Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., referee comment RC1
https://doi.org/10.5194/amt-2021-213-RC1, 2021

Comment on amt-2021-213

Anonymous Referee #1

Referee comment on "Calculating the vertical column density of O_4 during daytime from surface values of pressure, temperature, and relative humidity" by Steffen Beirle et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2021-213-RC1, 2021

1. General comments

The manuscript "Calculating the vertical column density of O_4 from surface values of pressure, temperature and relative humidity" presents in varying level of detail the derivation of, and the validation of, a daytime applicable method to calculate the O_4 vertical column density from surface values of pressure, temperature and relative humidity. Hence, the title is well chosen. The main application area of this approximation are parametrized profile inversion algorithms for MAX-DOAS measurements. The authors also mention possible benefits for optimal estimation based profile inversion algorithms, stemming from the high correlation of surface relative humidity and temperature effective lapse rate. I recommend this manuscript for publication after intermediate revisions. While there is certainly scientific value in the presented method, the quality of the presentation and the structure of the manuscript, have to be improved. The validation of the method is not quite sufficient and has to be extended. The degree of explicit derivation of equations varies from "very detailed" to "almost insufficient" and should be brought to a more "equal" level.

The presented derivation of Eq. 9 seems unnecessarily complicated: Starting at the usual barometric formula for the air density for an atmosphere with constant laps rate

$$\rho = \rho_0 \ (T_0/(T_0 + \Gamma(x-x_0)))^{1+(gM/R/\Gamma)}] = \rho_0 \ (1+ \Gamma x'/T_0)^{-(gM/R/\Gamma)}$$

($x'=x-x_0$) and integrating the square of this ρ multiplied by the oxygen volume mixing ratio (i.e. integrate the O_4 density) from the surface (0) to infinity, replacing ρ_0 by $p_0/T_0/R$, directly yields Eq. 9. Maybe the authors can comment on why they chose to over complicate things with introducing the ratio of h_{O2} and h_{O4}. (Likewise, choose the density formula for 0 lapserate and integrate the square, in order to arrive at the corresponding equation for 0 lapse rate). I highly recommend to streamline this. I cannot see any added benefit of the method used by the authors, but I do see a lot of unnecessary turns given. At several occasions in the manuscript, the authors refer to later sections or to, at that point, unproven and not referenced statements. This makes it impossible to read the manuscript in a linear fashion.The main line of reasoning should be clearly stated and followed. Several statements are made without proof or proper reference. Regarding style, the guidelines of AMT are, in several aspects, not followed. The quality of the plots is mostly ok but should also be improved before final acceptance (especially the readability

of axis labels).
Apart from the unfortunate structuring of the manuscript and the unnecessary turns given in order to arrive at the important equation, the biggest point of criticism is perhaps on the method validation and the lack of showing the improvement when using this new method over other methods to estimate the O_4 VCD, as well as the actual effect on the final product, the retrieved AOD.

Three rather limited data sets were used for validation. Each of these datasets needs to be extended. For one of the datasets (global model), half of the dataset was used to fit parameters in the model; still, that same half was also used for validation. This should really be avoided. It is advisable to add a separate day for parameter fitting. For the regional model dataset, the description seems to indicate that it consists of 2 months (May and June 2018, see line 179), although it appears that only a few days (beginning of May) were used to derive the statistics. It would be advisable to use at least 2 months covering different seasons (so instead of May and June, maybe June and December). For the third data set, data from radio sounding, I believe there are plenty of data available since meteorological services such as the MetOffice, launch weather balloons twice a day at several stations (I believe the DWD does the same). In a next step, these results need to be compared to a "current standard method" approximating O4 VCD. This is entirely missing in this manuscript.

I made a quick test using 3 years (2018, 2019 and 2020) of data from 6 stations (some overlap with stations that the authors use) from 12 UTC radiosonde launches (Cambourne [N=1061], Nottingham [N=315], Essen [N=468], Munich [N=1022], Lindenberg [N=1062], Lamont [N=430, 0 UTC for the latter to comply with the SZA requirement]). I find, using Eq.12 and Eq. 13, respectively, a high correlation (0.93, 0.94, 0.92, 0.93, 0.93, 0.93) for estimated and integrated O4 columns (confirming the authors' findings), and, using Eq. 15, low bias and low standard deviation (-1.8±1.0)%, (-1.5±1.1)%, (-1.0±1.3)%, (-1.7±1.1)%, (-1.4±1.2)%, (-0.8±1.7)% (again, confirming the authors' findings). If I do the same but, instead of using the approximation method, use climatology data, I get the following for mean bias and std: (1.6±2.0)%, (0.1±2.1)%, (0.2±2.1)%, (-3.9±1.7)%, (0.8±2.0)%, (-1.2±2.3)% (so 50 -- 100% worse standard deviation, but partly better mean). However, the correlations are certainly much lower (0.45, 0.51, 0.55, 0.65, 0.61, 0.53). I would like to see similar comparisons in the manuscript in order to show that using this method is in fact better than using climatology values [Is it actually worth it? Please show this!]. It would be advisable to include such statistics for at least 20--30 other well distributed weather balloon launching sites for a few years.

