Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-456-RC1, 2020


Atmospheric Measurement Techniques Discussions
Open Access



# *Interactive comment on* "Intercomparison of MAX-DOAS vertical profile retrieval algorithms: studies on field data from the CINDI-2 campaign" *by* Jan-Lukas Tirpitz et al.

**Anonymous Referee #1**

Received and published: 18 February 2020

Tirpitz et al presents trace gas concentration (NO2 & HCHO) and aerosol extinction profiles of 15 participating groups derived from MAX-DOAS measurements and implementing different retrieval algorithms during the CINDI-2 campaign. The authors attempt to validate profiles/partial columns using collocated observations. This is an important effort since there are several retrieval approaches using MAX-DOAS measurements, and even though MAX-DOAS measurements started a while ago still there are not harmonized approaches to retrieve gases and aerosols. Hence, this is an important work and likely suitable for the journal. However, I have major comments and foremost revisions are warranted before publication. In my opinion, the quality of the paper needs to be improved before publication.



Printer-friendly version

Discussion paper

Major Comments

- According with the manuscript the main goal “is to assess their consistency with respect to different conditions and to review strengths and weaknesses of the individual algorithms and techniques” and they use supporting collocated measurements to “validate” the retrieval algorithms. However, authors include primarily results of retrievals using “median dSCDs” obtained in a separate study (Kreher et al., 2019). I do completely understand the value of using the “median dSCDs” but I also see an extreme value in including detailed results using each participant’s dSCDs. The current approach seems quite unusual in a validation point of view. So far, section 3.8 describes briefly results using dSCDs of individual participant but needs to be expanded in the main body, abstract, and conclusions.

- The algorithms are assessed primarily with the root mean square difference. Authors focus primarily on this quantity, which is always positive, and definitively help to understand the comparisons, especially among instruments. However, I highly suggest to include a bias estimator to know the under-or overestimation with respect to the independent measurements. Figure 23 is key in the paper, and I highly suggest to include a similar figure but using bias in percent.

- I find very useful to include the three different type of algorithm approaches (OEM, PAR, and ANA). However, a thorough analysis of what technique yields the best results is missing, especially in the abstract. According with the results, OEM seems to be most appropriate/reliable, but ANA approaches might be ideal for near-real time analysis. I would include a section with main finding regarding the comparison of these methods.

- For the groups using OEM, they use same dSCDs and main retrieval parameters are prescribed, still there are extremely large differences among the groups using OEM. A thorough analysis of the reason is missing. Additionally, If I understand correctly, the recommended altitude grid for all participants was from the surface to 4km (20 layers of 200 m). This is quite unusual in transfer models, how is the atmosphere represented above 4km? If this is fact true I highly recommend having realistic information above 4km. Furthermore, I am surprise that for the retrieval settings all participants use average values of pressure, temperature, and O3 vertical profiles obtained in 2013-2015. However, the campaign was held in 2016. I believe pressure, temperature, water vapor, etc, might have an important effect in the forward model and foremost in the retrieval of aerosol extinction using O4. I do not understand why radiosondes (or even re-analysis) data obtained during the campaign are not used. If the goal is to validate profiles I highly suggest using the real atmospheric conditions during the campaign.

- It is well-known that sensitivity needs to be considered when comparing different measurement techniques. However, after reading the manuscript it sounds like you introduce new findings, e.g., last short paragraph in the abstract. I do not think it is assumed that integrated extinction profiles from MAX-DOAS and the AOD from the sun photometer should be comparable. In my opinion, this is not a finding or result in this paper. I suggest to re-write your findings accordingly, e.g, include that after smoothing (applying the “AOD correction”) comparisons yield better results. . . in fact, I think authors should describe that this correction (partial OAD correction) is related to the O4 scaling factor used in past studies (and here too for some groups). If I understand correctly, the “AOD correction” yields better results/comparisons because sensitivity in mainly in the lower troposphere, hence aerosol layers aloft are not captured with MAX-DOAS. In this context, after reading Ortega et al. (2016) this reference is not pointed out but offers some insights and should be included.

- It is mentioned that “The ceilometer aerosol extinction profiles should be consulted for qualitative comparison only” and I fully agree due that many assumptions are used to calculate extinction from backscatter measurements. In this context, the aerosol extinction derived from the ceilometer cannot be used to validate the profiles. However, I do believe they offer you additional information that can be further used, especially for OEM. In the manuscript, a priori extinction profiles for both aerosol and trace gas retrievals were exponentially-decreasing and of course OEM will converge, i.e., it is an ill-posed problem. However, if you use the aerosol extinction profiles as an a priori at least you estimate a better profile shape and the OEM technique might give you a better result. I highly recommend to use the ceilometer extinction profiles as a priori profiles and compare with the exponentially decrease profile. Several questions might arise: do sensitivity increase at higher layers? do AKs change? is the partial AOD correction still the same?

- Lastly, I do not agree that retrievals of NO2 in the UV and vis should give you same results, unless you proof homogeneity around the line of sight.

Specific Comments

P2, L1-6. This paragraph does not belong here, I suggest to move it to the introduction and expand the abstract based on major comments.

