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

Julien Chimot *et al.*

J.J.Chimot@tudelft.nl

We thank Referee #1 for his / her very comprehensive and valuable comments. They give us the opportunity to solidify our different interpretations based on our results and to improve our manuscript. Below we address them one by one (Referee #1 comments in blue, author and co-authors in black).

Review of 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

The authors present a study exploring the retrieval of an aerosol layer height parameter from OMI measurements of the O_2-O_2 band in the visual wavelength range. They follow basically the same approach as the OMI O_2-O_2 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 averages 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 diffversion 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.

We took into account all the comments and questions asked by Referee #1 below. We reformulated where necessary according to the remarks and question, in order to ensure a better clarification. More details on these reformulations are given below where appropriate.

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.

We thank the reviewer for this comment! We verified what we wrote in these mentioned pages. And we only found the reference to "ill-conditioning of the inverse problem" on p.7, l.30 "The Multilayer Perceptron (MLP) neural networks have been widely used and acknowledged for decades in the field of remote sensing (Atkinson and Tatnall, 1997). Indeed, most retrieval problems in this field are ill-posed and nonlinear". On p.8,l.16 and p.21,l.6 we do not address the ill-conditioning problems through the NNs but compare the advantages of Neural networks (NN) with respect (w.r.t.) Optimal Estimation Method (OEM) in terms of high speed processing and optimized interpolation techniques:

- "The choice of a NN approach relies on their advantages compared to more conventional methods such as linear regression, linear interpolation in a LUT or the Optimal Estimation Method (OEM). In particular, it enables 1) very fast computations with modern computers in spite of the number of required parameters, 2) optimized interpolation technique even in case of non linear problems as the NNs are better able to reproduce the curvature between the LUT nodes therefore lead to lower systematic biases compared to a linear interpolation, 3) reduced memory use compared to a LUT with a very high sampling" (*cf.* p.8, l.14-19).
- "In spite of the high computing time due to the learning database creation and the training of these algorithms, very fast operational processing is allowed. Such processing is much faster than approaches relying on the Optimal Estimation Method and employs more optimized interpolation techniques than a classical linear interpolation within a LUT" (*cf.* p.21, l.6).

Based on the many references available and cited here, these aspects are correct.
We do not mean that the use of a NN, like we did, explicitly resolve the ill-conditioning aspect of the inverse problem: "The Multilayer Perceptron (MLP) neural networks have been widely used and acknowledged for decades in the field of remote sensing (Atkinson and Tatnall, 1997). Indeed, most retrieval problems in this field are ill-posed and nonlinear, in particular the atmospheric ones. Thus, the associated inverse problems can only be addressed by including a priori information and relying on statistical analysis. Since aerosol retrieval from passive spectral measurements is well-known as a non-linear inverse problem, the MLP technique represents then a powerful approach to design a retrieval algorithm in a fast and robust way." (*cf.* from p.7, l.29 to p.8, l.2). However, at a high level, similarly to the OEM approach, it deals (or can deal) with this aspect by relying on prior information, either explicitly (*i.e.* prior input parameter given to the retrieval algorithm) or implicitly by assuming some assumptions: typically the choice of our NNs rely on our assumed Single Scattering Albedo (SSA) value, either 0.9 or 0.95. Nevertheless, the way of how to deal with all these prior information differs depending on the numerical approach which is employed: For example, in the OEM, prior parameters are generally given with associated prior errors while these errors are not used by NN or Look-Up-Table (LUT) approaches.

To be sure to avoid any confusion, we clarified this issue in our manuscript.

A conventional interpolation within a LUT is strongly sensitive to: 1) the interpolation technique (*e.g.* linear interpolation, spline cubic interpolation), 2) the resolution or sampling of your LUT. Regarding 1), there is in general very little arguments to support the interpolation technique selection. In general, the linear interpolation is likely the fastest interpolation technique and least dependent on (arbitrary) specifications such as which dimensions are interpolated first and which last. Regarding point 2), for a given problem, a very coarse LUT can lead to substantial numerical biases independent of your

interpolation technique, like with the previous version of the OMI cloud LUT [Chimot *et al.*, 2016; Veefkind *et al.*, 2016]. In principle, we would say that an infinite high resolution LUT could do as good as a Neural Network. But, such a LUT is impractical. If 1) the training database includes a high enough number of cases to learn (here OMI O_2-O_2 and aerosol spectra simulations), 2) it is representative enough of the most likely cases to be encountered by the retrieval system, 3) the number of cases is greater than the number of degrees of freedom of the problem and the variability characterized by the combination of all the cases, then a NN is expected to be more accurate than a coarse (or too limited resolution) LUT.

Inverse problems with a LUT should, in general, demonstrate that the resolution specified in all the dimensions (and the combined dimensions altogether) is high enough for a given problem. This is, in general, very difficult to establish.

- 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.

Thanks for your good remark here! This allows us to emphasize better the key points messages of this figure.

We modified Figure 16 in such a way that the impact of the mentioned parameters (detailed below) are now more explicitly and easily visible. We only have 4 panels, each panel showing results:

- for 1 specific NN retrieval algorithm, either based on the OMI continuum reflectance (475 nm) Rc (*cf.* NN_Rc_NsO_2-O_2), or based on MODIS aerosol optical thickness, or ζ (550 nm), Dark Target Land as a prior information (*cf.* NN_ ζ_NsO_2-O_2).

- for 1 specific prior surface albedo database: either OMLER, or the MODIS Black Albedo.

Each of these panels only represents 1 line with 4 points associated with the Aerosol Layer Height (ALH) of LIVAS per season and, over it, the spread of our spatial-seasonal averaged OMI ALH retrievals between our 2 aerosol Single Scattering Albedo (SSA or ω0) assumptions: *i.e* a lower limit of 0.9, an upper limit of 0.95. This spread is coloured in grey, with the upper limit depicted in red while the lower

limit is depicted in blue. All these panels depict retrievals where a temperature correction has been applied to the OMI O_2-O_2 slant column densities (SCD), named here after NsO_2-O_2 prior to the retrievals.

To set up the best retrieval setting for each individual OMI O_2-O_2 observation is, at this stage, impossible as it requests to use an accurate aerosol model. By an accurate aerosol, we mean here the one describing accurately (in particular) the optical properties of the particles (*e.g.* ω0). However, 1) no accurate ω0 information is available for each OMI individual observation pixel, and 2) ω0 likely significantly varies spatially and temporally. Instead, we designed and trained one NN algorithm for each assumed ω0 value. And then, these algorithms are applied directly over all the OMI pixels. This is equivalent of assuming a constant ω0 value for the ensemble of all OMI O_2-O_2 observations. Note that the chosen ω0 values are based on literature mentioned in our manuscript.

Assuming "only" two ω0 values is not representative enough to perform an average. Therefore, we think it is more appropriate to depict the gap of the spatial-seasonal OMI ALH average due to the 2 (lower-upper) assumed ω0, and the location of the LIVAS black line within this gap.

For these 4 panels, we can confirm all the following elements:

- assumptions on the forward aerosol model (for the creation of the supervision database) lead to the highest impacts on the ALH retrievals. This is mostly related to the ability of the corresponding NN to interpret the combined scattering and absorption of aerosols affecting the NsO_2-O_2. Assuming OMLER surface albedo, differences between average ALH retrievals with ω0=0.95 and ω0=0.9 are in the range of 540-1200 m with NN_Rc_NsO_2-O_2, and 560-660 m with NN_ ζ_NsO_2-O_2.

- assumptions on chosen surface albedo (OMLER *vs.* MODIS Black Sky Albedo) also change the spatial-seasonal averaged ALH retrievals. Related differences in the retrievals are in the range of 0-730 m with NN_Rc_NsO_2-O_2, and 0-180 m with NN_ ζ_NsO_2-O_2.

- Although not shown here anymore, we remarked that applying the temperature correction on NsO_2-O_2 is crucial as it corrects the retrievals between 50 m and 300 m.

In general, largest differences between the spatial-seasonal averaged ALH and LIVAS are found with ω0 = 0.95, in the range of 650-1140 m with NN_Rc_NsO_2-O_2, and 260-800 m with NN_ ζ_NsO_2-O_2. With ω0 = 0.9, differences are reduced to 180-310 m regardless the NN algorithm.

Independently of the chosen NN configuration, the impacts due to assumptions on ω0 (and in general, the aerosol model) are the most crucial ones. They lead to the highest biases on the ALH retrievals. The use of two different surface albedo dataset (cf. OMLER and MODIS Black Sky), although not negligible, has somehow reduced impacts.

Because of the reduced gaps between the ALH retrievals with ω0 = 0.9 or 0.95, the more reduced impacts related to the chosen surface albedo database, and the reduction of the encountered maximum differences with respect to (w.r.t.) LIVAS ALH, NN_ ζ_NsO_2-O_2 is actually the algorithm depicting the best ALH retrieval results. This is likely the case because this configuration uses more accurate prior aerosol ζ(550 nm) information (coming from MODIS ζ Dark Target land product). Please see more details about this specific point below (our answers to your point number 4).

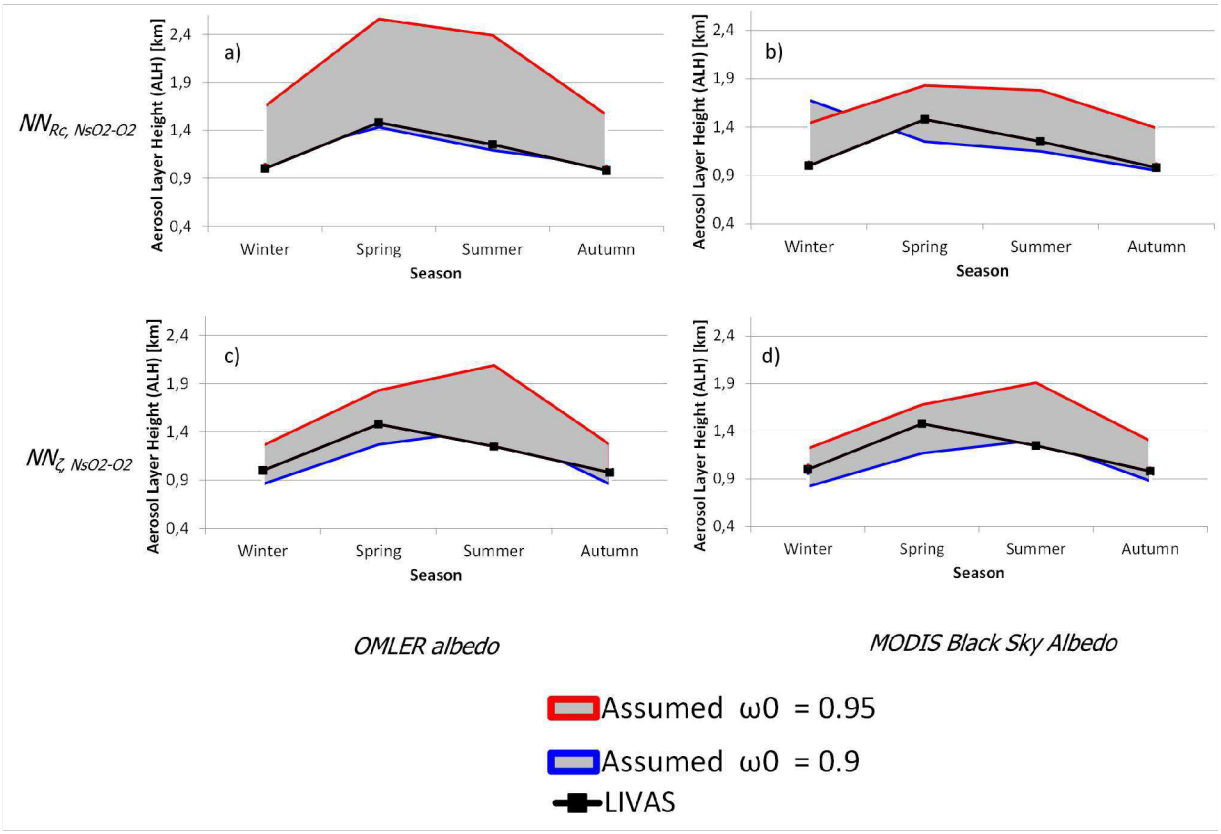


Figure 17: Comparison of the average of the OMI aerosol layer height (ALH) retrievals obtained over scenes with collocated MODIS ζ(550 nm) > 1.0, with the LIVAS ALH climatology database. The retrievals are achieved over North-East Asia for cloud-free scenes, over 3 years 2005-2007, and for the 4 seasons. OMI and MODIS Black Sky surface albedos are considered alternatively. The 2 NN OMI ALH algorithms are used (see Sect. 5.1): NN_Rc_NsO_2-O_2 based on OMI Rc(475 nm) and) NN_ ζ_NsO_2-O_2 based on MODIS ζ (550 nm). A temperature correction is applied to the OMI NsO_2-O_2 prior to the retrievals (*cf.* Sect 6.1): (a) NN_Rc_NsO_2-O_2 algorithm and OMLER surface albedo, (b) NN_Rc_NsO_2-O_2 algorithm and MODIS Black Sky surface albedo, (c) NN_ ζ_NsO_2-O_2 algorithm, and OMLER surface albedo, (d) NN_ ζ_NsO_2-O_2 algorithm and MODIS Black Sky surface albedo.

-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.

Because of the climatology nature of the LIVAS database and of our OMI ALH retrievals, their difference in terms of sampling (which leads then to additional representation errors) and their associated spatial-seasonal average, we do not think it is very appropriate to depict a RMS as you are proposing. One of the reasons, as mentioned above, is that applying a constant ω0 value to the whole ensemble of OMI observations may create specific spatio-temporal patterns in the retrievals. Instead, by emphasizing more the spreads in our averaged ALH retrievals (as done above, *cf.* Figure 17), are likely more representative of the stability and overall quality of the ensemble of our retrievals.

Furthermore, incorrect physical assumptions on surface albedo or aerosol properties would naturally lead to a systematic on an individual ALH retrieval (*i.e.* contrary to a random noise, the ALH error is reproducible). The integrated uncertainty related to the physical error sources and the instrument noise cannot be accurately estimated for each individual OMI pixel. Indeed, it depends on the cases, the real aerosol amounts and properties, their spatial patterns and temporal variability. The different academic cases done in Sec.5 allow to identify and quantify how each error source theoretically contributes to the retrieval error. The averaged real retrievals in Sec.6 and their comparison with LIVAS gives a first global idea of their consistency assuming then LIVAS is a reference. Overall, by combining the comparisons shown in Fig.17 above and the associated standard deviations in Fig.16 (below) allow to state that maximum differences are found on the seasonal and spatial averaged ALH with ω0 = 0.95, in the range of 260-800 m ± 1 km. Lower differences are in the range of 180-310 m, and similar variability, with ω0 = 0.9. This is a global estimate over the considered region as we may not have enough knowledge, at this stage, an adequate estimate of the retrieval error of each individual OMI pixel.

Figure 1 below shows, just as illustration here, how our OMI ALH retrievals spatially vary along the latitude for one day over North East China on 2006.10.02 while a strong aerosol pollution plume was present. We can see that the associated spatial variability is about 500 m on that day, with a maximum peak-to-peak of ~1 km. The accuracy of this synoptic variability is currently under analyses for another study.

Since the mentioned differences can be due to retrieval errors on both sides as well as to spatial representativity errors, we realize that using the term bias is not exact. We therefore removed in our updated manuscript this term. We mentioned now only the "differences" and the diverse error sources that contribute.

Please see our answers to your remarks on climatology *vs.* exceptional events below.

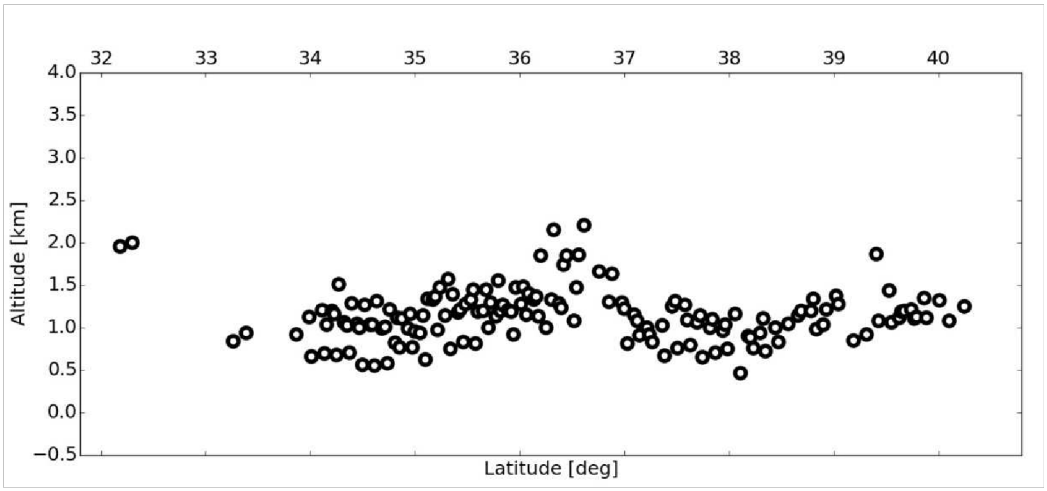


Figure 1: OMI Aerosol layer Height (ALH) retrieved on 2006.10.02 over North-East Asia along a CALIPSO track. These retrievals are obtained from OMI cloud-free scenes, with collocated MODIS ζ(550 nm) > 0.75 and within a distance of 50 km of CALIPSO track.

-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.

As extensively discussed in our manuscript, the possibility to derive an aerosol height from a satellite OMI-like O_2-O_2 visible absorption band is a quite challenging and important topic. In particular for the air quality retrieval community which has discussed and thought about it during the last 15 years (and perhaps even longer). If we leave aside the effective cloud parameters and their more or less sensitivity to the aerosol scattering effects and their altitude, at our knowledge, only Park *et al.*, (2016) with a Look-Up-Table (LUT) with about 10 case studies over ocean, and our manuscript with NNs over land in North-East Asia have retrieved an explicit aerosol height parameter from the OMI O_2-O_2 477 nm band.

Comprehensive and detailed investigations of the performances of our retrievals will require in-depth analyses combining different scales and strategies to fully understand our potential new OMI product. They would mix: global scale, long-time observation periods, focus on regional and/or local areas, day-to-day or pixel-by-pixel variability, statistical analysis to characterize the general behaviour of the algorithms over an ensemble of data, careful investigations of each single observation over a well-know and characterized case study.

In our view, addressing all these strategies together is out of the scope of this very first study and will require more than one manuscript. Instead, for our first paper presenting our algorithms, we chose to address an ensemble of 3-year OMI observations covering a somehow large industrialized land region (North-East Asia). This ensemble is analysed through a statistic strategy in order to characterize the general performance and main sensitivities of our product. Our general idea is that if a statistic analysis

already shows a good global performance and seasonal patterns, then this would encourage to continue this work by focusing on smaller scales (regional, local, reference daily case studies) and further investigating the behaviour of individual OMI retrievals on a pixel basis.

Overall, our manuscript on this topic shows first encouragements about the feasibility of retrieving aerosol height from OMI. Therefore, we are planning to publish follow-up studies on the same topic, but with complementary strategies. In fact, we can already say that we have started a study comparing OMI aerosol height retrievals over land, on some specific case studies (*e.g.* strong urban pollution & biomass burning events) with CALIPSO aerosol measurements along-track. We are planning to publish these results in the forthcoming months in a new journal paper.

Regarding your remark about MODIS RGB images, we do not think this is applicable. These images can tell us where are aerosol plumes and, potentially, how thick they are. But they do not give any information about aerosol altitudes.

-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.

Your remark about the need for an aerosol scattering correction in DOAS trace gas retrievals is very relevant. Analysing what we really need for that purpose (*i.e.* a climatology database of a daily pixel-by-pixel ALH variability) and whether we can use the current ALH retrievals are clearly another study in itself, and therefore beyond the scope of this manuscript.

We realize then, although our motivation statements are clear and correct, they are possibly more related to our long-term objectives than to the present manuscript. Therefore, we softened our statements as follow:

In the abstract:

" [...] This study presents a first step of this long-term objective which evaluates, from a statistic point-of-view, an ensemble of OMI ALH retrievals over a long-time period of 3 years covering a large industrialized continental region. […] "

In the introduction:

" […] Since aerosol altitude, in addition to aerosol optical thickness ζ , is one of the key parameters affecting the computation of AMF for trace gases retrievals such as NO_2 (Leitão *et al.*, 2010; Chimot *et al.,* 2016), our long-term motivation is to evaluate the capability of retrieving it from the satellite O_2 $-O_2$ absorption band at 477 nm. This exploratory study is the first step and statistically analyses an ensemble of OMI observations over 3 year period (2005-2007) and covering a large industrialized continental region (*i.e.* North-East Asia). […] "

In the conclusion:

"[…] The main objective of this work is first to evaluate the feasibility of a direct retrieval of this key aerosol parameter from a statistical point of view: *i.e.* over a long-time period and large industrialized continental area, and therefore a high number of observations. […] Our study indicates that it is worthwhile to design and evaluate aerosol height retrieval algorithm exploiting the satellites 477 nm O_2 $-O_2$ absorption band. Our long-term motivation is to evaluate the feasibility of replacing the effective clouds by more explicit aerosol parameters in the computation of trace gas AMF. This is relevant not only for OMI but for most of the UV-Vis satellite missions devoted to air quality monitoring. For that purpose, further analyses must be performed by focusing on significant geophysical variability cases: *e.g.* pixel-by-pixel variability over smaller regions. Furthermore, single OMI ALH retrievals should be compared with reference aerosol vertical profile measurements (ground-based and/or satellites) over some remarkable case studies."

Regarding the reviewer's remark about the ALH variability, we added in the updated Figure 16 (ex Figure 15) of our manuscript (see also below) the variability as extracted from the LIVAS database. The methodology for deriving LIVAS ALH and its variability are described further in this document, as response to one of the technical questions. On average, the spatial and climatology variability over North-East Asia is about 700 m. Our spatially and seasonally averaged OMI ALH shows similar variability in winter and autumn, but more variability and hence higher standard deviation in spring and summer (from 1100 to 1800 m). The OMI ALH variability is lower, and closer to LIVAS, when using the NN_ ζ_NsO_2-O_2 algorithm with the MODIS ζ(550 nm) Dark target product.

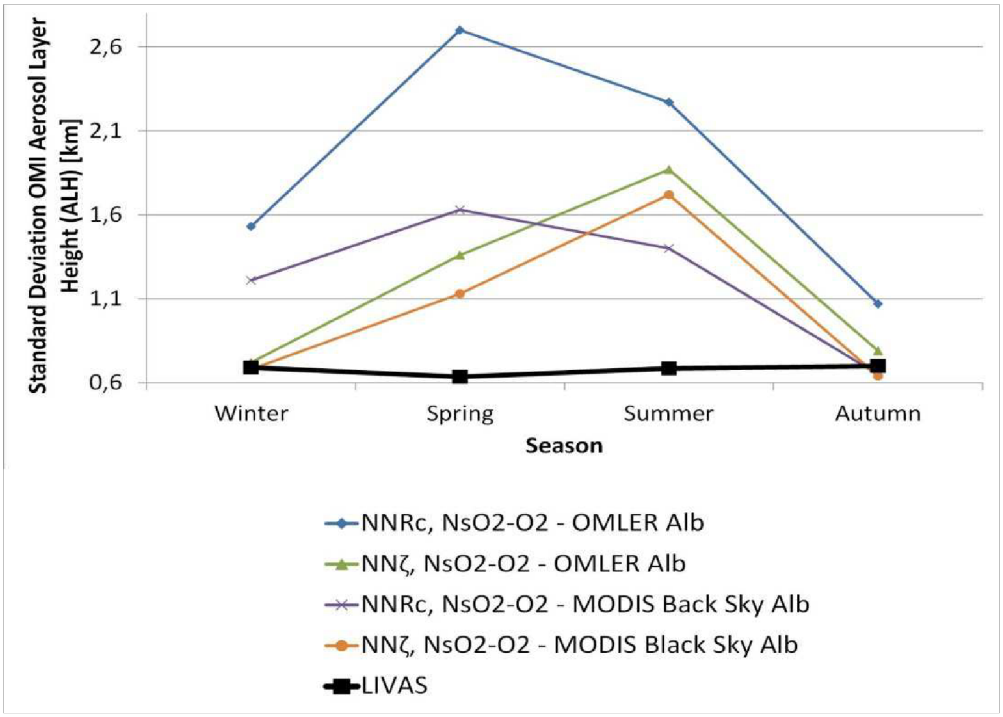


Figure 16: Standard deviation of the OMI aerosol layer height (ALH) retrievals obtained for the cloud-free scenes with MODIS τ (550nm) ≥ 1.0 over North-East Asia. The retrievals are done for cloud-free scene, over 3 years 2005-2007. Aerosol single scattering albedo ω0 = 0.95 is assumed. OMLER and MODIS black sky surface albedo, and the NN configurations (N NRc,NsO2 −O2 and NNτ,NsO2 −O2 with MODIS τ (550 nm) as input) are alternatively considered (*cf.* Sect.6.1).