Lastly, since the use of this approximation is clearly, as the authors point out, tailored towards parametrized MAX-DOAS inversion methods, it is highly desirable that the authors show some result of those: compare the retrieval results using the "previously standard method" to results obtained using this new method.

The manuscript frequently refers to Wagner et al. 2019 which makes it appear to be rather suited as an extension of that publication than a stand-alone publication. From my point of view, the innovative part of this manuscript is the empirical relation between the effective lapse rate and the surface relative humidity and the fact that this knowledge can be used in the formalism of Wagner et. al 2019 to replace the fixed lapse rate. This part however, I do find sufficiently important for publishing. However, the current format of the manuscript is not good enough. It over complicates things. It does not include sufficient validation data, nor does it include a test on the final product this method is thought to be used for. Hence, I recommend intermediate revisions of this manuscript.

Restructure the manuscript. Streamline and simplify the method section, especially how to arrive at Eq. 9. For validation, extend regional model data covering a larger time period, use a separate day for the parameter fitting for the global model, use many more datasets from radiosondes. Include comparisons with currently used methods (such as using climatologies). Include tests showing that the final product (AOD retrieved from MAX-DOAS measurements) in fact benefits from this new method to estimate O4 VCDs by comparing results ("old" and "new") to complementary measurements.

2. Specific comments

- line 28: vertical profiles of T, P and RH are also needed to calculate the refraction index and hence are anyway needed for radiative transfer simulations, aren't they?
- line 33 -- 39: Here, the reader starts to wonder about the impact of temperature inversions. Later on (line 157), the authors mention temperature inversions and that they are far more frequent during night. Further, they mention that the main application is MAX-DOAS measurements and the presented method is limited to daytime. This should be already mentioned here, otherwise the reader will immediately wonder how temperature inversions are dealt with.
- Section 2.1 seems largely unnecessary:
  line 59 - 60: Incomplete list, and a 1 sentence paragraph: skip this.
  line 61 -- 64: This should be skipped, does not add any new information (?) n_{O4,0)} == n^2_{O2,0} is stated explicitly in Eq. 5
  line 65 -- 68: No new information, all is contained in Table 1. Skip
- Table 2: The g here corresponds to which latitude and which altitude? Since g varies about 0.5% between the poles and the equator, it should be important since the authors claim an accuracy of the same order of magnitude. With the precision for C given in units of K Pa-2 mol2 m-5, the value for C in units of K hPa-2 molecules2 cm-5 equates to 6.73267e39. The authors should consider to give one more digit of precision for the former units, so the unit conversion actually is consistent (a numeric value of 0.01856455 using the former units, which is consistent with the value given, indeed results [to the given precision] in the numeric value given using the latter units. However, as written now, it is not consistent). Is it really necessary to give the same constant in two different units?
- As mentioned above, I recommend to rewrite everything up to Equation 9 and derive this equation directly from the integration of the density assuming either a constant lapse rate (Eq. 9) or 0 lapse rate (Eq.8). However, if the authors really want to stick to their complicated way of showing things (I highly discourage this!!), they should consider the following points:
  line 88 -- 89: Neither of equations 7 or 8 is based on Eq. 6. Equation 7 follows, as the authors write themselves, directly from the equation of hydrostatic equilibrium and the assumption of a constant O2 VMR (see comment below), Eq. 8 is simply the ideal gas law for quantities at the surface. Hence, this sentence should be moved to after line 100.
  Eq. 7: While the authors spoon feed the derivation of Eq. 6, they skip totally the derivation of Eq. 7 from the equation of hydrostatic equilibrium dp = -rho g dz. (e.g.

