown by dots, green dots using the SA described in this work and the pink dots using the SA-Matrix described by Wang et al, 2017. Details see text.

The points are calculated theoretically either using the SA we constructed from the ensemble of modeled profiles (green points, Slope_arr) or taken from Wang et al., 2017 (pink points, Slope_arr_aer).

The graph explains even quantitatively the underestimation and no other arguments are needed. Just for the average of very few layers, there seems to be a discrepancy, which might indicate that the SA-Matrix do not describe the variability and correlation of the lowest level correctly.

We learned a lot by this exercise, but we would prefer not to complicate the manuscript too much and suggest just to add

"and variability Sa" behind the averaging kernel and maybe add as well the two calculated values

10/Acknowledgements: Depending on the conditions of use, the sources of ancillary data included in your study should be acknowledged here.

Thank you for pointing this out, we will include appropriate acknowledgements.

Technical corrections:
1/Page 1, line 1: ‘. . .to retrieve profiles. . .’ -> ‘. . .to retrieve vertical profiles. . .’
corrected
2/Page 1, line 10: ‘. . .at the Universidad Nacional Autónoma de México (UNAM) campus.’
corrected
3/Page 1, line 20: ‘. . .The Multi-AXis Differential Optical Absorption Spectroscopy

(MAX-DOAS) technique. . ..’
corrected
4/Page 2, line 30: ‘. . .at the Universidad Nacional Autónoma de México (UNAM).’
corrected
5/Page 6, line 5: ‘Sect.4.1’ -> ‘Sect. 4.1’. This typo should be corrected throughout the manuscript; similar corrections also needed for ‘Fig.’ and ‘Eq.’.
Sect, Fig. And Eq. checked for space behind.
6/Page 6, line 9: ‘CINDI2’ -> ‘CINDI-2’
corrected
7/Page 10, line 14: ‘. . .Eq. 4:’; same on Page 11, line 4
corrected
8/Page 11, line 20: ‘extincion’ -> ‘extinction’
corrected
9/Page 16, legend of Fig. 4: (b,right)’ -> ‘(b, right)’; should be also corrected for (a, left).
corrected
10/Page 19, legend of Fig. 6: ‘total error’
corrected
11/Page 19, line 3: ‘algoritm’ -> ‘algorithm’
corrected
12/Page 20, line 4: ‘aprox.’ -> ‘approx.’
corrected
13/Page 21, line 11: ‘Curretly’ -> ‘Currently’
corrected
14/Page 24, end of line 23: A reference could be added here.
We added García-Franco et al. 2018

NO_2 vertical profiles and column densities from MAX-DOAS measurements in Mexico City

Martina Michaela Friedrich[1,2], Claudia Rivera[1,3], Wolfgang Stremme[1], Zuleica Ojeda[1], Josué Arellano[1,4], Alejandro Bezanilla[1], José Agustín García-Reynoso[1], and Michel Grutter[1]

[1]Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México
[2]Belgian Institute for Space Aeronomie (BIRA-IASB)
[3]Facultad de Química, Universidad Nacional Autónoma de México
[4]Instituto de Geofísica, Universidad Nacional Autónoma de México

Correspondence: M.M. Friedrich (martina@atmosfera.unam.mx)

Abstract. We present a new numerical code, Mexican Maxdoas Fit (MMF), developed to retrieve profiles of different trace gases from the network of MAX-DOAS instruments operated in Mexico City. MMF uses differential slant column densities (dSCDs) retrieved with the QDOAS (Danckaert et al., 2013) software. The retrieval is comprised of two steps, an aerosol retrieval and the trace gas retrieval that uses the retrieved aerosol profile in the forward model for the trace gas. For forward model simulations, VLIDORT is used (e.g. Spurr et al., 2001; Spurr, 2006, 2013). Both steps use constrained least square fitting, but the aerosol retrieval uses Tikhonov regularization and the trace gas retrieval optimal estimation. Aerosol optical depth and scattering properties from the AERONET database, averaged ceilometer data, WRF-Chem model data as well as temperature and pressure sounding data are used for different steps in the retrieval chain.

The MMF code was applied to retrieve NO_2 profiles with two degrees of freedom (DOFs=2) from spectra of the MAX-DOAS instrument located at the UNAM campus. We describe the full error analysis of the retrievals and include a sensitivity exercise to quantify the contribution of the uncertainties in the aerosol extinction profiles to the total error. A dataset comprised of measurements from January 2015 to July 2016 was processed and the results compared to independent surface measurements. We concentrate on the analysis of 4 single days and additionally present diurnal and annual variabilities from averaging the 1.5 years of data. ~~Even though the total erroris considerably large (~~The total error, depending on the exact counting is 14 ~~–20 %) this work still~~ –20 % and this work provides new and relevant information about NO_2 in the boundary layer of Mexico City ~~.~~

1 Introduction

Air pollution is a serious environmental problem due to its negative impacts on human health and ecosystems. Fast growing urban and industrial centers are continuously affected by bad air quality and in order to assess their current efforts to mitigate emissions and plan for more efficient strategies to lower the concentration levels of harmful contaminants, it is indispensable to have the proper tools to measure them not only at ground level but also throughout the boundary layer ~~. The~~ (García-Franco et al., 2018). The Multi-AXis Differential Optical Absorption Spectroscopy (MAX-DOAS) technique (e.g. Hönninger et al., 2004; Platt and Stutz, 2008) has rapidly developed in recent years and has proven extremely valuable in

tropospheric chemistry and air pollution studies, since it provides vertical distribution of trace gases with high temporal resolution.

This remote sensing technique is based on the spectroscopically resolved measurement of scattered sunlight at different elevation angles, allowing for the retrieval of total column amounts of aerosols and trace gases with profiling capability. Powerful applications of this technique have been demonstrated to provide useful information about the vertical distribution of aerosols (Frieß et al., 2006; Wang et al., 2016) and photochemically relevant species such as nitrogen dioxide (NO_2), formaldehyde (HCHO), glyoxal (CHOCHO) and nitrous acid (HONO) among other gases (e.g. Wittrock et al., 2004; Wagner et al., 2011; Ortega et al., 2015; Hendrick et al., 2014).

Photochemical reactions involving NO_2 play an important role in the formation of O_3 (Finlayson-Pitts and Pitts, 2000). The Mexico City Metropolitan Area (MCMA) has been particularly affected since the 1990's by high O_3 episodes threatening the population and forcing the authorities to impose strict restrictions ~~in~~ on the usage of motor vehicles (Molina and Molina, 2002). Measurements have been performed in the region using fixed and mobile DOAS zenith-scattered sunlight (Melamed et al., 2009; Johansson et al., 2009; Rivera et al., 2013)~~and in~~ . In 2014 a MAX-DOAS network, initially consisting of four instruments, was established in the Mexico City Metropolitan Area (MCMA) ~~and~~ which has been operating since (Arellano et al., 2016).

In this contribution, we describe the MMF (Mexican Maxdoas Fit) code that has been implemented to retrieve vertical distribution of aerosols and trace gases ~~giving~~ with emphasis on the errors and diagnostics of the results. An overview of the MAX-DOAS instruments is provided in ~~Section~~ Sect. 2 and the complete retrieval strategy from the measured spectra to vertical trace gas profiles is summarized in Fig. 1.

Radiative transfer simulations (constituting the forward model, yellow boxes in Fig. 1) are performed with VLIDORT (Spurr, 2013) to derive simulated differential slant column densities (dSCDs) at the middle of the corresponding wavelength interval used to derive dSCDs from measured spectra. The dSCD retrieval (blue boxes in Fig. 1) is performed with QDOAS (Danckaert et al., 2013) and is described briefly in Sect. 3. The orange parts in the figure refer to the aerosol retrieval while the green parts belong to the trace gas retrieval. Details on the forward model choice, the forward model input calculation (light blue box in Fig. 1) and processing of output quantities in the inversion algorithm are described in Sect. 4.

An error analysis has been included in Sect. 5, especially investigating the effect of the aerosol retrieval on the NO_2 results. Some examples of the NO_2 variability are provided from one of the network's stations and compared to surface concentrations in Sect. 6 and a summary of the work and an outlook on planned improvements to the retrieval code is presented in Sect. 7.

2 Instruments

An instrument based on the MAX-DOAS technique was designed and developed by the Center for Atmospheric Sciences at ~~UNAM~~the Universidad Nacional Autónoma de México (UNAM). It consists of two main parts: the scanner unit which collects the scattered light, and the acquisition/control unit, containing a spectrometer and computer that records and stores the measurements. The two components are connected by an optical fiber and a data connection cable. Both are described briefly in

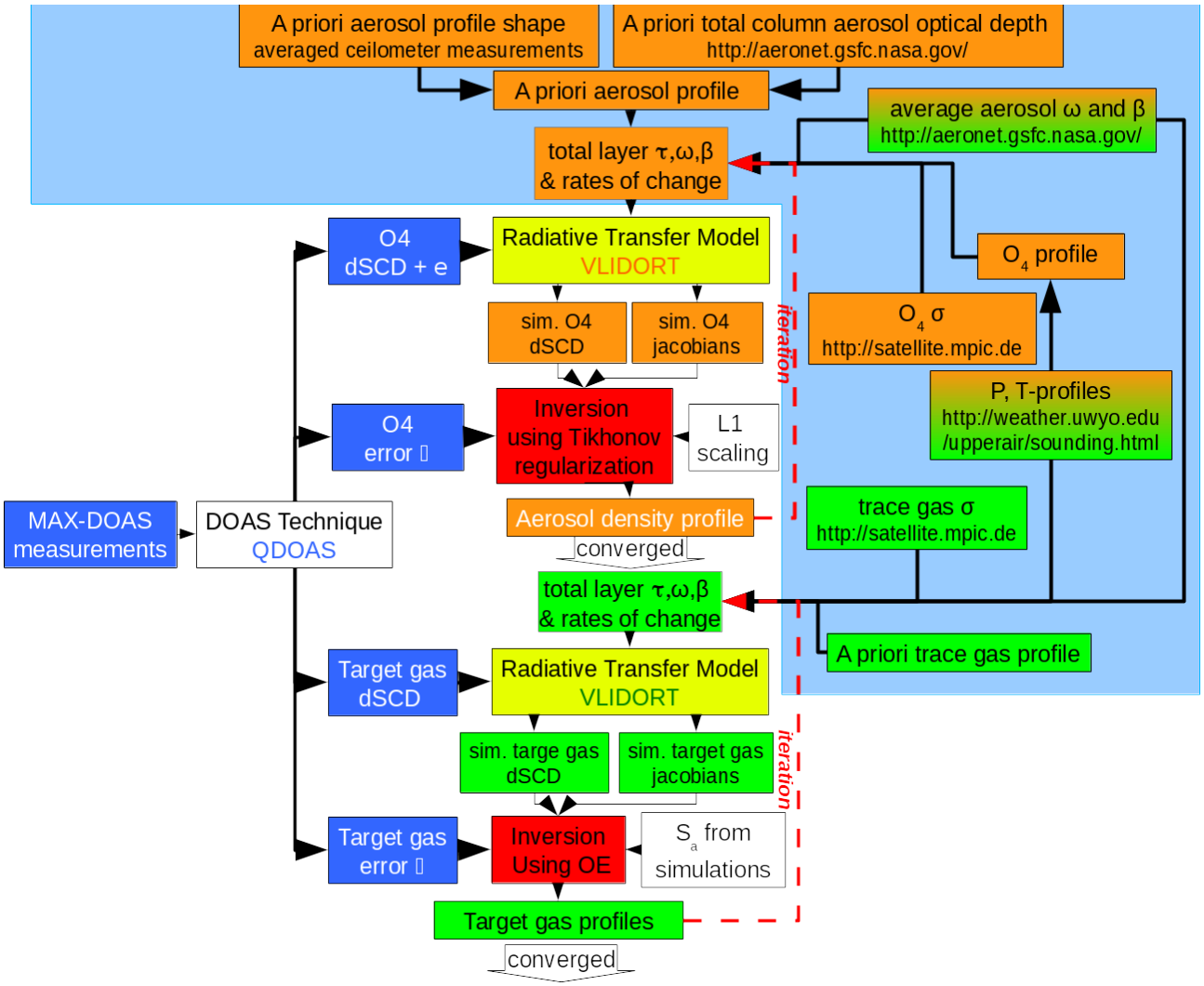

Figure 1. Flowchart of the complete trace gas retrieval algorithm. Orange boxes belong to the aerosol retrieval, green boxes to trace gas retrieval. The light blue box encompasses forward-model-input calculation. The dark blue boxes are in/ outputs of QDOAS. The yellow boxes represent the forward modelling steps. The red boxes are the inversion steps, using Tikhonov regularization for aerosol retrieval and optical estimation (OE) for tracegas retrieval.

the following sections. For a more complete description and a map of the area with the station locations and their measurement orientation, we refer the reader to Arellano et al. (2016).

2.1 The scanner

The scanner is composed of a plastic enclosure (NEMA-rated type 3) resistant to sunlight and hermetically sealed to protect the inner parts from water and bugs. This is important in order to assure ~~large~~ long measurement periods under harsh conditions (Hönninger et al., 2004; Galle et al., 2010; Arellano et al., 2016). The scattered light is collected by a ~~a~~ plano-convex quartz lens (Edmund Optics, $\oslash$ = 25.4 mm, f = 100 mm), focusing it into the entrance of an optical fiber (Fiber Tech Optica, quartz, 6 m long, $\oslash$ = 0.6 mm) which transfers the light into the acquisition/control unit. These optical parts are mounted inside a telescope housing constructed of Nylamid material. A shutter system is installed also within the telescope. It consists of a stepper motor (Mercury, 7.5° by step) that rotates a metal circular plate to prevent the passage of light into the optical fiber and an optical switch which is used to indicate the position of the plate. The shutter is used to make measurements of dark spectra in between scans.

A stepper motor (Oriental Motors, PK266-02A) inside the enclosure allows the movement of the scanner unit in a range of 180° with steps of 0.1°. A mechanical switch is used as a reference to indicate the starting position of each scan. The motor is controlled by an electronic board composed of an 8-bit microcontroller (AVR- architecture), a RS-232 port for communication with the acquisition/control unit, a temperature sensor (Maxim 18B20, accuracy $\pm 0.5^\circ$C), and a dual axis accelerometer (Analog Devices, accuracy $\pm 0.1^\circ$) that provides an accurate determination of the telescope's pointing elevation. The theoretical field of view (FOV) of the optical system is 0.31°.

2.2 The acquisition/control unit

The second part of the instrument consists of a metallic housing ~~receiveing~~ receiving the collected light through the optical fiber and sending it to the spectrometer (Ocean Optics, USB2000+). This commercial device has a crossed asymmetric Czerny-Turner configuration, diffraction grating (1800 lines/mm) and a slit size of 50 μm wide x 1 mm high, recording spectra in a wavelength range of 289 - 510 nm at a resolution of 0.69 nm (full width at half maximum). It uses a Charge-Coupled Device (CCD) detector array (Sony ILX511B) of 2048 pixels with an integration time adjustable between 1 ms to 65 s.