-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).

As discussed above and with our upgraded Figure 16, we confirm that we think our results derived with the NN_ ζ_NsO_2-O_2 algorithm using MODIS ζ(550 nm) Dark target land present overall the best quality.

Because of all the uncertainties mentioned in our answers to your major comment #2, we do not think it is fully appropriate to look only at individual lines in Figure 16 (now Figure 17!). The apparent minor changes of the line associated with ω0 = 0.9 and OMLER surface albedo from NN_Rc_NsO_2-O_2 and to NN_ζ_NsO_2-O_2 cannot indicate a degradation. Instead, the spreads between the ALH retrievals due to aerosols and surface albedo assumptions and from a statistic point-of-view (*i.e.* spatial-seasonal average) are the main criteria evaluating the robustness of each NN algorithm. A second criterion is the comparison of the OMI ALH retrievals w.r.t. LIVAS ALH over the same region (see updated Figure 15 above). Because these gaps strongly reduce and the OMI & LIVAS ALH variability are closer by employing NN_ζ_NsO_2-O_2 with MODIS ζ(550 nm), they confirm that an accurate knowledge of ζ is then necessary.

Such infomration reduces the freedom of the system and thus the impact of inaccurate assumptions driving the OMI O_2-O_2 observations.

In addition, we would like to highlight the following points:

- The ALH retrieval results from the NN_ ζ_NsO_2-O_2 algorithm statistically show a higher quality only because the MODIS ζ(550 nm) product is used as input. Indeed, as explained in Section 2.3, both ζ and ALH drive the NsO_2-O_2 magnitude. And these effects need to be distinguished. The use of the NN_ ζ_NsO_2-O_2 algorithm in itself does not guarantee a better performance than the NN_ Rc_NsO_2-O_2 algorithm.
- Since Rc is mainly driven by ζ, using NN_ Rc_NsO_2-O_2 is at a first approximation equivalent to using NN_ ζ_NsO_2-O_2 with the OMI ζ(550 nm) product (instead of MODIS) retrieved prior to ALH retrieval. However, as shown in our study, large uncertainties remain on our OMI ζ(550 nm) due to incorrect assumptions on aerosol model and surface albedo. Since these uncertainties are larger than MODIS ζ(550 nm) Dark target uncertainties, they impact the quality of our OMI ALH retrievals.
- Because uncertainties for our OMI ζ(550 nm) product are larger than for MODIS ζ(550 nm) Dark target product, performance of the NN_ Rc_NsO_2-O_2 algorithm is more limited. To increase this performance, our capability to interpret Rc(475 nm) must be improved by separating the effects of aerosol ζ(550 nm), ω0 and surface albedo. In summary, OMI ζ(550 nm) retrieval from the O_2-O_2 visible continuum must be further developed.

We emphasized all these points in our manuscript where necessary.

The MODIS ζ uncertainties are now discussed in the new Section 2.2 as follow:

"The expected error of MODIS DT ζ is about +/- 0.05+15% over land (Levy *et al.*, 2013). The "Deep Blue" retrieval algorithm has been developed to complement the DT algorithm by retrieving ζ over bright arid land surfaces (*e.g.* deserts). The typical associated uncertainties are about +/-0.03 on average (Sayer *et al.*, 2013).".

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.

By default, as mentioned by our references, a NN is always created based on an ensemble of training or supervision dataset. The term LUT is not really used on this topic for this specific need. We added some recent references related to the work of Crevoisier *et al.* on CH_4 and CO_2 retrievals from IASI using a NN approach, and Di Noia *et al.* for Ozone and aerosol retrievals, and the work of cloud retrievals with ROCINN.

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

An exercise comparison with a LUT would be sensitive to the LUT settings, in particular its resolution and sampling. We can state that a NN is a bit faster since using a LUT requires to identify the nodes, in each

dimension, where to perform the interpolation for each single observation. On the contrary, a NN directly "interpolates" or perform its regularization on the given observation.

Similarly, the memory consumption of a LUT depends on the number of dimensions and nodes within the LUT. Therefore, this is difficult to give a full estimate of increase speed and reduced memory.

What we can say is that performing 3-year ALH + ζ retrievals over the North East Asia takes about 20 min per season, by using only 1 single CPU. The memory is only consumed by the synaptic weights between each neuron (input - hidden layer number 1, hidden layer number 1 – hidden layer number 2, hidden layer number 2 – output layer). This takes between 100 and 300 Mo (a simple and rough estimate) when using 1 single NN algorithm.

However, training times are more expensive, but this is of less importance for an operational use.

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

Ok corrected

-p.2,l.5: clouds –> cloud

Done

-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)

In this paper, we numerically define the size of particles through the Angstrom parameter α. This parameter describes the spectral variation of the Aerosol Optical Thickness (AOT) ζ. The higher this value, the higher the dependence on the wavelength is the ζ. Typically, $\alpha < 1$ suggests an optical of coarse particle type (such as dust). $\alpha > 1$ suggests an optical dominance of fine particles (*e.g.* smoke). In our study, the entire supervision dataset used for training the neural networks contains simulations with $\alpha = 1.5$ (fine mode particles). The main reason is because of the assumed dominant presence of aerosols released by anthropogenic fossil-fuel based activities such as industries, power plants and traffic. Also, such a value is dominant in the MODIS aerosol product over the region of North-East Asia.

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

We changed it as follows: "… are still the largest uncertainties of the total RF estimate …" ("are" instead of "contributes").

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

Ok

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

Ok

-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?

Here, I mean that all the uncertainties in general included in the calculation of the tropospheric NO_2 AMF are the dominant source of errors in the retrieval of the tropospheric NO_2 Vertical Column Density. In addition, uncertainties related to the spectral fit of the slant column density also contribute to these retrieval errors. However, according to the mentioned literature, they are of lower magnitude than the errors associated with the AMF.

Of course, aerosols play a big part in the tropospheric NO_2 AMF errors. But, uncertainties related to the shape of the NO_2 vertical profile, clouds and surface albedo are also significant.

Overall, the magnitude of the error on the tropospheric NO_2 column retrieval is over polluted areas mostly determined by the uncertainty of the air mass factor, and not by the uncertainty of the slant column fit (Boersma *et al.*, 2007). We reformulated accordingly.

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

Ok

-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

We carefully checked that all abbreviations are only introduced once through the whole paper.

-p.2,l.35: O_2-O_2 spectral band –> O_2-O_2 absorption band?

Ok

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

The OMI effective cloud parameters are the OMI cloud products retrieved from the same O_2- O_2 visible absorption band assuming a simple opaque and Lambertian cloud model (single albedo = 0.8). They give the effective cloud fraction and effective cloud pressure. More can be read in Acarreta *et al.*, 2004; Boersma *et al.*, 2011; Veefkind *et al.*, 2016; and Chimot *et al.*, 2016.

-p.3,l.26-27: The 477 nm ... individual band. –> How many DFS does the O_4 band add compared to continuum reflectances according to this study? That's the interesting number here.

The statement just before, in the manuscript, mentions that the O_2 band at 477 nm adds between about 1 DFS according to the study of Veihelmann *et al.* (2007).

-p.4,l.5: area –> areas

Ok, done.

-p.4,l.7: emphasize –> emphasis

Ok, this is corrected.

-p.4,l.11: the North –> North

Ok

-p.4,l.3: expectation –> benefit?

Ok (I assume that l.13 was meant here, not l.3).

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

Ok

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

Ok

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

For avoiding any confusion, in particular with the continuum reflectance (475 nm), we removed this equation which not used in the next parts of this study.

-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

Ok

-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?

We removed all the details about OMAERUV that are not used or helpful for this study. We left only the mention of the existence of this algorithm and its general performance (*i.e.* for ζ retrieval).

-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.

In order to make it more clear, we reformulated the associated (new) Sect.3.1 as follow:

« In this paper, the aerosol Neural Network (NN) retrieval algorithms allow the conversion of the continuum reflectance Rc(475nm) and the O_2–O_2 SCD NsO_2–O_2 into ζ and ALP (in hPa). As a consequence, the NN retrievals rely on the way how the aerosol parameters modify these two variables and thus the photons average light path.

Prior to this conversion, a spectral DOAS fit must be performed to derive Rc(475nm) and NsO_2–O_2 from the OMI O_2–O_2 477 nm absorption band. The various DOAS techniques rely on the same key concept: a simultaneous fit of several trace gas slant column densities from the fine spectral features due to their absorption (*i.e.* the high frequency part) present in passive UV-Vis spectral measurements of atmospheric radiation (Platt and Stutz, 2008). Here, the DOAS fit follows the same approach as in the OMI O_2–O_2 cloud algorithm (Acarreta *et al.*, 2004; Veefkind *et al.,* 2016): i.e. the absorption cross-section spectrum of O_2–O_2 is fitted together with a first order polynomial:

$$-\ln(R(\lambda)) = \gamma 1 + \gamma 2 \cdot \lambda + NsO_2\text{–}O_2(\lambda) \cdot \sigma O_2\text{–}O_2 + NsO_3(\lambda) \cdot {}_{\sigma O3}$$

where $\gamma 1 + \gamma 2 \cdot \lambda$ defines the first order polynomial, σO_2 –O_2 and σO_3 are the O_2–O_2 absorption cross-section spectrum and the O_3 absorption cross section spectrum respectively, convoluted with the OMI slit function, and NsO_3 is the O_3 slant column density. σO_2 –O_2 is based on measurements of the cross

section made by C. Hermans (see http://www.aeronomie.be/spectrolab/o2.htm - file O4.txt). The O_3 cross section spectrum is included because it overlaps with the O_2–O_2 spectrum. The fitted parameters are γ1, γ2, $Ns_{O_2-O_2}$, and Ns_{O_3}. In the absence of absorbers, one may define the continuum reflectance Rc at the reference wavelength λ0 :

$Rc = \exp(-\gamma 1 - \gamma 2 \cdot \lambda 0)$. (2)

The reference wavelength is specified as the middle of the DOAS fit window at λ0 = 475 nm."

-p.5,l.22: extraction –> fit?

Ok

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

Added

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

Yes, this is. We added this clarification.

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

Since the reformulation of this section (see above), these two words are now removed.

-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 O_4 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.

The O_2-O_2 xsec that we have used is based on measurements of the cross section by Christian Hermans (http://www.aeronomie.be/spectrolab/o2.htm - file O4.txt).

According to your previous comments, the here mentioned Eq. 1 is now removed to avoid any confusion.

Yes indeed, we indeed fit the square of the oxygen slant column density because of the use of the O_4 xsec reference spectra.

-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.

Ok, this is removed.

-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

We carefully verified the equation label numbers and their use / mention though the entire manuscript. Thanks for your remark!

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

By "homogeneous", we mean that all the particles present in the scattering layer have exactly same properties (*e.g.* single scattering albedo, size through the Angstrom parameter etc..).

By "finite", we want to emphasize that our aerosol layer is approximated by a single box layer which has then, by definition, a finite geometric thickness of about 1 km. Therefore, we assume that no particles are present above and below this box layer. This is opposite to a full vertical profile, described over the entire atmospheric column, which could be visualized then as multiple box aerosol layers with diverse thickness and optical properties.

For more clarity, we reformulated this part as follow:

"In this paper, the aerosol layer is assumed to be one single scattering layer (*i.e.* "box layer") with a constant geometric thickness (about 1 km). All the particles included in this layer are supposed to be homogeneous (*i.e.* same size and optical properties). The aerosol layer height (ALH) is then expressed by the aerosol layer pressure (ALP), in (hPa), defined as the mid-pressure of this scattering layer."

-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.

Following your suggestion, we changed the figure 1 as follow (see below as well the new figure 1 please): we put the measurement variables into the color maps, while the physical variables (aerosol optical thickness and aerosol layer pressure - ALP) are out of them. In that way, the reader can better see how aerosols (both ζ and their altitude expressed in pressure) directly drive the OMI O_2-O_2 measurements. Furthermore, we added 2 panels with different aerosol SSA and surface albedo properties. Then, it becomes more clear how these properties, in addition of ζ and ALP, affect these measurements.

Finally, we reformulated in part the referred section to remove all the confusions as follow. The new section is now written as follows:

"Figure 1 illustrates how aerosol particles drive the OMI O_2–O_2 DOAS parameters at 477 nm assuming cloud-free space-borne observations. These effects are obtained from radiative transfer simulations including aerosols and no clouds. The detailed generation of such simulations is given in Sect. 3.2. The DOAS fit equations following Eq. (2) and Eq. (3) are then applied on these simulated spectra. In this paper, the aerosol layer is assumed to be one single scattering layer (*i.e.* "box layer") with a constant geometric thickness (about 1 km). All the particles included in this layer are supposed to be

homogeneous (*i.e.* same size and optical properties). The aerosol layer height (ALH) is then expressed by the aerosol layer pressure (ALP), in hPa, defined as the mid-pressure of this scattering layer.

Qualitatively, aerosols have two separate effects on the average light path, and therefore on the $O_2–O_2$ absorption signal at the top of the atmosphere (TOA). These two effects are similar as aerosols and clouds have on NO_2 absorption signal (Leitão *et al.*, 2010; Chimot *et al.*, 2016): 1) a shielding effect, *i.e.* a decreased sensitivity within and below the aerosol layer due to a reduced amount of photons coming from the TOA and reaching the lowest part of the atmosphere compared to an aerosol-free scene, 2) an enhancement (albedo) effect, *i.e.* an increased sensitivity within and above the aerosol layer as more photons are scattered back towards the sensors, the part of the atmosphere above the aerosol layer is then sampled by a larger fraction of detected photons. Shielding then leads to a reduced O_2-O_2 absorption while enhancement may increase the O_2-O_2 absorption especially for low cloud or aerosol layers. The overall effect (enhancement *vs.* shielding) depends on the aerosol optical properties, the total column ζ and ALP.

The OMI Rc(475 nm) is directly and primarily affected by the total column ζ of particles present in the observed scene. Indeed, Rc increases with increasing ζ independently of the ALP (*cf.* Fig. 1a). This mostly results from the influence of aerosols on the number of detected photons and on the additional scattering effects observed in the scene compared to an aerosol-free scene. However, the magnitude of this increase relies on aerosol optical properties and the surface brightness. As a consequence, Rc is also affected by aerosol ω0 , phase function, and the surface albedo A. Indeed, Rc decreases with decreasing ω0 and over a darker surface (*i.e.* smaller A value) for all the ζ values (*cf.* Fig. 1c and Fig. 1e). The importance of these parameters is further discussed in Sect. 4 and Sect. 5. Note that, in addition, the reflectance is also driven by the geometry angles: *i.e.* viewing zenith angles θ, solar zenith angles θ0 and relative azimuth angle φ − φ0. An increase of θ or θ0 will lead to longer average light path, and thus will amplify aerosol related additional scattering effects (for a given τ).

OMI $NsO_2–O_2$ relies on the $O_2–O_2$ absorption magnitude along the average light path in the whole atmosphere. It is driven by the overall shielding or enhancement effect of photons by the $O_2–O_2$ complex in the visible spectral range, due to the presence of particles. As depicted in Fig. 1b, $NsO_2–O_2$ decreases with decreasing ALP. This is a direct consequence of a larger shielding effect applied by aerosols located at higher altitudes (*i.e.* part of the $O_2–O_2$ complex located below the aerosol layers are shielded). Nevertheless, in case of low ζ values (*i.e.* ≤ 0.5), $NsO_2–O_2$ does not significantly vary with respect to ALP. This shows that a low amount of aerosols has very little impacts on $O_2–O_2$ absorption measurements.

However, as depicted in Fig. 1b, 1d and 1f, not only ALP but also τ directly influences the slant $O_2–O_2$ absorption since both parameters simultaneously affect the photon path distribution, and therefore the overall shielding or enhancement effect. An increase of ζ leads to a decrease of $NsO_2–O_2$. The slope of this decrease depends on the aerosol altitude (*i.e.* higher for particles at high altitude). Note that both ω0 and A also affect $NsO_2–O_2$, but this effect is smaller than ζ. For example, a reduced ω0 and A lead to a small decrease of $NsO_2–O_2$ (*cf.* Fig. 1d and Fig. 1f).

As a consequence:

– the single parameter $NsO_2–O_2$ contains information on both ζ and ALP. These parameters cannot be separated from this unique variable alone. As a consequence, if ζ is not accurately known, there will likely be an ambiguity when analysing $NsO_2–O_2$ to retrieve ALP.

– if an external or prior ζ estimate is not available, then the two parameters $Ns_{O_2-O_2}$ and Rc could be simultaneously combined to retrieve ALP provided that one can accurately and independently retrieve ζ from Rc. Then, in that condition, Rc may help to distinguish both ζ and ALP contributions in $Ns_{O_2-O_2}$. However, the simultaneous effects of aerosol ω0 and A on Rc (as discussed above), and therefore their associated uncertainties, will impact the feasibility of retrieving ζ from OMI measurements. It may then degrade the retrieved ALP performances.

– ζ , Rc and $Ns_{O_2-O_2}$ have a non negligible correlation. Indeed, an increase of ζ results in a simultaneous increase of Rc and $Ns_{O_2-O_2}$. Therefore, it has to be noted that these two last parameters are not independent and combining them does not provide with two independent pieces of information.

Overall, the impact of aerosol particles on the OMI O_2-O_2 spectral band is similar to cloud particles. This explains in part the difficulty to distinguish aerosols from clouds. In cases with a mix of aerosols and clouds, there is an ambiguity between Rc , ζ and the OMI effective cloud fraction on the one hand, and $Ns_{O_2-O_2}$, the aerosol layer pressure, ζ, the OMI effective cloud pressure and fraction on the other hand (Boersma *et al.*, 2011; Castellanos *et al.*, 2015; Chimot *et al.*, 2016). Therefore, this study only focuses on cloud-free reflectance to avoid this complexity."

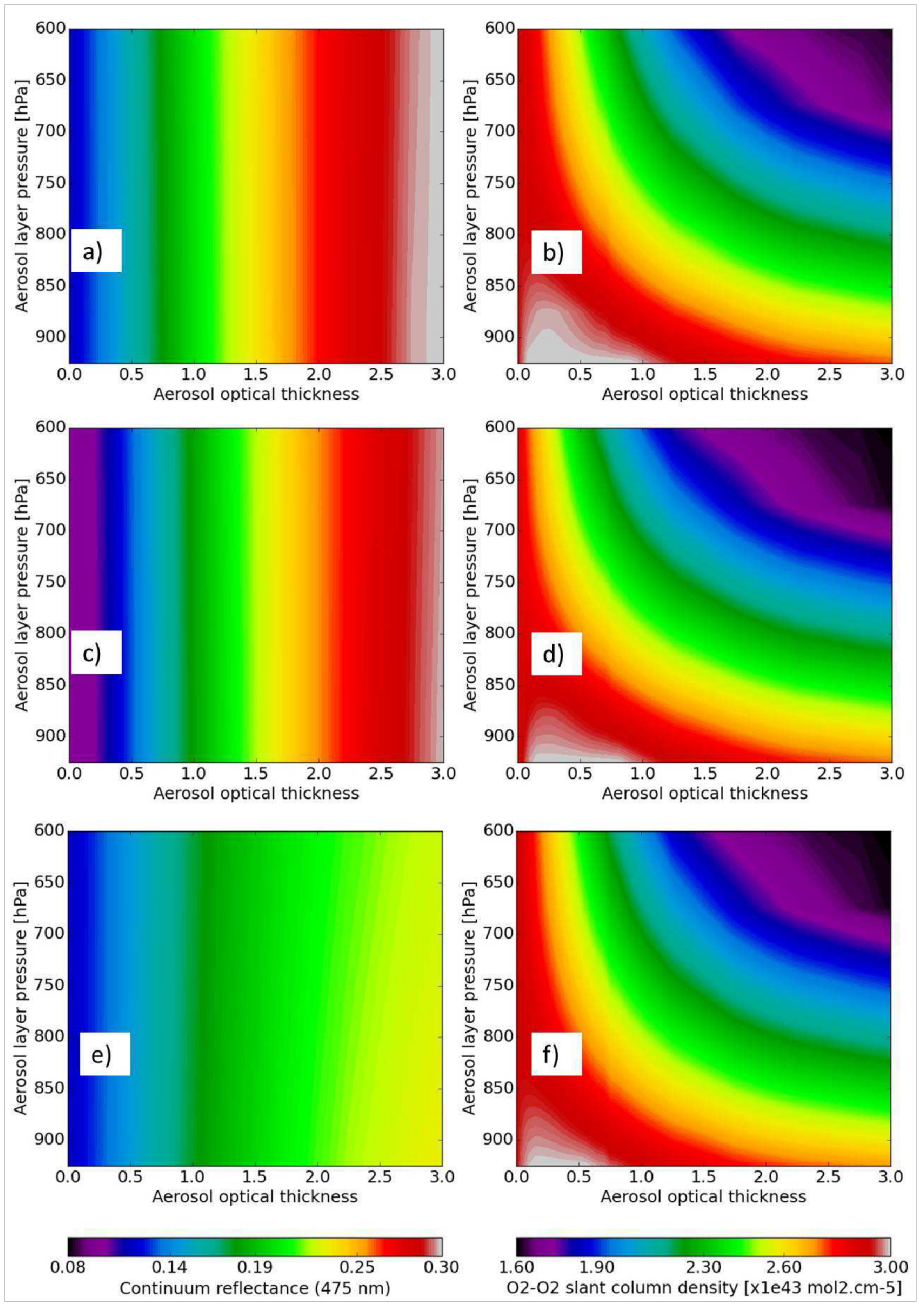


Figure 1: Continuum reflectance Rc (475nm) and O2 −O2 slant column density NsO2 −O2 as a function of τ and aerosol layer pressure for the following conditions: climatology mid-latitude summer temperature, NO2, O3 and H2 O profiles, θ0 = 32◦, θ = 32◦, surface pressure = 1010 hPa and fine aerosol particles (α = 1.5, g = 0.7): (a) and (b) surface albedo = 0.07 and aerosol ω0 = 0.95, , (c) and (d) surface albedo = 0.03 and aerosol ω0 = 0.95, (e) and (f) surface albedo = 0.07 and aerosol ω0 = 0.9.

-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!

We corrected the writing of numbers where needed. We checked the recent guidelines. We acknowledged that the word "Table" should never abbreviated. We corrected accordingly (*e.g.* Table 2).

We homogenized the way of describing ranges of numbers in the entire manuscript as follow: 2005-2007, or 0.03-0.05-0.07.

-p.7,l.23: from –> on

Corrected.

-p.8,l.8: Networks –> networks

Ok

-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.

In our sentence, "with accurate functions" we mean that no invertible analytical function exist that describes the dependence of slant columns and continuum reflectances on aerosols.

Instead, we can accurately simulate it through the processing chain radiative transfer model + spectral DOAS fit. But for the conversion to vertical columns one then employs an approximate inverse model such a conventional Look--Up-Table (LUT) or NNs (as we did).

In our sentence, we therefore replace "the link between the OMI O_2-O_2 measurements and aerosol parameters cannot be exactly described with accurate functions" by "no invertible analytical function exist that describes the dependence of slant columns and continuum reflectances on aerosols.".

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

We corrected this in the whole manuscript.

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

This sentence means that a MLP NN algorithm needs a well-known ensemble of dataset in order to perform an accurate training. We reformulated this sentence.

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

We corrected this in the whole 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

The notion of average light path, and how aerosols impact it, are detailed in the previous Section 2.3 through the so-called shielding and enhancement effects. Aerosols impact the length of the average light path depending on the altitude of atmospheric layers that are shielded and enhanced. Since this is clarified there, we don't think we need to repeat it here. We removed the word "distributions".

-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?

No. As described on p. 9 l5-6, while the NN_ ζ,_Ns configuration uses a ζ information as input for retrieving ALP, the NN_Rc_Ns uses the OMI continuum reflectance at 475 nm as input information, instead of ζ. Therefore, no ζ information is used with this algorithm.

Only the NN_ ζ _Ns algorithm uses an AOT information then. The choice of which AOT information use here is up to the user. Here, we made the choice of using MODIS AOT, as this provides the best information available. But, of course, it would be very possible to use the retrieved OMI AOTs as input instead.