by Ptop - Pbottom = - rho g h - (Ptop=0)--> -P0 = - rho g h --(rho = n*m/vol)--> P0/g = n m h / vol --(n h / vol == Vair)--> P0/ g/ m = Vair). Apart from assuming that, it is also rightfully assumed that the volume mixing ratio is constant; although implicitly clear by Table 2, it should be added.

line 98 -- 101 (+ Appendix A): All equations should have a number. Please add a number to the equation stated here. Also,reference the derivation of this directly as appendix A2. Regarding equation A3 from the appendix A2: Also here, a derivation of this equation is missing. Maybe start with the equation for polytropic atmosphere p/p0 = (T/ T0)**(gM/R/Gamma) and T = T0 + Gamma * z' and say that Eq. A3 follows from this together with the ideal gas law. I think there is something not going quite right with the signs: The authors define the laps rate to be negative. As such (since g is defined positive in Table 2), -alpha +1 = -gM/R/ Gamma would be (Gamma < 0) a positive quantity. With this, the term in round brackets in Eq. A6 for z' = ∞ would otherwise not disappear. I also wonder if there might be a minus sign missing in Eq. A6 in the last line (or z should be maybe going from 0 to -∞, this solves the problems, I think, please check) .... In line 429, what does "12" refer to? Eq. 12? It's easy to make mistakes, so it is of course also possible that I made a mistake in my checking. Please carefully check the signs anyway.

- line 116 -- 128: The authors refer to a section "in the future". This is very bad practice. Reorganize the structure of the article in order to make it possible for the reader to follow.
- line 117 -- 118: Please explain this statement.
- line 118 -- 119: Please refer to the equation where you show that this statement is true. "effective" laps rate is in fact not defined, what makes the lapse rate an "effective" one?
- line 127 -- 128: Please explain what you mean by this statement. Why "basically"?
- line 130: The authors refer to "the future". Please restructure the manuscript.
- line 163/ line 167: what does "truncation at T639"/ "truncation at T255" mean?
- line 165: The authors are using one and the same day for fitting the parameters and for validation. This has to be changed.
- line 168: How is this a "pre-processed" data set if the authors use the model out put and post-process it?
- line 195: Are the authors following some recommendations with this? If so, please cite. Otherwise, please explain why this choice justifies the statement "only retain the measurements of the highest quality".
- Fig. 1: Can the authors comment on the apparent differences between the models and the GRUAN measurements about the covered parameter space? The author mention the very high values, but do not comment on the many missing intermediate values. The authors use 2 months of WRF data but choose to show only a single day (in a). The authors say that this is in order to keep the figure readable. I suggest to include a second row showing all the data. From statements elsewhere in the manuscript, I have the impression that even the data used for the statistics is not the full 2 months, but only a few days from the beginning of May. Is this correct, if so, why?
- line 217 - 218: Please show this in a plot (see comment above, include all data points).
- Sect. 3.2: The little detail given is really not well structured. I suggest to include a thorough description as an appendix.
- line 219 -- 220: This statement is not quite correct. In terms of covered V, both WRF and GRUAN cover roughly â□□V = 0.7e43 molec^2/cm^5, where WRF covers the space more evenly, GRUAN has a gap of â□□V = 0.2e43 molec^2/cm^5.
- line 225: Where do you show that this statement is correct? There is no direct comparison made. Either include such a comparison or remove this statement.
- line 229: Please show this 0.5% explicitly.
- Figure 2: Why is the histogram only considering 8 days instead of the complete 2 month data set? Or do I understand this incorrectly?
- Figure 3: Maybe comment on great lakes in North America in summer.
- line 238 - 239: You do not show this. Either show it (appendix?) or remove this

comment.

