Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2016-352-RC1, 2016

Atmospheric Measurement Techniques Discussions

Interactive comment on "An exploratory study on the aerosol height retrieval from OMI measurements of the 477 nm O_2–O_2 spectral band, using a Neural Network approach" *by* Julien Chimot et al.

Anonymous Referee #1

Received and published: 16 December 2016

Review of An exploratory study on the aerosol height retrieval from OMI measurements of the 477 nm O2-O2 spectral band, using a neural network approach

The authors present a study exploring the retrieval of an aerosol layer height parameter from OMI measurements of the O2-O2 band in the visual wavelength range. They follow basically the same approach as the OMI O2-O2 cloud algorithm but use a neural network to replace the traditional look-up table. Sensitivities of the retrieval to assumptions related to the aerosol optical properties and the surface albedo are investigated. The algorithm setup has been applied to three years of OMI data and seasonal aver-

ages are compared with MODIS aerosol optical thickness and climatological aerosol heights from the LIVAS CALIPSO climatology.

In my opinion, the paper presents interesting and substantial work that in principle warrants publication in AMT. However, I feel that there are a number of major issues that should be addressed before I can commend publication. Major and minor comments are listed below. I would appreciate if these are addressed point-by-point and a diff-version of the manuscript is provided together with the replies. In addition, I struggled with the style and the structuring of the manuscript, which made it hard at times for me to quickly understand what the authors are trying to say. I give examples of this below. I realise that this is partly a matter of personal taste, but I strongly encourage the authors to use these comments to critically look at the entire manuscript. It would help the reader to better understand and appreciate the authors' valuable work.

Major comments

-A look-up table is generated and used to train a neural network, which is in turn used in the retrieval instead of the original look-up table. This lowers memory demands and increases computational speed. In my opinion this is a very good application of a neural network. However, in your extensive general discussion of neural networks you suggest that you use a neural network to resolve the ill-conditioning of the inverse problem. For example, p.7,l.30-p.8,l.2., particularly in p.8,l.16 and again in the conclusion on p.21,l.6. This is clearly not what you do. Your neural network is trained with simulated data, so all the physical relations between input and output contained in your neural network are already explicitly described in the RT model, and strong assumptions on the aerosol model, the vertical distribution of the aerosols and the ground surface are still a priori inputs for your retrieval. I think the text should be changed to really avoid giving the reader this wrong impression. Also, I don't quite see why interpolation in a look-up table should be less accurate than 'interpolation' in a look-up table with a neural network even in the case of non-linear data.

-An important figure is Figure 16. You are showing the seasonal dependence of ALH for your OMI retrievals and as derived from many years of CALIPSO retrievals. The amplitude of the seasonal variability in the LIVAS ALHs is only about 0.5 km and I am wondering whether that seasonal variability is picked up by your OMI retrievals. After all, biases due to wrong assumptions on the aerosol model are of the same order of magnitude or even larger. What makes Figure 16 a bit deceptive in my view is that you plot four lines showing OMI ALP retrievals for slightly different settings. These lines are obviously dependent and therefore correlate. At first sight, the overall consistency between the collection of plot lines suggested to me that you are indeed picking up the seasonal variability. But when having a closer look, I am not sure... the apparent consistency may be visually driven by these OMI ALP lines. In practice, you would pick some optimal retrieval setting based on sensitivity studies, literature etc. and do the retrieval, or you would perhaps run the retrieval for several retrieval settings and then take the average. I really think that for a proper comparison of LIVAS ALHs with your OMI ALHs, a figure like Figure 16 should therefore basically contain one panel with only two lines: one for LIVAS ALP and one for your best OMI ALP. Can you change Figure 16 (the effect of temperature corrections is not a key point of your paper and the effect of surface albedo and SSA could then be moved to a separate plot)? I fear that the agreement between OMI ALH and LIVAS ALH will look less convincing with only four pairs of points, but it is more realistic.