Because changes in temperature can affect the wavelength/pixel ratio and also modify the optical properties of the spectrometer like the alignment and the line shape (e.g. Carlson et al., 2010; Coburn et al., 2011), a temperature control system composed of a Peltier cell (Multicomp) and three temperature sensors (Maxim 18B20, accuracy $\pm 0.5^\circ$C) controlled by an electronic board were implemented. The cooling side of the Peltier cell was placed on top of the spectrometer and the heating side was attached to a heat sink. The Peltier cell and the spectrometer were wrapped in a styrofoam box to keep the temperature insulated from the outside. The three temperature sensors were placed on the heat sink, the Peltier cell and spectrometer to monitor temperature changes. The temperature control system was wrapped in an aluminum enclosure to prevent the heat spreading to other parts of the system and a fan was installed to extract the heat from the enclosure.

The electronic board for the temperature control is composed of an 8-bit microcontroller (AVR, architecture) and a RS-232 communication port. This board is responsible for obtaining the data from the sensors and adjusting the voltage in the Peltier cell to keep the temperature constant.

The acquisition/ control unit has a laptop computer (Dell, Latitude 2021) with Linux operating system contained within the enclosure. The program controlling the hardware is written in C++ using Qt libraries. A script is used to carry out the measurement sequence previously defined and to monitor the spectrometer temperature.

2.3 Measurement strategy

A complete scan consists of a sequence that begins with a measurement towards the zenith, (90 ° elevation angle), followed by ~~four~~ tree measurements between elevation angles 0 ° and 10 ° towards the west (for UNAM station, this is 85 ° azimuth angle), then ~~16~~ 12 measurements are taken with elevation angles between 10 ° elevation angle towards west and 10 ° elevation angle towards the east (crossing the zenith, but without taking a measurement), followed by ~~four~~ three measurements between 10 ° and 0 ° elevation angle towards the east. The same sequence is then repeated but in reverse order. At the end of this cycle, a dark spectrum with the closed shutter is taken. With this setup, a complete scan takes about 7 minutes. The measurement sequence (90°Zenith, 0°W, 2°W, 6°W, 13°W, 23°W, 36°W, 50°W, 65°W, 82°W, 82°E, 65°E, 50°E, 36°E, 23°E, 13°E, 6°E, 2°E, 0°E) is likely to change in the future to use longer integration times but at the same time reduce the number of viewing directions in the range between 10 ° elevation angle towards the west and 10 ° elevation angle towards the east in order to keep the total scan time roughly constant. All data presented in this manuscript uses this setup meaning that we include both westerly and easterly viewing directions in the same retrieval and hence the result represents ~~and~~ an average. This strategy leads to larger fitting errors since this likely deviates further from the assumption of horizontal ~~inhomogeneity is likely to be less true~~homogeneity. With our retrieval chain, it is possible to consider the easterly and westerly directions separately to investigate the differences in viewing directions which is subject of current investigation.

The output from the MAX-DOAS instruments consists of five files per complete scan: A file containing the meta data of each spectrum (e.g. accelerometer data for each measurement, time of the acquisition, temperatures within the acquisition unit), all the spectra in non-zenith directions, the dark spectrum measured with the shutter closed, a meta file for the dark spectrum and the first zenith reference measurement. These files are stored for further processing.

3 Differential Slant Column Densities (dSCDs) Retrieval

The spectra were evaluated using the QDOAS (version 2.105) software (Danckaert et al., 2013). As a pre-processing step before the QDOAS analysis, the dark signal was subtracted from each of the measurement spectra. A wavelength calibration was conducted in QDOAS by applying a non-linear least squares fit to a solar atlas (Kurucz et al., 1984).

For NO_2, the retrieval was conducted in the 405 to 465 nm wavelength range using the spectrum measured at the zenith position at the beginning of each of the measurement sequences as reference. For the analysis, differential cross-sections of NO_2 at 298 K (Vandaele et al., 1998), O_3 at 221 K and 241 K (Burrows et al., 1999) and the oxygen dimer (Hermans et al.,

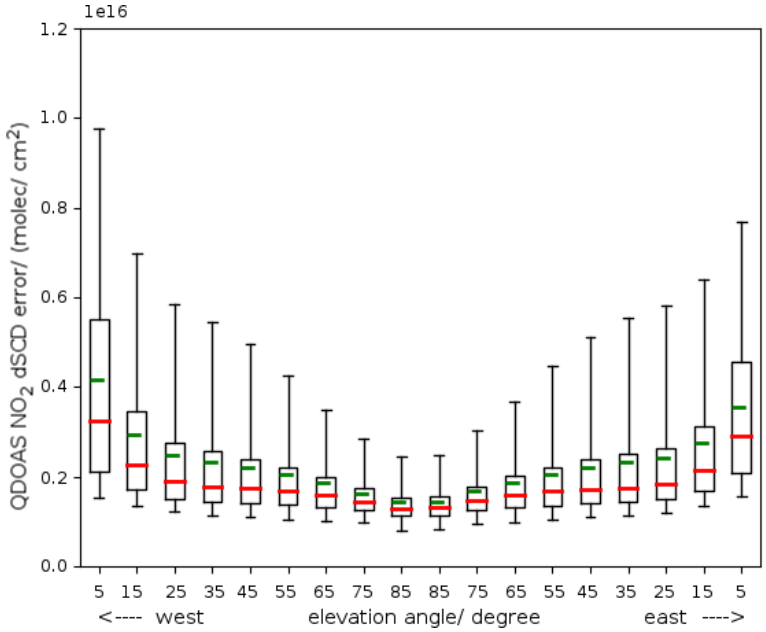

Figure 2. dSCDs measurement error statistics for NO_2 at the UNAM station as a function of elevation angle for data taken in the year 2016. The box encloses the 25-75 percentile, the whiskers are 5-95 percent, the green bar is the mean and the red bar the median.

1999) were convolved with the slit function of the spectrometer and a wavelength calibration file (created using a mercury lamp) and using the convolution tool of the QDOAS software. A Ring spectrum, generated at 273 K from a high resolution Kurucz file using the Ring tool of the QDOAS software (Danckaert et al., 2013), was also included in the analysis.

For O_4, the retrieval was conducted in the 336 to 390 nm wavelength range. Differential cross-sections of O_4 (Hermans et al., 1999), O_3 at 221 K and 241 K (Burrows et al., 1999), NO_2 at 294 K (Vandaele et al., 1998), BrO at 298 K (Wilmouth et al., 1999), HCHO at 298 K (Meller and Moortgat, 2000) and a Ring spectrum were included in the analysis. A 3rd degree polynomial was used for the retrievals.

~~Figure~~ Fig. 2 shows the dSCDs retrieval error statistics for NO_2 as a function of elevation angle. The plot shows results of 31,448 scans from January to December 2016. Larger dSCDs fitting errors are found for viewing angles closer to the horizon (smaller elevation angles), likely due to physical interferences during the measurements. As elevation angles approach the zenith, retrieval dSCDs errors decrease considerably.

4 Mexican Maxdoas Fit

The method for the trace gas retrieval from slant column densities using the MMF code is comprised of two parts, an aerosol retrieval using the known O_4 profile, and the trace gas retrieval (e.g. Platt and Stutz, 2008). Both parts consist of the same steps: A foward model and an inversion algorithm.

In Sect.4.1, details about the inversion strategy are given. Our choice of forward model, VLIDORT (e.g. Spurr et al., 2001; Spurr, 2006, 2013), and the input parameter calculation are detailed in Section 4.2. How forward model output quantities are processed for the inversion step is detailed in Sect. 4.3.

~~MMF has been participating in the Round-Robin comparison of different retrieval codes for the FRM4DOAS project (Frießet al., in preparation). It has also participated in the profile retrieval from dSCD from the CINDI2 campaign, both for NO_2 and HCHO (Tirpitz et al., in preparation) as well as for HONO (Wang et al., in preparation).~~ 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.

4.1 Inversion theory

The inversion strategy relies on the fact that the problem is ~~not too non-linear~~ 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.

$$x_{i+1} = x_a + (K_i^T S_m^{-1} K_i + R)^{-1} K_i^T S_m [(y - F(x_i)) - K_i(x_a - x_i)] \quad (1)$$

~~In a recent update of the code, implemented after the analysis presented here (i.e. not used for obtaining the results here) this GN iteration scheme was replaced by a slightly slower[2] but more stable Levenberg Marquardt (LM) iteration scheme in order to enable working in the logarithmic retrieval space which makes the problem more unlinear: Eq. 1 was replaced by Eq.A1 for more non-linear inversion problems (Rodgers, 2000)~~

$$x_{i+1} = x_i + \left[(1+\gamma)R + K_i^T S_m^{-1} K_i\right]^{-1} \left[K_i^T S_m^{-1} (y - F(x_i)) - R(x_i - x_a)\right]$$

~~In both equations~~Here, superscript T denotes transposed, superscript ~~-1~~ -1 denotes the inverse. The index i is the iteration index, the subscript ~~a indicates a-priori values. S_m~~ a indicates *a priori* values. Sm is the measurement error covariance matrix, y denotes the vector of measured differential slant column densities. $F(x_i)$ are the simulated differential slant column densities, calculated using the forward model with input profile x_i . Both y and $F(x_i)$ are vectors of dimension ~~(#~~ which is the number of telescope viewing angles~~). $K_i = \partial F(x)^l / \partial x^n$ is the jacobian~~ . $K_i = \partial F(x)^l / \partial x_n$ is the Jacobian matrix at the i-th iteration describing the change of simulated dSCD for viewing angle l when the profile x in layer n is varied.

[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

[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 is around 5 seconds.~~

In ~~case of the use of Eq.A1, the new x_{i+1} is only accepted if the cost function in Eq. A2 decreases w.r.t the previous cost-function~~

$$\text{cost} = \sum_{\text{k}}^{\text{angles}} \sum_{\text{j}}^{\text{angles}} (\text{y} - \text{F(x)})_{\text{k}} S_m^{-1\,kj} (y - F(x))_j + \sum_{k}^{layers} \sum_{j}^{layers} (x - x_a)_k Sa^{-1\,kj} (x - x_a)_j$$

~~If this is the case and $(1+\gamma)$ is not yet equals to 1, the factor $(1+\gamma)$ is halved for the next iteration. If however, the cost-function increases, the newly calculated x_{i+1} is discarded and the i-th calculation repeated with a factor $(1+\gamma)$ increased by a factor of 16.~~

~~In this study, the GN iteration scheme was used and the retrieval grid equals the simulation grid. Details about the layer distribution are given in Sect. 4.2.~~

~~In~~ the case of optimal estimation (OE), the regularization matrix R is equal to the inverse of the ~~a-priori~~ *a priori* covariance matrix, $R = S_a^{-1}$. OE regularization is used for trace gas retrieval. ~~The LM scheme of Eq. A1 has currently only been tested with this choice of regularization matrix.~~ Other regularization matrices are possible, see e.g. Steck (2002).

For the aerosol retrieval used in this study, we use the $L1$ operator (~~$R = L1^T \alpha L1$~~ $R = L1^T \alpha L1$) 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 ~~can~~ could be supplied, as well as a complete regularization matrix ~~R.~~

~~Tests using the logarithm of the partial layer vertical column density (for NO_2 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. The studies performed during the FRM4DOAS Round-Robin analysis of synthetic data (Frießet al., in preparation) and the CINDI2 retrieval exercises for HCHO and NO_2 (Tirpitz et al., in preparation) and IO (Wang et al., in preparation) used the retrieval in logarithmic space with LM iteration scheme~~R.

4.2 Forward model

Several radiative transfer codes have been developed to serve as forward models for this kind of retrieval algorithms. ~~There~~ For example, there are Monte Carlo (MC) codes, such as AMFTRAN (Marquard, 1998), TRACY (von Friedeburg et al., 2002; von Friedeburg, 2003) or PROMSAR (Palazzi et al., 2005). All these MC radiative transfer codes start with ejecting photons from the instrument and following the photon path backwards (see also e.g. Perliski and Solomon, 1993; Marquard et al., 2000), hence they are sometimes somewhat confusingly referred to as backward models. The advantage of MC codes is their high accuracy, the disadvantage is the rather long calculation time. SCIATRAN (e.g. Buchwitz et al., 1998; Rozanov et al., 2014, 2017) combines multi scattering radiative transfer modelling using the Picard-Iterative approximation, a chemistry and an inversion code. It has been heavily used in NO_2 (e.g. Vidot et al., 2010) and ozone (e.g. Rahpoe et al., 2013b, a) retrieval from satelite observations.

Another class of radiative transfer codes, such as VLIDORT and DISORT (Stamnes et al., 1988; Dahlback and Stamnes, 1991), uses the discrete ordinate algorithm. This finite difference method is based on finding solutions for an atmosphere

consisting of a number of homogeneous layers using gaussian quadrature approximations to the integro-differential radiative transfer equation expressed using Legendre polynomials for the phase function and Fourier series for the intensity.

The big advantage of the second class of codes is their ~~heigh~~ high speed. Another advantage of VLIDORT is that it calculates not only the intensity but also analytic Jacobians. Therefore, instead of ~~(~~the $2 \times$ ~~# layers)~~ the number of layers calls to the program to perform a finite difference approximation, one call is sufficient. MMF uses the intensity part of VLIDORT, version 2.7, released in August 2014. (V)LIDORT is configured to return intensity ~~jacobians~~ Jacobians w.r.t. gas absorption layer properties or aerosol total extinction layer properties.

Layers above the retrieval grid can be added for the forward model simulation. Both grids are supplied via an input file to the code together with the corresponding ~~a priori~~ *a priori* values of the trace gas concentration and the aerosol profile. In this study, the simulation grid is identical to the retrieval grid and the layer distribution of the retrieval grid is not equidistant. The grid consists of 22 layers up to 25 km. The layer thickness increases from 100 m at the lowest layer to 5 km in the upper-most layer. The exact height distribution can be seen in Fig. 5. The surface albedo used in this study was set to 0.07, this value can be passed to MMF in the configuration file. However, in practice the effect for downwelling intensity and hence for the dSCD calculation was found to be negligible.