- line 239 - 240: Can you prove this statement? Please show a plot of lapse rate vs. δ_Γ
- line 242: How is this considered (implicitly due to known profiles)? Also include a reference and an equation using the partial pressures of dry air and water vapour which makes the dependence on specific humidity apparent.
- line 243: Please define specific humidity (mass ratio of water vapour content and total mass of air parcel) and relate it to relative humidity (ratio of water vapour pressure and equilibrium water vapour pressure).
- Sect. 4.2.: The authors lack to clearly state the logical chain of causes here: relative humidity affects effective lapse rate. Lapse rate affects V.
- line 250 - 256: Make a plot as Fig. 3 using effective lapse rate calculated from the ECMWF model to show that this statement is correct.
- Sect. 4.3.:
  It is stated that ECMWF data from June, 18th 2018 was used for the fit. Further, the authors state (line 267) that they investigate June, 18th 2018. This should never be done. You cannot use the same day for fitting and verification. Please choose a third day for the fitting and use the same day (June 18 and December 18) only for verification.
- line 264: Please state clearly which figures to compare. Also, it seems that for certain regions (e.g. the Andes, central Europe around lunch time), the absolute value of δ_Γ increased. It might make sense to show a map of the relative improvements of δ_Γ. (? or maybe not...)
- Table 3: Include, in analogy to Eq. 14 and 15, an equation for δ_ECMWF. Include also a histogram for the data in Table 3 in analogy to the histograms in Figs. 2,3,4,6,7,8. Why do the authors not present a correlation plot of V true and V parametrized? I think this is could be very instructive (I made it for the aforementioned stations and it looks very nice).
- line 301: with "radiation shield", you mean a Stevenson screen?
- line 303: What is sufficient?
- line 319 - 320: Please explain what you mean by "and V_{O4}, RH is almost the same for WRF and ground stations".
- line 334: What is sufficiently here? Why do you judge it to be sufficiently?
- Figure 10: How did you choose the points to be plotted? Are the correlation values indicated still using all points?
- line 354: Where does the 3% estimate come from?
- Sect. 5.5: How do you measure S_{O2}?
- line 379: Why "basically"?
- line 386: Why is it sufficient?
- line 389: Inside a Stevenson screen I assume, otherwise the readings might be rather useless.

3. Technical comments

3.1 general

Since one of the co-authors is the chief-executive editor of AMT and the first author is an associate editor, I find it slightly worrying that the authors disregard so many of the AMT guidelines:

- The journal guidelines clearly state that the recommendations of the SI brochure and the IUPAX Green Book (links can be found here: https://www.atmospheric-measurement-techniques.net/submission.html#math) should be followed. This is largely neglected in the axis labels and table headers. Physical quantities and units should not be written as "quantity [unit]" but as "quantity/ unit". Consider SI brochure Sect. 5.4 or alternatively, page 3 of the IUPAC Green Book. Please adjust this throughout the manuscript.

- The journal guidelines clearly state that universal time should be indicated as "UTC". Please correct all "utc" (e.g. Figure 9) in the manuscript.
- The journal guidelines clearly state (https://www.atmospheric-measurement-techniques.net/submission.html#figurestables) that table should be written Table if followed by a number, please correct throughout the manuscript (e.g. line 26, 263, 283, 306).
- The journal guidelines state: "Coordinates need a degree sign and a space when naming the direction (e.g. 30° N, 25° E)". This is not done anywhere, please correct throughout the manuscript (e.g. in Fig. 2, 6, B1 and C1, in Table E1, line 175 no space is included).
- Inconsistent use of section (most of the manuscript) vs sect. (line 279, 217). Journal guidelines say "Sect." unless at the start of the sentence. Please correct.
- Inconsistent use of fig. (e.g. line 177, 197), Fig. (e.g. 216) and Figure (338). Journal guidelines say "Fig." unless at the start of the sentence. Please correct.
- Inconsistent use of equation (e.g. line 360) and eq. (Table 2, line 83, 88,...) Journal guidelines say "Eq." unless at the start of the sentence. Please correct.
- The journal guidelines clearly state not to use hyphens for ranges, but to use en dashes to indicate ranges (https://www.atmospheric-measurement-techniques.net/submission.html#english). However, the authors use sometimes hyphens (e.g. Fig.9 caption). Other times they use "to" which, according to https://www.atmospheric-measurement-techniques.net/submission.html#math is ok. Please change the hyphens to en dashes or "to".
- "data" should be considered plural (https://www.atmospheric-measurement-techniques.net/submission.html#english). Please correct throughout the manuscript. (e.g. 253, 169, 460)

3.2 specific

- line 14: What is absorbed is the light, not the O4, hence "O4 absorption in scattered light" is not a correct formulation. (Maybe add "pattern"?)
- line 15: "light path distributions in the atmosphere" seems also not quite correct.
- line 15: "light path increases" should be "light path length increases"
- line 16: "cloud heights": be more specific: cloud top heights or cloud base heights?
- line 29: "[...] measured profiles [...] do not provide continuous temporal coverage [...]" The profiles do not provide temporal coverage of what? (of itself?) This a somewhat awkward formulation, please reformulate
- Table 1: "Relative deviation between of parametrized and true O4 VCD" --> remove "of"
- Table 2: replace e+39 by "$\times 10^{39}$ "
- line 70 -- 82: Why so wordy? Eq. (1) -- (3) can be summarized in one line of equation without the unnecessary text.
- line 86: add coma: "So far, .."
- line 117: "on first glance" --> "at first glance"