-In your analysis you assume that climatological CALIPSO profiles are the truth and you focus then on biases. I am willing to follow your assumption for the moment but I think presenting the story like this is too optimistic on the error in retrieved ALH for an individual pixel (which is what you would need for scatter corrections). (Again, I am making a point of this because you repeat bias values in the abstract, p.1.,l.9-10.) In the sensitivity analysis you identify several error sources that affect retrieved ALP (aerosol optical properties, surface albedo, neural network implementation). Together they form something more like a random error. When calculating biases as you do these random errors are averaged out. In my view, you should also report the root-mean-square

and the standard deviation of the differences (as in: rms**2 = bias**2 + std**2). This gives a more complete estimate and breakdown of the ALH error as derived from the evaluation with LIVAS ALHs. In addition, a scatter plot of OMI ALH vs LIVAS ALH would definitely help the reader a lot here. As a final note, I would need to do some more thinking on the implication of assuming climatological (i.e. averaged values) for your analysis. You make a strong point of ALH retrieval to support scatter corrections in trace gas retrievals. What would be needed then are ALH retrievals that capture exceptional events not represented by climatological averages, right? See point below.

-It is a pity that you only show only seasonal three-year averages. The temporal variability in such an average is small (discussed in a previous point). Figure 14 suggest that the std in ALP across pixels can easily be 200 hPa or higher. It would help if you could provide ALP maps for some selected scenes that show that within a particular scene there is spatial consistency but that the variability emerges when comparing different scenes. Perhaps together with some MODIS RGB images or other info..? I have to admit that I am not convinced at this stage that your ALP retrievals have indeed sufficient sensitivity to show geophysical variability - then again, it is an exploratory study. Such an addition would really strengthen the paper, but I realise that this requires substantial work. This is a strong recommendation but I will eventually leave it to the authors to decide whether or not they follow it up.

-In the introduction you quite extensively discuss the uncertainty in DOAS trace gas retrievals due to aerosols. This is fine with me. But in the abstract you clearly state that the 'main motivation of this study is to evaluate the possibility of retrieving ALH for potential future improvements of trace gas retrievals' (p.1.,l.4-5; repeated at the end of the intro and in conclusion). This is not what you do. Either statements concerning this claim should be softened (throughout) or a thorough and critical discussion of your results from the perspective of trace gas retrievals should be provided. The reason I make a point of this is the following: Figure 16 suggests that at least your seasonal average ALH is probably something like a boundary layer top height that perhaps follows the expected seasonal dependence. But for trace gas corrections you need ALHs on a pixel level. Given the large uncertainties on individual retrievals I doubt whether the OMI ALH will then be useful: 0.9 km +/- 1 km in winter and 1.4 km +/- 1 km in summer does not appear to me as a tighter constraint for trace gas retrievals than simply assuming climatological boundary layer heights or something similar. Of course, there is Figure 14 showing that individual ALHs show quite some variability, but the question is whether the ALH variability for AOT below about one is really geophysical variability or indicates a large retrieval error. And of course, you also retrieve AOT itself, which may be useful for the trace gas retrievals, but also the AOT seems to be very sensitive to the aerosol model you assume.

-Finally, you state in the conclusion that accurate knowledge of AOT is needed for a good ALH retrieval. I am not yet convinced, because I think the comparison with LIVAS ALH is not conclusive in this respect (see comment above). The simulations however do seem to point in this direction (Figure 6). However, also MODIS AOT has an associated error which is well documented in the literature but not discussed at all. If you want to retain the conclusion that a retrieval setup with external AOT input is the way forward, then a discussion of the uncertainty in MODIS AOT should definitely be provided and also a small test on the sensitivity of your retrieved ALH to this AOT uncertainty should be added to section 4 (can be done fairly quickly I think).

Minor comments:

-The use of a neural network trained by a look-up table has been done before. I am aware of the ROCINN cloud algorithm, but I guess there must be more references. Please add some references.

-Can you provide quantitative estimates of the increase in speed and the reduced memory needs compared to the original look-up table?