For each ~~atmospheric input~~ simulated atmospheric layer, the following input needs to be supplied to VLIDORT[2]:

1. total layer optical depth τ
2. single scattering albedo ω
3. phase function expansion coefficients β

Since Jacobians with respect to changes of quantity χ in each layer are required, the normalized derivatives of the total primary optical quantities, τ, ω and β with respect to this quantity need to be supplied as well. In case of trace gas retrieval, $\chi = a_{\text{gas}}\Delta h$, and in case of aerosol retrieval, $\chi = (\sigma_{\text{aer}} + a_{\text{aer}})\Delta h$. Here, a_{gas} and a_{aer} are the gaseous and aerosol absorption coefficients, respectively and σ_{aer} is the aerosol scattering coefficient. Δh is the layer thickness. Hence, for each layer the following input is additionally required:

4. rate of change of τ w.r.t. χ
5. rate of change of ω w.r.t. χ
6. rate of change of β w.r.t. χ

In order to calculate this input, we first need to calculate the separate properties for the trace gases and the aerosols. Afterwards, these quantities are combined to yield the total layer quantities. ~~The part of the input calculation is inclosed in Fig. 1 in the light blue box.~~

The calculation of the contribution from the trace gas and the air density (through Rayleigh scattering) for the layer inputs (1-6) is presented in Sect. 4.2.1. The aerosol contribution calculation is outlined in Sect. 4.2.2. These two sections also contain

[2]We only use the LIDORT part of VLIDORT, meaning that we only consider the total intensity and not the polarization

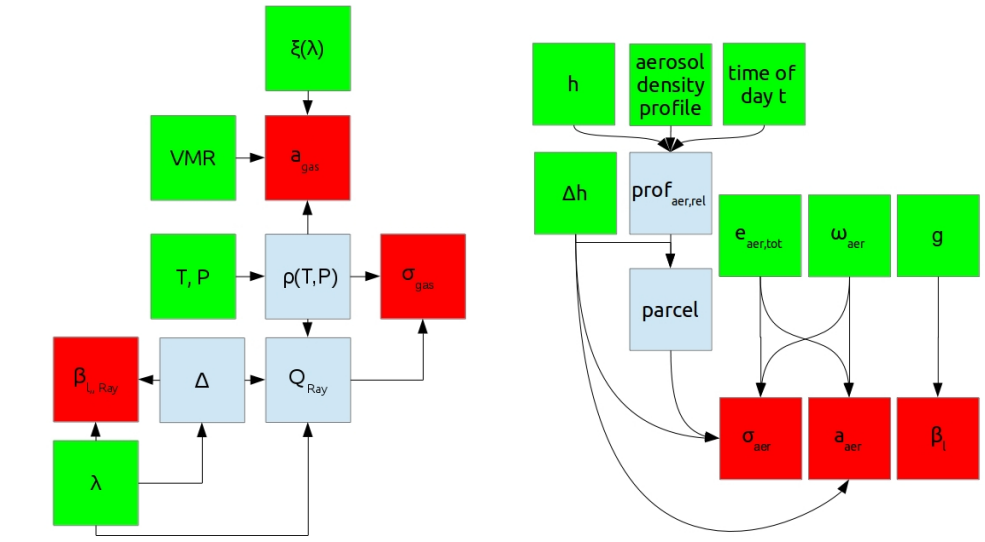

Figure 3. Calculating the gas and air (a) and aerosol (b) contribution to the layer input parameters for VLIDORT. The green boxes are primary input to MMF, blue boxes are intermediate quantities and red boxes final output that are combined for VLIDORT layer input. VMR denotes here the ~~a-priori~~ *a priori* volume mixing ratio. Instead of a VMR, an ~~a-priori~~ *a priori* trace gas density can be supplied, too.

information on the source of information for the respective ~~a-priori~~ *a priori* values, as well as, for gas, the error covariance matrix.

The calculation of the contributions from gas and aerosol to yield the VLIDORT inputs (1-6) is detailed in Sect. 4.2.3.

4.2.1 Input for gases

In Fig. 3(a), the steps to calculate the contribution of the trace gas and the contribution from Rayleigh scattering in each layer ~~is~~ are outlined. Red boxes represent the trace gas (and air) contribution to VLIDORT input quantities, green boxes are primary inputs to the inversion code, light blue boxes are intermediate quantities that are neither direct input to MMF, nor final input for the forward model.

The ~~a-priori volume mixing ration~~ *a priori* volume mixing ratio (VMR) profile and the covariance matrix for this study are calculated from simulations covering the year 2011 using WRF-Chem V3.6. The model domain covers Mexico and surrounding seas, the domain has 200 by 100 square grid cells with approximately 28 km width and 35 vertical layers. The parameterizations used were for Micro Physics WRF single moment 3-class (4), for PBL the Yonsei University scheme (1), for Cumulus the Grell 3D ensemble scheme (5), for radiation the Goddard shortwave scheme (1) and the Dudhia long wave scheme (4), time-varying sea surface temperature on, grid analysis nudging on, the unified Noah Land Use surface model (2), off an urban canopy model and the 2008 emissions inventory were used.

MMF takes pressure and temperature profiles on a separate height grid, internal interpolation to the provided retrieval grid is performed. In the processing chain implemented in Mexico, the temperature and pressure profiles from radiosonde data for the specific day are downloaded from the University of Wyoming (http://weather.uwyo.edu/upperair/sounding.html) from the Mexico City Airport, station number 76679. These temperature and pressure profiles are used to calculate the dry air number density ρ, see Appendix B.

The absorption cross section ξ_λ is taken at a wavelength in the middle of the wavelength interval used for the QDOAS retrieval (see Sect. 3.), for aerosol retrieval it is 361 nm and for NO_2 it is 414 nm. The same cross-section tables as for QDOAS dSCD retrievals (see Sect. 3) are used for NO_2 and O_4. However, since the cross section cancels out at the end of the calculation, its exact value is not too important.

There are analytical fits to the temperature dependence, using e.g. linear (e.g. Vandaele et al., 2003, for NO_2) or more complicated (Kirmse et al., 1997, e.g.) temperature dependences. For a comparison study of different fitting functions, see Orphal (2002). These could be implemented, however, the cross-section is currently not temperature dependent, in agreement with the assumption made for defining dSCDs.

The trace gas absorption coefficient is calculated according to Eq. 2~~.~~:

$$a_{gas} = \mathrm{VMR} \cdot \xi_\lambda(T) \cdot \rho \tag{2}$$

For the calculation of the depolarization ratio Δ, the main contributions are from N_2, CO_2, O_2 and Ar. Our implementation follows Bates (1984). Details are given in Appendix B. For the calculation of the Rayleigh cross-section Q_{Ray}, we follow the implementation of Goody and Yung (1989) (their equation 7.37), see also Platt et al. (2007) (their equation 19.5). The air scattering coefficient can then be calculated according to Eq. 3:

$$\sigma_{air} = \mathrm{Q}_{\mathrm{R}ay} \cdot \rho \tag{3}$$

and the Rayleigh scattering expansion coefficients β_l (e.g. Spurr et al., 2001), according to Eq. 4:

$$\beta_{\mathrm{air}0} = 1,\ \beta_{\mathrm{air},1} = 0,\ \beta_{\mathrm{air},2} = (1-\Delta)/(2+\Delta). \tag{4}$$

4.2.2 Input for aerosols

The (~~a-priori~~*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 ~~taken~~ 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.

The first part of MMF, the aerosol retrieval, is limited to the aerosol density profile and the total aerosol optical depth. The average single scattering albedo ω and asymmetry parameter g are not subject to retrieval and are assumed to be constant in all layers.

Fig. 3 (b) outlines the strategy how the aerosol contribution to the VLIDORT layer input parameters ~~are~~ is calculated.

The hourly averages of (relative) aerosol density ~~profile (arbitrary units)~~ profiles in arbitrary units from ceilometer measurements between November 2009 to February 2013 are interpolated at measurement day time t and at middle heights h ~~and~~ to provide a relative aerosol profile. This profile, turned into a partial optical depth per layer by multiplying with the layer thickness, is scaled to match the total aerosol extinction from AERONET τ_{aer}. ~~In each layer, the profile is converted~~ The profile is then converted back into an intensive[3] quantity by division by layer thickness, usually known as aerosol extinction profile (extinction per unit length, AE). This is then used to calculate the aerosol scattering coefficient σ_{aer} in each layer by multiplying the layer aerosol ~~extincion~~ extinction with the aerosol single scattering albedo ω_{aer}.

The aerosol absorption coefficient a_{aer} is the layer aerosol extinction times $(1-\omega_{\mathrm{aer}})$ and the aerosol phase function coefficients $\beta_{\mathrm{aer,l}}$ can be calculated via the Henyey Greenstein phase function and the asymmetry parameter g, see Eq. 5, e.g. Hess et al. (1998), where l denotes the moment,

$$\beta_{\mathrm{aer},l} = (2l+1)\cdot g^l. \tag{5}$$

4.2.3 Calculating final VLIDORT input

As mentioned in the introduction to this section, the layer input parameters for VLIDORT are total optical depth τ and single scattering albedo ω in the layer, as well as the phase function coefficients β_l. The quantities which have been calculated so far are the aerosol and air contributions to β_l and the air and aerosol scattering coefficients σ_{air} and σ_{aer}, as well as the absorption coefficients a_{gas} from the trace gas in question and a_{aer} from the aerosol in each VLIDORT input layer (red boxes in Fig. 3). These quantities need to be combined to total layer input parameters.

The total layer optical depth is simply the product of the layer thickness and the sum of all extinction and scattering coefficients:

$$\tau = \Delta h \cdot (\sigma_{\mathrm{air}} + \sigma_{\mathrm{aer}} + a_{\mathrm{gas}} + a_{\mathrm{aer}}) \tag{6}$$

The combined single scattering albedo can be calculated as

$$\omega = (\sigma_{\mathrm{air}} + \sigma_{\mathrm{aer}}) \cdot \Delta h/\tau \tag{7}$$

The combined expansion coefficients are calculated as follows

$$\beta_l = (\beta_{\mathrm{air,l}} \cdot \sigma_{\mathrm{air}} + \beta_{\mathrm{aer,l}} \cdot \sigma_{\mathrm{aer}})/(\sigma_{\mathrm{air}} + \sigma_{\mathrm{aer}}) \tag{8}$$

Since Jacobians with respect to changes of quantity χ in each layer are required, the normalized derivatives of the total primary optical quantities, τ, ω and β with respect to this quantity need to be supplied. In case of trace gas retrieval, $\chi = a_{\mathrm{gas}}\Delta h$ and in case of aerosol retrieval, $\chi = \sigma_{\mathrm{aer}}\Delta h + a_{\mathrm{aer}}\Delta h$. Therefore, what needs to be done is to calculate the normalized derivatives of Eq. 6 – Eq. 8 with respect to these quantities. It should be remembered that β is a vector.

[3] bulk property that does not change when changing the size of the system

The three quantities for aerosol are:

$$\frac{\chi}{\tau}\frac{d\tau}{d\chi} = \frac{(\sigma_{\text{aer}}+a_{\text{aer}})\Delta h}{\tau}\frac{d\tau}{d(\sigma_{\text{aer}}+a_{\text{aer}})\Delta h} = \frac{\sigma_{\text{aer}}+a_{\text{aer}}}{e} \tag{9}$$

$$\begin{aligned}\frac{\chi}{\omega}\frac{d\omega}{d\chi} &= \frac{(\sigma_{\text{aer}}+a_{\text{aer}})\Delta h}{\omega}\cdot\left[\frac{\partial\omega}{\partial\sigma_{\text{aer}}}\frac{d\sigma_{\text{aer}}}{d[(\sigma_{\text{aer}}+a_{\text{aer}})\Delta h]}+\frac{\partial\omega}{\partial a_{\text{aer}}}\frac{da_{\text{aer}}}{d[(\sigma_{\text{aer}}+a_{\text{aer}})\Delta h]}\right] \\ &= \frac{\sigma_{\text{aer}}a_{\text{gas}}-a_{\text{aer}}\sigma_{\text{air}}}{\omega e^2}\end{aligned} \tag{10}$$

$$\frac{\chi}{\beta}\frac{d\beta}{d\chi} = \frac{\sigma_{\text{aer}}}{\beta}\frac{\beta_{\text{aer}}-\beta}{\sigma_{\text{aer}}+\sigma_{\text{air}}} \tag{11}$$

Here, e is the total extinction ~~coeficient~~coefficient: $e = \tau/\Delta h$

For the trace gas ~~jacobian~~ Jacobian calculation, the corresponding ~~quantities~~ normalized derivatives are:

$$\frac{\chi}{\tau}\frac{d\tau}{d\chi} = \frac{a_{\text{gas}}\Delta h}{\tau}\frac{d\tau}{da_{\text{gas}}} = \frac{a_{\text{gas}}}{e} \tag{12}$$

$$\frac{\chi}{\omega}\frac{d\omega}{d\chi} = \frac{a_{\text{gas}}\Delta h}{\omega}\frac{\omega}{a_{\text{gas}}} = -\frac{a_{\text{gas}}}{e} \tag{13}$$

$$\frac{\chi}{\beta}\frac{d\beta}{d\chi} = 0 \tag{14}$$

4.3 Calculating dSCDs and weighting functions

The forward model outputs, intensities and intensity ~~jacobians~~Jacobians, need to be converted into differential slant column densities and corresponding ~~jacobians~~Jacobians. For each set of dSCD, we have to run the forward model 2 times: Once with the gas absorption included and once without the gas absorption. Each of these sets consists of simulations in the desired telescope directions and one simulation towards the zenith. In the following, the intensity and ~~jacobian~~ Jacobian of the simulation towards the desired angles without gas absorption are denoted I_0^α and K_0^α, the intensity (~~jacobian~~Jacobian) with gas absorption towards the zenith I_g^z (K_g^z), the intensity (~~jacobian~~Jacobian) with gas absorption towards the desired angle I_g^α (K_g^α) and the intensity (~~jacobian~~Jacobian) without gas absorption towards the zenith I_0^z (K_0^z). The dSCD and the corresponding weighting function K can be calculated as

$$\text{dSCD} = \log\ln\left(\frac{\mathrm{I}_0^\alpha\cdot\mathrm{I}_\mathrm{g}^\mathrm{z}}{\mathrm{I}_\mathrm{g}^\alpha\cdot\mathrm{I}_0^\mathrm{z}}\right)/(\xi_\lambda \tag{15}$$

and

$$K = \frac{\left(K_0^\alpha I_g^z I_g^\alpha I_0^z + I_0^\alpha K_g^z I_g^\alpha I_0^z - I_0^\alpha I_g^z K_g^\alpha I_0^z - I_0^\alpha I_g^z I_g^\alpha K_0^z\right)}{\left(I_0^\alpha I_g^z I_g^\alpha I_0^z \xi_\lambda\right)}. \tag{16}$$

If the retrieval is to be performed in log-space, i.e. to work with $ln(x)$ instead of with x as the retrieval parameter, K in Eq. 16 is multiplied by x and the uncertainty covariance matrix S_a in Eq. A1 needs to be left and right multiplied by $1/x$. Note that the output from VLIDORT is already Kx, hence the immediate output from VLIDORT has to be divided by x in case of linear retrieval and no extra processing step has to be performed for logarithmic retrieval.

5 Error analysis

The errors of gas profile retrievals in remote sensing applications can be divided into the following four different types (Rodgers, 2000):

1. Smoothing error, arising from the constraint of the NO_2 and aerosol profiles,
2. Measurement noise error, arising from the noise in the spectra,
3. Parameter error, including different parameters as aerosol profiles and cross-sections, and the
4. Forward model error, that originates from the simplification and uncertainties in the radiative transfer algorithm:

The ~~Gain~~ Gain matrix, as defined by Eq. 17 describes the change in retrieved x when measurement y changes and can hence be used to map the uncertainty in measurement space, to a state-vector uncertainty.

$$\text{~~Gain~~Gain} := \frac{\partial \mathrm{x}}{\partial \mathrm{y}} = (\mathrm{K_i^T S_m^{-1} K_i + R})^{-1} \mathrm{K_i^T S_m} \tag{17}$$

The error analysis in general ~~produce~~ produces an error pattern ϵ_y in the measurement space of vectors of the dSCDs retrieved by QDOAS and simulated by MMF. With help of the ~~Gain~~ Gain matrix this error pattern is then mapped as ϵ_x into the space of the solution state.