One could argue that the NN_Rc_Ns implicitly uses the OMI ζ information to retrieve ALP. Indeed, as explained in Section 2.3, when analyzing the O_2-O_2 SCD, it is crucial to be able to separate the effects due to aerosol height (or ALP) on the one hand, and those due to ζ on the other hand. If a prior ζ information from an independent instrument is not available, then the OMI continuum reflectance (475 nm) may be used to retrieve ζ. The accuracy however strongly relies on assumptions about surface albedo and aerosol optical properties which drive as well the magnitude of Rc. If these assumptions are accurate enough, then one can retrieve a good OMI ζ which could be then used by NN_ ζ_NsO_2-O_2 for an accurate ALP retrieval.

Since therefore, retrieving a very accurate OMI AOT theoretically means that one is able to precisely interpret the continuum reflectance (475 nm), one could in principle directly retrieve in one step OMI ALP by using the NN_Rc_Ns configuration. That would replace a 2-step approach consisting of firstly retrieving OMI AOT using the NN_Rc_Ns ζ retrieval algorithm, and then secondly using the NN_ ζ _Ns ALP retrieval algorithm with the retrieved OMI AOT as prior information.

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

The thickness of the aerosol layer, considered in all the retrievals is 1 km. We described it.

-p.11,l.1: most relevant –> optimal?

This is changed.

-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?

We just meant that computed error here has, by definition, a positive sign (because of the sum-of-squared-residuals). We removed this term to avoid any confusion.

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

The equation given here, of the error minimization function during the training process of the Neural Network algorithm follows the more general equation as defined and critically specified by Rumelhart *et al.*, (1986) through the description of the Error back-propagation algorithm (described in the present previous sections).

The general error minimization equation given by Rumelhart *et al.*, (1986) is the following:

$$E = \frac{1}{2}\sum_{c}\sum_{j} y_{j,c} - d_{j,c}$$

Where c is an index over the cases (input-output pairs) contained in the training database, j is an index over the output neurons, y is the retrieved value of an output neuron while d is the true expected output. This general equation differs by the one given in our paper by the number of sums, two sums here (over the output neurons and the training cases) while only one in our manuscript (over the training cases).

In theory, a Neural Network algorithm can contain more than one output neuron: *i.e.* a single NN algorithm can simultaneously retrieve more than one single parameter. In practice, because retrieving ζ and ALP from the OMI O_2-O_2 absorption band does not exactly require the same input parameters, we made the choice of one unique output neuron for each of the developed NN algorithm: *i.e.* each algorithm only retrieves one parameter, either AOT or ALP. As a consequence, only one sum remains in our equation (the sum over the output neurons implicitly disappears because the upper limit of the index j is 1). Therefore, our equation keeps the sum over the training cases and the factor ½ from the more general error minimization equation.

Leaving out the factor 1/2 does not lead to different results, but that it (probably) originates from more general cases where the 1/2 weights this term compared to another addition to E.

-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?

No. The number 70% (mentioned in line 24) refers to the percentage of the initial learning dataset (*i.e.* 460 000 spectral simulations as described in section 3.2.) that is randomly extracted and then used for the training of each designed NN. So, about 322 000 simulations (=70% x 460 000) are then used to train the NNs. The verification and evaluation steps are achieved with the remaining learning dataset (15% each) to ensure independency. The training, step is about finding the optimal weights of the synaptic connections of each NN configuration.

As described in section 3.3., the training step is iteratively performed. Prior to the 1st iteration, the synaptic weights are randomly initialized. Then, at each iteration, they are changed through the back-propagation approach. By experience, the error function (cf. Eq. 7) is maximal at the first iteration and decreases with increasing number of iterations. Generally, through all the experiments, this error strongly decreases during the first 50 iterations. After that, the decrease magnitude is relatively low, although not negligible. Therefore, it is not expected (it is even extremely unlikely) that overfitting occurs over the first 15 iterations. Generally, the overfitting, when identified, only occurs after the first 100 iterations. Then, when the training system identified an overfitting (*i.e.* verification step), the training process is stopped.

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

Yes. The so-called valuation score is the error as defined in the Error Equation computed over the evaluation dataset (= 15% of the entire learning database).

-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)?

No. As written in the caption of the figure, the direct output (or retrieved) values from the NN algorithms are normalized between -1 and 1. Because of the nature of the sigmoid functions, used by our algorithms in the hidden layer, all the input parameters are normalized between -1 and 1 to avoid the saturation of these activation functions. Therefore, the outputs are also similarly normalized.

As a consequence, in the NN error back-propagation algorithm, all the trainings and error calculations are performed over normalized values. Then, the values of the total quadratic error E is based on same orders of magnitude. The plotted scores here do show then that ALP errors are, on average, larger than ζ errors. This confirms than the ALP training score is, in theory, less reliable (and thus more complex) than the ζ training score.

Note that we did not really emphasize too much the normalization step and its reason too much in the text of our manuscript, to avoid to dilute too much information. We think this is a technical detail which, except in the caption of figure 3, does not add any value in our analyses. However, if the associate editor thinks otherwise, we will add it then in Section 3.

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

Yes, we agree. If the assumed asymmetry parameter is overestimated (*i.e.* g = 0.7, while the "true" g is 0.6), then we would expect to observe in the signal an increased forward scattering. Since, the true g is lower, there is actually less forward scattering (*i.e.* more back-scattering) than expected. Therefore, the retrieved aerosol optical thickness is biased negatively: i.e. the retrieved AOT is lower than the true AOT.

To ensure clarity, we slightly reformulated these sentences as follow:

"An overestimation of g (i.e. assumed g = 0.7 while true g = 0.6) leads to an increased retrieved ζ value (*i.e.* positive bias) as a result of less photons scattered towards the surface, and therefore more photons scattered towards the satellite sensor, compared to what is theoretically assumed. Reciprocally, an underestimation of g (*i.e.* assumed g = 0.7 while true g = 0.8) leads to a decreased retrieved ζ value (*i.e.* negative bias) as a result of less photons scattered towards the satellite sensor, and more towards the surface compared to the assumption."

-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.

We removed the box-plots and reproduced impacts of surface albedo on retrieved AOT and ALP similarly to the previous ones (see new Figures 4 and 8).

However, as explained now in our upgraded manuscript, we kept the approach of box-whisker plots for evaluating the impacts of O_2-O_2 SCD uncertainties. Indeed, these uncertainties are precision problem (*i.e.* random error by opposite to systematic error or bias). They are mostly (but not only) related to the affects of instrument noise when fitting the O_2-O_2 SCD. Therefore, they will lead to a random error on a the retrieval of ALP, and thus affect the ALP precision. We explained this accordingly in our text.

We added these elements in our upgraded manuscript:

Sect. 5.1:

"Errors in surface albedo also lead to biases in retrieved τ (*cf.* Fig. 4c). Overall, biases are larger over scenes with small τ values. The reason is the dominance of surface reflection in this regime. In cases of high amount of aerosols, then aerosol scattering signals are dominant and surface reflection uncertainties have less impacts. An underestimated (overestimated) surface albedo results in a negative (positive) retrieved τ bias. This is directly related to the change in the measured OMI Rc (*cf.* Fig. 1). Surface albedo uncertainties in the range of 0.025-0.05 lead to biases, in absolute, close to 0.5 for τ in the range of 0.0-0.5, smaller than 0.25 for τ = 2.0. Typical differences in climatological surface albedo from the total ozone monitoring spectrometer (TOMS) and the global ozone monitoring experiement (GOME) (Koelemeijer *et al.*, 2003), or between the Lambertian equivalent reflectance (LER) of OMI and the MODIS black sky albedo (Kleipool *et al.*, 2008), are known to be up to 0.02."

Sect.5.2:

"Surface albedo contributes to the length of the average light path and thus affects NsO_2 $-O_2$. Retrieved ALP biases are maximum (several hundreds hPa) for ζ ≤ 0.5 (cf. Fig. 8a and Fig.8b). For ζ in the range of 0.5-1.0, retrieved ALP have lower absolute values (between 100 hPa and 200 hPa on average) with NN_ ζ _Ns, while they remain too high with NN_Rc_Ns. Over scenes with ζ ≥ 1.0, biases are reduced to 0-50 hPa since aerosol scattering signals dominate over surface reflection. The main cause of all these improvements is that using an accurate prior ζ information (or at least more accurate than retrieved OMI ζ from Rc) allows a better distinction of ζ and ALP effects on the O_2–O_2 slant column density.

An accuracy better than 0.2 must be required on prior ζ information (*cf.* Fig. 8c). Indeed, a ζ bias of 0.25 can impact, in absolute, the retrieved ALP up to 50 hPa for τ in the range of 0.5-1.0. For τ ≥ 1.0, ompact on ALP almost becomes null. Therefore, using MODIS τ as prior to NN_ ζ _Ns is likely expected to show retrieved ALP with a higher quality O_2–O_2 than with NN_Rc_Ns. Indeed, the current retrieved OMI ζ from Rc are below the MODIS τ accuracy."

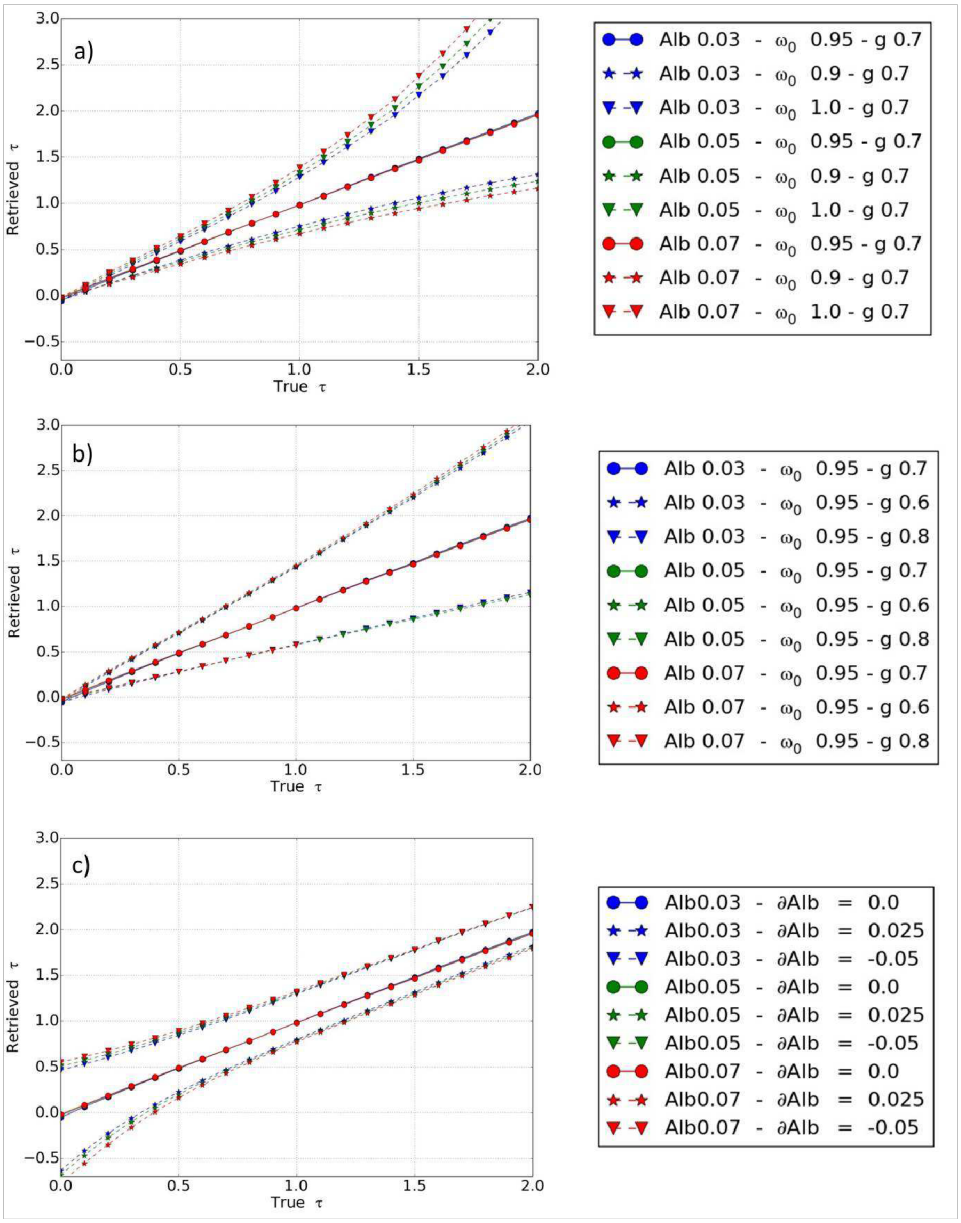


Figure 4: Simulated τ (550 nm) retrievals, based on noise-free synthetic spectra with aerosols, as a function of true τ . The assumed geophysical conditions are: temperature, H2O, O3 , and NO2 from climatology mid latitude summer, θ0 = 25◦ and θ = 25◦, P s = 1010 hPa. All particles are located between 800 and 900 hPa and α = 1.5. Note that the scenarios with lines and similar symbols general tend to fall on top of each other. The reference aerosol scenario is plotted with continuous lines and circle symbols and includes consistent aerosol optical properties with the supervision dataset used to train the neural network algorithm: *i.e.* ω0 = 0.95, g = 0.7. All the retrievals are achieved with the NN algorithm trained with ω0 = 0.95: (a) Sensitivity of τ (550 nm) retrievals to the aerosol single scattering albedo (ω0 = 0.95, 0.9 or 1.0), (b) Sensitivity of τ (550 nm) retrievals to the aerosol asymmetry parameter (g = 0.6, 0.7 or 0.8), (c) Sensitivity of τ (550 nm) retrievals to a surface albedo bias (∂Alb = 0.0, -0.025, 0.05) with ω0 = 0.95, g = 0.7.

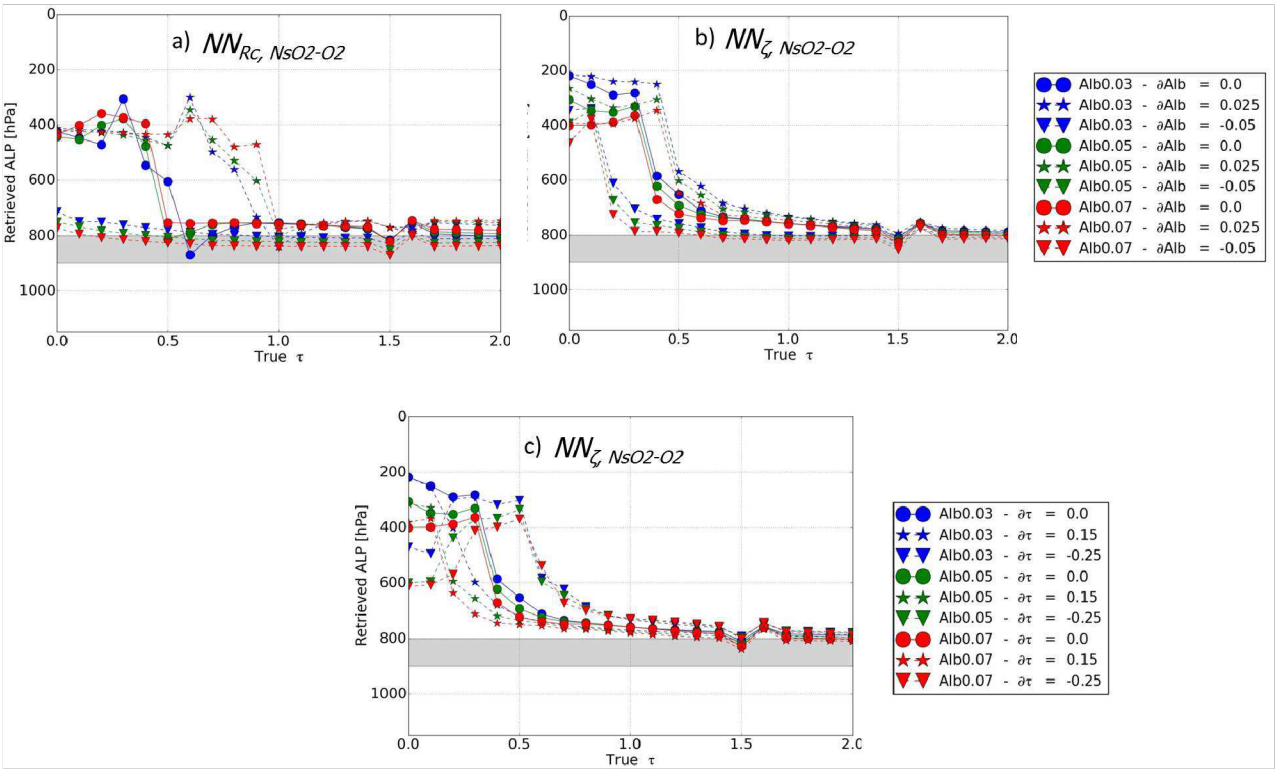


Figure 8: Simulated ALP retrievals, based on noise-free synthetic spectra with aerosols, as a function of true τ . The retrievals are achieved with the NN configurations (N NRc,N s and / or N Nτ,N s) (see Sect. 3.1) trained with ω0 = 0.95. The assumed geophysical conditions are: temperature, H2 O, O3 , and NO2 from climatology mid latitude summer, θ0 = 25◦ and θ = 25◦ , Ps = 1010 hPa. The reference aerosol scenario assumes fine scattering particles (α = 1.5, ω0 = 0.95, g = 0.7) located between 800 and 900 hPa: (a) and (b) Sensitivity of ALP retrievals to a surface albedo bias (∂Alb = 0.0, -0.025, 0.05) with N NRc ,N s and N Nτ,N s , (c) Sensitivity of ALP retrievals to a τ bias (∂τ = 0.0, -0.025, 0.05) with NNτ,NsO2 −O2.

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

Please note that, according to the previous remark, we changed these figures 5a, b and c.

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

Here we are only referring to cases where no error on aerosol properties and surface albedo are included (*cf.* continuous lines with circles in the plots which are associated with reference cases). The biases up to 300 hPa in this ζ range appear when including errors in the assumed aerosol SSA. We move this note from the end to the beginning statement.

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

No. We really mean here the entire learning database (*i.e.* training-verification-evaluation, *cf.* section 3.3) which is used to train the NNs, avoid overfitting and identify the optimal NN architecture. We added the clarification.

-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?

We think that being in a "closed-loop" condition is not enough in this case. Approaches like machine Learning (*e.g.* neural Networks) or interpolation within a Look-Up-Table (LUT) can be seen as a retrieval method based on an off-line forward (radiative transfer) model. This means that the explicit use of a full physics radiative transfer model for each single synthetic spectrum is here replaced by an approximate model, which tries to predict the behaviour of the real model for a given set of input parameters. The retrieval approach relies then more on a regularization process which resolves the retrieval problem from a statistic point of view (*i.e.* the most likely output associated with the set of all combined inputs), with an approximate (or more simple) model without iteration.

Bayesian retrieval approaches, such as the Optimal Estimation Method (OEM), might be able to retrieve the truth on academic cases (noise-free spectra and very low ζ) as 1) they use an exact and full physics forward model directly on-line for each synthetic case study, 2) prior knowledge and parameters are weighted with their corresponding uncertainties, 3) the non-linear problem is linearized around the prior information, 4) iteration is generally performed until the residuals are low enough (below a given threshold). Contrary to approaches based on off-line model, such an on-line forward model approach tries to fit all the details / fine structure present in the combination of input set of parameters (by opposite to regularization techniques).

However, OEM approaches may encounter other problems: *e.g.* the iterative process (in cases of very low ζ) may not converge, and thus no solution may be delivered to the addressed retrieval problem. Such a so-called divergence problem is not accessible by regularization processes (*e.g.* NNs).

NNs can be seen, at a high level, as an optimal interpolator (compared to a more conventional interpolation within a LUT). The large (negative) biases in ALP retrievals over cases with low AOTs, can be interpreted as our inability to interpret very minor changes in O_2-O_2 SCD. This is very similar to the problems of OMI effective cloud pressure in cases of very low effective cloud fraction. As described in different studies, OMI effective cloud pressures, which are retrieved through a LUT, are generally largely (negatively) biased in these conditions (Acarreta *et al.*, 2004; Chimot *et al.*, 2016; Veefkind *et al.*, 2016). The negative sign of the retrieval biases in those cases may be related to the specific curvature of the solution space in those regimes.

-p.14,l.3: on –> in

Ok

-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).

It is true that using an accurate prior ζ value as input, there is no systematic improvement ALP retrievals. However, major improvements occur when there were already large biases (more than 50 hPa) when using the continuum reflectance (475 nm) Rc. When biases with the use of Rc are less major (*i.e.* smaller than 50 hPa), the improvements are not significant anymore.

However, because, the use of AOT already helps to significantly improve cases with large biases, we can conclude that it is better (and therefore recommended) to use AOT instead of Rc as input for ALP retrieval.

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

We carefully and corrected, where necessary, all the equation and section numbers.

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

These sentences summarize and give all the major arguments justifying the use of a prior accurate ζ as input for a correct ALP retrieval. Because these information are not exactly given, or only partially, in the previous analyses, we would prefer to leave these final and important conclusion at this place.

-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.

The O_2-O_2 SCD precision numbers that we considered come from the most typical values encountered in the OMI NASA O2-O2 product (we made a basic histogram analyse to identify their most likely ranges).

The O_2-O_2 SCD precision is typically related to the impact of the OMI noise instrument on the O_2-O_2 fit. Of course, it also includes other uncertainties, such as those related to the employed DOAS approach in the O_2-O_2 absorption band. We replaced "ALP biases" by "ALP uncertainties". We do not see the reason to take into account a noise error in the other input parameter. The impact of surface albedo and ζ errors are already analysed previously. For these parameters, we assumed them more as geophysical parameter errors which may have specific spatial and/or temporal patterns (contrary to a noise which, by definition, presents a random pattern). Therefore, these errors would lead directly to a bias on the ALP retrieval instead of a precision impact. An illustration is given in Section 6 with the OMI surface albedo which include some aerosol residuals. These residuals lead to some specific patterns in the ALP retrievals as observed in Fig. 11.

We removed the remark about the temperature correction, since this is mentioned later on, in the Section 5.

-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.

The considered area is defined by the range of latitude 25-40 deg North, and longitude 110-130 deg East excluding the part over the Gobi desert which presents a too bright surface (as explained in Sect. 4.2 and 6.1).

As already mentioned in Sect. 6.3 (former Section 5.3), when comparing OMI ALH results with LIVAS, about 17% of the pixels passed the selection in summer and spring, and between 5% and 6% in winter and autumn. This leads to about 77 000 OMI observations in spring and summer, and about 20 000 in winter and spring.

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

Ok, this is removed.

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

By applying simultaneously OMI and MODIS cloud fraction threshold, we maximize the probability of filtering out OMI cloudy pixels. According to past studies [Boersma *et al.*, 2011; Castellanos *et al.*, 2015; Chimot *et al.*, 2016], we know that OMI observation pixels with an effective cloud fraction larger than 0.1 are more likely to contain clouds than aerosols. If, the MODIS cloud fraction is lower than 0.1, but OMI effective cloud fraction is larger than 0.1, we preferred to reject this pixel. Such cases might happen since OMI and MODIS pixels do not always strictly overlap, and thus apparent "inconsistencies" between OMI and MODIS cloud detection might be observed.

-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.

We shortened our discussion on the temperature correction and mostly included it when analysing the spatially and seasonal ALH average retrievals. Ideally, for each single O_2-O_2 SCD measurements, a NCEP temperature profile (the spatially closest one) has to be used to compute the so-called gamma factro and then apply it for the temperature correction. That being said, theoretically, by definition of the gamma factor, this correction would depend on aerosol properties (typically ζ and ALP).

However, we found out that in cases of aerosols (contrary to effective cloud retrievals as performed in the OMI O_2-O_2 cloud algorithm), the gamma values does not vary (or very barely) with respect to ζ and ALP variability, at least in the ζ range of 0.0-2.0. It is mostly changing with NCEP temperature profiles. The reasons are double:1) the magnitude of aerosol optical thickness (between 0 and 2) is way lower than typical cloud optical thickness (at least 1 order less), 2) the range of 0-2 of aerosol optical thickness somehow depicts low changes (and then more higher stability) than changes in cloud optical thickness. Therefore, for convenience and low computing time, we considered for all the OMI observations similar aerosol properties (typical optical thickness of 1 and ALP = 800 hPa).

Also, as indicated in our written text, for practical reasons (implementation and computing time), we considered a spatial-seasonal average NCEP temperature profile which is then applied for computing the gamma factor for each single ALP / ALH retrieval. This may have some impacts (order of about 50 m) on single retrieval result. But the impacts on the average ALH results are expected to be minimal.

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

We assumed that you meant p.16 here. This is changed.

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

This is to justify why we considered here the MODIS Black Sky albedo, and not the White Sky albedo.

-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?