P2, L2. Change “boundary layer and the lower troposphere” with “lower troposphere”

P2, L3. Change “radiation” with “absorption”

P2, L5. I would explicitly say that you retrieve aerosol extinction concentration for profiles.

P2, L10. Include all the supporting observations and remove others in the parenthesis.

P2, L15. Do you mean magnitude instead of intensity?

P2, L15-20. Results are shown in root mean square, however, in order to have a more quantitative description please also include the bias in percentage, or the rmsd in percent. Otherwise, it is hard to interpret the magnitude of the differences.

P2, L21-23. It is well-known that different sensitivity needs to be considered when comparing different measurement techniques. I do not think it is assumed that integrated extinction profiles from MAX-DOAS and the AOD from the sun photometer should be comparable. In my opinion, this is not a finding or result in this paper. There is nothing new on this short paragraph. I suggest to remove this paragraph or re-write your findings accordingly, e.g, include that after smoothing (applying the AOD correction) comparisons yield better results due that similar air masses are compared.

P2, L26-28. Transport is missing in your description of chemical composition in the PBL.

P3, L5. I agree that MAX-DOAS is a well-established technique with information of absorption signature of trace gases. However, it is misleading because the whole point of these type of studies is that MAX-DOAS is NOT a well-established technique to measure accurately gas concentration.

P3, L6, It is mentioned that MAX-DOAS infers information in the boundary layer and free troposphere. Please include some references for both cases.

P3, L8. I would remove "from the top of the atmosphere (TOA) to the instrument"

P3, L10. Change "Detectable gases are nitrogen dioxide (NO2), formaldehyde (HCHO). . ." with "Gases that have been analyzed in the UV and visible spectral range are nitrogen dioxide (NO2), formaldehyde (HCHO). . ."

P3, L18. Change "radiative transport models" with "radiative transfer models".

P3, L19. Change "such" with "of"

P3, L23. What do you mean by different conditions?... Weather conditions, pollution conditions?

P3, L30. Again, add all supporting instruments and remove "others ". Otherwise, remove "others".

P5, L16. Mention shortly what other effects, otherwise remove this.

P5, L27. I do not see see how Apituley et al fits in this study.

P5, L28 – P6, L9. As mentioned above, I see the value of using the “median dSCDs”, but I strongly suggest to include in detail (and not in the supplement) the retrieval results using their own dSCDs. In fact, I recommend the “median dSCDs” to be included in the supplement if authors believe the manuscript will be lengthily.

P6, L22. How is water vapor profile included in the forward model? is it important? Also, remove the dots after aerosol microphysical properties.

P6, L25. What is p?. Also, I’m surprise to see 4 DOF, for what gas?, is there a referene?

P7, L3. The short OEM description seems awkward. Remove “filling”. In general, you have an ill-posed problem and the solution is constrained by an a priori state vector.

P7, L7. It is mentioned that PAR require more memory, and the sentence sounds like this is a limitation. How much memory is needed for such a short campaign? Satellites use look up tables.

P7, L13. “The M3 algorithm by LMU appears as an additional algorithm in our study” looks awkward. What do you mean? Re-write this sentence. Why its description is included in the supplement?

P7, L25. As mentioned in the general comments. I highly suggest using real atmospheric conditions instead of average PTW from other years.

P7, L27. See my comment above regarding the altitude grid, it is not clear what was used above 4km.

P7, L33. My understanding is that the AERONET angstrom exponent (440-675 nm) derived from a single day (14 Sep) is used to extrapolate to 360nm for all days during CINDI-2, is this correct? If this is correct, please explain why you use a single day and not coincident measurements. I expect the angstrom exponent changing unless you have similar aerosol composition.

P8, L25. Remove the “. . .” in the sentence in parenthesis. Check many other sentences like this along the manuscript.

P9, L11. Change “true aerosol extinction” with “aerosol extinction”. Many assumptions are carried out for the creation of extinction profiles and might not be the true aerosol extinction.

P9, L23. What mean error does the 0.03 RMSD represent?

P9, L25. At the end of section 2.2.2 it is pointed out that “the ceilometer aerosol extinction profiles should be consulted for qualitative comparison only”, which I fully agree since many assumptions are carried out to derive extinction profiles. In this case, the retrieval of extinction profiles cannot be fully validated during CINDI-2.

P9, L25. It is mentioned that NO2 profiles from sondes and lidars were carried out sporadically, but include a description of how often. How many sondes were launched?

P12, L15. For the “different observations” do you mean MAX-DOAS and supporting measurements?, or different groups using MAX-DOAS?. Please clarify.

P12, L18. IS xref,t measurement from a reference measurement?, i.e., collocated supporting observation?. Clarify.

P12. While the root mean square difference is useful, this is always positive. I highly recommend to include a bias to see the sign of bias with respect to collocated observations. Simply, use something like this: bias = median(max-doas-reference)/reference when comparing to collocated supporting observations.

P13, L12. It is mentioned that UV and Vis dSCDs should be the same. I disagree, light path in the UV and Vis might be different. Hence, different dSCDs.