5.1 Smoothing error

The retrieved atmospheric state vector x only represents a smoothed version of the true real state. How a change in the true atmospheric state vector x_{true} is translated into changes in the retrieved state vector x is described by the partial column Averaging Kernel matrix $\mathrm{AK_{pcol}}$.

$$x - x_a = \mathrm{AK_{pcol}(x_{true} - x_a)} + \epsilon_\mathrm{x} \tag{18}$$

The Averaging Kernel of partial column and total column ($\mathrm{AK_{tot}}$) describe how the retrieved solution profile depends on the real atmosphere x_{true} and have to be taken into account if the profile and the column is used. The part which is not explained by $\mathrm{AK_{pcol}}$ is explained by the retrieval error ϵ_x.

As MMF uses ~~constraint~~ constrained least square fitting, the AK_{pcol} is calculated analytically by MMF itself for each scan according to Eq. ~~(19)~~19.

$$AK_{pcol} := \underbrace{\frac{\partial x}{\partial x_{true}}}_{\text{definition}} = \underbrace{\mathrm{Gain \cdot K}}_{\text{analytical calculation}} \tag{19}$$

The Averaging Kernel matrix is in general ~~not symmetrical~~ asymmetrical and the columns are representing response functions related to a perturbation in a certain layer while the sensitivity of the NO_2 partial column of a certain layer (or concentration in a certain layer) to all the different true partial columns or concentrations are expressed by the rows of the matrix. If the $\mathrm{AK_{pcol}}$ is expressed in units of partial columns (as here indicated by the subscript $_{pcol}$), the sum of the rows is the total column Averaging Kernel and represents the sensitivity of the vertical column density (VCD) to the anomalies in different heights. How the Averaging Kernel changes under transformation of the units is expressed by two matrix multiplications, one with the diagonal matrix containing the partial air column in its diagonal $\mathrm{U_{aircols}}$ from one side and its inverse $\mathrm{U^{-1}_{aircols}}$ from the other side.

$$\mathrm{AK_{VMR}} = \mathrm{U^{-1}_{aircols}} \cdot \mathrm{AK_{pcol}} \cdot \mathrm{U_{aircols}} \tag{20}$$

The trace of the $\mathrm{AK_{tot}}$ matrix represents the DOF, the number of independent pieces of information in the profile retrieval, which is around 2 for the NO_2 retrieval.

Equation (18) separates the smoothing effect described with AK ~~matrices~~ matrices from the others errors. Before comparing the retrieved profiles of MMF with profiles from models or other measurements, such as satellite measurements, it is necessary to smooth the profile from the other source with the AK from MMF (Rodgers and Connor, 2003). Alternatively, a smoothing error S_{smooth} for the profile can be calculated according to Rodgers (2000) from the ~~a-priori~~ *a priori* covariance matrix S_a and the Averaging Kernel. The latter is describing how the true atmospheric state varies, as a best estimate, while $\mathrm{AK} = \frac{\partial \mathrm{x}}{\partial \mathrm{x_{true}}}$ describes how sensitive the retrieved atmospheric state vector depends on the true atmospheric state. Both quantities can be given in different representations adjusted to the atmospheric state vector either as a fraction of the ~~a-priori~~*a priori*, the VMR ($_{VMR}$), as number densities or partial column profiles ($_{pcol}$) as used above.

$$S_{smooth} = (\mathrm{AK_{VMR}} - \mathbf{1})\mathrm{S_a}, \mathrm{VMR}(\mathrm{AK_{VMR}} - \mathbf{1})^{\mathbf{T}} \tag{21}$$

The smoothing error for the VCD is then calculated from the full covariance matrix of the smoothing error S_{smooth} using a total column operator.:

$$\sigma_{smooth} = \sqrt{g^T \cdot \underbrace{(\mathrm{AK_{pcol}} - \mathbf{1})\mathrm{S_{a,pcol}}(\mathrm{AK_{pcol}} - \mathbf{1})^{\mathbf{T}}}_{\mathrm{S^{pcol}_{smooth}}} \cdot g}, \tag{22}$$

where **g** is the total column operator for partial column profiles: $g^T = (\,1, 1, 1, 1, 1, 1, 1, 1, 1, 1,...1)$.

As described before, the NO_2 retrieval code uses a single ~~a-priori~~ *a priori* and covariance matrix taken from the chemical transport model WRF-Chem, which results in the sensitivities and resulting smoothing errors represented in Fig. 4(a).

The aerosol retrieval, however, uses a ~~Tihkonov~~ Tikhonov constraint and no covariance matrix is available. The ~~a-priori~~ *a priori* for the aerosol extinction profile are reconstructed by the actual total aerosol optical depth reported by the daily AERONET measurements and the average vertical distribution, reconstructed from ceilometer measurements for each hour of

the day. The covariance matrix for the aerosols $Sa_{aerosol}$ is obtained by assuming a 100% variability of the used ~~a-priori~~ *a priori* profile and an exponential correlation length of η =500 m between the different layers, as in the recent work by Wang et al. (2017).

$$Sa_{aerosol}[i,j] = (100\% \cdot \mathrm{AE_a[i]} \cdot \mathrm{AE_a[j]})\exp^{(-|z[i]-z[j]|/\eta)} \quad (23)$$

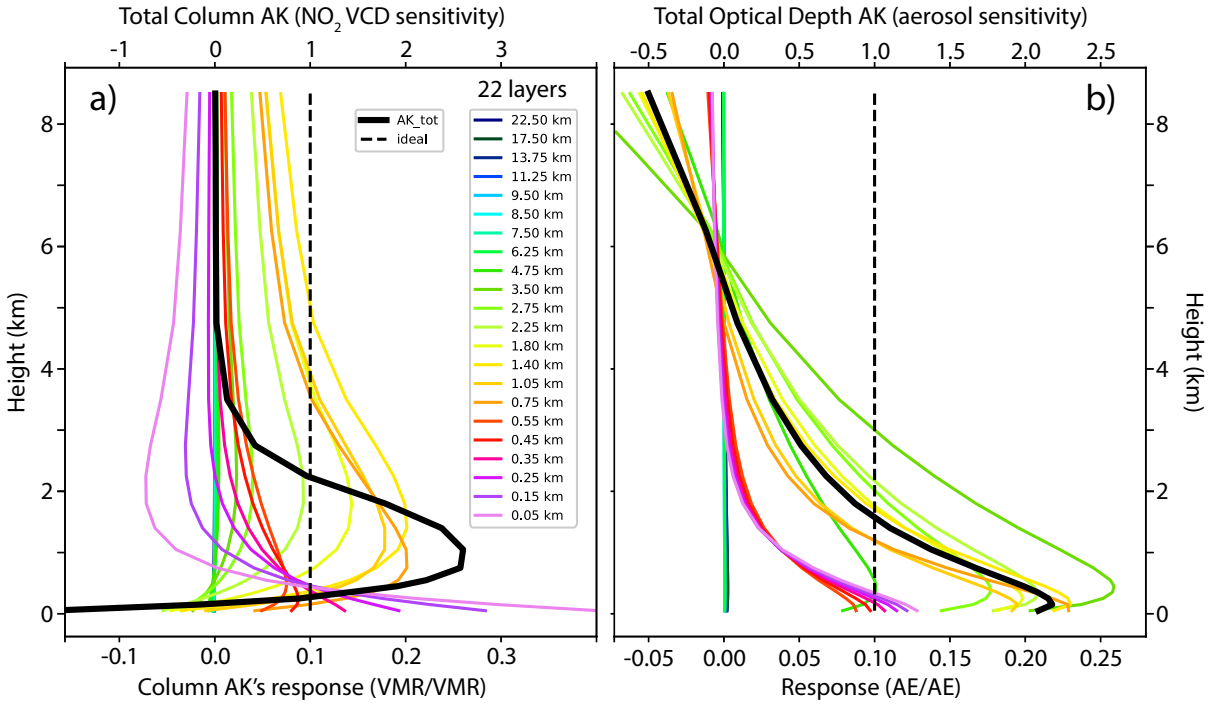

Figure 4. Averaging Kernel columns: (a, left) in VMR representation for NO_2 profile retrieval and (b, right) in the corresponding intensive quantity for the aerosol extinction (AE) profile. Each coloured line shows the expected response on perturbation in a certain layer and belongs to the lower axis. The total column Averaging Kernel: (a, left) for the NO_2-VCD and (b, right) for the total optical depth is given as a black line and belongs to the upper axis. The vertical dashed line in each panel indicate ~~a~~ an ideal sensitivity of 1.0 in all altitudes.

In the cases where no vertical aerosol extinction profile (AE-profile) retrieval is available, a NO_2 profile can still be retrieved by using an ~~a-priori~~ *a priori* AE-profile instead of the retrieved one. Since in that case, no information of the O_4 absorption is used, the estimated smoothing error S_{smooth} is then directly given by the ~~a-priori~~ *a priori* covariance matrix Sa for the AE-profile. 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.

For the calculation we have assumed a constant sensitivity AK and a constant apriori covariance matrix Sa. The AK of the trace gas profile indeed depends strongly on the aerosol profile and even slightly on the trace gas profile itself and the Sa

covariance matrix of the aerosol extinction profile should be given for each hour as the a priori profile. Therefore the estimation of the smoothing error as it is calculated here, gives just a rough general idea about the size of the smoothing error.

5.2 Measurement noise

The measurement noise error can be calculated from the ~~Gain~~ Gain matrix and the measurement noise matrix (S_m) (Rodgers, 2000). The diagonal Matrix (S_m) is already used in the optimal estimation of the profile retrieval to weight the different dSCDs of a scan according to the square of the errors from the QDOAS retrievals. The statistics of the measurement dSCDs errors and its dependence to the elevation angle has been presented in Fig. 2.

The measurement noise error matrix for both 1) the NO_2 profile and 2) aerosol extinction profile are calculated directly during the retrieval and are available from the MMF output of each measurement sequence as a full covariance matrix S_{noise}, given in units of partial columns and partial optical depths. The NO_2 total column error is then calculated from this and the total column operator, and amounts to around 2.4% of the VCD (see Table.1). The errors in the profile are shown in Fig. 6.

5.3 Parameter errors

The parameter errors originate from all uncertainties in input-parameters in the forward model that are not properly fitted, such as the cross-sections and for the NO_2-profile retrieval also the used aerosol extinction profile.

5.3.1 Parameter error from Spectroscopy

The error originating from an uncertainty in the cross-section ~~is estimated by assuming that the column amount regarding to the used cross-section has a uncertainty of~~ of 3~~% (Wang et al., 2017)~~% (Wang et al., 2017) is also around 3.0% in the vertical column (see Table.1) and similar in the vertical profile. Therefore, the error can be calculated using generalized measurement error covariance matrix $S^y_{spectroscopic}$ in the measurement space assuming that all retrieved dSCD are 3% too high (or low) and using the outer product of the measurement state vector containing the retrieved dSCDs.

$$S^y_{spectroscopic} = (0.33.0\%)^2 \cdot yy^T \tag{24}$$

$$S^x_{spectroscopic} = \mathrm{Gain} \cdot \mathrm{S^y_{spectroscopic}} \cdot \mathrm{Gain^T} \tag{25}$$

From the $S^x_{spectroscopic}$ matrix the error in the total column and profile ~~are~~ is calculated as shown earlier in Sec.5.1. As expected the error for the total column is 3% and the errors in the NO_2 VMR-profile are shown in Fig. 6. The spectroscopic error is a purely systematic error and affects all retrieval in the same manner.

5.3.2 Parameter error from aerosol profile

The AE-profile is crucial for the NO_2-profile retrieval because of its strong contribution to the airmass factor. The uncertainties in this vector of input parameters arise from errors in the aerosol extinction profile retrieved from the spectral signature of the

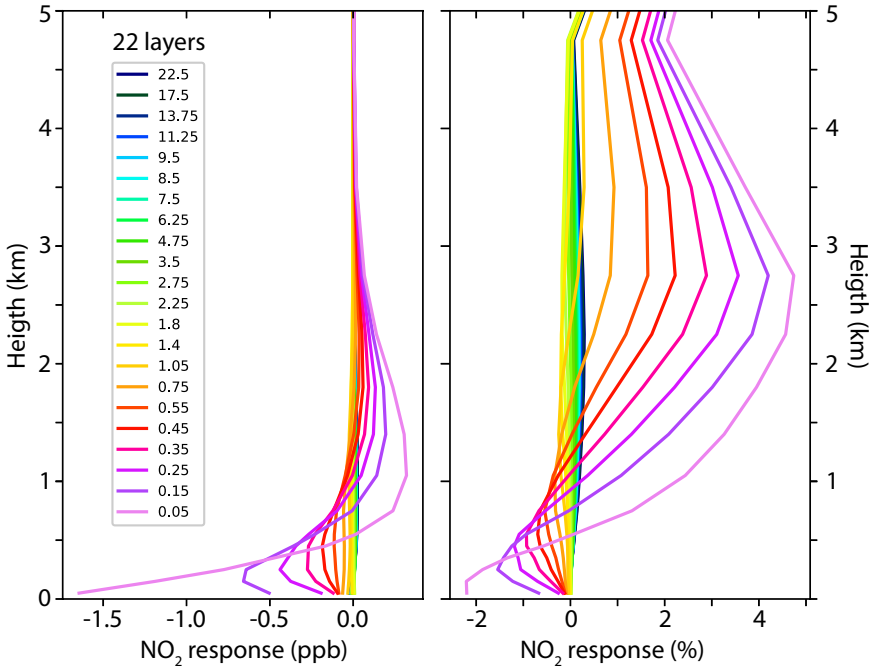

Figure 5. Plots showing the results of how sensitive the NO_2 profile is to changes in the aerosol extinction profile. Left: columns of the matrix D created to describe the response in the NO_2 VMR profile (the differences with respect to the original retrieval). Right: same as left but the response to the perturbation in the aerosol profile is given as fraction of the retrieved NO_2 profile.

O_4-dimer. For the propagation of the error of the O_4-AE-retrieval into the NO_2-retrieval, we assume here that the total error in the AE-profile depends on just the two contributions: a) smoothing error and b) measurement noise errors.

The effect that the assumed aerosol content in each layer has on the retrieved NO_2 profile is calculated by a sensitivity study. For this, 22 NO_2 retrievals are performed for the same measurement sequence (dSCDs) assuming slightly modified aerosol extinction profiles. First, a normal retrieval with the best estimated aerosol extinction profile is retrieved. Then, the aerosol extinction profile is modified so that the partial extinction of the i-th layer is increased by 1% of the total optical depth with respect to the original aerosol extinction profile.

The difference between the perturbed and original NO_2 VMR profiles are combined into the matrix D, describing thus how the NO_2 VMR profile responds to changes in the aerosol at different heights. The result is presented in Fig. 5.