-p.2,l.1: source (typo)

-p.2,l.5: clouds –> cloud

-p.2,l.5: Throughout the paper you often talk about fine particles, and it is not clear to me what you mean. Do you mean fine-mode particles? But also coarse-mode particles can act as CCNs? (throughout paper)

-p.2,l.8: contributes to –> contributes

-p.2,l.18: vertical distribution –> the vertical distribution

-p.2,l.19: affects the computation of Air Mass Factor –> affects trace gas Air Mass Factors

-p.2,l.22-23: uncertainties in the computed tropospheric NO2 AMF for OMI are the dominant source of errors –> you mean: uncertainties associated with aerosols?

-p.2,l.27: depending if –> depending on whether

-p.2,l.30: PBL –> throughout paper abbreviations are introduced several times (should be only first time), or abbreviations are introduced that are not further used (distracting); please check paper

-p.2,l.35: O2-O2 spectral band –> O2-O2 absorption band?

-p.3,l.1: what are effective cloud parameters?

-p.3,l.26-27: The 477 nm ... individual band. –> How many DFS does the O4 band add compared to continuum reflectances according to this study? That’s the interesting number here.

-p.4,l.5: area –> areas

-p.4,l.7: emphasize –> emphasis

-p.4,l.11: the North –> North

-p.4,l.3: expectation –> benefit?

-p.4,l.19-20: during daylight –> can be left out

-p.4,l.20: two dimensional –> two-dimensional

-p.4,l.24: usually –> but is this also how refl. is defined in this paper?

-p.4,l.27-30: explanation of row numbers, row-anomaly and its progression can be completely left out, because it is not relevant; just say that all OMI data used in this study are from before the row-anomaly

-p.5,l.1-20: Discussion of OMAERUV out of place; should be completely moved to introduction. But why discuss it so extensively? Are you using the data in your study?

-p.5,l.25: The initial purpose of this [sic] algorithm... –> You should describe your algorithm here. Later I understood that you use slant columns and continuum reflectances from the cloud product, but that is not the algorithm that you use for the sensitivity studies, right? Just describe what you are doing and say that you follow the same approach as the OMI cloud product.

-p.5,l.22: extraction –> fit?

-p.5,l.23: fine spectral features –> add: due to absorption

-p.5,l.24: 460-490 nm –> is this the fit window?

-p.5,l.5: initial purpose –> what is the main purpose then? what is the purpose now?

-p.5,eq.2: Please give references for xsecs. I guess you are fitting the square of the oxygen slant column: I only know of O4 xsec reference spectra that have the equilibrium constant between [O2] and [O4] included. Why define reflectance explicitly in Eq. 1 when R in Eq. 2 is basically a sun-normalized radiance? I know that I am being nitpicky here (pi/mu_o disappears in the polynomial) but I prefer not to read unnecessary info.

-p.6,l.10-15: This entire para can be left out. You mention MAX-DOAS measurements,

the Ring effect, radiance measurements without further explanation so this doesn't add anything and only distracts the reader.

-p.6,l.18: Eq. (3) and Eq. (4) –> eqs 2 and 3 ! Please check entire manuscript on references to equations and figures in text (preferably before submission): there are more examples of wrong references, for example on p.14 and p.17

-p.6,l.20: what do mean with homogeneous and finite? An infinite(ly thick?) layer doesn't seem an option.

-p.6-7,sect.2.3: I have difficulty with this section and with figure 1. I think it is good to briefly describe the effect of aerosol parameters on slant columns and continuum reflectances. But the text and the figure are confusing. First, you are making forward references a couple of times. Please discus all relevant effects here, including effects of aerosol optical properties, which I think is important. Second, I cannot quite follow your use of the terms shielding and enhancement (see for example p.6,l.30-31, which I just don't understand) but this may be because I am not too familiar with these terms. Third, you are discussing slant column and continuum reflectance as a function of AOT, ALH, and aerosol optical properties. But these dependencies are difficult to recognize in figure 1 because you put the independent variable into the color map (and sometimes you again don't). Can you please make an alternative figure 1 that clearly illustrates the observations in the text.