We reformulated where necessary this section to make it clear. Please read the new Sect.~6.2. We only illustrate the impact of surface albedo for the autumn season. The impacts are somehow similar for all the other seasons.

As explained at the end of our Sect.6.2., OMLER and MODIS BSA can be expected to present some differences and thus independences: OMLER likely includes some aerosol and cloud residuals, MODIS BSA is the integration over the full hemisphere while OMLER includes most (and only) the viewing angles encountered by OMI, and the assumed solar reference zenith angle in MODIS BSA is likely inconsistent with the true OMI local observation time.

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

Yes, indeed it did not include polarization since the OMI measurements are depolarized (because of the use of a scrambler). However, we wanted to acknowledge that not using polarized simulations, in the learning database, may lead to some errors (likely of the order of 2% in the Visible according to reference studies). This should be investigated in a future work, since this is out the scope of the present study.

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

Yes, we added the following references about the overall quality of MODIS ζ in the new Section 2.2 as follow:

"The expected error of MODIS DT ζ is about ±0.05+15 % over land (Levy *et al.,* 2013). The "Deep Blue" retrieval algorithm has been developed to complement the DT algorithm by retrieving ζ over bright arid land surfaces (*e.g.* deserts). The typical associated uncertainties are about ±0.03 on average (Sayer *et al.*, 2013)."

-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.

I am not sure that, by just looking at this figure alone, we can say this is obvious that using MODIS ζ as input should decrease the ALP variability. Or at least, this only becomes obvious thanks to the analyses achieved on synthetic cases as described and summarised in Sect. 5.2. Only thanks to these analyses, we know that retrievals at low ζ presents some instabilities due to a very low O_2-O_2 absorption signal.

To have an idea about the true geophysical variability included in LIVAS, we added the thick black line: it shows the expected variability of ALH as seen by LIVAS over the same area. This variability seem to be independent of seasons and is about 700 m. ideally, we should expect then our ALH retrievals to show varibility as close as possible to this black line. We discussed it in our updated manuscript and as response to one of your main remarks above.

The methodology of deriving the LIVAS variability is given below (next remark / answer).

-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.

The LIVAS ALH is derived from the given averaged vertical profile of aerosol extinction (532 nm) σ(l) over each vertical layer l defined by its altitude h(l) as follow:

$$ALH(LIVAS) = \frac{\sum_{1} h(l)\sigma(l)}{\sum_{1} \sigma(l)}$$

Since LIVAS also provides the standard deviation associated with each averaged vertical profile of aerosol extinction (532 nm) ∂σ(l), the equivalent standard deviation of each LIVAS ALH ∂ALH is derived as follow:

$$\partial ALH(LIVAS) = \frac{\sum_{1} \partial h(l)\partial\sigma(l)}{\sum_{1} \sigma(l)}$$

where ∂z(l) is the geometric thickness of each vertical layer l.

We added these equations in our new Sect.2.3.

We did not take into account the potential discrepancy between the number of OMI pixels and CALIPSO pixels included in LIVAS, and therefore their specific sampling. Of course, satellite observation resolution and samplings should have an impact. However, fully addressing in this unique study the problem of representation error between OMI and LIVAS (which is derived from CALIPSO observations) is a real work in itself which is currently out the scope of this project.

-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.

I don't think that, in general we have the right to focus on 1 single colour line precisely is appropriate. What really matters here is the spread of the overall / global retrievals due to assumptions on aerosol optical properties (in particular single scattering albedo) which are the most crucial assumptions impacting the spread of our results. The chosen surface albedo being of lower order of magnitude.

Indeed, we stated that, on average, a ω0 of 0.9 (in particular in spring, autumn and winter) is likely a more reasonable value than 0.95 (upper limit in this region). Since, there is not accurate aerosol ω0 value available for each individual OMI observation pixel, we do not have the choice but to use a constant ω0 value to all the OMI observations. However, using a constant ω0 value leads to substantial errors in the ALH results (biases and precisions). Therefore, even if the green line looks to overlap with the LIVAS

line when using NN_Rc_Ns, I don't think we can directly conclude that the retrievals are here very good in this framework.

What mainly matters here, because again of how assumed SSA values are employed here, is the spread between the lines associated with ω0 0.9 and 0.95 when using NN_Rc_Ns and NN_ ζ _Ns algorithms. This is because this spread significantly reduces from NN_Rc_Ns to NN_ζ_Ns, that I think we can conclude that the NN_ζ_Ns fits ALP / ALH better. But this algorithm also fits better because we use an accurate ζ prior information, provided by MODIS. If we are able to retrieve very accurate ζ from every OMI single pixels, the NN_Rc_Ns algorithm would probably correctly fit ALP / ALH as well.

-p.29,fig.1: What is azimuth difference (throughout)? Does H_2O absorb in your fit window? Then why isn't it included in the DOAS fit? At what wvl should I interpret TAU?

The relative azimuth difference is the difference between the viewing azimuth (of the OMI satellite) and the solar azimuth angle. This is already defined in our paper in Sect.~3.2.

It is known that the H_2O absorption is very low in this spectral window. Since, O_3 and O_2-O_2 are the dominant absorption gases here, we did not consider the fit of H_2O SCD. Note that, similarly, this gas is not considered in the OMI cloud algorithm.

As indicated in the text, and in most of the captions, ζ (*i.e.* AOT) should always be interpreted at the reference wavelength 550 nm. We verified that this information is always clarified though the entire manuscript.

-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?

The case with ω0 = 1.0 is now added.

By testing cases with g = 0.6 and g =0.7, while the NNs are trained over a database with g = 0.7, we already indeed considered positive and negative errors in the asymmetry parameter. We found out that making the axes such that the 1-1 line would correspond to the diagonal does not lead to nice plots. Indeed, depending on the magnitude of the retrieved ζ biases, we may have parts of the plots empty / white on the right or left part of the figures which would give strange and diverging feeling to the reader.

Instead, we added the remark, in all the captions, that continuous lines with the circle symbols should be considered as the reference. Indeed, the aerosol properties are consistent between the tested cases and the NNs trainings. And they do not include any bias on the input surface albedo. Therefore, any gap between these lines and the others, refer to a (negative or positive) bias in the plotted retrievals.

-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?

We remove the two plots at the top. Showing the impact of aerosol residuals in the OMI surface albedo for only one season (*e.g.* Autumn) is enough. Similar effects are found for the other seasons.

-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?

There are 2 basic reasons why to split the figures as you mentioned:

- putting all the 4 seasons (x 4 panels) on the same page via LaTex leads to a reduction of their global size, and thus make them less visible for the reader.

- Also, since summer seem to depict better retrieved AOT results assuming SSA = 0.95, it may make sense to separate it from the other seasons where assuming SSA = 0.9 seems to be more accurate.

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

Julien Chimot *et al.*

J.J.Chimot@tudelft.nl

We thank Referee #2 for his / her valuable comments. Below we address them one by one (Referee #2 comments in blue, author and co-authors in black).

The paper examines retrievals of aerosol optical thickness and layer pressure (height) from OMI spectral measurements (or as applicable to other instruments) of the O_2-O_2 absorption band near 477 nm using a neural network approach. A detailed analysis is carried out using simulated data. The approach is then applied in different ways using OMI and MODIS data over land areas of Asia with relatively high aerosol loading and compared with a lidar-based data set (LIVAS).

This is a detailed paper that should be published in AMT. The paper is in general clearly written though there are a number of typos and grammatical issues that it is assumed will be caught during the copy-editing of the manuscript. Only a very few are listed below. I agree with the comments of reviewer 1 and add some additional comments for minor revisions below (some of these may be duplicates).
We thank you for your recommendations and encouragements. We took into account and answer the comments and questions of Referee #1. You can see them in our corresponding document on the AMT website, below the link to the reviews of Referee #1.

We carefully reviewed the new version of our manuscript and verified & corrected the typos and grammatical issues where necessary.

The last sentence of the abstract - "This study shows the first encouraging aerosol layer height retrieval results over land from satellite observations of the 477 nm O_2-O_2 spectral band." - is correct as written. However, the authors may mention here that a previous study examined case study retrievals over ocean. This sentence may stick in the reader's head as this is a "first" implementation with real data (references are later given and it becomes more obvious that these are the first results shown with data over land). I had to go back and reread the sentence to find the over land part, which makes it correct.
Following the mention of the reference of Park *et al.*, (2016) in the other parts of our manuscript, we added the following statement in the abstract: "Following the previous work of Park *et al.*, (2016) over ocean, our study shows the first …."

It might be better to include up front a Data section with the various satellite data sets used (OMI, MODIS, LIVAS) rather than to mention them in different places (and not referenced at the first mention - LIVAS). As it is OMI is mentioned in its own section section with MODIS mentioned below in the Methodology section. It seems that MODIS is an important part of this study as it is important to get accurate ALH and perhaps it deserves more attention.
Thanks for your suggestion. We created a specific Section 2 where are gathered and described all the satellite dataset: aerosol measurements and products (OMI, MODIS AOT Collection 6, and LIVAS climatology), and the surface albedo databases (OMLER and MODIS Black Sky Albedo). These descriptions are mostly all the parts that were written in our former manuscript version (ex-Section 6.1, now renamed as section 7.1). This section is now reduced, mostly focusing on the process of collocation of OMI-MODIS, cloud filtering and the applied temperature correction on the O_2-O_2 slant column density (SCD).

P. 2, L. 34, here and also elsewhere suggest to add *e.g.* before references as there are others not in this list.
Ok, this is done on P.2, L.34 and other pages.

P. 3., L. 20, please add Torres *et al.*, 1998 before de Graaf reference, also suggest to add Torres *et al.*, 2013 reference here and explain that monthly climatology of CALIOP aerosol heights are currently in use for determination of aerosol parameters from OMI UV measurements.
We included the reference Torres *et al.*, (1998) prior to de Graaf *et al.*, (2005). In the introduction, we added a mention about the study of Torres *et al.*, (2013) that has integrated monthly climatology CALIPSO aerosol heights to update the OMAERUV algorithm

P.4, L. 2, please add appropriate references here (altogether, though they are listed above).
Section 2.3, 1st par., This information may go better in the introduction. It's not clear why MAX-DOAS is mentioned specifically here (this sentence seems out of place and not necessary). May be useful also to mention the work of using O_2 A-band to retrieve aerosol height (over ocean by *e.g.*, Dubuisson et al. 2009) and discuss possible advantages of the O_2-O_2 band (lower surface albedos over land?) and also discuss availability of these bands on various sensors such as OMI, OMPS, GOME-2, TROPOMI.
The statement on P.4 L.2 states that no aerosol height retrieval exists (at our knowledge) from the OMI O_2-O_2 visible absorption band over land. Because of the nature of this statement, we do not have specific references here. Regarding the previous statements, mentioning the past works analyzing the nature of this spectral band, the rereferences are already mentioned (*e.g.* Veihelmann *et al.*, (2007), Dirksen *et al.*, (2009), Park *et al.*, (2016)).
We removed our discussion about MAX-DOAS. It was just an information about existing O_2-O_2 ground-based measurements. But it is true that we do not use them then further in the manuscript.

P. 9, L. 6, I think "either" should be removed as it is confusing. Sect. 5.1, Perhaps I missed it but I do not see where the area of North-East Asia is defined. Please give the latitude-longitude of the area studied and/or show it on a map. I believe the highly industrialized areas (where there is heavy aerosol loading) used in this study may also be referred to as South-East Asia, thus it can be confusing.

We removed the word "either". We added the description of the area as follow:
"The considered area is defined by the range of latitude 25-40 deg North, and longitude 110-130 deg East excluding the part over the Gobi desert which presents a too bright surface (as explained in Sect. 4.2 and 6.1)."

P. 15, The discussion of the pairing of OMI and MODIS is confusing. It should be made clear that the resolution of MODIS is 1 km or better, but that collection 6 aerosol products are available at either 3 or 10 km resolution and that you are using 10 km.
We clarified in the new Section 2.2 as follow: "The MODIS instrument, launched on the NASA EOS-Aqua platform in May 2002, is a spectrometer delivering continuous images of the Earth in the visible, solar infrared and thermal infrared approximately 15 min prior to OMI on-board EOS-Aura. The considered MODIS Aqua Level 2 (L2) aerosol product is the collection 6 of M Y D04L 2, based on the Dark Target (DT)
Land algorithm with a high enough quality flag (Xiao *et al.*, 2016). While the MODIS measurement is available at the resolution of 1 km, the MODIS aerosol product is available at both 3 km x 3 km and 10 km x 10 km. Since this last one is relatively close to the OMI nadir spatial resolution, it is then used ine the work below (*cf.* Sect. 7). The improved calibration of MODIS Aqua instrument is included in the reprocessing of the collection 6 aerosol product (Levy *et al.,* 2013; Lyapustin *et al.*, 2014)."

Please mention that you are using dark target only if this is the case and mention the specific product name, *e.g.*, MYD04_L2. I believe that data are provided within the 10 km grids if they have some amount

of cloud free pixels, so there may still be clouds present within the given area even when MODIS data are reported. Also, the 10 km x 10 km areas should not be referred to as a pixel as this can be confused with native MODIS pixels. Does the MODIS geometric cloud fraction used come from the aerosol product? This should be clarified as there are multiple MODIS cloud detection algorithms.
We added the mention of the Dark Target algorithm where the MODIS aerosol product comes from over land. We clarified in the text that the MODIS measurement is acquired over a pixel observation of 1 km, but the grid cell of the used MODIS Level 2 (L2) aerosol product has a resolution of 10 km.

P. 15, L 24, change of to or.
Please check the English meaning of referent. It is defined in most English dictionaries as a noun but used in the paper primarily as an adjective. I believe the word reference may serve better in most instances and also perhaps the word default.
Change is done.
The word "reference" is more appropriate indeed. We modified where necessary in the text.

P. 17, L. 19, accuracy of OMI tau retrievals with respect to MODIS.
We added in Sect.2.2 the following:
"The expected error of MODIS DT τ is about ±0.05+15 % over land (Levy *et al.*, 2013). The "Deep Blue" retrieval algorithm has been developed to complement the DT algorithm by retrieving τ over bright arid land surfaces (*e.g.* deserts). The typical associated uncertainties are about ±0.03 on average (Sayer e*t al.*, 2013)."

P. 18, L. 29, remarkable agreement with respect to seasonal mean values Table 1: What happens for surface pressures < 963 hPa. The neural net will be extrapolating. How well will it do this? Does this occur with the real data? Likewise, why not add a node with single scattering albedo of unity? Also, the maximum value of VZA for OMI is > 44.2 degrees, so why not include the full range?
Somehow, when the input parameters are outside of the range of the training ensemble dataset, the NN will by default extrapolation. The activation function associated with the neurons in the hidden layer is a sigmoid function. Because of the shape of this function (by opposite to a Dirac for instance), the extrapolation is possible over a limited range of the dynamic of each parameter. Therefore, we can expect the NN to correctly extrapolate. This is confirmed by the ensemble results and our statistical analyses in this manuscript. No results show strange extrapolation features or saturated values because of extrapolation beyond the shape of the sigmoid function.
But, further analyses should be done over single and specific cases to ensure the quality and characterize the degree of this extrapolation.

Because of the specified geometric thickness of our aerosol layer, and its nature, we could not define a proper simulation in the training dataset with an ALP (Aerosol Layer Pressure) below 963 hPa. The radiative transfer simulations would have been unrealistic and would have greatly affected the NN training performances.
The reason not include OMI VZA > 44.2 degrees is due to the high time consuming to increase the number of simulations in the supervision dataset. We had already 400 000 simulations which we think is already a large enough number to ensure the robustness of the training step. Furthermore, most of the OMI observation pixels have statistically a VZA lower than 44 deg. But we do recognize that we probably selected pixels with larger angles…

Table 2: When providing values for delta NsO_2-O_2 it would be good to provide a percentage error for these for a given scenario (readers will not have a good idea as to how large these values are). Perhaps I missed it but doing a search of "table" doesn't turn up a reference to Table 2 from the text. The aerosol optical thickness error is quite large for a change in surface albedo of 0.05. If surface albedo errors are more of the order of 0.02 (as stated in the text) then perhaps this would be a more appropriate value to use.

The percentages may vary, but as a first order numbers in the range of 2-7% are probably a good approximation. We added them.
The impact of surface albedo uncertainties on retrieved AOT over synthetic cases is updated in Figure 4, following the comments of Referee #1. The retrieved AOT biases are shown per AOT and for 2 typical surface albedo uncertainties, with opposite signs: a high uncertainty of 0.05, and a lower of 0.025. The second one is then more in line with a more appropriate value as discussed in our manuscript.

All figures in general would benefit from larger fonts. Also the tau in the figures looks different enough from the tau in the text to be somewhat confusing. Figure 3 caption is confusing. What exactly is the supervised data set (training-validation-test)?
We have enlarged fonts where captions, legends, ax labels where not readable enough.

Figure 4, again lines and symbols hard to distinguish. It would be helpful to mention in the caption that the scenarios for the lines with the dot symbols tend to tall on top of one another.
Ok, we added this remark in the caption of Figure 4.

Figure 6: Something should be mentioned in the caption about the range of surface albedos (Alb) used (same for several other figures).
The range of surface albedos is already mentioned in the captions. The Albedo parameter is defined by the notation A. We verified that this notation is already properly introduced throughout the entire manuscript.

Figure 9: I don't see where it is stated that the dotted line is the 1:1 line.
Thanks! We added this remark in the caption of Figure 9.

Fig. 10: There are many lines on these plots. It would help the reader if the 1:1 lines were made thicker to distinguish them. There are backward brackets in the legends at the ends of lines.
Following your suggestion and the one from referee #1, we changed this figure by leaving only 1 season (autumn), so then the reader has only a limited number of lines to visualise.

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

Julien Chimot[1], J. Pepijn Veefkind[1,2], Tim Vlemmix[1], Johan F. de Haan[2], Vassilis Amiridis[3], Emmanouil Proestakis[3,4], Eleni Marinou[3,5], and Pieternel Felicitas Levelt[1,2]

[1]Department of Geoscience and Remote Sensing (GRS), Civil Engineering and Geosciences, TU Delft, the Netherlands
[2]Royal Netherlands Meteorological Institute, De Bilt, the Netherlands
[3]Institute for Astronomy, Astrophysics, Space Applications and Remote Sensing, National Observatory of Athens, Athens 15236, Greece
[4]Laboratory of Atmospheric Physics, Department of Physics, University of Patras, 26500, Greece
[5]Department of Physics, Aristotle University of Thessaloniki, Thessaloniki, 54124, Greece

Correspondence to: J. Chimot (j.j.chimot@tudelft.nl)

Abstract. This paper presents an exploratory study on the ~~retrieval of~~ aerosol layer height (ALH) retrieval from the OMI 477 nm O_2-O_2 spectral band. We have developed algorithms based on the Multilayer Perceptron (MLP) Neural Network (NN) approach and applied them on 3-year (2005-2007) OMI cloud-free scenes over North-East Asia, collocated with MODIS-Aqua aerosol product. In addition to the importance of aerosol altitude for climate and air quality objectives, ~~the main motivation of this study~~ our long-term motivation is to evaluate the possibility of retrieving ALH for potential future improvements of trace gas retrievals (e.g. NO_2, HCHO, SO_2, etc..) from UV-Vis air quality satellite measurements over scenes including high aerosol concentrations. ~~ALH~~ This study presents a first step of this long-term objective and evaluates, from a statistic point-of-view, an ensemble of OMI ALH retrievals over a long-time period of 3 years covering a large industrialized continental region. This ALH retrieval relies on the analysis of the O_2-O_2 slant column density (SCD) and requires an accurate knowledge of the aerosol optical thickness τ. Using ~~the~~ MODIS-Aqua ~~aerosol optical thickness at 550~~ $\tau(550\text{nm})$ as a prior information, ~~comparison with~~ maximum differences between the LIdar climatology of vertical Aerosol Structure for space-based lidar simulation (LIVAS) ~~shows that ALH average biases~~ and average OMI ALH, over scenes with MODIS $\tau \geq 1.0$, are in the range of 260-800 m depending on the seasons. These results depend on the assumed aerosol single scattering albedo (sensitivity up to ~~600~~ 660 m) and the chosen surface albedo (variation less than 200 m). Scenes with $\tau \leq 0.5$ are expected to show too large biases due to the little impacts of particles on the O_2-O_2 SCD changes. In addition, NN algorithms also enable aerosol optical thickness retrieval by exploring the OMI reflectance in the continuum. Comparisons with collocated MODIS-Aqua show agreements between -0.02 ± 0.45 and -0.18 ± 0.24 depending on the season. Improvements may be obtained from a better knowledge of the surface albedo, and higher accuracy of the aerosol model. ~~This~~ Following the previous work over ocean of Park et al. (2016) , our study shows the first encouraging aerosol layer height retrieval results over land from satellite observations of the 477 nm O_2-O_2 absorption spectral band.

1 Introduction

The ability to monitor air quality and climate from UltraViolet-Visible (UV-Vis) satellite spectral measurements requires accurate trace gas (e.g. NO_2, SO_2, HCHO, O_3) and aerosol observations. Aerosols and trace gases often share similar anthropogenic sources~~sources~~, and their concentrations, as shown by the satellite observations, often exhibit significant correlations (Veefkind et al., 2011). The reason is that trace gases are often precursors for aerosols. The importance of measuring vertical distribution of atmospheric aerosols on a global scale is triple. Firstly, aerosols directly impact the radiation budget of the Earth-atmosphere system through the scattering and absorption of solar and terrestrial radiation (Feingold et al., 1999). High concentrations of fine particles lead to reduced ~~clouds~~ cloud droplet size, enhanced cloud reflectance (Twomey et al., 1984), and reduced precipitation (Rosenfeld, 2000; Ramanathan et al., 2001; Rosenfeld et al., 2002). Therefore, large uncertainties of aerosol optical properties limit our climate predictive capabilities (IPCC: Solomon et al., 2007). In spite of more robust climate predictions in the last years, radiative forcing (RF) induced by aerosols ~~still contributes to~~ are still the largest uncertainty to the total RF estimate (IPCC: The Core Writing Team Pachauri and Meyer, 2014). The vertical distribution and relative location are determining factors of aerosol radiative forcing in the long-wave spectral range (Dufresne et al., 2002; Kaufman et al., 2002). Secondly, aerosols play a significant role in air quality, in particular near the surface. Due to the rapid growth of both population and economic activity, such as in Asian region, increase in fossil fuel emissions gives rise to concerns about fine particles formation and dispersion. Aerosols include a variety of hazardous organic and inorganic substances, reduce visibility, lead to reductions in crop productivity and strongly affect health of inhabitants in urban regions (Chameides et al., 1999; Prospero, 1999; Eck et al., 2005).

Thirdly, slant column densities (SCD) of trace gases, derived from UV-Vis air quality space-borne sensors, have a high sensitivity to aerosol heights. For partly cloudy conditions, clouds are the main error source of trace gas measurements. But, in the absence of clouds, vertical distribution of aerosols, combined with their scattering and absorbing properties, modifies the length of the average light path of the detected photons, and therefore affects ~~the computation of~~ trace gas Air Mass ~~Factor~~ Factors (AMF). The application of the AMF is crucial for the conversion of slant column densities (SCD) from satellite line-of-sight measurements into vertical column densities. Then, aerosols strongly contribute to the uncertainties of trace gas retrievals from space-borne observations. For example, ~~uncertainties in the computed tropospheric AMF for the~~ the magnitude of the error on the Ozone Monitoring Instrument (OMI) ~~are the dominant source of errors in the retrieved~~ tropospheric NO_2 ~~column~~ retrieval is, over polluted areas, mostly determined by the AMF uncertainty, not by the SCD ucnertainty. It results from the combination of aerosols, clouds and the shape of the NO_2 profile (Boersma et al., 2007). Negative biases on OMI tropospheric NO_2 columns, between $-26\,\%$ and $-50\,\%$, are found in urban and very polluted areas in cases of high aerosol pollution and particles located at elevated altitude (Shaiganfar et al., 2011; Ma et al., 2013; Kanaya et al., 2014). HCHO AMF for GOME-2 and SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY) shows about 20-50 % sensitivity to aerosols, depending ~~if~~ whether they are located within or above the boundary layer (Barkley et al., 2012; Hewson et al., 2015). Dust aerosols (large particles, with strong absorption in UV) can reduce the AMF in the SO_2 wavelengths (310 - 330 nm) by half, thus doubling the retrieved SO_2 (Krotkov et al., 2008). This impacts the ability of sensors like OMI

to monitor Planetary Boundary Layer (PBL) SO_2 with a sensitivity to local anthropogenic sources. Over regions of enhanced columns, aerosols highly contribute to the total SO_2 AMF error (Lee et al., 2009). Therefore, aerosol parameters (or retrievals) are a pre-requisite before retrieving trace gas vertical column densities.