With the help of the matrix $\mathrm{D_{VMR}}$, the different errors in the aerosol extinction profile as measurement noise, smoothing or even due to the algorithm error according to Wang et al. (2017) can be propagated to the NO_2 profile and its VCD.

$$S^{aerosol}_{NO2-VMR} = \mathrm{D_{VMR}Sa_{aerosol}D_{VMR}}^{\mathrm{T}} \tag{26}$$

$$error_{NO2-VCD}^{aerosol} = \sqrt{g^T \mathrm{U_{aircols}} \cdot \mathrm{D_{VMR}} \cdot \mathrm{Sa_{aerosol}} \cdot \mathrm{D_{VMR}}^{\mathrm{T}} \mathrm{U_{aircols}} \cdot \mathrm{g}}, \quad (27)$$

The propagation of the smoothing (4.6%) and measurement noise (2.2%) errors of the O_4 retrieval into the NO_2-retrieval results in a 5.1% error in the NO_2 VCD, while if no O_4-retrieval is performed successfully the error would be in our example 9.8 % (see Table.1). In case we would include the algorithm error (7.8%) introduced by Wang et al. (2017), the error when a O_4-retrieval is performed successfully would be ~~9.4~~9.3%. However the algorithm-error is calculated from the resulting residual of the fit and is ~~not independent~~ dependent on the other error sources as mentioned earlier.

5.4 Forward model error

The forward model error could be evaluated if an improved forward model was available, which ~~it~~ is not the case. However, errors in the forward model would result in systematic structures in the residual and a larger residual than expected from the error calculated by QDOAS for the slant columns. The ~~most rigorous~~ least supported simplification in the forward model is the assumption of a horizontal homogeneity of gas and aerosols. A horizontal inhomogeneity leads to a set of slant columns in a scan which cannot be simulated by ~~the~~ VLIDORT and the residuals of measured minus calculated slant columns indicate this error. Following Wang et al. (2017), we calculate the variance in the residuals as a function of viewing angle and hence the algorithm error. But as all errors increase the residual the so called *algorithm error* is not identical to the forward model error. However, it is maybe a good way to check empirically if there are important error contributions missing in the analytical analysis.

5.5 Total error

The results from the error estimations are summarized in Table. 1. The overall error in the NO_2 VCD is estimated to be around 14.1% (20.3% including the ~~algoritm errors (~~algorithm errors for NO_2 and O_4 ~~)~~ calculated from the residuals). The contributions are 12.5% from smoothing, 2.4% from measurement noise, 3% from spectroscopy and 5.1 % from errors in the aerosol extinction profile. The ~~algorithmn~~ algorithm error are 12.3% from the NO_2 retrieval itself and 7.8% from the O_4-AE retrieval. These ~~results are fairly~~ errors are similar to those reported by Wang et al. (2017).

The error in the vertical column is ~~relatively~~ smaller than the errors in the VMR profile for almost all layers (Fig. 6)~~as there is~~. This can be explained by an anti-correlation in different partial column errors indicated by ~~to~~ the full error covariance matrix. While the measurement noise seems to play a minor role, the smoothing and the aerosol profile are the main sources of error. Even though the error might seem large, the retrievals still provide new and relevant information of NO_2 within the boundary layer of Mexico City.

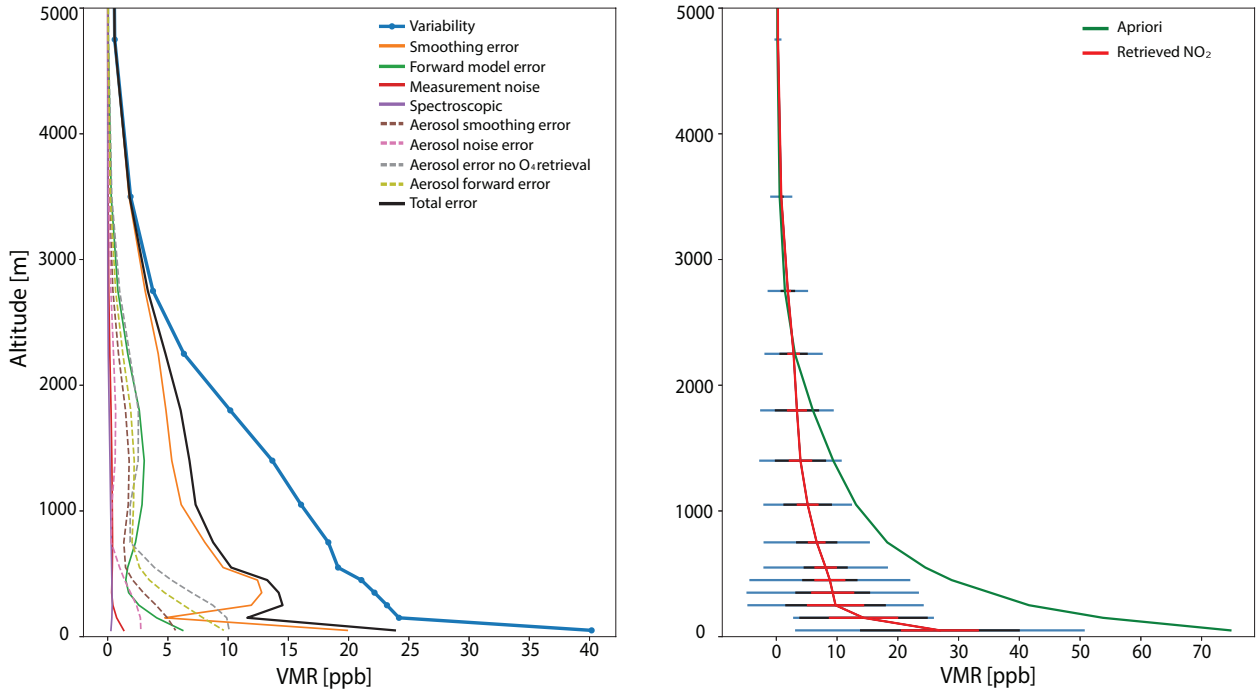

Figure 6. Altitude-dependent errors in a NO_2 VMR profile measured at 13:58 on May 20, 2016 (UTC-6h). a) The square root of the diagonal in the covariance matrix describing the individual error contributions in the VMR profile and b) the retrieved NO_2 VMR profile with corresponding error bars (total ~~.~~error and variability). The left panel also includes the *a priori* profile in green.

6 Results

In this section, we present results of the NO_2 variability measured in one of the stations comprising the MAX-DOAS network operated in Mexico City, and compare them with *in situ* measurements performed at the surface. The location of the MAXDOAS station, see Arellano et al. (2016) for a map, coincides with the aeronet station and the location of the surface sensor (19.32 N, 99.18 W). The launch location of the radio-sondes is close to the airport, around 15 km north-east from UNAM at 19.40 N, 99.20 W.

A data set from January 2015 to July 2016 (~~aprox~~ approx. 19 months) is considered in this study, in which both MAX-DOAS and *in situ* surface measurements were available. A total number of 2531 coincidences of hourly averages were found in this period, considering that at least 6 measurement sequences of the MAX-DOAS instrument in each reported hour needed to be available. The complete time series is presented in Fig. 7, showing hourly NO_2 averages of the *in situ* surface concentration in red, the total vertical column from the MAX-DOAS in blue, and the average VMR of its first 6 layers closest to the surface in green. This plot includes all MAX-DOAS results regardless of the sky conditions.

	Error	Description	NO_2, %
1	Smoothing	Variability from model	12.6
2	Measurement noise	Fig. 2	2.4
3	Spectroscopy	3 %	3
4	from NO_2 residuals	20 residuals (diagonal)	(12.3)
5	Aerosol noise	from QDOAS	2.2
6	Aerosol smoothing	Wang et al.,(2017)	4.6
7	Aerosol variability	Wang et al.,(2017)	(9.8)
8	from O_4 residuals	20 residuals (diagonal)	(7.8)
9	Total aerosol	5,6	5.1
10	Total aerosol	5,6,8	9.3
11	Total, including	1,2,3,5,6	**14.1**
12	Total, including	1,2,3,7 (without O_4 retrieval)	16.4
13	Total error (all)	1,2,3,4,5,6,8	20.3

Table 1. Errors in NO_2 vertical column density as fraction of a retrieved VCD of 3.2e16 molec./cm^2 measured at 13:58 on May 20 of 2016 (UTC-6h). The total error is calculated as the square root of the sum of the squares of different independent errors. Different total errors are calculated by including or not including the algorithm error (Wang et al., 2017) and for the two cases that an aerosol extinction profile is retrieved from O_4-absorption or just an ~~a-priori~~ *a priori* guess is used. For the assumed variability of the aerosol extinction the error due to the uncertainty remaining after an aerosol extinction profile retrieval is 5.6% instead of 9.8%. ~~But~~ However, if the algorithm error according to Wang et al. (2017) is included, the remaining error due to the uncertainty in the aerosol profile is slightly better: 9.3% instead of the 9.8% without O_4 retrieval.

Apart from the large data gaps in the beginning of 2016, the time series is quite complete and the fitted annual periodic functions (a Fourier series constrained to a fixed seasonal cycle shape) applied to the three data sets show a similar pattern. VMR values from the surface measurements are clearly higher than the VMR's detected in the lowest layers of the MAX-DOAS retrievals.

The 2531 coincidences were correlated averaging different number of layers starting from the ground and the resulting Pearson's coefficients (R) are plotted in Fig. 8 (top). Currently, the integration times in the spectra from which the O_4 dSCDs are calculated, are not long enough to ensure an O_4 dSCD error resulting in DOF larger than 1 for the aerosol retrieval. Since we use a ~~Tihkonov~~ Tikhonov regularization for aerosol retrieval, this means that we can basically retrieve the total aerosol extinction. Due to the topographical circumstances in Mexico City, where the boundary layer rises steadily during the day, using an ~~a-priori~~ *a priori* that is calculated from hourly averages of ceilometer data will generally provide a good ~~a-priori~~ *a priori* profile shape. However, for days that likely have a very different aerosol profile from our ~~a-priori~~ *a priori* profile, the aerosol profile included in the trace gas forward model accounts for far more than the estimated error in 5.3.2 since this estimation uses a fixed percentage of the ~~a-priori~~ *a priori* profile in the S_a matrix. For cloudy days, for example, the form of the profile will look considerably different than the ceilometer averages. We are currently not able to retrieve such profiles. Hence,

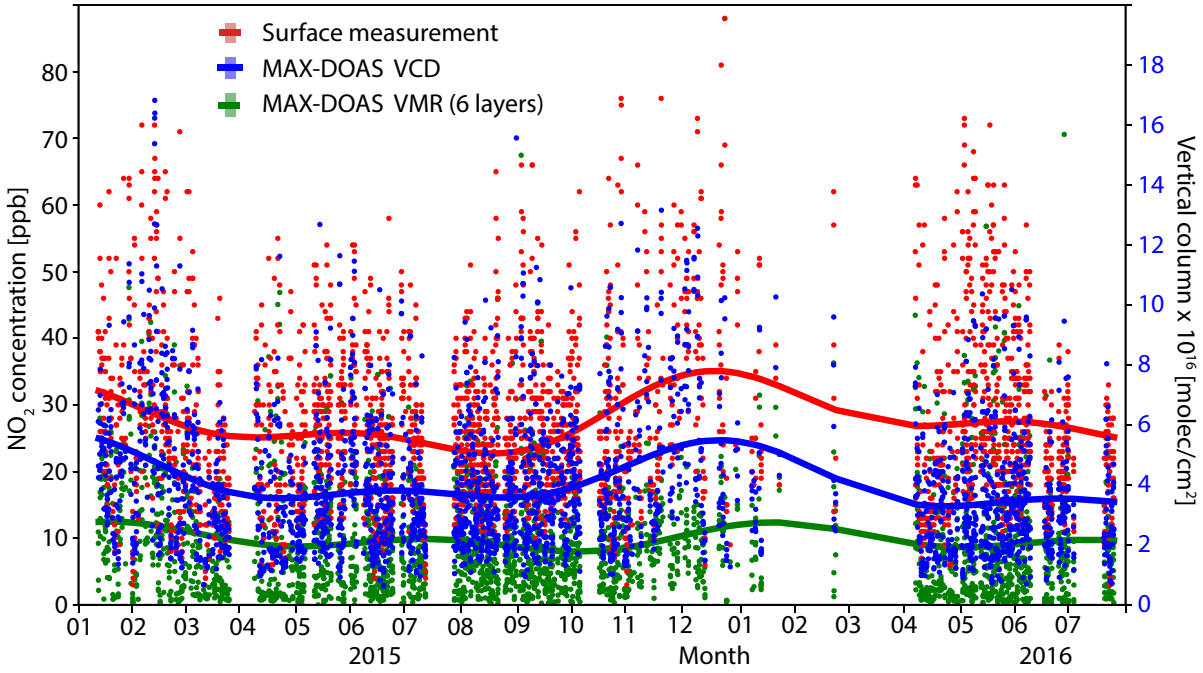

Figure 7. Hourly NO_2 averages of *in situ* surface concentrations (red) in ppb, total VCD's from the MAX-DOAS instrument (blue) in molecules/cm^2 and the average VMR of the first 6 layers (up to 550 m) in the retrieved MAX-DOAS profiles (green), also in ppb.

we limit our analysis to cloud free days. ~~Currelty~~Currently, the retrieval chain has no cloud-screening algorithm included. As an ad-hoc solution, we use the presence of AERONET data on that day as an estimator for cloud free days. For our entire data set, ~~about in~~ for about half of the coincident measurements (1270) there was at least one AERONET measurement available on that particular day while for the rest of our data (1261), no AERONET data was available on that day.

Generally, a ~~good~~ better correlation is found when averaged MAX-DOAS VMR's are compared to *in situ* surface measurements. Average R values larger than 0.7 are obtained in all cases when the first 6-8 layers are considered in the correlation. These correlations decrease rapidly when less layers are considered due to increased errors in the lowermost part of the profile. Naturally, the R value also decreases towards larger number of layers averaged since different air masses are measured at higher altitudes. As expected, slightly larger correlations are obtained for MAX-DOAS retrievals using a measured aerosol input from the AERONET (blue traces in Fig. 8), which correspond to clear sky conditions.

Linear regressions as the ones shown in Fig. 9 were generated for each data set, in order to gain more information as to how the number of layers averaged relates to the VMR comparison with the *in situ* surface measurements. The slope of all coincident data using the first 7-8 layers is 0.45, and it can reach values above 0.6 considering the first 6 layers in clear sky conditions (AERONET data available within ± 1 h of the measurement time). We thus decide to use in this study the 6-layer VMR averages of our MAX-DOAS retrievals, reaching a height of 550 m above the ground level, in the comparisons we present with surface concentrations.