-p.7,l.8: 2 –> two. Please check the AMT manuscript preparation guidelines before submitting, not only here, but throughout the manuscript (I also read 'Tab. 2'). And check for typos. I see various typos which I have no time to correct. I have seen three different ways of describing ranges ('2005-2007', '[260:800]', '[0.03-0.05-0.07]'). Also I read in several captions '(2005, 2006,2007)'. In my pdf, the greek symbol for tau in the plots is not recognizable as tau!

-p.7,l.23: from –> on

-p.8,l.8: Networks –> networks

-p.8,l.12-13: Don't agree: The dependence of slant columns and continuum reflectances on aerosol parameters can be accurately simulated with an RT model.

-p.8,l.17,l.22: use non-linear consistently (not: non linear) throughout manuscript

-p.9,l.14: The MLP ... data set –> don't understand this sentence

-p.9,l.30: referent –> reference (throughout manuscript)

-p.10,l.1: you have been referring several times now to average light path distributions, maybe you should define it because I don't quite know how exactly ALH affects this distribution and what this means for your retrieval

-p.10,l.2-3: The second ... quality. –> I am very confused: Are you following a two step approach for the NN_Rc_Ns retrieval? Fit AOT first which you then use as a fixed input for the ALP retrieval?

-p.10,l.15: what is the geometric thickness of your aerosol layer?

-p.11,l.1: most relevant –> optimal?

-p.11,l.4: positive-definite –> What does this mean? Your error function is just a plain sum-of-squared-residuals, right? Nothing special here?

-p.11,eq.7: why factor 1/2?

-p.11,l.18: you mean the last 15 iteration of 70% * 460000 iterations in total? What if overfitting is already present before the last 15 points?

-p.11,l.22: best evaluation score –> you mean for the evaluation set?

-p.11,l.28: The ALP retrieval scores are significantly larger. –> But this makes sense because pressures are two orders of magnitude larger than AOTs? So ALP performs even better (by a factor of hundred)?

-p.12,l.28-30: An overestimation ... surface. –> I am confused: a higher asymmetry parameter means more forward scattering (not: backward)...?

-p.13,l.3-17: I find it confusing that the way of analysing and presenting results is different when discussing the sensitivity to the error in the surface albedo. I would have preferred a plot in Figure 5 similar to Figures 4 and 6. Also, I don't understand the advantage of a box-whisker plot here: a single box-whisker is based on only three data points, right (three angles)? In addition, I would definitely like to know the direction of effects, particularly the effect of over- and underestimations of the surface albedo on ALP. If you want to retain the box-whisker plot (not preferred by me), then at least mention direction of effects in the text.

-p.13,l.19: You haven't discussed figures 5b and 5c yet!

-p.13,l.20-21: For scenes ... 100 hPa. –> But in this AOT range I see biases up to 300 hPa?

-p.13,l.25: entire learning database –> you mean: the independent data set?

-p.13,fig.6: Why are there large biases when there are no forward model errors (no error in SSA or g)? These are closed-loop, noise-free simulations so the truth should be retrieved?

-p.14,l.3: on –> in

-p.14,l.8-16: I don't really see a systematic improvement in ALP retrievals when using a priori AOT in case the aerosol model has a bias (for example, when g=0.8 yes but when g=0.6 no).

-p.14,l.18: eq. 11? Please check all figure ad equation numbers, there are more wrong references.

-p.14,l.22-26: This is an example of repetition; I would leave sentences out.

-p.14,l.26-28: Where do these SC precision values come from? Give ref. Are these values typical for OMI instrument noise? But then for each scenario there is only one typical SC error as it only depends on the radiance and no range should be considered, right? How can a random error in SCs lead to ALP biases? Why don't you take into account the noise error in your other input parameter? Your remark on the temperature correction seems out of place.