State-of-the-art trace gas retrieval algorithms correct for aerosol effects either explicitly using modeled aerosol vertical profiles (e.g., (Barkley et al., 2012, 2013; Kuhlmann et al., 2015; Lin et al., 2014, 2015)), or alternatively implicitly via cloud algorithms. For example, the OMI O_2-O_2 ~~spectral~~ absorption band at 477 nm has been widely exploited to derive cloud information (Acarreta et al., 2004; Sneep et al., 2008). However, the OMI cloud algorithm is sensitive to aerosols, and thus the retrieved effective cloud parameters are modified in their presence (Boersma et al., 2007; Castellanos et al., 2015; Chimot et al., 2016). The OMI O_2-O_2 spectral band at 477 nm contains significant information on aerosol properties and height. The retrieved effective clouds are then used to correct the computed AMF (de Smedt et al., 2008; Boersma et al., 2011). In spite of these well considered perturbations, the use of the effective cloud parameters, assuming that the opaque Lambertian cloud model can reproduce the distribution of scattering fine particle effects, does not yet completely correct for the aerosol effects when computing the AMF, in particular for the tropospheric NO_2 columns (Castellanos et al., 2015; Chimot et al., 2016).

Characterizing the aerosol vertical distribution, in addition to the associated optical properties, using passive space-borne measurements is challenging due to the absence of spectral features in the aerosol optical properties and the combined influences of surface and cloud reflection. Contrary to effective cloud retrievals, aerosol retrieval is a more complex problem mainly because of the variability of particle microphysical properties and the lower optical thickness (typically 1-2 orders of magnitude). As a consequence, methods assuming large multiple scattering contributions, such as a simple cloud model assuming Lambertian properties, cannot be used. Passive radiometers like Moderate Resolution Imaging Spectroradiometer (MODIS) can only retrieve a limited amount of independent information from their measurements, usually aerosol optical thickness τ and the extinction Ångström exponent α, as a proxy for the particle size distribution (Levy et al., 2007, 2013). The near-UV technique has been widely used to map the daily global distribution of UV-absorbing aerosols such as desert dust particles as well as carbonaceous aerosols generated by anthropogenic biomass burning and wildfires. It allows to retrieve τ, Single Scattering Albedo ω_0 and the qualitative Aerosol Absorbing Index (AAI) in the 330-388 nm of the Total Ozone Mapping Spectrometer (TOMS) and OMI sensors (Torres et al., 1998; Torres et al., 2002; Torres et al., 2007). However, this technique is highly affected by the dependency of the measured radiances on the height of the absorbing aerosol layer ~~(de Graaf et al., 2005) .~~ (Torres et al., 1998; de Graaf et al., 2005) . OMAERUV has been upgraded by integrating a monthly climatology of CALIOP aerosol heights to retrieve aerosol parameters from OMI UV measurements (Torres et al., 2013) . The Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) has been providing vertical profiles of aerosols but with limited spatial coverage because of its measurements characteristics (Omar et al., 2009). Park et al. (2016) evaluated the sensitivity of the O_2-O_2 slant column density to changes in aerosol layer height over ocean. It is demonstrated that the O_2-O_2 spectral band at 477 nm is the most sensitive to the aerosol layer effective height (compared to the O_2-O_2 absorption bands at 340, 360 and 380 nm), due to the largest O_2-O_2 absorption and reduced Rayleigh scattering. Veihelmann et al. (2007) determined that the complete OMI UV-Vis reflectance measurements contain between 2 and 4 Degrees of Freedom of Signal (DFS). The 477 nm O_2-O_2 band adds about 1 degree and therefore contains more information than any other individual band. This relative large number

of DFS for UV-Vis satellite solar backscatter observations is explained by the sensitivity of the reflectance to the aerosol layer height. Detailed O_2-O_2 radiative transfer simulations performed by Dirksen et al. (2009) revealed the availability of the altitude information about smoke aerosol plume, released by intense forest fires and transported over long distance, under specific conditions: high AAI and no clouds. In spite of all these efforts, no aerosol height retrieval has been done at this moment from O_2-O_2 satellite measurements at 477 nm over land.

Since aerosol altitude, in addition to ~~aerosol optical thickness~~ τ, is one of the key parameters affecting the computation of AMF for trace gases retrievals such as NO_2 (Leitão et al., 2010; Chimot et al., 2016), ~~the motivation of this exploratory study~~ our long-term motivation is to evaluate the capability of retrieving it from the satellite O_2-O_2 absorption band at 477 nm. This ~~study~~ exploratory study is the first step and statistically analyses an ensemble of OMI observations over a three year period (from 2005 to 2007) and covering a large industrialized continental region (i.e. North-East Asia). This study follows the conclusions of previous works focused on the sensitivity of this spectral band and the observed links between the O_2-O_2 effective cloud retrievals and aerosol parameters. In this paper, quite a few algorithm concepts are developed, based on the Neural Network (NN) approach, and then tested on a high number of OMI observations over land. ~~The primary focus of this exploratory study~~ Our primary focus is the retrieval performance of aerosol layer pressure (ALP) associated with scattering and fine particles over large urban, industrialized and highly polluted area and cloud-free scenes. In addition, the sensitivity of the algorithms to τ knowledge is investigated and, therefore, the capability of τ retrievals from the same OMI band is evaluated. The considered ~~OMI observations~~ satellite observations and input dataset are described in Sect. ~~2, with a particular emphasize~~ 2. Sect. 3 focuses on the available OMI O_2-O_2 Differential Optical Absorption Spectroscopy (DOAS) parameters and their link with ~~ALH~~ Aerosol Layer Height (ALH) and τ. The development of the different NN algorithms are described in Sect. ~~3.~~ 4. Their performances are evaluated in Sect. ~~4~~ 5 on synthetic and independent data set with a characterization of the main limiting factors. Finally, these algorithms are applied ~~to 3 years (2005-2007) of~~ in Sect. 6 to cloud-free OMI observation over the North-East Asia ~~,~~ where large amounts of aerosols are emitted from both natural and anthropogenic sources (Lee et al., 2012). They are then compared with other observation products, namely MODIS Aqua τ and the LIdar climatology of vertical Aerosol Structure for space-based lidar simulation (LIVAS). ~~This allows to address the potential expectation of aerosol layer height (ALH) retrievals for future AMF computations.~~

2 ~~OMI DOAS analysis and aerosols~~

2 Aerosol and surface albedo satellite data

In this section, are described the three main aerosol satellite data that are used in this study: OMI Visible measurements, MODIS aerosol product and LIVAS climatology database. In addition, the two considered surface albedo databases: OMI Lambert Equivalent Spectral Surface Reflectance (OMLER) and MODIS Black Sky Albedo are also described.

2.1 OMI satellite data

The Dutch-Finnish mission OMI (Levelt et al., 2006) is a nadir-viewing push-broom imaging spectrometer launched on the National Aeronautics and Space Administration (NASA) Earth Observing System (EOS)-Aura satellite. It provides daily global coverage of key air quality components through observations of the backscattered solar radiation that are captured ~~during daylight~~ in the UV-Vis spectral domain. Based on a ~~two dimensional~~ two-dimensional detector array concept, radiance spectra are simultaneously measured on a 2600 km wide swath within a nadir pixel size of 13x24 km^2 (28 x 150 km^2 at extreme off-nadir). OMI has a higher spatial resolution than any other UV-Vis hyperspectral spectrometers. It measures in the wavelength range of 270 nm to 500 nm with a spectral resolution of 0.45 nm in the UV-2 band (310-360 nm) and 0.63 nm in the visible band (360-500 nm). ~~Retrieval of atmospheric parameters are usually obtained from the reflectance spectrum defined as:~~

~~$$R(\lambda) = \frac{\pi \cdot E(\lambda)}{cos(\theta_o) \cdot F(\lambda)},$$~~

~~with λ the wavelength, $E(\lambda)$ and $F(\lambda)$, the radiance and the solar irradiance respectively, and θ_o the solar zenith angle.~~

Starting mid-2007, ~~an external obstruction to the sensor's field of view~~ the so-called "row anomaly" has been perturbing OMI measurements of the Earth-shine radiance at all the wavelengths. ~~At this moment, about half of the sensor's sixty viewing positions are affected by this so-called "row anomaly", referring to the row numbers of the viewing positions on the CCD detectors. Details on the onset and progression of this anomaly~~ Details are given on the site http://www.knmi.nl/omi/research/product/rowanomaly-background.php. For practical reasons, this study only used the OMI data acquired during 2005-2007, i.e. before the development of this anomaly.

OMI has not been optimized for aerosol monitoring. However, ~~various studies demonstrated the sensitivity of the reflectance measurements to aerosol properties in the UV. The main advantage of this spectral domain is the reduced impacts due to uncertainties in surface reflectance, partly due to the low values of both ocean and land surface albedo in the UV (Torres et al., 1998) . Absorbing aerosols are detected and distinguished from the other aerosol types by analyzing their spectral contrast to Rayleigh scattering. The~~ the OMI near-UV aerosol algorithm (OMAERUV) independently retrieves atmospheric total columns of ~~aerosol optical thickness~~ τ and ~~single scattering albedo~~ ω_0 from the 2 UV wavelengths, 354 nm and 388 nm (Torres et al., 2007, 2013). ~~The OMI multi-wavelength algorithm (OMAERO) uses up to 19 wavelength bands between 331 and 500 to derive aerosol characteristics (i.e. AOD, aerosol type, AAI as well as ancillary information) from the OMI cloud-free spectral reflectance, based on the best fitting aerosol type models provided. Three main aerosol types are considered: dust, carboneceous aerosol associated with biomass burning, or weakly absorbing sulfate based aerosol. The wavelengths were selected such that they are essentially free from strong Raman scattering. The collision complex at 477 is included to enhance the sensitivity to the aerosol layer height (Veihelmann et al., 2007) .~~

In comparison with 44 Aerosol Robotics Network (AERONET) sites, evaluated OMAERUV τ yield a root mean square error (RMSE) of 0.16 and a correlation coefficient of 0.81 over the years 2005-2008 (Ahn et al., 2014). About 65 % of these retrievals lie within the expected uncertainty. The OMAERUV ~~SSA products agree~~ ω_0 product agrees with AERONET to

within 0.03 in 46,% of the collocated pairs, and to within 0.05 in 69 % of the cases (Jethva et al., 2014). ~~Park et al. (2016) used~~

The OMI O_2-O_2 477 nm absorption band is currently operationally exploited by the OMI O_2-O_2 cloud algorithm to derive effective cloud fraction and pressure (Acarreta et al., 2004; Veefkind et al., 2016) . Park et al. (2016) applied a Look-Up-Table (LUT) approach ~~to retrieve, from the OMI 477 spectral band,~~ on this band to retrieve aerosol effective height over ocean, close to East Asia, within the error range of 1 km (compared to CALIOP). This approach was applied to 7 case studies, each of them covering a few days. No ~~aerosol optical thickness~~ τ was retrieved. No study has yet explicitly used this satellite band to directly retrieve ~~aerosol layer height and~~ ALH and τ ~~over land~~over land. This band is not only available on OMI but also on various sensors such GOME-2, OMPS and the next Tropospheric Ozone Monitoring (TROPOMI) space-borne.

2.2 MODIS aerosol product

The MODIS instrument, launched on the NASA EOS-Aqua platform in May 2002, is a spectrometer delivering continuous images of the Earth in the visible, solar and thermal infrared approximately 15 min prior to OMI on-board EOS-Aura. The considered MODIS Aqua Level 2 (L2) aerosol product is the collection 6 of $MYD04_L2$, based on the Dark Target (DT) Land algorithm with a high enough quality flag (Xiao et al., 2016) . While the MODIS measurement is acquired at the resolution of 1 km, the MODIS aerosol product is available at both 3 km x 3 km and 10 km x 10 km. Since this last one is relatively close to the OMI nadir spatial resolution, it is then used in the work below (c.f. Sect. 7). The improved calibration of MODIS Aqua instrument is included in the reprocessing of the collection 6 aerosol product (Levy et al., 2013; Lyapustin et al., 2014) .

The availability of the MODIS aerosol products is generally a good confirmation of cloud-free scenes as MODIS Aqua average τ variable is exclusively given provided a high amount of cloud-free sub-pixels is available (i.e. the MODIS measurement resolution of 1 km).

The expected error of MODIS DT τ is about $\pm 0.05 + 15\%$ over land (Levy et al., 2013) . The "Deep Blue" retrieval algorithm has been developed to complement the DT algorithm by retrieving τ over bright arid land surfaces (e.g. deserts). The typical associated uncertainties are about ± 0.03 on average (**?**) .

2.3 ~~DOAS analysis~~ LIVAS climatology database

The LIdar climatology of Vertical Aerosol Structure for space-based lidar simulation studies (LIVAS) is a 3-D multi-wavelength global aerosol and cloud optical database (Amiridis et al., 2015) . This database provides averaged profiles of aerosol optical properties over 9 years (1 january 2007 - 31 December 2015) from the Cloud Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) data on a uniform grid of 1° x 1°. LIVAS addresses the wavelength dependency of aerosol properties for many laser operating wavelengths including 532 nm. LIVAS data set has been evaluated against AERONET in Amiridis et al. (2015) showing realistic and representative mean state aerosol optical depth values in 532 nm making this data set ideal for synergistic use with other satellite products.

The LIVAS ALH is derived from the given averaged vertical profile of aerosol extinction (532 nm) $\sigma(l)$ over each vertical layer l defined by its altitude $h(l)$ as follow:

$$ALH(LIVAS) = \frac{\sum_{l} h(l)\sigma(l)}{\sum_{l} \sigma(l)} \quad (1)$$

Since LIVAS also provides the standard deviation associated with each averaged vertical profile of aerosol extinction (532 nm) $\partial\sigma(l)$, the equivalent standard deviation ∂ALH of each LIVAS ALH is derived as follow:

$$\partial ALH(LIVAS) = \frac{\sum_{l} \partial h(l)\partial\sigma(l)}{\sum_{l} \sigma(l)} \quad (2)$$

where $\partial z(l)$ is the geometric thickness of each vertical layer l.

2.4 Surface albedo dataset

The standard and reference product is the OMLER climatology derived from several years of OMI observations at the spatial resolution of 0.5 ° x 0.5 ° longitude-latitude grid for each calendar month (Kleipool et al., 2008) . The OMLER algorithm is based on temporal histograms of the observed Lambert Equivalent Spectral Surface Reflectance (LER) values per grid box. Potential small residual cloud and aerosol contaminations are expected to remain in the OMLER product. As an alternative, the global and spatially complete MODIS black sky surface albedo in the band 3 (459-479 nm) is considered. It is defined as the directional hemispherical reflectance and is a function of solar zenith angle θ_0 (Schaaf et al., 2002) . It is derived by integrating the atmospheric corrected Bidirectional Reflectance Distribution Function (BRDF), derived from combined MODIS-Aqua and Terra observations over every 16-day period. The downwelling flux in the MODIS black sky albedo has no diffuse component. The Collection 6 of ~~the OMI~~ MCD43C3 product is given on a 0.05° (5.6 km) latitude/longitude Climate Modeling Grid (CMG). Note that (Kleipool et al., 2008) demonstrated that the OMLER data set is closer to the black sky than to the white sky by evaluating the ratio between diffuse and direct illumination.

3 OMI O_2-O_2 DOAS analysis and aerosols

3.1 DOAS analysis of the OMI O_2-O_2 477 nm absorption band

In this paper, the aerosol Neural Network (NN) retrieval algorithms allow the conversion of the continuum reflectance $R_c(475\text{nm})$ and the O_2-O_2 SCD $N^s_{O_2-O_2}$ into τ and ALP (in hPa). As a consequence, the NN retrievals rely on the way how the aerosol parameters modify these two variables and thus the photons average light path.

Prior to this conversion, a spectral DOAS fit must be performed to derive $R_c(475\text{nm})$ and $N^s_{O_2-O_2}$ from the OMI O_2-O_2 477 nm ~~spectral band~~

~~The~~ absorption band. The various DOAS techniques rely on the same key concept: a simultaneous ~~extraction~~ fit of several trace gas slant column densities from the fine spectral features due to their absorption (i.e. the high frequency part) present in passive UV-Vis spectral measurements of atmospheric radiation (Platt and Stutz, 2008). ~~The OMI cloud algorithm exploits the 460-490spectral band, focused on the absorption line at 477. The initial purpose of this algorithm is the retrieval of cloud parameters, namely effective cloud fraction and effective cloud pressure (Acarreta et al., 2004; Veefkind et al., 2016) . The first step of this algorithm is a DOAS spectral fit in which~~ Here, the DOAS fit follows the same approach as in the OMI O_2-O_2 cloud algorithm (Acarreta et al., 2004; Veefkind et al., 2016) : i.e. the absorption cross-section spectrum of O_2-O_2 is fitted together with a first order polynomial:

$$-\ln(R(\lambda)) = \gamma_1 + \gamma_2 \cdot \lambda + N^{\mathrm{s}}_{\mathrm{O_2-O_2}}(\lambda) \cdot \sigma_{\mathrm{O_2-O_2}} + N^{\mathrm{s}}_{\mathrm{O_3}}(\lambda) \cdot \sigma_{\mathrm{O_3}}, \tag{3}$$

where $\gamma_1 + \gamma_2 \cdot \lambda$ defines the first order polynomial, $\sigma_{\mathrm{O_2-O_2}}$ and $\sigma_{\mathrm{O_3}}$ are the O_2-O_2 absorption cross-section spectrum ~~(at 253)~~ and the O_3 absorption cross section spectrum respectively, convoluted with the OMI slit function, and $N^{\mathrm{s}}_{\mathrm{O_3}}$ is the O_3 slant column density~~and $N^{\mathrm{s}}_{\mathrm{O_2-O_2}}$ is the slant column density.~~ . $\sigma_{\mathrm{O_2-O_2}}$ is based on measurements of the cross section made by C. Hermans (see http://www.aeronomie.be/spectrolab/o2.htm - file O4.txt). The O_3 cross section spectrum is included because it overlaps with the O_2-O_2 spectrum. The fitted parameters are γ_1, γ_2, $N^{\mathrm{s}}_{\mathrm{O_2-O_2}}$, and $N^{\mathrm{s}}_{\mathrm{O_3}}$. In the absence of absorbers, one may define the continuum reflectance R_{c} at the reference wavelength λ_0:

$$R_{\mathrm{c}} = \exp(-\gamma_1 - \gamma_2 \cdot \lambda_0). \tag{4}$$

The reference wavelength is specified as the middle of the DOAS fit window at $\lambda_0 = 475\,\mathrm{nm}$.

~~The OMI cloud algorithm employs, in a second step, a LUT to convert the retrieved $N^{\mathrm{s}}_{\mathrm{O_2-O_2}}$ and R_{c} into effective cloud parameters. In this paper, the aerosol retrieval algorithms differ by replacing this second step by a Neural Network (NN) approach, described in Sect. 3. This approach allows the conversion of the DOAS fit parameters into τ and ALP (in). As a consequence, the NN retrievals rely on the way how these 2 aerosol parameters modify the DOAS variables and thus the photons average light path.~~

3.2 On the impact of aerosols on R_{c} and O_2-O_2 SCD

~~The feasibility to exploit the UV and Vis spectral bands to derive aerosol information has been widely investigated the last 15 years by the ground-based Multi-Axis Differential Optical Absorption Spectroscopy (MAX-DOAS) community. Three types of passive UV-Vis spectral measurements can be related to the presence of aerosols: radiance measurements (Frieß et al., 2006) , measurements of absorption by the complex (Wagner et al., 2004; Vlemmix et al., 2010) , and Ring absorption measurements (Wagner et al., 2009) . The sensitivity to aerosols and their vertical distribution is different for these three types of measurements, due to differences in the underlying physical processes.~~

Figure 1 illustrates how aerosol particles directly drive the OMI O_2-O_2 DOAS parameters at 477 nm assuming cloud-free space-borne observations. These effects are obtained from radiative transfer simulations including aerosols and no clouds. The detailed generation of such simulations is given in Sect. ~~3.2~~4.2. The DOAS fit equations following Eq. (3) and Eq. (4) are then

applied ~~on~~ to these simulations. In this paper, the aerosol layer ~~height (ALH) is expressed by the aerosol layer pressure (ALP)~~is assumed to be one single scattering layer (i.e. "box layer") with a constant geometric thickness (about 1 km). All the particles included in this layer are supposed to be homogeneous (i.e. same size and optical properties). ALH is then expressed by ALP, in hPa, defined as the mid-pressure of ~~an homogeneous and finite aerosol layer .~~

~~The OMI continuum reflectance at 475 R_c is directly~~ this scattering layer.

Qualitatively, aerosols have two separate effects on the average light path, and therefore on the O_2-O_2 absorption signal at the top of the atmosphere (TOA). These two effects are similar to aerosols and clouds have on NO_2 absorption signal (Leitão et al., 2010; Chimot et al., 2016) : 1) a shielding effect, i.e. a decreased sensitivity within and below the aerosol layer due to a reduced amount of photons coming from the TOA and reaching the lowest part of the atmosphere compared to an aerosol-free scene, 2) an enhancement (albedo) effect, i.e. an increased sensitivity within and above the aerosol layer as more photons are scattered back towards the sensors, the part of the atmosphere above the aerosol layer is then sampled by a larger fraction of detected photons. Shielding then leads to a reduced O_2-O_2 absorption while enhancement may increase the O_2-O_2 absorption especially for low cloud or aerosol layers. The overall effect (enhancement vs. shielding) depends on the aerosol optical properties, the total column τ and ALP.

OMI $RC(475\text{nm})$ is directly and primarily affected by the total column τ of ~~fine~~ particles present in the observed scene~~as well as their optical properties (e. g. single scattering albedo ω_0). Indeed, reflectance values increase~~ . Indeed, R_c increases with increasing τ independently of the ALP (cf. Fig. 1a). This mostly results from the influence of aerosols on the ~~ensemble~~ number of detected photons and on the additional scattering effects observed in the scene compared to an aerosol-free scene. However, the magnitude of this increase relies on aerosol optical properties and the surface brightness. As a consequence, R_c ~~does not only depend on the total τ, but results as well from the optical properties of the particles, described for instance by~~ is also affected by aerosol ω_0 ~~and the~~ , phase function, and the surface albedo A. Indeed, R_c decreases with decreasing ω_0 and over a darker surface (i.e. smaller A value) for all the τ values (cf. Fig. 1c and Fig. 1e). The importance of these parameters is further discussed in Sect. ~~4~~ 5 and Sect. ~~5. In~~ 6. Note that, in addition, the reflectance is also driven by the geometry angles: i.e. viewing zenith angles θ, ~~solar zenith angles~~ θ_0 and relative azimuth angle defined as the difference between viewing and zenith azimuth angles $\phi-\phi_0$. An increase of θ or θ_0 will lead to longer average light path, and thus will amplify aerosol related additional scattering effects (for a given τ).

OMI $N^s_{O_2-O_2}$ relies on the O_2-O_2 absorption ~~measurements determine the~~ magnitude along the average light path in the whole atmosphere. It is driven by the overall shielding or enhancement effect of photons by the O_2-O_2 complex in the visible spectral range, due to the presence of particles. ~~absorption increases with increasing path lengths.From that perspective, the aerosol layer height directly influences the absorption since particles at higher altitude partly shield the complex located below~~ As depicted in Fig. 1b, $N^s_{O_2-O_2}$ decreases with decreasing ALP. This is a direct consequence of a larger shielding effect applied by aerosols located at higher altitudes (i.e. ~~decrease of the length of the average light path). Such an effect has a direct impact on the~~ part of the O_2-O_2 ~~slant column density (SCD) $N^s_{O_2-O_2}$ values . As observed in Fig. 1 and also reported by Wagner et al. (2004); Park et al. (2016); Chimot et al. (2016)~~ complex located below the aerosol layers are shielded). Nevertheless, in case of low τ values (i.e. ≤ 0.5), $N^s_{O_2-O_2}$ ~~generally increases with increasing aerosol layer pressure~~

~~(or decreasing aerosol layer height).~~ does not significantly vary with respect to ALP. This shows that a low amount of aerosols has very little impacts on O_2-O_2 absorption measurements.

However, as depicted in Fig. ~~1c,~~ 1b, 1d and 1f, not only ALP but also τ directly influences the slant O_2-O_2 absorption since both parameters simultaneously affect the ~~photon path distribution~~avergae path followed by the photons, and therefore the overall shielding or enhancement effect. An increase of τ ~~, for particles with $\omega 0 \leq 0.95$,~~ leads to a decrease of $N^{s}_{O_2-O_2}$. The slope of this decrease depends on the aerosol altitude (i.e. higher for particles at high altitude). ~~This means that τ also drives the shielding effect, while the aerosol height does not significantly affect R_c. As a consequence: the 2 parameters R_c and~~ Note that both ω_0 and A also affect $N^{s}_{O_2-O_2}$ ~~have a non negligible correlation (see~~, but this effect is smaller than τ. For example, a reduced ω_0 and A lead to a small decrease of $N^{s}_{O_2-O_2}$ (cf. Fig. ~~1b) as they share, in part, same information: the length of the average light path and the aerosol optical thickness τ.They are not two independent variables and combining them does not provide with two independent pieces of aerosol information.~~ 1d and Fig. 1f).