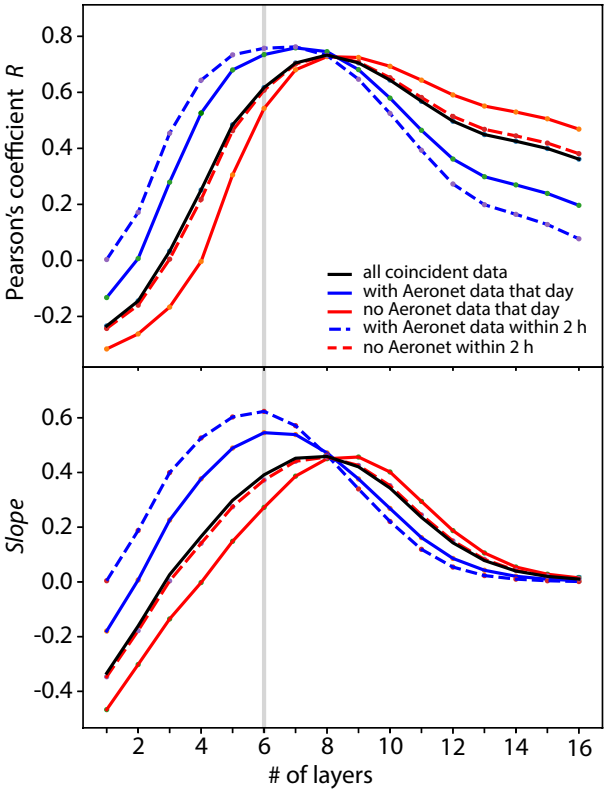

Figure 8. Pearson's correlation coefficients (top) obtained from hourly surface and MAX-DOAS coincident measurements as a function of the number of layers considered in the VMR averages. The slopes obtained from the linear regressions, performed forcing the intercept to be at zero to the same data sets, are shown in the lower panel. The layers below layer 6 are very likely inside the mixed boundary layer, while layers above layer 6 might be on the edge or above the boundary layer during some hours in the morning ~~(Franco-García et al., 2018)~~(García-Franco et al., 2018). The blue dashed lines only show data when AERONET data are avaible within 2 hours of the scan and hence limit the measurements to cloud free conditions. This selection criteria improves the correlation significantly.

To investigate the difference in our data set between clear and cloudy days, a correlation plot is presented in Fig. 9 showing the linear regressions produced from the MAX-DOAS 6-layer VMR averages for the retrievals with AERONET availability on that day in blue and those retrievals using a default aerosol ~~a-priori~~ *a priori* in red. A significant improvement in the correlation coefficient going from 0.54 to 0.74 is evident in the plot. ~~The *R* and slope values when~~ When all the coincident data is considered, ~~indistinctly~~ 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 ~~.~~

Examples of how the averaged 6-layer VMR's and VCD compare to the ground level concentrations in individual days are presented in Fig. 10. These examples were chosen to depict distinct diurnal patterns that occur in different times of the year. Again, the MAX-DOAS 6-layer product is consistently lower than the NO_2 concentrations measured at the surface. These,

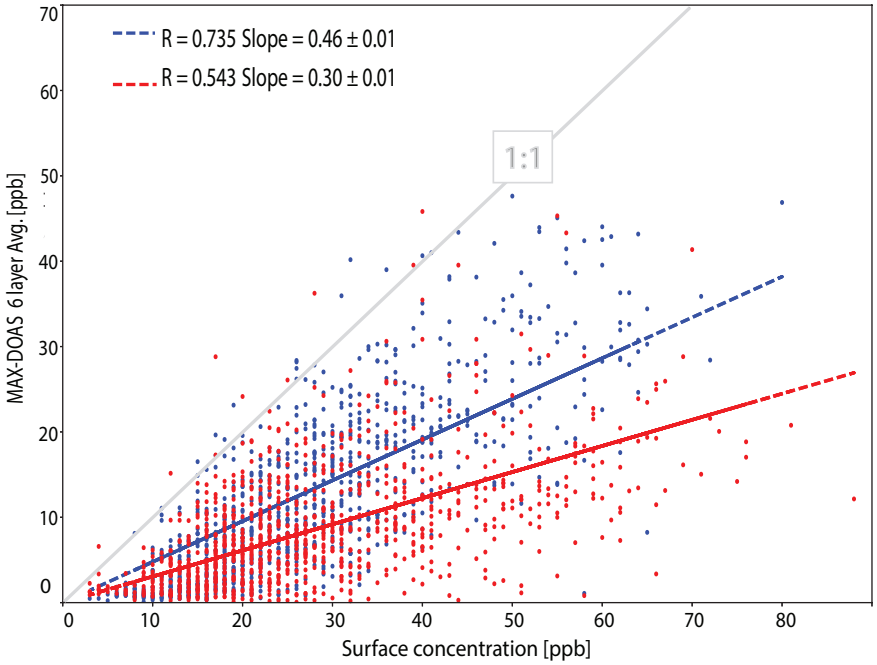

Figure 9. Correlation plot of the MAX-DOAS 6-layer VMR averages vs. the surface measurements shown separately for clear sky measurements in which aerosol data on that day was available (blue) and for MAX-DOAS retrievals with no AOD availability from AERONET, corresponding most likely to cloudy days (red).

together with the corresponding VCD's plotted on a different y-axis, follow the pattern of the surface measurements quite well. The MAX-DOAS instrument captures the features observed at the surface but not without some interesting differences.

May 4, 2016 was in the middle of a high pollution episode in Mexico City, in which a restriction on the use of private vehicles was declared during four days. Ozone had surpassed, at least in one of the stations from the monitoring network operated by the city government, the 165 ppb 1-hour average concentration. It is interesting to see in the May 4 plot, that indeed the surface concentrations of nitrogen dioxide were high in the morning, a condition favoring ozone production, but there is a large NO_2 peak just after noon in the MAX-DOAS VCD and 6-layer product data which is not detected at the surface. This is more evident from the individual measurements shown as dots than from the hour averages.

Other differences have to do with the evolution of the mixed layer height, which has been shown to have a rapid growth in the late morning, strongly influencing the surface concentrations (García-Franco et al., 2018). This is for example evident on the December 22 plot, where a peak is observed at 11 h and a second one at 14 h. This peak is strongest at 11 h in the case of the surface measurement, but the total column shows the peak at 14 h to be dominant. The mixed layer has grown in the mean time so that the registered 14 h surface concentration is ~~relatively lower despite that there are~~ lower despite there being more NO_2 molecules in the atmosphere.

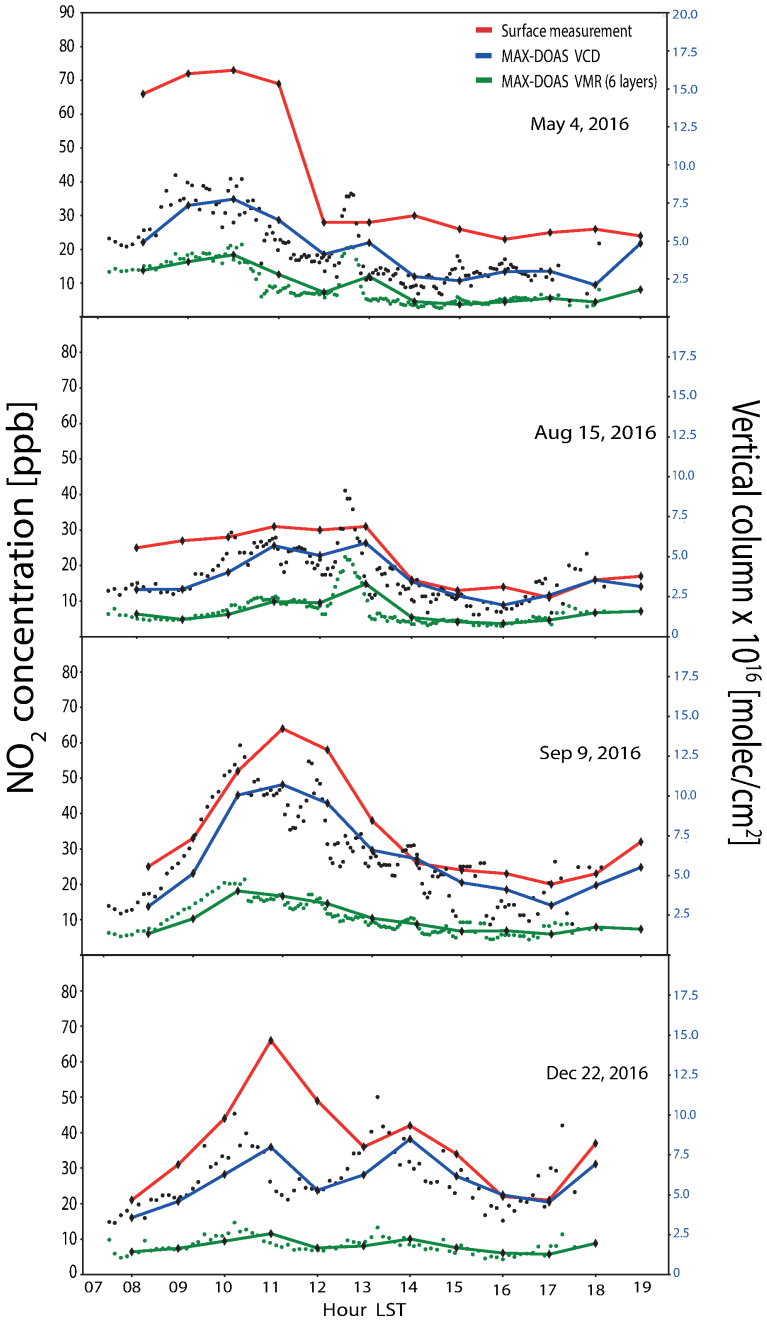

Figure 10. Example days showing the NO_2 variability in both VMR from the MAX-DOAS (green, < 550 m) and surface measurements (red), as well as the total vertical column densities (blue).

In order to see the diurnal and seasonal patterns from both MAX-DOAS and surface measurements, all coincident measurements between January 2015 to July 2016 including days with clouds were averaged to a specific hour or month, respectively. For consistency, only the coincident 1-hour data was considered in order to make the data sets comparable. The mean diurnal evolution with standard deviations as vertical bars are shown in Fig. 11. The maxima of both surface and MAX-DOAS 6-layer products are at around 10 h, while that of the vertical column is shifted towards noon. This is expected from what is known from the emissions, growing mixed layer and ventilation patterns in Mexico City (e.g. García-Franco et al., 2018). Despite the fact that the offset in the curves for surface- and MAX-DOAS measurements appears to be nearly constant throughout the day, it would be interesting to investigate further how this offset varies in different seasons particularly when vertical mixing is not favoured.

Fig. 12, on the other hand, presents the compiled data considered in this study as monthly means in order to observe the seasonal variability and compare the data sets from these two measurement techniques. The three data products coincide in having the highest monthly mean values between April and June. A difference which is probably interesting to note is during December and January, when the concentrations from the surface measurements are relatively higher than those from the MAX-DOAS products. This might have to do with the fact that during these winter months the mixed layer is shallower than during the rest of the year, and the sensitivity of the MAX-DOAS instrument is reduced for the lowest layers as described above.

7 Conclusions

In this contribution we describe the methodology used to analyze the data produced by the network of MAX-DOAS instruments (Arellano et al., 2016) operating in the Mexico City Metropolitan Area. In general terms, MMF is a retrieval code developed to process the data acquired by the instruments converting the measured spectra together with other input parameters to vertical profiles of aerosols and ~~target~~ certain trace gases. We have tested the performance of the code for NO_2 ~~in~~ at one of the stations and present some diagnostics and a full error analysis of the results. In the case of NO_2, all error sources amount to about 14.1%, the smoothing error being the dominating one (12.5%) followed by the error due to the uncertainties in the aerosol extinction profile (5.1%).

Both the resulting total vertical column densities and a product consisting of the average VMR in the first 6 layers (< 550 m above the ground level), were compared to ground level NO_2 concentrations. Good correlations were obtained between the 6-layer product and the values from the surface measurements, with R values between 0.6 and 0.7. However, the MAX-DOAS systematically underestimates the ground level concentrations by a factor of about ~~0.4~~0.6. This is consistent with the total column averaging kernels ~~reported in section~~ and variability Sa reported in Sect. 5.1, meaning that the MAX-DOAS instrument has a significantly lower sensitivity near the surface and is most sensitive at a height of around 1 km. It is shown, however, that this underestimation is less for clear sky conditions as suggested from comparing data when aerosol optical depths where available from independent measurements.

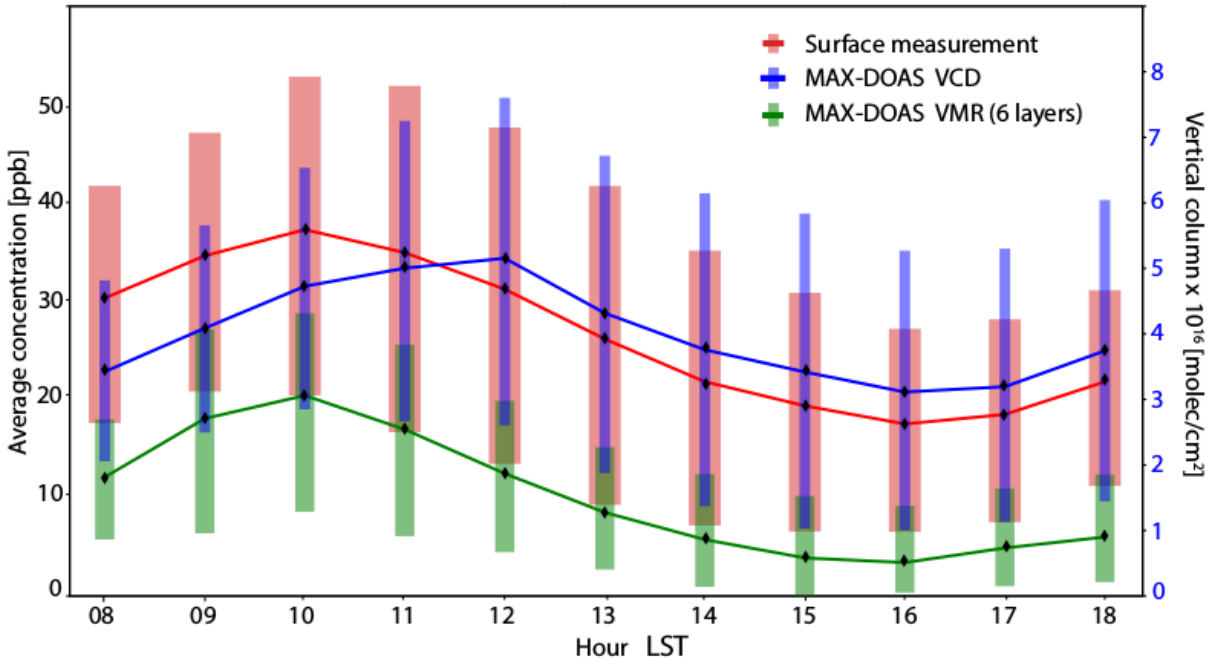

Figure 11. Diurnal NO_2 variability from hourly means of the entire data set spanning from January 2015 to July 2016. Only the data with coincident surface and MAX-DOAS measurements are included for consistency in the comparison. Vertical bars correspond to the standard deviations.

Although results are shown only from one instrument located in the southern part of the city (UNAM), several years of data are available from three other stations at different locations within the metropolitan area (Acatlán, Vallejo and Cuautitlán) that are being analyzed and used for studying the spatial and temporal variability of NO_2 ~~'s~~ in conjunction of several satellite products.

In this work we present a new and competitive retrieval code which has been developed for retrieving trace gas profiles from MAX-DOAS measurements. ~~It~~ Since this study, it has been further improved to perform retrievals in logarithmic space to avoid unphysical negative partial columns and oscillations; it uses the more stable Levenberg-Marquardt iteration scheme and decouples the retrieval and simulation grids.

Further efforts include the improvement in the measurement noise by increasing the integration times during spectral acquisition, which will allow us to improve the retrieval of other gases such as formaldehyde (HCHO) and other weak absorbers. This will also enable us to use OE as retrieval strategy for aerosol retrieval and possibly increase the DOF to comparable values as for the NO_2 retrieval. This, together with an inclusion of a cloud-screening algorithm in the retrieval chain, will make the retrieval less ~~dependend~~ dependent on the availability of AERONET data.