-p.15,l.4: Can you give some more details here about the area that you selected (latitude and longitude? How many pixels passed selection? How many ALP data points went into the seasonal averages.

-p.15,l.19: Applying ... scenes. –> repetition

-p.15,l.21-22: However... as well. –> then why do you apply the OMI cloud fraction threshold...?

-p.15,l.25-: I find your discussion of the temperature correction and later sensitivity tests completely out-of-place (particularly as you show results of it in the final most important figure). Also, I don't understand how you do the temperature correction. Is Ns_O4-meas the slant column you get from the OMI cloud product? How do you use NCEP temperature profiles to calculate gamma? This factor clearly depends on the aerosol conditions, which you are trying to retrieve? I guess the physics of this temperature correction is well understood, apparently not part of the OMI cloud product and Veefkind et al. 2016 describe that users should apply it. So applying the temperature correction is always better and doesn't present a source of uncertainty for your ALH retrieval.

-p.15,l.8: geophysical location –> grid box?

-p.15,l.15: Note that ... –> So what does this imply?

-p.15-16,sect.5.2: I stopped reading this section in detail. A few general remarks: There are several forward references to figures hat have not yet been explained. Some figure numbers are incorrect. When investigating the effect of the surface albedo input

on the retrieved AOT, I don't see the reason why you should split the analysis per season (perhaps personal taste). Are MODIS AOT and MODIS BSA truly independent?

-p.17,l.27: polarization –> Your RT calculations for the LUT excluded polarization effects?

-p.17,l.31-32: MODIS tau ... OMI tau. –> Can you provide refs?

-p.18,l.4-9: You discuss the variability in your set of ALH pixels, but what does this tell you? Isn't it obvious that the variability decreases when you use the setup with fixed AOT: there is one free parameter less? This doesn't show that those retrievals are more accurate. Isn't it obvious that the variability decreases with higher AOT as the aerosol signal is more stronger...? But there is an interfering effect of the geophysical variability! Perhaps low-AOT aerosols are typically closer to the ground (BL pollution) while high-AOT aerosols are extreme events that reach into the free troposphere? Don't know.

-p.18,l.20: Spatial average ... performed. –> How do you compute ALH from extinction profiles? How do you compute the spatial average? Do you take into account the sampling of your OMI pixels? I mean, if for a given location you have, say, twice as many OMI pixels in summer than in winter, do you take this into account when calculating the LIVAS average. Please at least explain exactly what you did.

-p.20,l.14-15: You state that the NN with AOT as a fixed input performs better. But comparing figure 16b and 16d, this holds for the red and purple lines (SSA 0.95), but the opposite is true for the green line (SSA 0.9) and the blue line is undecided. Yet, previously you argued based on the comparison of AOTS that an SSA of 0.90 is probably the more accurate value in three seasons. So I think this conclusion cannot yet be drawn.

-p.29,fig.1: What is azimuth difference (throughout)? Does H2O absorb in your fit window? Then why isn't it included in the DOAS fit? At what wvl should I interpret TAU?

-p.30,table2: I don't understand this figure. For example, the ALP test error for a given NN configuration is the sum-of-squared differences (eq. 7) for all the scenarios in the test set (the 15%). This is one number. But then you repeat the training three times. So you have three numbers. Did you calculate box-whiskers from only three points? That doesn't seem right.

-p.33,fig.4: In left panel, can you add also SSA = 1.0 simulation? You have tested positive and negative errors in the asymmetry parameter as well? Can you make axes in the right panel such that the 1-1 line corresponds to the diagonal?

-p.39,fig.10: I see many more lines in the plot than in the legend? Can you make the plot line colors the same for the same surface albedos in the top and bottom row? I don't think the differences between the OMLER and MODIS BSA are significant, do you?

-p.40-41,fig.12-13: It took me some time to realise that -for unclear reasons- these figures are split whereas they should be merged? The color map labels are not visible? The tau symbol is not a tau symbol?

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2016-352, 2016.