As a consequence:

- the ~~variable~~ single parameter $N^{s}_{O_2-O_2}$ ~~contain both information ,~~ contains information on both τ and ALP~~that cannot be easily separated . As a consequence~~. These parameters cannot be separated from this unique variable alone. Therefore, if τ is not accurately known, there will likely be an ambiguity when analysing $N^{s}_{O_2-O_2}$ to retrieve ALP~~since contributions from both aerosol height and absorption or scattering of the sunlight due to aerosol amounts cannot be separated. Moreover, potential uncertainties in surface albedo and aerosol optical properties may amplify this ambiguity and thus make the retrieval problem more complex. To avoid these constraints, it may be then relevant to test the ALP retrievals by replacing R_c with a prior .~~

- if an external or prior τ estimate is not available, then the two parameters $N^{s}_{O_2-O_2}$ and $R_c(475nm)$ could be simultaneously combined to retrieve ALP provided that one can accurately and independently retrieve τ from $R_c(475nm)$. Then, in that condition, OMI $R_c(475nm)$ may help to distinguish both τ and ALP contributions in $N^{s}_{O_2-O_2}$. However, the simultaneous effects of aerosol ω_0 and A on R_c (as discussed above), and therefore their associated uncertainties, will impact the feasibility of retrieving τ from OMI measurements. It may then degrade the retrieved ALP performances.

- τ, ~~provided this information is known with a high enough accuracy (e. g. MODIS~~ R_c and $N^{s}_{O_2-O_2}$ have a non negligible correlation. Indeed, an increase of τ ~~).~~ results in a simultaneous increase of R_c and $N^{s}_{O_2-O_2}$. Therefore, it has to be noted that these two last parameters are not independent and combining them does not provide with two independent pieces of information.

Overall, the impact of aerosol particles on the OMI O_2-O_2 spectral band is similar to cloud particles. This explains in part the difficulty to distinguish aerosols from clouds. In cases with a mix of aerosols and clouds, there is an ambiguity between R_c, ~~the~~ τ and the OMI effective cloud fraction on the one hand, and $N^{s}_{O_2-O_2}$, ~~the aerosol layer pressure~~ALP, τ, the OMI effective cloud pressure and fraction on the other hand (Boersma et al., 2011; Castellanos et al., 2015; Chimot et al., 2016). Therefore, this study only focuses on cloud-free reflectance to avoid this complexity.

4 Design of the neural network retrieval algorithms

The retrieval algorithms developed for this paper are based ~~from~~ on the Pybrain software (Schaul et al., 2010). Pybrain is a versatile machine learning library written in Python designed to facilitate both the application of and research on premier learning algorithms such as recurrent NNs. It includes several functions such as supervised learning algorithms, feed forward network design and error back-propagation computations. Only the main developments specific to the present study are explained and discussed in the next sub-sections. For more details related to the Pybrain specificities, the reader is encouraged to read Schaul et al. (2010).

The Multilayer Perceptron (MLP) neural networks have been widely used and acknowledged for decades in the field of remote sensing (Atkinson and Tatnall, 1997). Indeed, most retrieval problems in this field are ill-posed and nonlinear~~, in particular the atmospheric ones~~. Thus, the associated inverse problems can only be addressed by including a priori information and relying on statistical analysis. Since aerosol retrieval from passive spectral measurements is well-known as a non-linear inverse problem, the MLP technique represents then a powerful approach to design a retrieval algorithm in a fast and robust way. The basic idea is to build an optimal interpolator system making the link between the OMI 477 nm $\mathrm{O_2}-\mathrm{O_2}$ measurements and the retrieved ~~aerosol layer pressure and optical thickness (i. e. between the ellipses of~~ ALP and τ (cf. Fig. 1). However, knowledge must be acquired by the network by means of a supervision database. The following sections summarize then the design of the developed algorithms (see Sect. ~~3.1~~4.1), the generated supervision database (see Sect. ~~3.2~~4.2) and the employed learning process (Sect. ~~3.3~~4.3).

4.1 Multilayer Perceptron network approach: application to the OMI $\mathbf{O_2-O_2}$ aerosol retrievals

Artificial neural Networks are a family of models related to the machine learning and the artificial intelligence domain (Luger and Stubblefield, 1998). They are used to reduce the number of calculations of functions requiring a large number of inputs and being generally unknown (or not well defined). The idea is to approximate them by parameterized and more simple functions. Input and output signals are then interconnected by a set of activation functions and a set of weights associated with each of them (Luger and Stubblefield, 1998). In the context of this work, ~~the link between the OMI measurements and aerosol parameters cannot be exactly described with accurate functions~~ no invertible analytical function exist that describes the dependence of slant columns and continuum reflectances on aerosols (see Sect. 2.3). Thus, the idea of developing neural networks here is to identify input-output relationships directly from a well-known training ensemble. The choice of a NN approach relies on their advantages compared to more conventional methods such as linear regression, linear interpolation in a LUT or the Optimal Estimation Method (OEM). In particular, it enables 1) very fast computations with modern computers in spite of the number of required parameters, 2) optimized interpolation technique even in case of ~~non linear~~ non-linear statistical modelling and so potentially lower systematic biases compared to a linear interpolation, 3) reduced memory use compared to a LUT with a very high sampling.

As illustrated in Fig. 2, the designed NNs rely on a multi-layer architecture, based on the Multilayer Perceptron technique~~(Rumelhart et al., 1986)~~, composed of parallel processors (i.e. neurons) organized in distinct layers

(Rumelhart et al., 1986) . Such an architecture allows to separate non-linear data, and generally consists of 3 (or more) types of layers. The first layer includes all the required input variables. The last layer includes all the desired output data (or here retrievals). The intermediate layers are usually referred as hidden layers and contain the activation functions. All these layers are connected via neural links: two nodes or neurons i and j between two consecutive layers have synaptic connections associated with a synaptic weight ω_{ij}. Each neuron j computes a weighted sum of its N x_i information sent from the neurons of the previous layer. Then, this weighted sum is transported through a non-linear mathematical function before being passed to the next layer. Here use is made of the classical sigmoid function:

$$\sigma(x) = \frac{1}{1+\exp(-x)}. \tag{5}$$

The output z_j of the neuron j in the hidden layer is thus given by:

$$z_j = \sigma(\sum_{i=1}^{N} \omega_{ij} \cdot x_i). \tag{6}$$

The overall set W of synaptic weights ω_{ij} contain all the information about the network (i.e. its neural architecture defined by a specified number of layers, neurons and connections). When the transport reaches the output layer, it forms the network output.

The chosen neural architecture is the following (see Fig. 2). The input layer is composed of 7 parameters that include (1) θ, θ_0, $\phi - \phi_0$, surface pressure Ps, surface albedo A, and (2):

- either ~~R_c~~ $R_c(475\text{nm})$ and $N^s_{O_2-O_2}$ for ~~τ~~ $\tau(550\text{nm})$ and ALP retrievals: this configuration is named $NN_{R_c,N^s_{O_2-O_2}}$,
- or ~~τ~~ $\tau(550\text{nm})$ and $N^s_{O_2-O_2}$ for ALP retrieval: this configuration is named $NN_{\tau,N^s_{O_2-O_2}}$.

The output layer is, for each NN retrieval algorithm, composed of only one output variable: either τ, or aerosol layer pressure ALP. In total, three NN retrieval algorithms are then selected and used at the end: one $NN_{R_c,N^s_{O_2-O_2}}$ for ~~τ~~ $\tau(550\text{nm})$ retrieval, one $NN_{R_c,N^s_{O_2-O_2}}$ and one $NN_{\tau,N^s_{O_2-O_2}}$ for ALP retrieval.

The choice to use either $NN_{R_c,N^s_{O_2-O_2}}$ or $NN_{\tau,N^s_{O_2-O_2}}$ will impact the accuracy of the ALP retrieval results (see Sect. ~~4.2,~~ 5.2 and 5.3).

4.2 Generation of the supervision database: aerosol properties and simulations

The MLP neural networks must be ~~trained in order to build models that can learn~~ accurately trained from a well-known data set~~, and then~~ . They are then able to generalize the inverse problem by predicting the aerosol retrievals from input observations that have never been seen before. For that purpose, a learning database must be carefully designed and generated. It must be representative of the entire distribution of (input-output) values that can likely be encountered in the OMI observations. As a consequence for the MLP algorithms, a large quantity of data is often required for the learning process. However, very large learning data set can be extremely time consuming in terms of generation and then NN training.

~~The NNs are here trained with simulated data~~Training a neural network based on a large ensemble of synthetic dataset has been widely employed in atmospheric retrieval science such as for CO_2 and CH_4 (**??**) , aerosol (**?**) and cloud properties

(???) . This study ~~uses~~ created our own training dataset based on simulations from the Determining Instrument Specifications and Analyzing Methods for Atmospheric Retrieval (DISAMAR) software of KNMI (de Haan, 2011). DISAMAR includes a radiative transfer model and different retrieval methods. The radiative transfer model is based on the Doubling Adding KNMI (DAK) model (de Haan et al., 1987; Stammes, 2001) and thus computes the reflectance and transmittance in the atmosphere using the adding/doubling method. This method calculates the internal radiation field in the atmosphere at levels to be specified by the user and takes into account Rayleigh, aerosol and cloud scattering and trace gas and aerosol absorption. Scattering by aerosols is simulated with a Henyey–Greenstein scattering phase function $\Phi(\Theta)$ (Hovenier and Hage, 1989):

$$\Phi(\Theta) = \frac{1-g^2}{(1+g^2-2g\cos\Theta)^{3/2}}. \quad (7)$$

where Θ is the scattering angle. The phase function is then parameterized by the asymmetry parameter g, which is the average of the cosine of the scattering angle, such its variation $-1 \leq g \leq 1$ ranges from back-scattering through isotropic scattering to forward scattering. Following the DISAMAR configuration, τ values in the simulations are specified at the ~~referent~~ reference wavelength of 550 nm. The Ångström exponent α describes the spectral dependence of τ.

ALP is the main target parameter since this is one of the main parameters describing the average light path distribution in the tropospheric NO_2 AMF computation. The second target is τ since this information may be requested for a good ALP retrieval quality. We thus assume we do not need at this level to define more realistic aerosol models for every aerosol scene. With a ~~referent~~ reference asymmetry parameter of $g = 0.7$, intermediate value typically observed (Dubovik et al., 2002), the Henyey-Greenstein function is known to be smooth and reasonably well reproduce the Mie scattering functions for most of aerosol types. This approach is also used for the preparation of the operational aerosol layer height retrieval algorithm from Sentinel-5 Precursor (Sanders et al., 2015) and for explicit aerosol corrections in the AMF calculation when retrieving trace gases such as tropospheric NO_2 (Spada et al., 2006; Wagner et al., 2007; Castellanos et al., 2015).

The ensemble of parameters and associated values used for generating the learning database is detailed in Table 1. About 460 000 spectral simulations, over the O_2-O_2 spectral band (460-490 nm), were generated, assuming different satellite viewing and solar geometries, ~~surface albedo, surface pressure~~ A, Ps and aerosol pollution levels. Scenes with too large angles (i.e. $\theta_0 \geq 65°$) and too bright surfaces (i.e. $A > 0.1$) are excluded. For each of these simulations, ~~R_c~~ $R_c(475\text{nm})$ and $N^s_{O_2-O_2}$ were deduced from the DOAS fit equations Eq. (~~2~~3) and Eq. (~~3~~4). Aerosols are specified for a standard case, assuming fine particles with a unique value of ~~Ångström coefficient~~ $\alpha = 1.5$ and $g = 0.7$. Aerosol profiles are parameterized by scattering layers with constant aerosol volume extinction coefficient and ~~aerosol single scattering~~ ω_0 and with a fixed pressure thickness. ~~The ALP corresponds to the mid-pressure of the assumed scattering aerosol layer.~~ In order to investigate the assumptions related to the single scattering albedo properties ω_0, two typical values are considered: $\omega_0 = 0.95$ and 0.9. Contrary to the other variables, ω_0 is not known for each OMI pixel and thus cannot be used as an explicit input parameter by the designed NNs. Moreover, it cannot be retrieved from this band since it is supposed to affect ~~R_c~~ $R_c(475\text{nm})$ and $N^s_{O_2-O_2}$, similarly as τ. Therefore, 2 sets of NN models are developed for different purposes:

- one set of three MLP NN algorithms ($NN_{R_c,N^s_{O_2-O_2}}$ for τ retrieval, $NN_{R_c,N^s_{O_2-O_2}}$ and $NN_{\tau,N^s_{O_2-O_2}}$ for ALP retrieval) is trained assuming $\omega_0 = 0.95$;

– one set of three MLP NN algorithms ($NN_{R_c,N^s_{O_2-O_2}}$ for τ retrieval, $NN_{R_c,N^s_{O_2-O_2}}$ and $NN_{\tau,N^s_{O_2-O_2}}$ for ALP retrieval) is trained assuming $\omega_0 = 0.9$.

The choice to use one of these sets will impact the accuracy of the retrieval results.

4.3 Optimization of the learning process and selection of the best NN architecture

Prediction of the optimal number of neurons and the hidden layers is generally not possible as these values are strongly specific to the given problem (Atkinson and Tatnall, 1997). While it has been theoretically proven that a single hidden layer MLP networks with ~~non linear~~ non-linear activation functions may represent any nonlinear continuous function (Haykin, 1999), a 2-hidden layer MPL may approximate any function to any degree of non-linearity taking also into account discontinuities (Sontag, 1992). To identify the best NN architecture for each aerosol retrieval parameter (τ and ALP) and for each configuration ($NN_{R_c,N^s_{O_2-O_2}}$ or $NN_{\tau,N^s_{O_2-O_2}}$, see Sect. 3.1), several architectures are trained and then evaluated: one single hidden layer with a variable number of neurons (between 9 and 70) and two hidden layers (between 15 and 70 neurons on the 1st layer, and between 10 and 40 on the 2nd layer). Then, the ~~most relevant~~ optimal NN architecture is selected based on the best computed evaluation score. In total, about 96 different MLP architectures, for each configuration, were evaluated.

For one given NN architecture, the training process is the optimization technique that estimates the optimal network parameters W of synaptic weights σ_{ij} (see Sect. 3.1). For that purpose, ~~a positive-definite~~ an error function E must be minimized. This error function measures, for a set of p representative situations, for which inputs and outputs (i.e. τ and ALP) are known, the mismatch between the neural network outputs o_i and the true outputs t_i as follow:

$$E = \frac{1}{2}\sum_{i=1}^{p}(o_i - t_i)^2. \tag{8}$$

This error function minimization follows here the Error Back-Propagation algorithm as specified by Rumelhart et al. (1986). It is a stochastic steepest descent algorithm well adapted to the MLP hierarchical architecture. The learning step is made sample by sample, iteratively and stochastically selected in the training data set. The network is initialized with random synaptic weights. At each iteration, the error is computed and then propagated backwards from the output layer. The processes of error back-propagating and feeding forward signals are repeated iteratively until the error function is minimized or the maximum number of iterations is reached (i.e. 500).

During the training phase, the considered network architecture must obtain an optimal generalization performance: i.e. the network performance should not degrade significantly when data set other than the training one is analyzed. Standard NN architectures, like the fully connected MLP, generally have a too large parameter space, and are prone to over-fitting. Although the network performance seems to constantly improve on the training sets at each iteration, it can actually begin to worsen (in terms of errors) on unseen data set. Therefore, a verification step is performed, over the last 15 iterations, to detect this overfitting moment (i.e. no significant variation of E) and stop the training phase. This process is called early stopping. Finally, to ensure that the system is not trapped in local minima during the ~~cost~~ error function minimization, the learning phase (training+verification) is repeated three times, the synaptic weights being randomly initialized at the beginning of each training

phase. The network system presenting the best evaluation score (see Eq. ~~7~~8) is then selected. All these precautions are carried out by randomly splitting the learning data (see ~~Tab.~~Table 1) into 3 independent sets: training, verification and evaluation. They respectively consisted of 70 %, 15 % and 15 %.

Figure 3 depicts the box-whisker plots of the square of errors obtained over the ensemble of training-verification-evaluation data set for τ and ALP retrieval and for $NN_{R_c,N^s_{O_2-O_2}}$ configuration, assuming ~~aerosol ω_0 of 0.95~~$\omega_0 = 0.95$. For τ retrievals, although the NNs with 40 and 70 neurons on one hidden layer do reasonably well, the scores show improved values when two hidden layers are used. The ALP retrieval scores are significantly larger than for τ (a factor of 100). This is a direct consequence that ALP is less well constrained by the spectral measurements: lower pieces information are available compared to τ, in particular for scenes dominated by low τ values. While the NNs with one hidden layer do not show any significant improvements when increasing number of neurons, better scores are only obtained with 2 hidden layers. Overall, the similar behavior of training-verification-evaluation scores validate that the trained NNs are generalized enough to be able to reproduce similar variation of the scores on other independent data set. The identified best NN algorithms are thus found with 2 hidden layers, ~~with~~ including between 25 and 70 neurons on the first layer, and between 10 and 20 neurons on the second layer depending on the considered configuration (see Sect. 2.3) and retrieved parameter.

5 Sensitivity analyses on synthetic data set

The robustness of the trained and selected NN retrieval algorithms is assessed by applying them to independent simulations, not present in the learning (training-verification-evaluation) database. Simulated ~~reflectance~~ spectra are noise-free and only include aerosol particles (no clouds). The sensitivity of ~~τ~~ $\tau(550\text{nm})$ and ALP retrievals is verified for different ~~surface albedo,~~ A and aerosol properties (ω_0, τ, ALP). ~~R_c~~ $R_c(475\text{nm})$ and $N^s_{O_2-O_2}$ are derived from the spectra and provided as inputs to the NNs. The impact of uncertainties on surface albedo, aerosol model and $N^s_{O_2-O_2}$ are analysed. Consistent geophysical conditions (temperature, NO_2 and O_3 profiles) are considered between these simulations and those included in the learning database. All the analyses performed here are summarized in ~~Tab.~~Table 2.

5.1 Aerosol Optical Thickness retrievals

Figure 4 ~~depicts the derived τ values from the simulated spectral measurements, compared with the true values, and how much inaccurate assumptions about the aerosol single scattering albedo~~ compares the retrieved to the true $\tau(550\text{nm})$ values of the simulated spectra, and how uncertainties on ω_0~~or the asymmetry parameter~~ , g and A degrade the retrieval quality. ~~Figure 4a shows that~~ Overall, retrieved and true ~~τ~~ $\tau(550\text{nm})$ values are very well correlated for all the types of surface, assuming no error in the assumed surface albedo and aerosol properties. This confirms the success of the learning process implemented in Sect. ~~3.2~~ 4.2 and Sect. ~~3.3~~ 4.3 and the use of the NN approach.

The assumed aerosol properties (~~single scattering albedo~~ ω_0 and phase function through ~~the asymmetry parameter~~ g), and so the choice of the trained NN algorithm, are of high importance. They change the slope between retrieved and true τ values and drastically affect the retrieved τ accuracy. If the ~~true aerosol~~ assumed ω_0 ~~value is lower ($\omega_0 = 0.9$) than in the simulations~~

of (i.e. 0.95) through the simulations in the learning database ($\omega_0 = 0.95$is too high (true $\omega_0 = 0.9$), retrieved τ is smaller.then underestimated: i.e. a bias of -0.8 for $\tau(550\text{nm}) = 2$, and -0.1 for $\tau(550\text{nm}) \leq 0.5$ (see Fig. 4a). This is a direct result consequence of reduced scattering efficiency and thus as more photons are absorbed instead of being scattered back towards the satellite sensor(see . The measured $R_c(475\text{nm})$ is then lower (cf. Fig. 4a). The bias reaches a maximum value of about 0.8 for τ1). Reciprocally, an assumed too low aerosol ω_0 (i.e. true $\omega_0 = 1.0$) leads to an underestimation of retrieved $\tau(550\text{nm})$: i.e. a bias of 1.0 for $\tau(550\text{nm}) = 2$, and is lower than 0.1 for $\tau \leq 0.5$ (see (cf. Fig. 3b). This is directly related to the impact of aerosols and their optical properties on the measured R_c. 4a).

Figure 4b illustrates retrieved τ $\tau(550\text{nm})$ bias due to the difference between the assumed asymmetry parameter g in the learning database and in the synthetic spectra. This gives the direct impact of uncertainties associated with the aerosol scattering phase function characterization. While $g = 0.7$ is the referent reference value for most of aerosols, scenes with $g = 0.6$ are related to finer and weakly absorbing particles with a somewhat reduced forward scattering direction such as carbonaceous aerosols, desert dust and volcanic dash models as given by the ESA aerosol CCI-project (de Leeuw et al., 2013). Values of $g = 0.8$ are associated with larger particles and an increased forward scattering direction such as cirrus (Sanders et al., 2015). An overestimation of g (i.e. true assumed g = 0.6 while assumed 0.7 while true g = 0.70.6) leads to an increased retrieved τ value (i.e. positive bias) as a result of because of less photons scattered towards the surface, and therefore more photons scattered back towards the satellite sensor, and less towards the surfacecompared to what is theoretically assumed. Reciprocally, an underestimation of g (i.e. true assumed g = 0.8 while assumed 0.7 while true g = 0.70.8) leads to a decreased retrieved τ value (i.e. negative bias) as a result of due to less photons scattered back towards the satellite sensor, and more towards the surface compared to the assumption. Absolute bias values can exceed 0.5 for $\tau = 1.5$ $\tau(550\text{nm}) = 1.5$ while they stay close to 0.25 for $\tau = 0.5$. Overall, uncertainties in aerosol models can drastically impact the retrieved τ accuracy and modify the slope between retrieved and true τ values $\tau(550\text{nm}) = 0.5$.

Figure 5a depicts the box-whisker distribution of τ biases due to uncertainties Errors in surface albedo . Estimations are obtained for fine and scattering particles (α=1.5, ω_0=0.95, g=0.7). The also lead to biases in retrieved $\tau(550\text{nm})$ (cf. Fig. 4c). Overall, biases are larger over scenes with small $\tau(550\text{nm})$ values. The reason is the dominance of surface reflection in this regime. In cases of high amount of aerosols, then aerosol scattering signals are dominant and surface reflection uncertainties have less impacts. An underestimated (overestimated) surface albedo results in a negative (positive) retrieved $\tau(550\text{nm})$ bias. This is directly related to the change in the measured OMI Rc (cf. Fig. 1). Surface albedo uncertainties in the range of 0.025-0.05 lead to absolute biases close to 0.5 for τ uncertainty is obtained from the half of τ differences between adding and deducting uncertainties of the variables as follows:

$$\epsilon(\partial x) = \frac{1}{2} \mid \tau(x + \partial x) - \tau(x - \partial x) \mid .$$

where $\epsilon(\partial x)$ is the τ bias due to the variable error source x (here surface albedo), ∂x is the uncertainty applied to x. $\epsilon(\partial x)$ values are computed for all combinations of surface albedo 0.03-0.05-0.07and θ_0-θ=.