- Eq. 12: inconsistent accuracy of C (c.f. Table 3).
- line 156: add "profiles" after "daytime".
- line 159: "selecting data for ..." wrong preposition.
- line 164: add a coma after "here"
- line 177: Insert coma after "Vertically"
- line 180: Reverse the sentence. Start with "For constraining.... we use..." otherwise it is confusing why you start again with ERA5 data.
- line 180: include "of" after resolution
- line 184: "The selection of SZA < 85°" is not correct. You do not select the SZA, you select data at times of the day at which SZA < 85°. Please reformulate.
- line 185: replace "reach up to a pressure level" by "extend to a pressure of" or similar.
- line 211: "if" --> "of"?
- line 212: Please reformulate this sentence (especially "apply eq. 9 in section...")
- Sect. 4.1. title: add "a" before function.
- Fig. 1: axis labels are too small. Figures are too small, extend to page width. Legend box partly covers line. "mountaineous" --> "mountainous" or better: high altitude. If the authors choose to use only y-tick labels and y labels on the first subplot, the hspace should be 0.
- line 218: add "the" in front of "highest"
- line 219: "matching to a lapse rate" needs reformulation
- line 230: The choice of δ_Γ is not the best, it seems to indicate that the δ is w.r.t Γ while it is w.r.t. V. Please consider renaming. Is it really "parametrized" yet? As I understand, the authors apply Eq. 9 with the constant lapse rate. So I do not see any parametrization here.
- Figure 2: Please repeat the meaning of σ and μ (from line 147) in the figure caption. Swap color bar and histogram, include ticks on the right hand side of the histogram. Please consider putting the coordinates at the axis instead of in the middle of the figure (same for Fig. 4).
- line 235: "Also for ECMWF data..." should really be something like "Considering ECMWF data as the basis for calculating δ_Γ, also results in δ_Γ values close to 0 for the area covering Germany". Please reformulate.
- line 237: "For continents..."needs reformulation.
- line 238: It is advisable to stick to either abbreviations or the symbolic notation, do not mix.
- line 239: insert "the" in front of "same".
- Figure 3: Add year after "18 June" (or reformulate to "the same day").
- Figure 4: Add year after "18 December" (or reformulate, see above).
- line 242: I think it is more correct to refer to the density instead of the weight of air. Please reformulate.
- line 244: Is "compared" really the correct verb to use here?
- line 252: "Subsidence" of what?
- line 257 -- 259: Please reformulate and clearly state that you used Eq. 11 to fit parameters a and b and the result is Eq. 12. As it currently reads, it is hardly comprehensible.
- line 270: "low values are improved"? Reformulate, e.g. "Areas where δ_Γ showed large negative values, show less extreme δ_{RH}".
- line 274: Why "basically"?
- line 276: "weather condition [...] are usually not considered in MAX-DOAS retrieval". This sentence does not make sense, what the authors want to say is that days with such weather conditions are usually not considered for MAX-DOAS retrieval.
- line 296: I think that the 3% is not supported by the plots Fig. 8 (Fig. 7 has to be disregarded because that day was used to fit the parameters). Values of up to ±6% seem more correct, but it is hard to see from the figures.
- line 300: this seems to have the wrong indentation, please check.
- line 315: add a comma after "this"
- line 317: "for" --> "at"

- line 321: Do not start sentences with "But"
- line 320 - 320: insert "order of" between "same" and "magnitude".
- line 320 - 321: This sentence is incomprehensible
- Figure 9: Axis labels too small. "[...] all cycles are referred to the mean value [...]". This sentence does not make sense. Please reformulate.
- Figure 10: Axis labels and Axes tick labels are far too small. Please adjust the range of the x-axis of panel (c). Please check the y-axis label: Should this not be δ_{RH}?
- line 368: refer to Eq. A9 instead of Appendix A.
- Appendix B and Appendix C are never mentioned in the text.
- line 387: remove comma after "measurements".
- Figure D1: axes labels are not readable. "DWD--> " and "WRF -->" seem to indicate a direction as displayed. Maybe remove the arrow, it is misleading.
- line 442: add comma after "maintain".
- line 443 - 444: "This is a consequence of the spatial resolution of the WRF simulations of 1 km not resolving single mountains". This sentence seems incorrect. Reformulate.
- Figure E1: station labels are not very well readable in the figure.