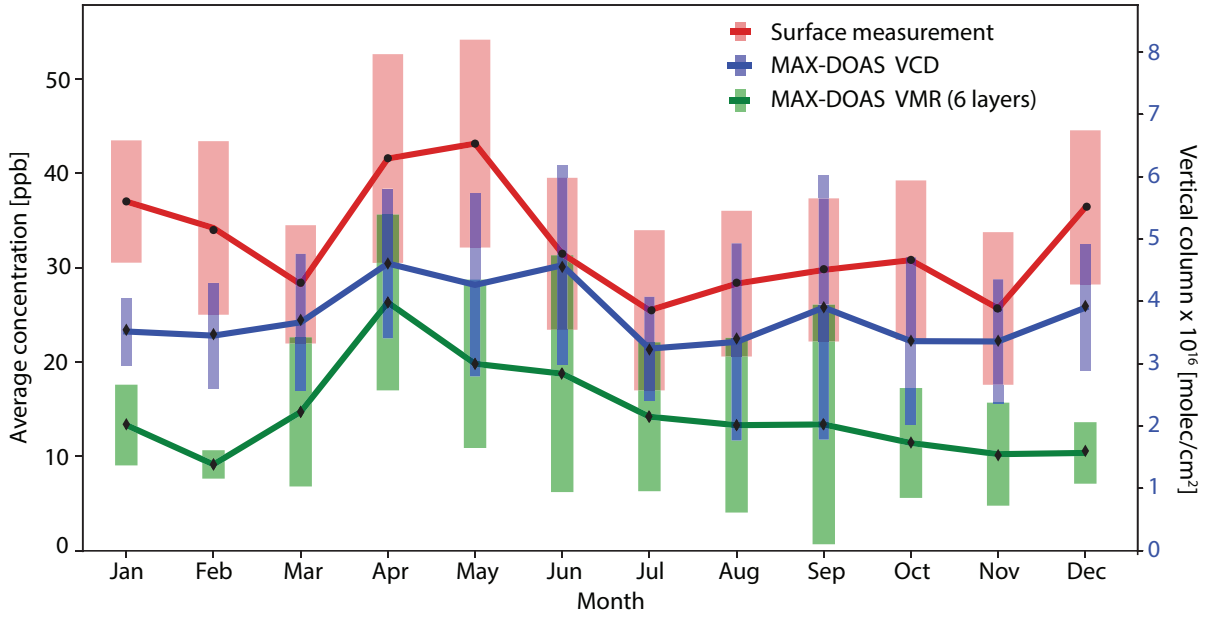

Figure 12. Annual NO_2 variability resulting from the monthly mean coincident data (surface and MAX-DOAS) between January 2015 and July 2016. The standard deviations are given in vertical bars.

Appendix A

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, i.e. Eq. 1 was replaced by Eq. A1 for more non-linear inversion problems (Rodgers, 2000)

$$x_{i+1} = x_i + \left[(1+\gamma)R + K_i^T S_m^{-1} K_i\right]^{-1} \left[K_i^T S_m^{-1}(y - F(x_i)) - R(x_i - x_a)\right] \quad \text{(A1)}$$

The symbols mean the same as in Eq. 1: superscript T denotes transposed, superscript -1 denotes the inverse, subscript a indicates *a priori* values. S_m is the measurement error covariance matrix, y denotes the vector of measured differential slant column densities, $F(x_i)$ are the simulated differential slant column densities at iteration i, calculated using the forward model with input profile x_i, K_i is the Jacobian matrix at the i-th iteration.

The new x_{i+1} is only accepted if the cost function in Eq. A2 decreases w.r.t the previous cost-function

$$\text{cost} = \sum_{\text{k}}^{\text{angles}} \sum_{\text{j}}^{\text{angles}} (\text{y} - \text{F(x)})_{\text{k}} S_m^{-1kj} (y - F(x))_j + \sum_{k}^{layers} \sum_{j}^{layers} (x - x_a)_k Sa^{-1kj} (x - x_a)_j. \quad \text{(A2)}$$

If this is the case and $(1+\gamma)$ is not yet equal to 1, the factor $(1+\gamma)$ is halved for the next iteration. If however, the cost-function increases, the newly calculated x_{i+1} is discarded and the i-th calculation repeated with a factor $(1+\gamma)$ increased by a factor of 16.

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 is around 5 seconds.

Tests using the logarithm of the partial layer vertical column density (for NO_2 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., 2018). It has also participated in the profile retrieval from dSCD from the CINDI-2 campaign, both for NO_2 and HCHO (Tirpitz et al., in preparation).

The LM scheme of Eq. A1 has currently only been tested with OE and not with Tikhonov regularization, i.e. the aerosol retrieval was also performed using OE.

Appendix B: Calculations for dry air density

In this appendix, we outline the calculation of Z, the molar mass of air M_a and the depolarization ratio Δ. We follow Ciddor (1996). If T is given in K (temperature in C is denoted as T_C), σ the wavenumber in 1/μm, pressure P in Pa, volume mixing ratio of CO_2 in ppm, the constants in Eq. B1 - Eq. B14 have the following values:

$k_0 = 238.0185$	$a_0 = 1.58123 \times 10^{-6}$	$b_0 = 5.707 \times 10^{-6}$	$w_0 = 295.235$
$k_1 = 5792105.0$	$a_1 = -2.9331 \times 10^{-8}$	$b_1 = -2.051 \times 10^{-8}$	$w_1 = 2.6422$
$k_2 = 57.362$	$a_2 = 1.1043 \times 10^{-10}$	$c_0 = 1.9898 \times 10^{-4}$	$w_2 = -0.032380 - 0.03238$
$k_3 = 167917.0$	$\alpha = 1.00062$	$c_1 = -2.376 \times 10^{-6}$	$w_3 = 0.004028$
$d = 1.83 \times 10^{-11}$	$\beta = 3.14 \times 10^{-8}$	$A = 1.2378847 1.2379 \times 10^{-5}$	$C = 33.9371104733.937$
$e = -0.765 \times 10^{-8}$	$\gamma = 5.6 \times 10^{-7}$	$B = -1.9121316 - 1.912 \times 10^{-2}$	$D = -6.3431645 - 6.343 \times 10^{3}$

First, the saturation vapor pressure $\mathcal{S}$ is calculated according to Eq. B1 then the enhancement factor for water vapour in air, $\mathcal{F}$, is calculated according to Eq. B2 and with this, the molar fraction of water vapor in moist air x_w is calculated according to

Eq.B3 where P_w is partial water vapor pressure in Pa.

$$\mathcal{S} = \exp\left(AT^2 + BT + C + D/T\right) \tag{B1}$$

$$\mathcal{F} = \alpha + \beta P + \gamma T_C^2 \tag{B2}$$

$$x_w = \mathcal{F} P_w / P \tag{B3}$$

The air density in dry air $n_{\text{air,dry,450}}$ (0 % humidity) at standard conditions (15 °C, 101.325 Pa, xCO_2 = 450 ppm) can be calculated as in Eq. B4 and from this, the dry air density at xCO_2 is calculated according to Eq. B5.

$$(n_{\text{air,dry,450}} - 1) \times 10^8 = \frac{k_1}{k_0 - \sigma^2} + \frac{k_3}{k_2 - \sigma^2} \tag{B4}$$

$$(n_{\text{air,dry,xCO}_2} - 1) = (n_{\text{air,dry,450}} - 1) \times 10^8 \left(1.0 + 0.534 \times 10^{-6} \cdot (xCO_2 - 450.0)\right) \tag{B5}$$

The equivalent quantity for water vapour, n_{wv}, is calculated as follows:

$$n_{\text{wv}} = 1.022 \cdot (w_0 + w_1\,\sigma^2 + w_2\,\sigma^4 + w_3\,\sigma^6) \tag{B6}$$

The molar mass of dry air, M_a with xCO_2 is calculated as:

$$M_a = 10^{-3}\left(28.9635 + 12.011 \times 10^{-6}(x\text{CO}_2 - 400.0)\right) \tag{B7}$$

The compressibility of dry air, Z_a, and pure water vapour, Z_w, are calculated as:

$$Z_a = 1.0 - \frac{101325.0}{288.15}\,(a_0 + 15\,a_1 + 225\,a_2) \tag{B8}$$

$$Z_w = 1.0 - \frac{1333.0}{293.15}\,(a_0 + 20\,a_1 + 400\,a_2 + b_0 + 20\,b_1 + c_0 + 20\,c_1) + \left(\frac{1333.0}{293.15}\right)^2 (d + e) \tag{B9}$$

Using Eq. B10 once with the values for Z, R and T for standard water vapour and once with the values for standard air, one can calculate ρ_{ws} and ρ_{axs}, respectively.

$$\rho = (pM_a/(ZR_{gas}T)\,(1 - x_w(1 - M_{\text{mwv}}/M_a)) \tag{B10}$$

Here, $R_{gas} = 8.3144621$ J/mol/K is the gas constant, $M_{\text{mwv}} = 0.018015$ kg/mol is the molar mass of water vapour. Z, the compressibility of moist air at T, x_w and P can be calculated as

$$Z = 1.0 - (P/T)(a_0 + a_1 T_C + a_2 T_C^2 + (b_0 + b_1 T_C)x_w + (c_0 + c_1 T_C)x_w^2) + (P/T)^2\,(d + e\,x_w^2) \tag{B11}$$

Using again Eq. B10, with the actual values for P, x_w and Z, the density of the dry air component can be calculated according to Eq. B12 and the density of the water vapour component can be calculated according to Eq. B13

$$\rho_{\text{dry air}} = P\,M_a\,(1 - x_w)/(Z\,R_{\text{gas}}\,T) \tag{B12}$$

$$\rho_{\text{wv}} = P\,M_{\text{mwv}} x_w/(Z\,R_{\text{gas}}\,T) \tag{B13}$$

Table B1. depolarization ratios and VMR for main air constituents

gas	VMR/ppm	Δ gas
N_2	780840.0	see Eq. B16
O_2	209460.0	see Eq. B16
Ar	9340.0	0.0
CO_2	400.0	0.0814

Finally, the number density of the air ρ and the refractive index n are calculated as:

$$\rho = \frac{\rho_{\text{dry air}}}{M_a} A_v \tag{B14}$$

$$n = \frac{\rho_{\text{dry air}}}{\rho_{\text{axs}}} (n_{\text{air,dry,xCO}_2} - 1) + \frac{\rho_{\text{wv}}}{\rho_{\text{ws}}} (n_{\text{wv}} - 1) \tag{B15}$$

Here, $A_v = 6.02214129 \times 10^{23}$ is the Avogadro number.

We use the N_2, O_2 and CO_2 depolarization factors from Bates (1984) and fixed VMR as summarized in Table B1 and calculate the wavelength dependent depolarization factor via Eq. B16 where F is the so called King factor and is calculated for N_2 (λ in μm) and O_2 according to Eq. B18 and Eq. B19, respectively. The complete depolarization factor is then calculated according to Eq. B19

$$\Delta = 6.0 \cdot (1.0 - F1 - F)/(-7.0 - 7F - 3.03) \tag{B16}$$

$$F(\text{N}_2) = 1.034 + 3.17 \times 10^{-4}/\lambda^2 \tag{B17}$$

$$F(\text{O}_2) = 1.096 + 1.385 \times 10^{-3}/\lambda^2 + 1.448 \times 10^{-4}/\lambda^4 \tag{B18}$$

$$\Delta = \sum (\Delta_i \text{VMR}_i), \quad i = \text{N}_2,\text{O}_2,\text{Ar, CO}_2 \tag{B19}$$

For the Rayleigh scattering cross section, we implement equation 7.37 from Goody and Yung (1989) and equation 19.5 from Platt et al. (2007), see Eq. B20.

$$Q_{\text{Ray}} = \frac{24.0\pi^3(n^2-1)^2}{\lambda^4 \rho^2 (n^2+2)^2} \frac{6.0+3.0\Delta}{6.0-7.0\Delta} Q_{\text{Ray}} = \frac{24\pi^3(n^2-1)^2}{\lambda^4 \rho^2 (n^2+2)^2} \frac{6+3\Delta}{6-7\Delta} \tag{B20}$$

Author contributions. MMF: part of Abstract and Section 1, Figure 1, Figure 2, Figure 3, Section 4, Appendix, retrieval code development and testing, retrieval chain setup from spectra to profiles, retrieval parameter choice for MMF, software support CR: parts of Section 2, Section 3. QDOAS setup, dSCD retrieval parameter choice, technical support WS and ZO: parts of Abstract, Section 5, Figures 4 – 12, part of Sect. 6 and 7. data retrieval, data analysis and interpretation, sensitivity tests and complete error analysis. JA: Section 2, instrument construction, installation and maintenance AB: parts of Sect. 2, technical support and data management support, instrument maintenance

JAGR: part of Sect. 4.2.1, WRF-Chem simulations that serve as a-priori data for NO_2 retrieval MG: Sect. 1, Sect 6 and Sect. 7., data analysis and interpretation.

Competing interests. The authors declare that they have no conflict of interest

Acknowledgements. We would like to thank the financial support provided by CONACYT (grants 275239 & 290589) and DGAPA-UNAM (grants IN107417,IN111418 and IA100716 & M.M. Friedrich's postdoctoral fellowship). SEDEMA is acknowledged for providing the *in situ* surface concentrations from their monitoring network (RAMA) as well for partially supporting the MAX-DOAS network, and A. Krueger, A. Rodríguez, H. Soto and D. Flores are thanked for their technical assistance. We thank ~~Thomas Danckaert (thomas.danckaert@aeronomie.be),~~ C. Fayt (caroline.fayt@aeronomie.be) and M. Van Roozendael (michelv@aeronomie.be) and Thomas Danckaert for the free use of the QDOAS software and we thank R. Spurr for free use of the VLIDORT radiative transfer code package. We thank the Solarimetric Mexican Service for their effort in establishing and maintaining the AERONET Mexico City site. We thank the University of Wyoming Department of Atmospheric Science for providing the sounding data on http://weather.uwyo.edu/upperair/sounding.html.

References

Arellano, J., Krüger, A., Rivera, C., Stremme, W., Friedrich, M., Bezanilla, A., and Grutter, M.: The MAX-DOAS network in Mexico City to measure atmospheric pollutants, Atmosfera, 29, 157–167, http://www.revistascca.unam.mx/atm/index.php/atm/article/view/ATM.2016.29.02.05, 2016.

Bates, D. R.: Rayleigh scattering by air, Planet. Space Sci., 32, 785–790, https://doi.org/10.1016/0032-0633(84)90102-8, 1984.

Buchwitz, M., Rozanov, V. V., and Burrows, J. P.: Development of a correlated-k distribution band model scheme for the radiative transfer program GOMETRAN/SCIATRAN for retrieval of atmospheric constituents from SCIMACHY/ENVISAT-1 data, in: Satellite Remote Sensing of Clouds and the Atmosphere III, vol. 3495 of Proc. SPIE, pp. 171–186, https://doi.org/10.1117/12.332681, 1998.