Errors in surface albedo lead to biases in derived τ values. Retrieved τ bias lies in the range of 0.04-0.48 on average, for surface albedo uncertainties ∂x_{alb} between 0.005 and 0.05 (see Fig. 5a). To limit the τ bias values to an average of 0.2, ∂x_{alb}

~~should not exceed 0.025. Typical differences of~~ 0.0-0.5, smaller than 0.25 for $\tau = 2.0$. Typical differences in climatological surface albedo ~~between that obtained~~ from the total ozone monitoring spectrometer (TOMS) and the global ozone monitoring experiement (GOME) (Koelemeijer et al., 2003), or between ~~the Lambertian equivalent reflectance (LER) of OMI~~ OMLER and the MODIS black sky albedo (Kleipool et al., 2008), are known to be up to 0.02. ~~It was observed that surface albedo uncertainty mostly impacts cases with low τ (lower than 0.5), as emphasized as well by Park et al. (2016) . The measured R_c over scenes with low τ is a combination of aerosol scattering effects and surface reflection. But, in case of high τ, aerosols strongly attenuate the surface contribution, and their scattering proprieties dominate. Differences in the standard deviation lie in the range of 0.07-0.17 depending on the surface albedo uncertainty.~~

5.2 Aerosol Layer Pressure retrievals

Figure ~~6~~ 5 and Figure ~~7~~ 6 depict the expected performances of the developed NN algorithms for Aerosol Layer Pressure (ALP) retrievals. ~~For scenes dominated by τ~~ Over scenes with $\tau(550\text{nm})$ in the range of 0.5-1.0, and assuming no error in prior A and the employed aerosol model, ALP retrievals are ~~quite~~ relatively stable presenting biases close to 100 hPa. Only for ~~$\tau \geq 1.0$~~ $\tau(550\text{nm}) \geq 1.0$, biases are smaller than 50 hPa. The accuracy of the retrieved ALP generally increases with increasing τ. Indeed, assuming true ALP of 850 hPa ~~and no error in the employed surface albedo and aerosol model~~ (see Fig. ~~7~~6), positive biases larger than 400 hPa are found for ~~$\tau \leq 0.5$~~ $\tau(550\text{nm}) \leq 0.5$. Note that this behavior is observed for all the NN configurations ($NN_{R_c,N^s_{O_2-O_2}}$ and $NN_{\tau,N^s_{O_2-O_2}}$). A box-whisker plot, in Figure ~~8~~7, illustrates the variability of the ALP NN biases as a function of τ over all the simulations contained in the entire learning database (as defined and used in Sect.3). This confirms that, in spite of the strict training-verification-evaluation process achieved in Sect. ~~3.3~~4.3, the NN ALP retrievals are not expected to be accurate for small τ values, especially below 0.5. The reason is directly linked to the magnitude of the O_2-O_2 shielding effect and its ~~dependence on the~~ combined dependence on aerosol amount (~~in addition to the aerosol altitude). Low~~ or τ) and aerosol altitude (cf. Fig. 1 and Sect. 3.2). Because low amount of aerosols have very limited effects on the O_2-O_2 absorption (see ~~Fig. 1). Thus~~Sect. 3.2), even advanced interpolation techniques like NNs have difficulties for these cases. When τ increases, the O_2-O_2 shielding effect amplifies and the algorithms are more able to link ~~it to the ALP variation~~the O_2-O_2 absorption signal to ALP. Overall, even for small ~~τ~~ $\tau(550\text{nm})$ values (like 0.5), the retrieved aerosol pressures correlate with the aerosol layer height in spite of very poor accuracy (see Fig. ~~7a~~ 6a and Fig. ~~7c~~6c) .

~~The advantages of using~~ A very accurate prior τ information as input ~~instead of R_c(475) are multiple, provided this information is given with a high enough accuracy~~is required to generally improve the ALP retrieval performances. As explained in Sect. 3.2, using $NN_{R_c,N^s_{O_2-O_2}}$ implicitly relies on the τ retrieval capability from OMI $Rc(475\text{nm})$. Using $NN_{\tau,N^s_{O_2-O_2}}$ with the true τ value as input (thus no error) is expected to deliver higher performances. Firstly, it allows to improve the accuracy of the ALP retrieval ~~for cases dominated by~~ over scenes with low τ and particles located at high altitude (above 800 hPa or 2 km). Indeed, ~~on~~ in Fig. ~~6b~~5b, for $\tau = 0.5$ and ALP between 750 and 850 hPa, ALP biases are reduced from ~~250:350to 150:250~~250-350 hPa with $NN_{R_c,N^s_{O_2-O_2}}$ to 150-250 hPa with $NN_{\tau,N^s_{O_2-O_2}}$ and true $\tau(550\text{nm})$ value. For particles higher than 650 hPa (or 3.5 km), no improvements are however observed. The low sensitivity to retrieve ALP when particles are

located at a very high altitude is directly due to the O_2-O_2 complex and its vertical distribution. This was demonstrated by Park et al. (2016): O_2-O_2 concentration exponentially decreases with increasing atmospheric altitude.

Secondly, impacts due to uncertainties ~~related to~~ on the chosen surface albedo and aerosol model are reduced. Assumptions on aerosol ω_0 drive the interpretation of the shielding of the O_2-O_2 dimers by aerosols. ~~If the assumed~~ ω_0 ~~is too high (i.e. assumed ω_0 = 0.95 while ω_0 true = 0.9),~~ can perturb ALP retrievals obtained with ~~the~~ $NN_{R_c,N^s_{O_2-O_2}}$ ~~configuration are perturbed by~~ more than 100 hPa (see Fig. ~~6a~~5a). These perturbations are reduced to the range of 0-100 hPa over scenes with high τ values (larger than 1) only for particles close to the surface, i.e. true $ALP \geq 850$hPa (see Fig. ~~7a~~ 6a and Fig. ~~7c~~6c). Using ~~the~~ $NN_{\tau,N^s_{O_2-O_2}}$ ~~configuration, assuming true τ as input,~~ with true $\tau(550\text{nm})$ value helps to mitigate theses biases. ~~The ALP retrievals are almost superposed to those with true ω_0 assumptions and~~ All the ALP retrievals present the same behaviors with respect to the particles altitude and τ and biases lie in the range of 0-50 hPa(see Fig. ~~6b~~5b). Similar conclusions are observed regarding uncertainties on ~~the asymmetry parameter. Higher~~ g (cf. Fig. 5c and Fig. 5d). Too high g values impact the ALP retrievals from $NN_{R_c,N^s_{O_2-O_2}}$ over scenes with $\tau \leq 1.0$. Such a bias is largely reduced with the $NN_{\tau,N^s_{O_2-O_2}}$ configuration.

Surface albedo ~~variability contributes to~~ contributes the length of the average light path and thus ~~leads to changes in~~ affects $N^s_{O_2-O_2}$. Retrieved ALP ~~bias (Eq. 11) varies between 23~~ biases are maximum (several hundreds hPa~~and 141 on average, for surface albedo uncertainties ∂x_{alb} in the range of 0.005-0.05 with the $NN_{R_c,N^s_{O_2-O_2}}$ configuration (see~~) for $\tau(550\text{nm}) \leq 0.5$ (cf. Fig. ~~6b). To limit the values below 100~~ 8a and Fig.8b). For $\tau(550\text{nm})$ in the range of 0.5-1.0, retrieved ALP are impacted by lower absolute values (between 50 hPa ~~on average, ∂x_{alb} should not exceed 0.025. Using the $NN_{\tau,N^s_{O_2-O_2}}$ configuration with true τ value, not only reduces the average but also the standard deviation of the bias: i.e. for ∂x_{alb} = 0.05, the aerosol pressure bias decreases from 141 ± 96~~ and 100 hPa on average) with $NN_{\tau,N^s_{O_2-O_2}}$, while they remain too high with $NN_{R_c,N^s_{O_2-O_2}}$ ~~to 69 ± 79~~. Over scenes with $\tau(550\text{nm}) \geq 1.0$, biases are reduced to 0-50 hPa ~~with $NN_{\tau,N^s_{O_2-O_2}}$. The reasons~~ since aerosol scattering signals dominate over surface reflection. The main cause of all these improvements ~~are: 1) the strong correlation between R_c and $N^s_{O_2-O_2}$ provides with a limited number of independent pieces of information to the retrieval system (see Sect. 2.3), 2) uncertainties about the surface albedo and aerosols ω_0 impactthe interpretation of both of R_c and $N^s_{O_2-O_2}$ and mislead then the interpolation process, 3) using~~ is that using an accurate prior τ ~~as input is more consistent with the definition of aerosol models in the learning database and allows a better distinction the effects of τ and ALP on the~~ information (or at least more accurate than retrieved OMI $\tau(550\text{nm})$ from $Rc(475\text{nm})$) allows a better distinction of τ and ALP effects on the O_2-O_2 slant column density.

An accuracy better than 0.2 must be required on prior τ information (cf. Fig. 8c). Indeed, a $\tau(550\text{nm})$ bias of 0.25 can impact, in absolute, the retrieved ALP up to 50 hPa for $\tau(550\text{nm})$ in the range of 0.6-1.0. For $\tau(550\text{nm}) \geq 1.0$, impact on ALP becomes almost null. Therefore, using MODIS τ as prior to $NN_{\tau,N^s_{O_2-O_2}}$ is likely expected to show retrieved ALP with a higher quality than with $NN_{R_c,N^s_{O_2-O_2}}$. Indeed, the current retrieved OMI $\tau(550\text{nm})$ from $Rc(475\text{nm})$ does not present a better accuracy than MODIS $\tau(550\text{nm})$.

Figure 9 depicts the box-whisker distribution of ALP precision $\epsilon(\partial N^s_{O_2-O_2})$ due to $N^s_{O_2-O_2}$ precision. Estimations are obtained for fine and scattering particles ($\alpha = 1.5$, $\omega_0 = 0.95$, $g = 0.7$). $\epsilon(\partial N^s_{O_2-O_2})$ is obtained from the half of ALP

differences between adding and deducting uncertainties of the variables as follows:

$$\epsilon(\partial N^{s}_{O_2-O_2}) = \frac{1}{2} \mid ALP(N^{s}_{O_2-O_2} + \partial N^{s}_{O_2-O_2}) - ALP(N^{s}_{O_2-O_2} - \partial N^{s}_{O_2-O_2}) \mid . \quad (9)$$

where $\partial N^{s}_{O_2-O_2}$ is the uncertainty applied to $N^{s}_{O_2-O_2}$. $\epsilon(\partial N^{s}_{O_2-O_2})$ values are computed for all combinations of surface albedo 0.03-0.05-0.07 and θ_0-θ = [25°-25°, 50°-25°, 25°-45°]. The reason to use this aproach here is that, since $N^{s}_{O_2-O_2}$ precision is a random error (opposite to systematic), it will directly impact the retrieved ALP precision instead of leading to a systematic bias. A precision of $N^{s}_{O_2-O_2}$ lying in the range of 0.05-0.25 10^{-43} $\mathrm{mol^2 cm^{-5}}$ ~~results in similar ALP bias~~ (i.e. at a first order, 2-7 % of $N^{s}_{O_2-O_2}$) results in ALP uncertainties between 19 ± 29 hPa and 57 ± 31 hPa on average for both NN configurations (see Fig. 5c). ~~Note a temperature correction, for real observations, must be taken into account to correctly interpret $N^{s}_{O_2-O_2}$ (see Sect. 5.1).~~

Overall all the estimated NN retrieval uncertainties are in line with the theoretical sensitivity analyses of Park et al. (2016) who found that the O_2-O_2 at 477 nm is significantly influenced by aerosol optical properties (including ~~single scattering albedo~~ω_0), τ, particle size and ~~surface albedo~~A. In particular, a ω_0 uncertainty of 10 % was demonstrated to lead to the aerosol effective height (AEH) retrieval error ranging from 270 to 1440 m, depending on the aerosol types. Errors were found larger for high particle altitude and low τ cases. A surface albedo uncertainty of 0.02 was expected to impact AEH retrievals between 154 m and 434 m on average. AEH error was frequently larger only for low ~~τ~~ $\tau(550\mathrm{nm})$ (≤ 0.4) and high AEH (≥ 1 km).

6 Application to OMI observation measurements

6.1 Methodology

Aerosol retrievals, as described in the previous sections, are performed on the OMI O_2-O_2 observations over large industrialized continental areas in North-East Asia over 3 years~~(, 2005-2007)~~, and cloud-free scenes. All the associated results are summarized in ~~Tab.~~Table 2. The considered North-East Asia area is defined by the range of latitude 25-40 °North, and longitude 110-130 °East excluding the part over the Gobi desert which presents a too bright surface (as further explained below).

Only OMI observations collocated with MODIS Aqua ~~Level 2 (~~L2 ~~)~~ aerosol product collection 6 ~~(Xiao et al., 2016) are considered . Their spatial resolution, 10 x 10 , is relatively close to the OMI nadir spatial resolution.The MODIS instrument, launched on the NASA EOS-Aqua platform in May 2002, is a spectrometer delivering continuous images of the Earth in the visible, solar infrared and thermal infrared approximately 15 min prior to OMI on-board EOS-Aura. The improved calibration of MODIS Aqua instrument is included in the reprocessing of the collection 6 aerosol product (Levy et al., 2013; Lyapustin et al., 2014) . The reason to consider the MODIS aerosol products is here~~ are considered (cf. Sect. 2.2). The reason is triple: 1) to maximize the probability of the selection of cloud-free OMI observation pixels dominated by aerosol pollution, 2) to evaluate the retrieved OMI ~~τ~~ $\tau(550\mathrm{nm})$ products by comparing with collocated MODIS ~~τ(550)~~$\tau(550\mathrm{nm})$, 3) to use the MODIS ~~τ(550)~~ $\tau(550\mathrm{nm})$ as input of the $NN_{\tau, N^{s}_{O_2-O_2}}$ algorithm for retrieving the OMI ALP prod-

uct, assuming then this is the most accurate τ information available for each collocated ~~OMI-MODIS observation pixel~~OMI observation pixel-MODIS aerosol grid cell.

MODIS data are paired on a OMI pixel-by-pixel basis if the distance between ~~pixel centers~~ OMI pixel center and MODIS aerosol grid cell is smaller than 5 km and if both ~~observations~~ observation and product are acquired within 15 min. A threshold of 0.1 is applied to both OMI and MODIS cloud fraction: i.e. if the OMI effective cloud fraction and/or the MODIS geometric cloud fraction (given in the same MODIS aerosol product at 10 km resolution) has a cloud fraction value higher than 0.1, the OMI pixels are filtered out. ~~Applying such a threshold increases the probability of identifying cloud-free scenes. Furthermore, the availability of the MODIS aerosol products is a good confirmation of cloud-free scenes as MODIS Aqua τ variable is exclusively retrieved over sufficient cloud-free pixels.~~ However, since the OMI ~~effecive~~ effective cloud fraction is sensitive to the scattering aerosols, it is well recognized that cloud-free observations with large presence of scattering aerosols are frequently excluded as well. In addition, a threshold of 0.1 is applied ~~on the OMI surface albedo (i.e. OMLER database)~~ to the OMLER database in order to filter out too bright surfaces (either desert ~~of~~ or snow covered pixels).

The NN retrieval algorithms developed and selected in Sect. 3 are used and evaluated here: $NN_{R_{\tau},N^{s}_{O_2-O_2}}$ for $\tau(550\text{nm})$ and ALP retrievals, $NN_{\tau,N^{s}_{O_2-O_2}}$ with MODIS $\tau(550\text{nm})$, from DT algorithm over land, as input for ALP retrieval. For reminder, retrieving ALP from $NN_{R_{\tau},N^{s}_{O_2-O_2}}$ is implicitly similar than from $NN_{\tau,N^{s}_{O_2-O_2}}$ but with the retrieved OMI $\tau(550\text{nm})$, as input (cf. Sect. 3.2 and 5.2). All the NNs designed and trained with the two different ω_0 (i.e. 0.9 and 0.95) are considered in order to investgate the impact of ω_0 assumptions.

These algorithms are applied on the OMI DOAS O_2-O_2 observations, available in the OMCLDO2 product (Acarreta et al., 2004) which can be downloaded here: $http://disc.sci.gsfc.nasa.gov/Aura/data-holdings/OMI/omcldo2_v003.shtml$. As explained in Veefkind et al. (2016), ~~the slant column density~~ $N^{s}_{O_2-O_2}$ depends on the temperature profile due to the nature of dimers of which the absorption scales with the pressure-squared instead of being linear with pressure. Therefore, a simple temperature correction is here applied by using seasonal mean temperature profiles given by the National Centers for Environmental Prediction (NCEP) analysis data. This correction is performed through the computation of the γ factor (Veefkind et al., 2016):

$$\gamma = \frac{N^{\text{sRef}}_{O_2-O_2}(\lambda)}{N^{\text{sMeas}}_{O_2-O_2}(\lambda)}. \tag{10}$$

with $N^{\text{sRef}}_{O_2-O_2}$, the O_2-O_2 SCD associated with the reference temperature profile employed in the learning database and $N^{\text{sMeas}}_{O_2-O_2}$, the measured O_2-O_2 SCD related to the actual temperature conditions. As a first and simple ~~assumption~~approximation, no prior knowledge on aerosols is considered here. The main reason is the little sensitivity to aerosol loading and altitude of this γ factor, for $\tau \leq 2.0$, compared to the change of temperature profiles for the considered OMI observations.

Finally, retrievals are performed based on different assumed surface albedo databases~~. The standard and referent product is the OMLER surface reflectance climatology derived from several years of OMI observations at the spatial resolution of 0.5 ° x 0.5 ° longitude-latitude grid for each calendar month (Kleipool et al., 2008) . The OMLER algorithm is based on temporal histograms of the observed LER values per geophysical location. Potential small residual cloud and aerosol contaminations~~

~~are expected to remain in the OMLER product. As an alternative, the global and spatially complete MODIS black sky surface albedo in the band 3 (459-479) is considered. It is defined as the directional hemispherical reflectance and is a function of θ_0 (Schaaf et al., 2002) .It is derived by integrating the atmospheric corrected Bidirectional Reflectance Distribution Function (BRDF), derived from combined MODIS-Aqua and Terra observations over every 16-day period. The downwelling flux in the MODIS black sky albedo has no diffuse component. The Collection 6 of MCD43C3 product is given on a 0.05° (5.6) latitude/longitude Climate Modeling Grid (CMG), and is then~~ : either OMLER or MODIS Black Sky Albedo (cf. Sect. 2.4). The MODIS Black Sky Albedo is resampled to match the OMI pixel resolution by calculating the average of all MODIS pixels falling within the processed OMI pixel. ~~Note that (Kleipool et al., 2008) demonstrated that the OMLER data set is closer to the black sky than to the white sky by evaluating the ratio between diffuse and direct illumination.~~

6.2 Aerosol optical thickness accuracy: on the importance of the surface albedo and the assumed aerosol properties

Figure ~~9 shows~~ 10 compares collocated retrieved OMI and MODIS τ ~~over North-East Asia in 2005-2007 and for cloud-free scenes. As seen in~~ (550 nm). Similarly to the analyses on synthetic cases (cf. Sect. ~~4.1~~5.1), the change of ~~aerosol~~ assumed ω_0, mostly perturbs retrievals of high τ values, and thus the slope between OMI and MODIS τ(550 nm). Increasing ω_0 from 0.9 to 0.95 reduces the retrieved τ ~~retrieval~~ values of about 0.5 for MODIS τ = 1.5. ~~This is in line with the analyses on simulations in Sect. 4.1.~~ Overall a very good agreement is obtained assuming ω_0 = 0.9 for the seasons spring, autumn and winter (see Fig. ~~9~~ 10 and Fig. 13): differences (OMI-MODIS) of τ(550 nm) lie between -0.18 ± 0.24 in winter and -0.02 ± 0.45 in spring. In summer time, the best agreements are found assuming ω_0 = 0.95 with differences in the range of -0.06 ± 0.31 (see Fig. 8 and Fig. 12).

From the end of autumn to spring, westerly winds transport mineral dust from the Taklimakan and Gobi deserts in northern China and Mongolia region. These dust particles are then frequently mixed with the local anthropogenic aerosols released from the industrial activities, vehicle emissions and coal burning (Eck et al., 2005). Southeast Asia is affected in spring by biomass-burning activity (mostly over the peninsular) which is a major source of carbonaceous aerosols in the world. ~~Overall,~~ Jethva et al. (2014) show that AERONET and OMAERUV retrieve aerosol ω_0 values on average between 0.9 and 0.95 in these regions: most of the sulphate particles have ω_0 close to 0.95, while smoke and dust present lower values (closer to 0.9, even below in some cases). These ~~observations~~ analysis confirm that the ~~employed aerosol model with~~ assumption of $\omega_0 = 0.95$ should be considered as an upper limit for the OMI τ retrievals in autumn, winter and spring times, while a lower ω_0 (i.e. 0.9) is likely more appropriate and thus allows, on average, more reliable τ retrievals. In summer time, because of reduced amounts of dust particles, τ values are more representative of local anthropogenic urban pollution, with ~~more daily variabilities~~ a higher daily variability in the optical and scattering properties. Lee et al. (2007); Lin et al. (2015) also found higher ω_0 values over ~~North-Eas~~ North-East Asia in summer (0.95-0.96) and lower for the other seasons (0.88-0.92). Overall, assuming same constant value (i.e. average) for all the acquired OMI pixels probably lead to some errors since aerosol scattering and absorption properties likely vary day-to-day, even month-to-month.

Figure 11 depicts ~~a~~ the dependence of the retrieved OMI ~~τ to~~ τ(550nm) on the OMLER surface albedo values. As discussed in Sect. ~~4.1~~5.1, error in surface albedo directly creates a bias on the retrieved τ. ~~In summer and spring, higher values are~~

found Most of the retrievals depict higher values over darker surfaces(i.e. OMLER surface albedobetween 0.05 and 0.06), or lower surface albedo. In the range of OMI surface albedo values 0.05-0.1, such behavior should not be observed assuming no systematic bias on the surface properties. Moreover, OMI τ $\tau(550\text{nm})$ shows too small values for scenes with MODIS $\tau \leq 0.4$ $\tau(550\text{nm}) \leq 0.4$ in autumn and winter using OMLER (see Fig. 11 and Fig. 13). Using the 14), MODIS black sky surface albedo allows to reduce this dependence in summer and spring (see Fig. 11d) and reasonably increases OMI τ retrievals over scenes with low MODIS τ values in autumn and winter (cf. Fig. 11b, Fig. 13c and Fig. 13g)in autumn and winter14). Furthermore, standard deviation of differences (OMI-MODIS) τ(550 nm) shows a net improvement of the retrievals precision, with a general mean reduction of 0.05 from OMLER to MODIS black sky . (cf. Fig. 12). An exception is however noticed in winter, which may be due to remaining snow covered pixels in spite of the applied filtering(see Sect. 5.1). Overall. Overall, spatial patterns better match between collocated MODIS and OMI pixels products when employing MODIS black sky albedo with higher values over the high density population areas (i.e. North-East and South-West of selected Chinese region, South-West of Korea) and lower values over South-East of China (see Fig. 12 13 and Fig. 1314). These improvements may be due to a more accurate atmospheric correction in the MODIS black sky surface albedo and potential remaining aerosol residuals present in the OMLER database.

In spite of these improved precisions, using the MODIS black sky albedo does not always improve the accuracy of the OMI τ retrievals. In particular, summer and spring seasons present too high τ values compared to the use of OMLER. This emphasizes that applying the MODIS black sky albedo to OMI measurements may be not fully optimal as: 1) MODIS albedo is the integral value over the full hemisphere which is not in line with the range of angles (θ_0 and θ) encountered by OMI, 2) the MODIS black sky albedo is valid for local solar noon zenith angle of each location which does not match the 1345 ascending node equator crossing time of OMI. An ideal surface albedo database should be aerosol and cloud free, and representative of the viewing and solar angles encountered by the space-borne sensor. Problems related to uncertainties in surface albedo climatologies climatology for the aerosol retrieval problem are well known, and has recently been highlighted by Sanders et al. (2015), although a different spectral band is used (O_2-A at 758-770 nm).

As a conclusion, mostly because uncertainties on assumed aerosol ω_0 parameter and prior surface albedo are dominant, our retrieved OMI $\tau(550\text{nm})$ show lower performances than MODIS $\tau(550\text{nm})$ from the DT algorithm over land in Collection 6 (cf. Sect. 2.2). Furthermore, errors in the phase function or not taking into account the effect of polarization, can play a role. These aspects should be further investigated.

6.3 Long-term analyses of the aerosol layer pressure retrievals

Figure 14 shows the behavior of the 15 shows the retrieved OMI aerosol layer pressure as a function of collocated MODIS τ(550 nm). MODIS τ $\tau(550\text{nm})$ is considered for OMI ALP retrievals since, at this stage, they are considered as the best prior information available with higher accuracy than OMI τ (see $\tau(550\text{nm})$ (cf. Sect. 5.16.1). While ALP retrievals over scenes with MODIS τ $\tau(550\text{nm}) \leq 0.5$ exhibit large variability (more than 400 hPa) and are systematically very high, they start converging to more realistic values with increasing MODIS τ. At MODIS $\tau \geq 1.0$ $\tau(550\text{nm}) \geq 1.0$, retrieved ALP lies in the range of 800-1000 hPa depending on the season, with lower variability (between 50 hPa and 200 hPa maximum). As discussed

in Sect. ~~4.2~~5.2, scenes with $\tau \leq 0.5$ are expected to present substantial large biases because of the minor impacts on the O_2-O_2 changes. Part of the variability can be related to uncertainties of surface albedo and non-constant and inhomogeneous aerosol properties from OMI pixel-to-pixel (e.g. aerosol ω_0 in the OMI observations).