P14, Section 2.3.3. I believe you can quantify the spatial mismatch between sonde-MAXDOAS by using the sonde gps information. It might be interesting to see the actual spatial difference.

Section 3.1.

P15, L12. "Figure 2 visualizes the average AVK matrices"... what do you mean by average AVK?. Are these averages of a single group using OE, or average of all groups?

P15, L13. I agree with this "Note, that the AVKs do not necessarily represent the real/ total sensitivity and information content of MAX-DOAS observations as they only consider the gain of information with respect to the a priori knowledge" and I think some literature is missing, e.g., Friess et al (2006) showed that aerosol extinction above 3km can be retrieved using O4 dSCDs measured at different wavelengths. Ortega el al (2016) showed that elevated aerosol layers modify O4 dSCDs, hence some sensitivity of aerosols aloft. In my opinion, this is a clear effect of an ill-posed issue, where an appropriate a priori information is important. In this case, I do not agree with authors claiming that there is not sensitivity of layers aloft, but it is difficult to retrieve layers aloft due to assumptions and less-ideal a priori information.

P15, L27. It is mentioned that "the presence of clouds can increase the sensitivity to higher layers due to multiple scattering and thus light path enhancement in the clouds". If clouds can enhance the sensitivity at higher altitudes, aerosols might have a similar effect, correct?

P28, L3, I would add if a priori information is not reliable at the end of this sentence: "high-altitude abundances of trace gases and aerosol typically cannot be reliably detected by ground- based MAX-DOAS observations "

P28, L11. If I understand correctly, in addition to the description provided, the ratio from equation 11 provides you the fraction of the aerosol retrieved by OEM. So, a factor of 0.8 means that about 20% extinction should be aloft, is my interpretation correct? If so, I think this is a very important result and should be further explained. Furthermore, could this fraction be related with the correction factor?

P29, L3. It is mentioned that "a scaling of the measured O4 dSCDs prior to the retrieval with SF $\approx f_{\mathcal{T}}$ might be used to at least partly account for the PAC for MAPA and probably other PAR and ANA algorithms (see Supplement S3), even though the physical reason for PAC and SF are different.", please explain further and provide the physical differences between PAC and SF. Would it be possible that past correction factors are used due that they miss aerosols aloft, which if I understand correctly might be in agreement with findings in Ortega et el. (2016)?.

P29, L9. "underprivileged" sounds weird, please change it.

P30. figure 14. Please add bias in % (negative/positive) as mentioned above. Additionally, light vs opaque are not distinguishable, maybe using other colors might help? Furthermore, symbols on the two right column plots are not shown in the legend, maybe you meant to use the same symbols?

P30, L10. As suggested above, please include the bias in percent here, in addition to the rmsd.

P31, L1-2. In the text, it would be handy to describe the group (as in Figure 14) and in parenthesis the approach/name) in order to avoid going to table 2 every time. For example, PRIAM is mentioned in line 2 but this is not in figure 14 and table 2 needs to be checked.

P31, L7. KNMI/ MARK and NASA/ Realtime are mentioned as high rmsd, but I also see MPIC being high but not included in the text. So, all parameterization approaches show high rmsd.

P31, L9-12. It seems like the correction factor improves the agreement, but further description is missing. According with your "partial AOT correction" this might be due that PAR approaches miss layers aloft?. I consider this an important finding but is not described.

P35, Section 3.7. I do not agree that NO2 Vis and UV should yield similar results, unless you show with independent measurements that there is homogeneity in the sensitivity range (vertical/horizontal). Rather than an "intrinsic consistency check" I would use this section to actually assess inhomogeneity. On the other hand, the manuscript is long enough and I would consider removing this section.

P37, Section 3.8. This section is important and deserves more description. A bunch of figures have been thrown in in Supplement S10 but not a complete description. In my opinion, this is a key section to show how reliable are the MAX-DOAS products, hence I also recommend a thorough description of the bias per participant, and not only rmsd.

P38, L11. Please include the approaches. Some people only read conclusions. Profiles are not really assessed, especially for trace gases. I recommend to explicitly describe that lower tropospheric columns are assessed. Figure 23. It is difficult to track what algorithm is used for each group. I suggest to include the algorithm next to the group name, maybe in parenthesis. I suggest to include another figure, similar as Fig. 23, but for the bias in percent.

P40, L20. It is mentioned that "O4 scaling and PAC were found to have similar impact on the MAX-DOAS AOT results." In my opinion, this is a major finding. It is shown that sensitivity needs to be considered when comparing two different remote sensing techniques, and here you have shown that the lower tropospheric column of extinction agrees well with Total column of AERONET when "corrected". This "PAC" is the same as the O4 scaling factor and by reading Ortega et al. (2016) might be due that aerosol layers aloft are normally neglected. I highly recommend to further describe this.

U. Friess, P.S. Monks, J.J. Remedios, A. Rozanov, R. Sinreich, T. Wagner, U. PlattMAX-DOAS O4 measurements: a new technique to derive information on atmospheric aerosols: 2. Modeling studies, J Geophys Res (Atmos), 111 (2006), p. D14203, 10.1029/2005JD006618

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-456, 2020.