Burrows, J. P., Richter, A., Dehn, A., Deters, B., Himmelmann, S., Voigt, S., and Orphal, J.: Atmospheric remote-sensing reference data from GOME. 2. Temperature dependent absorption cross sections of O_3 in the 231-794 nm range., J. Quant. Spec. Radiat. Transf., 61, 509–517, https://doi.org/10.1016/S0022-4073(98)00037-5, 1999.

Carlson, D., Donohoue, D., Platt, U., and Simpson, W. R.: A low power automated MAX-DOAS instrument for the Arctic and other remote unmanned locations, Atmospheric Measurement Techniques, 3, 429–439, 2010.

Ciddor, P. E.: Refractive index of air: new equations for the visible and near infrared, Appl. Opt., 35, 1566, https://doi.org/10.1364/AO.35.001566, 1996.

Coburn, S., Dix, B., Sinreich, R., and Volkamer, R.: The CU ground MAX-DOAS instrument: characterization of RMS noise limitations and first measurements near Pensacola, FL of BrO, IO, and CHOCHO, Atmospheric Measurement Techniques, 4, 2421–2439, https://doi.org/10.5194/amt-4-2421-2011, 2011.

Dahlback, A. and Stamnes, K.: A new spherical model for computing the radiation field available for photolysis and heating at twilight, Planetary Space Science, 39, 671–683, https://doi.org/10.1016/0032-0633(91)90061-E, 1991.

Danckaert, T., Fayt, C., Van Roozendael, M., De Smedt, I., Letocart, V., Merlaud, A., and Pinardi, G.: QDOAS Software user manual, Belgian Institute for Space Aeronomy, 2013.

Finlayson-Pitts, B. J. and Pitts, J. N.: Chemistry of the upper and lower atmosphere: theory, experiments, and applications, Academic Press, San Diego, CA, 2000.

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, D14203, https://doi.org/10.1029/2005JD006618, 2006.

Frieß, U., Beirle, S., Alvarado Bonilla, L., Bösch, T., Friedrich, M. M., Hendrick, F., Piters, A., Richter, A., van Roozendael, M., Rozanov, V. V., Spinei, E., Tirpitz, J.-L., Vlemmix, T., Wagner, T., and Wang, Y.: Intercomparison of MAX-DOAS Vertical Profile Retrieval Algorithms: Studies using Synthetic Data, Atmospheric Measurement Techniques Discussions, 2018, 1–40, https://doi.org/10.5194/amt-2018-423, 2018.

Galle, B., Johansson, M., Rivera, C., Zhang, Y., Kihlman, M., Kern, C., Lehmann, T., Platt, U., Arellano, S., and Hidalgo, S.: Network for Observation of Volcanic and Atmospheric Change (NOVAC) A global network for volcanic gas monitoring: Network layout and instrument description, Journal of Geophysical Research, 115, D05 304, https://doi.org/10.1029/2009JD011823, 2010.

García-Franco, J., Stremme, W., Bezanilla, A., Ruiz-Angulo, A., and Grutter, M.: Variability of the Mixed-Layer Height Over Mexico City, Boundary-Layer Meteorol., https://doi.org/10.1007/s10546-018-0334-x, 2018.

Goody, R. M. and Yung, Y. L.: Atmospheric radiation : theoretical basis, 1989.

Hendrick, F., Müller, J.-F., Clémer, K., Wang, P., De Mazière, M., Fayt, C., Gielen, C., Hermans, C., Ma, J. Z., Pinardi, G., Stavrakou, T., Vlemmix, T., and Van Roozendael, M.: Four years of ground-based MAX-DOAS observations of HONO and NO_2 in the Beijing area, Atmospheric Chemistry and Physics, 14, 765–781, https://doi.org/10.5194/acp-14-765-2014, 2014.

Hermans, C., Vandaele, A. C., Carleer, M., Fally, S., Colin, R., Jenouvrier, A., Coquart, B., and Mérienne, M. F.: Absorption cross-sections of atmospheric constituents: NO_2, O_2, and H_2O, Environmental Science and Pollution Research, 6, 151–158, 1999.

Hess, M., Koepke, P., and Schult, I.: Optical Properties of Aerosols and Clouds: The Software Package OPAC., Bulletin of the American Meteorological Society, 79, 831–844, https://doi.org/10.1175/1520-0477(1998)079<0831:OPOAAC>2.0.CO;2, 1998.

Hönninger, G., Bobrowski, N., Palenque, E. R., Torrez, R., and Platt, U.: Reactive bromine and sulfur emissions at Salar de Uyuni, Bolivia, Journal Geophysical Research Letters, 31, L04101, https://doi.org/10.1029/2003GL018818, 2004.

Johansson, M., Rivera, C., de Foy, B., 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, https://doi.org/10.5194/acp-9-5647-2009, 2009.

Kirmse, B., Delon, A., and Jost, R.: NO_2 absorption cross section and its temperature dependence, J. Geophys. Res., 102, 16, https://doi.org/10.1029/97JD00075, 1997.

Kurucz, R. L., Furenlid, I., Brault, J., and Testerman, L.: Solar flux atlas from 296 nm to 1300 nm, National Solar Observatory Atlas No. 1, 1984.

Marquard, L.: Modellierung des Strahlungstransports in der Erdatmosphäre für absorptionsspektroskopische Messungen im ultravioletten und sichtbaren Spektralbereich, http://books.google.se/books?id=1hm8cQAACAAJ, 1998.

Marquard, L. C., Wagner, T., and Platt, U.: Improved air mass factor concepts for scattered radiation differential optical absorption spectroscopy of atmospheric species, Journal of Geophysics Research, 105, 1315–1327, https://doi.org/10.1029/1999JD900340, 2000.

Melamed, M. L., Basaldud, R., Steinbrecher, R., Emeis, S., Ruíz-Suárez, L. G., and Grutter, M.: Detection of pollution transport events southeast of Mexico City using ground-based visible spectroscopy measurements of nitrogen dioxide, Atmospheric Chemistry and Physics, 9, 4827–4840, https://doi.org/10.5194/acp-9-4827-2009, 2009.

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

Molina, L. T. and Molina, M. J.: Air quality in the Mexico megacity : an integrated assessment, Kluwer Academic Publishers, Dordrecht; Boston, 2002.

Orphal, J.: A critical review of the absorption cross-sections of O_3 and NO_2 in the ultraviolet and visible, JPPA, 157, 185–209, https://doi.org/10.1016/S1010-6030(03)00061-3, 2002.

Ortega, I., Koenig, T., Sinreich, R., Thomson, D., and Volkamer, R.: The CU 2-D-MAX-DOAS instrument - Part 1: Retrieval of 3-D distributions of NO_2 and azimuth-dependent OVOC ratios, Atmospheric Measurement Techniques, 8, 2371–2395, https://doi.org/10.5194/amt-8-2371-2015, 2015.

Palazzi, E., Petritoli, A., Giovanelli, G., Kostadinov, I., Bortoli, D., Ravegnani, F., and Sackey, S. S.: PROMSAR: A backward Monte Carlo spherical RTM for the analysis of DOAS remote sensing measurements, Advances in Space Research, 36, 1007–1014, https://doi.org/10.1016/j.asr.2005.05.017, 2005.

Perliski, L. M. and Solomon, S.: On the evaluation of air mass factors for atmospheric near-ultraviolet and visible absorption spectroscopy, Journal of Geophysics Research, 98, 10 363, https://doi.org/10.1029/93JD00465, 1993.

Platt, U. and Stutz, J.: Differential Optical Absorption Spectroscopy, Springer-Verlag Berlin Heidelberg, https://doi.org/10.1007/978-3-540-75776-4, 2008.

Platt, U., Pfeilsticker, K., and Vollmer, M.: Radiation and Optics in the Atmosphere, p. 1165, https://doi.org/10.1007/978-0-387-30420-5_19, 2007.

Rahpoe, N., von Savigny, C., Weber, M., Rozanov, A. V., Bovensmann, H., and Burrows, J. P.: Error budget analysis of SCIAMACHY limb ozone profile retrievals using the SCIATRAN model, Atmospheric Measurement Techniques, 6, 2825–2837, https://doi.org/10.5194/amt-6-2825-2013, 2013a.

Rahpoe, N., von Savigny, C., Weber, M., Rozanov, A. V., Bovensmann, H., and Burrows, J. P.: Error budget analysis of SCIAMACHY limb ozone profile retrievals using the SCIATRAN model, Atmospheric Measurement Techniques Discussions, 6, 4645–4676, https://doi.org/10.5194/amtd-6-4645-2013, 2013b.

Rivera, C., Barrera, H., Grutter, M., Zavala, M., Galle, B., Bei, N., Li, G., and Molina, L. T.: NO_2 fluxes from Tijuana using a mobile mini-DOAS during Cal-Mex 2010, Atmospheric Environment, 70, 532–539, https://doi.org/10.1016/j.atmosenv.2012.12.026, 2013.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific Publishing Co, https://doi.org/10.1142/3171, 2000.

Rodgers, C. D. and Connor, B. J.: Intercomparison of remote sounding instruments, Journal of Geophysical Research (Atmospheres), 108, 4116, https://doi.org/10.1029/2002JD002299, 2003.

Rozanov, V. V., Rozanov, A. V., Kokhanovsky, A. A., and Burrows, J. P.: Radiative transfer through terrestrial atmosphere and ocean: Software package SCIATRAN, J. Quant. Spec. Radiat. Transf., 133, 13–71, https://doi.org/10.1016/j.jqsrt.2013.07.004, 2014.

Rozanov, V. V., Dinter, T., Rozanov, A. V., Wolanin, A., Bracher, A., and Burrows, J. P.: Radiative transfer modeling through terrestrial atmosphere and ocean accounting for inelastic processes: Software package SCIATRAN, J. Quant. Spec. Radiat. Transf., 194, 65–85, https://doi.org/10.1016/j.jqsrt.2017.03.009, 2017.

Spurr, R.: VLIDORT: A linearized pseudo-spherical vector discrete ordinate radiative transfer code for forward model and retrieval studies in multilayer multiple scattering media, Journal of Quantitative Spectroscopy & Radiative Transfer, pp. 316–342, 2006.

Spurr, R.: User's Guide VLIDORT Version 2.6, RT Solutions, Inc., 2013.

Spurr, R. J. D., Kurosu, T. P., and Chance, K. V.: A linearized discrete ordinate radiative transfer model for atmospheric remote-sensing retrieval, Journal of Quantitative Spectroscopy and Radiative Transfer, 68, 689–735, https://doi.org/10.1016/S0022-4073(00)00055-8, 2001.

Stamnes, K., Tsay, S.-C., Jayaweera, K., and Wiscombe, W.: Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media, Applied Optics, 27, 2502–2509, https://doi.org/10.1364/AO.27.002502, 1988.

Steck, T.: Methods for determining regularization for atmospheric retrieval problems, Appl. Opt., 41, 1788–1797, https://doi.org/10.1364/AO.41.001788, 2002.

Tirpitz, J.-L., Frieß, U., Hendrick, F., Alberti, C., Allaart, M., Apituley, A., Bais, A., Beirle, S., Benavent, N., Berkhout, S., Bösch, T., Bognar, K., Bruchkouski, I., Chan, K. L., Chengxin, Z., den Hoed, M., Donner, S., Drosoglou, T., Friedrich, M. M., Frumau, A., Gast, L., Gomez, L., Gielen, C., Hao, N., Hensen, A., Henzing, B., Hoque, S., Irie, H., Jin, J., Koenig, T. K., Kreher, K., Kuhn, J., Kumar, V., Lampel, J., Li, A., Liu, C., Ma, J., Merlaud, A., Mishra, A., Nieto, D., Peters, E., Pinardi, G., Piters, A., Pöhler, D., Postylyakov, O., Richter, A., van Roozendael, M., Schmitt, S., Sinha, V., Spinei, E., Stein, D., Swart, D., Tack, F., Vlemmix, T., van der Hoff, R., Vonk, J., Wagner, T., Wang, S., Wang, Y., Wang, Z., Wenig, M., Wiegner, M., Wittrock, F., Xie, P., Xing, C., Xu, J., and Zhao, X.: DOAS Vertical Profile Retrieval Algorithms: Studies on Field Data from the CINDI-2 Campaign, Atmospheric Measurement Techniques, in preparation.

Vandaele, A. C., Hermans, C., Simon, P. C., Carleer, M., Colin, R., Fally, S., Mérienne, 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.

Vandaele, A. C., Hermans, C., Fally, S., Carleer, M., Merienne, M.-F., Jenouvrier, A., Coquart, B., and Colin, R.: Absorption cross-sections of NO_2: simulation of temperature and pressure effects, J. Quant. Spec. Radiat. Transf., 76, 373–391, https://doi.org/10.1016/S0022-4073(02)00064-X, 2003.

Vidot, J., Jourdan, O., Kokhanosvky, A. A., Szczap, F., Giraud, V., and Rozanov, V. V.: Retrieval of tropospheric NO_2 columns from satellite measurements in presence of cirrus: A theoretical sensitivity study using SCIATRAN and prospect application for the A-Train, J. Quant. Spec. Radiat. Transf., 111, 586–601, 2010.

von Friedeburg, C.: Derivation of Trace Gas Information combining Differential Optical Absorption Spectroscopy with Radiative Transfer Modelling, Ph.D. thesis, Ruperto-Carola University of Heidelberg, Germany (1993), 2003.

von Friedeburg, C., Wagner, T., Geyer, A., Kaiser, N., Vogel, B., Vogel, H., and Platt, U.: Derivation of tropospheric NO_3 profiles using off-axis differential optical absorption spectroscopy measurements during sunrise and comparison with simulations, Journal of Geophysical Research (Atmospheres), 107, 4168, https://doi.org/10.1029/2001JD000481, 2002.

Wagner, T., Beirle, S., Brauers, T., Deutschmann, T., Frieß, U., Hak, C., Halla, J. D., Heue, K. P., Junkermann, W., Li, X., Platt, U., and Pundt-Gruber, I.: Inversion of tropospheric profiles of aerosol extinction and HCHO and NO_2 mixing ratios from MAX-DOAS observations in Milano during the summer of 2003 and comparison with independent data sets, Atmospheric Measurement Techniques, 4, 2685–2715, https://doi.org/10.5194/amt-4-2685-2011, 2011.

Wang, S., Cuevas, C. A., Frieß, U., and Saiz-Lopez, A.: MAX-DOAS retrieval of aerosol extinction properties in Madrid, Spain, Atmospheric Measurement Techniques, 9, 5089–5101, https://doi.org/10.5194/amt-9-5089-2016, 2016.

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, https://doi.org/10.5194/acp-17-2189-2017, 2017.

Wilmouth, D. M., Hanisco, T. F., Donaue, N. M., and Anderson, J. G.: Fourier Transform Ultraviolet Spectroscopy of the $A^2\prod_{3/2} \leftarrow X^2\prod_{3/2}$, Journal of Physical Chemistry A, 103, 8935–8945, 1999.

Wittrock, F., Oetjen, H., Richter, A., Fietkau, S., Medeke, T., Rozanov, A., and Burrows, J. P.: MAX-DOAS measurements of atmospheric trace gases in Ny-Ålesund - Radiative transfer studies and their application, Atmospheric Chemistry and Physics, 4, 955–966, https://doi.org/10.5194/acp-4-955-2004, 2004.