When considering ~~the~~ $NN_{\tau,N^s_{O_2-O_2}}$ ~~configuration~~ with MODIS $\tau(550\ \mathrm{nm})$, from ~~Dark target Algorithm~~ DT algorithm over land, as input, the retrievals globally show a reduced variability, especially for τ values in the range of 0.6~~:2.0~~-2.0 compared to the $NN_{R_c,N^s_{O_2-O_2}}$ configuration (see Fig. ~~14~~16). Over scenes with MODIS $\tau \geq 1.0$~~(see Fig. 15)~~, the variability of the OMI Aerosol Layer Height (ALH), derived from Eq. ~~10~~ 11 as explained in the next subsection, greatly decreases from the range of 1.1~~:2.7~~-2.7 km ($NN_{R_c,N^s_{O_2-O_2}}$) to 0.7~~:1.9~~-1.9 km ($NN_{\tau,N^s_{O_2-O_2}}$) depending on the season. When the OMLER is replaced by the MODIS black sky albedo database, the ALH variability ~~continue~~ continues to decrease of about 0.1 km ~~.~~(cf. Fig. 16).

6.4 Comparison of OMI aerosol layer height with LIVAS climatology

The results of 3 years of OMI ALP retrievals over North-East Asia can be statistically compared to a climatology. ~~The LIdar climatology of Vertical Aerosol Structure for space-based lidar simulation studies (LIVAS) is a 3-D multi-wavelength global aerosol and cloud optical database (Amiridis et al., 2015) . This database provides averaged profiles of aerosol optical properties over 9 years (1 january 2007 - 31 December 2015) from the Cloud Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) data on a uniform grid of 1° x 1°. LIVAS addresses the wavelength dependency of aerosol properties for many laser operating wavelengths including 532 nm. LIVAS data set has been evaluated against AERONET in Amiridis et al. (2015) showing realistic and representative mean state aerosol optical depth values in 532 making this data set ideal for synergistic use with other satellite products.~~ Although the years of the OMI "climatology" and LIVAS do not strictly overlap, ~~it is assumed that~~ the average aerosol layer height (ALH) ~~does not significantly change~~ is assumed not to change significantly between both periods. The comparison is done per season. Spatial average of LIVAS ALH is done over the same area where retrievals are performed. Since large biases are expected at low τ, only OMI retrievals acquired for MODIS $\tau(550\mathrm{nm}) \geq 1.0$ are taken into account and then spatially and temporally averaged per season. About 17 % in summer and spring, and between 5 % and 6 % in winter and autumn, of the OMI retrievals over the 3 years were then selected. As a first and simple approximation, OMI ALP retrievals are converted into ALH in km above sea level, assuming the atmosphere is in hydrostatic balance, scale height of 8 km and a surface pressure at the sea level of 1013 hpa:

$$ALH(OMI) = -8 * \ln(ALP/1013). \quad (11)$$

~~Figure 16 illustrates that the sensitivity to the choice of the aerosol ω_0 is higher when using the $NN_{R_c,N^s_{O_2-O_2}}$ configuration. A minimum bias of 650 (in winter) and a maximum of 1140 (in Autumn) compared to LIVAS are found with $\omega_0 = 0.95$. On the other hand, ALH retrievalsassuming~~ Assumptions on the forward aerosol model (used then in the creation of the supervision database) lead to the highest impacts on the spatial-seasonal averaged ALH retrievals. This is mostly related to the ability of the corresponding NN algorithm to interpret the scattering vs. absorption aerosol effects on the $N^s_{O_2-O_2}$. Assuming OMLER surface albedo, differences between average ALH retrievals with $\omega_0 = 0.95$ and $\omega_0 = 0.9$ ~~already show a remarkable agreement with LIVAS (see Fig. 16b). The OMI ALH can differ up to 1200~~ are in the range of 540-1200 m ~~due to the chosen aerosol~~

~~ω_0. The choice of the surface albedo mostly perturbs the OMI ALH, between 400 and almost 500~~ with $NN_{R_c,N^s_{O_2\sim O_2}}$, and 560-660 m ~~, when assuming $\omega_0 = 0.95$. Using the MODIS black sky surface albedo , instead of OMLER , allows to decrease the maximum ALH bias from 1140~~ with $NN_{\tau,N^s_{O_2\sim O_2}}$ using MODIS $\tau(550\text{nm})$. Assumptions on prior surface albedo (OMLER vs. MODIS Black Sky Albedo) also affect the spatial-seasonal averaged ALH retrievals. Related differences are in the range of 0-730 m ~~to 500~~ with $NN_{R_\omega,N^s_{O_2\sim O_2}}$, and 0-180 m ~~(see Fig. 16b). As discussed in Sect. 5.2, an exception is noticed for the winter results with MODIS black sky albedo that may be due to remaining~~ with $NN_{\tau,N^s_{O_2\sim O_2}}$. Note that a remarkable change is noticed using MODIS Black Albedo in winter with $NN_{R_c,N^s_{O_2\sim O_2}}$ which is likely due to non-filtered snow covered pixels. Associated impacts are then lower than those related to assumed ω_0. Since only scenes with MODIS $\tau(550\text{nm}) \geq 1.0$ are selected, aerosol scattering signal dominate over the surface brightness. Although not shown here, we remarked that applying the temperature correction on $N^s_{O_2\sim O_2}$ is crucial as it correct the retrievals between 50 m and 300 m.

~~Overall, retrievals from $NN_{\tau,N^s_{O_2-O_2}}$ configuration, with~~ Because MODIS τ(550 nm) ~~as input, exhibit lower biases with $\omega_0 = 0.95$: in the range of 260:800(see Fig. 16d). The choice of the aerosol model also has reduced impacts with a maximum dispersion of 600 , between $\omega_0 = 0.9$ and $\omega_0 = 0.95$. The assumptions about surface albedo show very little perturbation on the retrievals, less than 200 . This result looks more consistent. Indeed, since selected OMI cases include a very high amount of particles, surface reflection is expected to be largely attenuated. Mostly aerosol scattering and shielding effects should dominate the scenes. If the $NN_{R_c,N^s_{O_2-O_2}}$ configuration is used, any error in the surface albedo leads to a misinterpretation of R_c and rises ambiguities when analysing $N^s_{O_2-O_2}$ without information on the amount of particles or their optical thickness. Using MODIS τ as input allows to mitigate this problem.~~ has a better accuracy than OMI τ(550 nm), seasonal-spatial average ALH retrievals from $NN_{\tau,N^s_{O_2\sim O_2}}$ combined with this first product shows a clear reduced impact and higher stability with respect to uncertainties on surface albedo and aerosol model. This is because using the most accurate prior information on aerosol amount provides then with the best ALH retrieval performance.

~~Applying the temperature correction on $N^s_{O_2-O_2}$ (see~~ By comparing these best ALH results with the seasonal spatial averaged LIVAS ALH values, maximum differences in the range of 260-800 m are obtained, depending on the seasons, assuming $\omega_0 = 0.95$ and OMLER surface albedo. These differences are reduced to 180-310 m with $\omega_0 = 0.9$ regardless the prior surface albedo. Furthermore, when comparing with the LIVAS ALH variability (cf. Eq. ~~12) leads to changes between 50 and 300 on the ALH. Major impacts are, however, expected on cases with lower~~ 2) spatially averaged over the area, the results from $NN_{\tau,N^s_{O_2\sim O_2}}$ with MODIS τ ~~cases associated with smaller $N^s_{O_2-O_2}$ changes (Veefkind et al., 2016) . Another benefit of this temperature correction is to allow a more consistent seasonal patterns between~~ (550 nm) and prior MODIS black sky albedo have the closest variability with a spread in the range of 0.0-1.0 km (cf. Fig. 16).

Consistent seasonal patterns can be observed between averaged OMI and LIVAS ALH ~~,~~ with higher values in spring and summer, probably due to long-range transport ~~aerosols during the maximum~~ during the dust activity from the ~~deserts (see~~ desert (cf. Sect. ~~5.2) ,~~ 6.2) and smaller values in ~~winter and autumn (see Fig. 16c and~~ autumn and winter (cf. Fig. ~~16d~~17). Nevertheless, while the LIVAS ALH ~~depict that aerosols should be at higher altitude~~ depicts higher aerosol heights in spring than in summer, the OMI ALH ~~show~~ shows the opposite. Several explanations are possible: 1) exclusion of OMI scenes with strong aerosol ~~pollutions~~ pollution episodes because of a too strict threshold applied on the OMI effective cloud fraction (see

Sect. ~~5.1~~6.1), 2) a more rigorous temperature correction should be applied on measured $N^{s}_{O_2-O_2}$ (assuming daily instead of seasonal temperature profiles), 3) inaccuracies ~~in the~~ of MODIS τ ~~values~~, etc... All these elements should be further investigated.

~~These~~ Assuming that LIVAS ALH are here the reference, these results seem to present higher accuracy than the exercise of Park et al. (2016) showing a bias of 1 km of between retrieved OMI and the CALIPSO AEH values over ocean, during the Asian dust event on 31 March 2007. The reasons can be multiple: use of NNs instead of linear interpolation within a LUT, aerosol retrieved over land instead of ocean surfaces, consideration of variable surface albedos as inputs instead of a single value, application of a temperature correction on $N^{s}_{O_2-O_2}$, use of longer data records etc...

7 Conclusions

In this study, different Multilayer Perceptron Neural Network (NN) algorithms were developed and evaluated in order to retrieve aerosol layer height (ALH) over land from the OMI 477 nm O_2-O_2 spectral band. The aerosol height was here retrieved as aerosol layer pressure (ALP) and defined as the mid-pressure of an homogeneous ~~and fixed scattering layer~~ scattering layer with a constant geometric thickness. The focus was on North-East Asia and cloud-free scenes dominated by scattering aerosol ~~particles with~~ fine particles with Angstrom coefficient $\alpha = 1.5$, single scattering albedo ω_0 in the range of 0.9-0.95 and asymmetry parameter $g = 0.7$. The algorithms were trained with a large ensemble of synthetic data set and several precautions were taken into account to avoid ~~problems such as~~ over-training or local minima problems. The key concept of OMI ALP retrievals is the link between the measured O_2-O_2 slant column density (SCD) $N^{s}_{O_2-O_2}$ and the aerosol altitude as a ~~result~~ consequence of shielding effect applied by the particles on the O_2-O_2 dimer complexes that are at lower altitudes. ALP was retrieved on 3 years ~~(~~2005-2007 ~~)~~ of OMI cloud-free observations collocated with MODIS-Aqua aerosol product in North-East Asia. The main objective of this work is first to evaluate the feasibility of a direct retrieval of this key aerosol parameter ~~necessary to calculate air mass factors for trace gas retrievals.~~from a statistical point of view: i.e. over a long-time period and large industrialized continental area, and therefore a high number of observations. All the evaluated performances are summarized in ~~Tab.~~Table 2

~~Accurate~~ Analysis show that a good ALP retrieval requires an accurate prior knowledge of aerosol optical thickness τ ~~is required for a good ALP retrieval~~as input information. Indeed, both τ and ALP parameters simultaneously contribute to the shielding of O_2-O_2 dimers. The analyses of ~~the measured SCD~~ $N^{s}_{O_2-O_2}$ alone leads to an ambiguity since aerosol extinction and aerosol altitude cannot be distinguished. ~~Without knowledge of τ, the continuum reflectance (475nm) R_c can be used as an alternative. However, this configuration is not optimal as 1) both SCD and R_c contain overlapping information (i. e.~~

Because low amount of aerosols have very little impacts on $N^{s}_{O_2-O_2}$ changes, large biases are expected over scenes including aerosol particles with $\tau(550\,\mathrm{nm}) \leq 0.5$. This τ ~~and the length of the average light path) and 2) other parameters such as surface albedo, aerosol properties (e. g. aerosol single scattering albedo ω_0 and phase function), and their associated uncertainties, impact as well R_c.~~ value should be considered as a threshold for a good ALP retrieval quality. Moreover, the algorithms are expected to present a very low sensitivity to particles located at an altitude higher than 4 km. This is because of the nature of the O_2-O_2 complex of which the absorption scales with the pressure-squared instead of being linear with pressure.

Different NN configurations were tested. Sensitivity analysis on ~~simulations~~ synthetic cases show that ALP accuracy with the best NN configuration (i.e. $NN_{\tau,N^s_{O_2-O_2}}$ algorithm with true τ value as input) lies in the range of ~~50-100~~ 50-200 hPa (i.e. about 500 m and 1 km) over aerosol scenes with ~~$\tau \geq 1.0$. Using accurate~~ $\tau(550\text{nm})$ in the range of 0.5-1.0. The accuracy is improved to 0-50 hPa over scenes with $\tau(550\text{nm}) \geq 1.0$. Using the best available τ information ~~, instead of R_c, reduces~~ is theoretically expected to limit the impact of uncertainties due to 1) aerosol model: bias ~~from 100~~ in the range of 0-100 hPa ~~to almost zero,~~ if $\tau \leq 1.0$, for a difference of 0.05 in ω_0 or for a difference of 0.1 in g, 2) surface albedo ~~(bias from 141 ± 96 to 69 ± 79~~ uncertainty in the range of 0.025-0.05 which leads to absolute ALP biases in the range of 50-100 hPa for ~~a surface albedo uncertainty of 0.05). When comparing the 3-year retrievals with the LIVAS climatology database, the Aerosol Layer Height (ALH) results using~~ τ between 0.5 and 1.0, and 0-50 hPa for $\tau \geq 1.0$. Real ALH retrievals were performed over 3-year of OMI O_2-O_2 visible observations over cloud-free scenes with MODIS $\tau(550\text{nm}) \geq 1.0$ and using $NN_{\tau,N^s_{O_2-O_2}}$ combined with MODIS-Aqua $\tau(550\text{ nm})$~~, over cloud-free scenes with $\tau \geq 1.0$, show the best accuracy: maximum biases~~ . Comparison of seasonal and spatial averages with the LIVAS climatology database shows maximum (minimum) differences in the range of 260-800 (180-310) m, depending on the season and assuming $\omega_0 = 0.95$ ~~. Assumed ω_0 (either 0.9 or 0.95) impacts ALH retrievals up to 600~~ ($\omega_0 = 0.9$). Aerosol model assumptions (in particular ω_0) are the most critical as, in that configuration, they impact OMI seasonal-spatial averaged ALH in the range of 560-660 m~~on average, while changes~~ . Changes due to the ~~the assumed~~ prior surface albedo database (OMLER or MODIS black sky) ~~do not exceed~~ have a second order impact, up to 200 m.

In addition, algorithms should take into account that the O_2-O_2 SCD precision, resulting from the DOAS spectral fitting, affects the ALP retrieval. O_2-O_2 SCD precision lying in the range of 0.05-0.25 10^{-43} $\text{mol}^2\text{cm}^{-5}$ leads to ALP ~~bias~~ precision between 19 ± 29 and 57 ± 31 hPa. Due to the nature of the O_2-O_2 collision complex, a temperature correction must be applied ~~on~~ to the SCD prior to retrievals. ~~This impacts the retrieved ALH between 50 and 300 and enables to reproduce more consistent seasonal patterns.~~ Other parameters should be further investigated such as polarization effects and assumptions about the vertical distribution of particles.

~~Because low amount of aerosols have very little impacts on SCD changes, large biases are expected over scenes with~~ An accuracy better than 0.2 must be required on prior τ ~~(550)$\leq$ 0.5. This~~ information. Indeed, a $\tau(550\text{nm})$ bias of 0.25 is expected to bias the retrieved ALP up to 50 hPa for τ ~~value should be considered as a threshold for a good ALP retrieval quality. Moreover, the algorithms are expected to present a very low sensitivity to particles located at an altitude higher than 4 . This is because of the nature of the complex of which the absorption scales with the pressure-squared instead of being linear with pressure,~~

~~The capability of deriving a~~ in the range of 0.6-1.0. For $\tau(550\text{nm}) \geq 1.0$, related ALP impacts almost become almost null. If no prior accurate τ information, such as from the MODIS aerosol Dark Target algorithm, is available, then this input parameter can be replaced by the OMI continuum reflectance $Rc(475\text{nm})$ (cf. the $NN_{Rc,N^s_{O_2-O_2}}$ algorithm). Indeed, this parameter is primarily affected by the aerosol amount and therefore contains information on $\tau(550\text{nm})$. It may then help to analyse $N^s_{O_2-O_2}$ for retrieving ALP provided that we can retrieve τ with a good quality.

Similarly to ALP retrieval, a NN algorithm was also developped to retrieve $\tau(550\text{nm})$ information from the OMI 477 nm O_2-O_2 spectral band~~was also investigated. Accuracy of τ retrievals relies~~. However, its capability is strongly affected by uncertainties on the assumed ~~parameters affecting R_c~~aerosol model. An overestimation of aerosol single scattering albedo, from $\omega_0 = 0.9$ to 0.95, induces a negative bias of 0.8 for $\tau = 2$. The impact is much lower for smaller τ (lower than 0.1 for τ~~s~~ ≤ 0.5). Similar conclusions were found regarding uncertainty of the asymmetry parameter and thus the phase function characterization. Another major challenge when retrieving aerosol properties from passive satellite sensors is to separate the atmospheric and surface contributions in the total observed reflectance. Similarly ~~as~~ to aerosol ω_0, an overestimation of surface reflection leads to an underestimation of retrieved τ. Surface albedo uncertainty ~~should be limited to~~ below 0.025 ~~to ensure τ~~ should limit OMI retrieved $\tau(550\text{nm})$ bias ~~≤ 0.2~~smaller than 0.5 for $\tau(550\text{nm})$ in the range of 0.0-0.5, 0.25 for $\tau(550\text{nm}) = 2.0$. Comparisons of OMI retrievals with collocated MODIS τ show agreements between -0.02 ± 0.45 and -0.18 ± 0.24 depending on the seasons. Further improvements should be made before being able to use these OMI τ products as prior information to ALP retrievals.

Using the $NN_{R_c,N^s_{O_2-O_2}}$ algorithm for ALP retrieval is, in practice, similar to $NN_{\tau,N^s_{O_2-O_2}}$ combined with retrieved OMI $\tau(550\text{nm})$. Since, the retrieved OMI τ accuracy is lower than MODIS τ accuracy from the Dark target Land algorithm, $NN_{R_c,N^s_{O_2-O_2}}$ shows reduced performances and higher sensitivity to aerosol model and surface albedo uncertainties. Associated 3-year OMI ALH retrievals over North-East Asia are impacted by ω_0 uncertainties (0.9-0.95) in the range of 540-1200 m, and by surface albedo (OMLER vs. MODIS Black sky albedo) up to 730 m.

The NN approach presents, at this stage, quite promising results for a future operational processing of the OMI O_2-O_2 spectral band and the next UV-Vis satellite missions such as the TROPOspheric Monitoring Instrument (TROPOMI) (Veefkind et al., 2012). In spite of the high computing time due to the learning database creation and the training of these algorithms, very fast operational processing is allowed. Such processing is much faster than approaches relying on the Optimal Estimation Method and employs more optimized interpolation techniques than a classical linear interpolation within a LUT. For future processing of the OMI data, the OMLER climatology database should be optimized by filtering out small aerosol residuals.

~~The results described in this paper indicate~~ Our study indicates that it is worthwhile to design and evaluate aerosol height retrieval algorithm exploiting the satellites 477 nm O_2-O_2 ~~spectral band. This could lead to a replacement of~~ absorption band. Our long-term motivation is to evaluate the feasibility of replacing the effective clouds by more explicit aerosol parameters in the computation of trace gas AMF. This ~~has to be evaluated in the context of all~~ is relevant not only for OMI but for most of the UV-Vis satellite missions devoted to air quality monitoring. For that purpose, further analyses must be performed by focusing on significant geophysical variability cases: e.g. pixel-by-pixel variability over smaller regions. Furthermore, single OMI ALH retrievals should be compared with reference aerosol vertical profile measurements (ground-based and/or satellites) over some remarkable case studies.

Acknowledgements. This work was funded by the Netherlands Space Office (NSO) under the OMI contract. The authors thank Piet Stammes, Folkert Boersma and Maarten Sneep from KNMI for the discussions about aerosol simulations and measurements, and sharing their experience with respect to the DISAMAR software.

We thank all the developers of the PyBrain software. PyBrain is a joint project of PostDocs, PhD and master level students. The core programmers are (or were) students of Prof. Jurgen Schmidhuber at the Dalle Molle Institute for Artificial Intelligence in Switzerland and the Technische Universitat Munchen in Germany.

The LIVAS products have been collected from the LIVAS database (http://lidar.space.noa.gr:8080/livas), and were produced by the LIVAS team under the European Space Agency (ESA) study contract No. 4000104106/11/NL/FF/fk. We thank all the developers who have provided us with all the data set and for sharing their discussions and recommendations.

The developers of LIVAS would like to acknowledge support through the project MarcoPolo under grant agreement n° 606953 from the European Union Seventh Framework Program (FP7/2007-2013) and the research program ACTRIS-2 under grant agreement no. 654109 from the European Union's Horizon 2020 research and innovation program.

References

Acarreta, J. R., de Haan, J. F., and Stammes, P.: Cloud pressure retrieval using the O_2-O_2 absorption band at 477 nm, Journal of Geophysical Research: Atmospheres, 109, n/a–n/a, doi:10.1029/2003JD003915, http://dx.doi.org/10.1029/2003JD003915, d05204, 2004.

Ahn, C., Torres, O., and Jethva, H.: Assessment of OMI near-UV aerosol optical depth over land, Journal of Geophysical Research: Atmospheres, 119, 2457–2473, doi:10.1002/2013JD020188, http://dx.doi.org/10.1002/2013JD020188, 2013JD020188, 2014.

Amiridis, V., Marinou, E., Tsekeri, A., Wandinger, U., Schwarz, A., Giannakaki, E., Mamouri, R., Kokkalis, P., Binietoglou, I., Solomos, S., Herekakis, T., Kazadzis, S., Gerasopoulos, E., Proestakis, E., Kottas, M., Balis, D., Papayannis, A., Kontoes, C., Kourtidis, K., Papagiannopoulos, N., Mona, L., Pappalardo, G., Le Rille, O., and Ansmann, A.: LIVAS: a 3-D multi-wavelength aerosol/cloud database based on CALIPSO and EARLINET, Atmospheric Chemistry and Physics, 15, 7127–7153, doi:10.5194/acp-15-7127-2015, http://www.atmos-chem-phys.net/15/7127/2015/, 2015.

Atkinson, P. M. and Tatnall, A. R. L.: Introduction Neural networks in remote sensing, International Journal of Remote Sensing, 18, 699–709, doi:10.1080/014311697218700, http://dx.doi.org/10.1080/014311697218700, 1997.

Barkley, M. P., Kurosu, T. P., Chance, K., De Smedt, I., Van Roozendael, M., Arneth, A., Hagberg, D., and Guenther, A.: Assessing sources of uncertainty in formaldehyde air mass factors over tropical South America: Implications for top-down isoprene emission estimates, Journal of Geophysical Research: Atmospheres, 117, n/a–n/a, doi:10.1029/2011JD016827, http://dx.doi.org/10.1029/2011JD016827, d13304, 2012.

Barkley, M. P., De Smedt, I., VanRoozendael, M., Kurosu, T. P., Chance, K., Arneth, A., Hagberg, D., Guenther, A., Paulot, F., Marais, E., and Mao, J.: Top-down isoprene emissions over tropical South America inferred from SCIAMACHY and OMI formaldehyde columns, Journal of Geophysical Research: Atmospheres, 118, 6849–6868, doi:10.1002/jgrd.50552, http://dx.doi.org/10.1002/jgrd.50552, 2013.

Boersma, K. F., Eskes, H. J., Veefkind, J. P., Brinksma, E. J., van der A, R. J., Sneep, M., van den Oord, G. H. J., Levelt, P. F., Stammes, P., Gleason, J. F., and Bucsela, E. J.: Near-real time retrieval of tropospheric NO_2 from OMI, Atmospheric Chemistry and Physics, 7, 2103–2118, doi:10.5194/acp-7-2103-2007, http://www.atmos-chem-phys.net/7/2103/2007/, 2007.

Boersma, K. F., Eskes, H. J., Dirksen, R. J., van der A, R. J., Veefkind, J. P., Stammes, P., Huijnen, V., Kleipool, Q. L., Sneep, M., Claas, J., Leitao, J., Richter, A., Zhou, Y., and Brunner, D.: An improved tropospheric NO_2 column retrieval algorithm for the Ozone Monitoring Instrument, Atmospheric Measurement Techniques, 4, 1905–1928, doi:10.5194/amt-4-1905-2011, http://www.atmos-meas-tech.net/4/1905/2011/, 2011.

Castellanos, P., Boersma, K. F., Torres, O., and de Haan, J. F.: OMI tropospheric NO_2 air mass factors over South America: effects of biomass burning aerosols, Atmospheric Measurement Techniques, 8, 3831–3849, doi:10.5194/amt-8-3831-2015, http://www.atmos-meas-tech.net/8/3831/2015/, 2015.

Chameides, W. L., Yu, H., Liu, S. C., Bergin, M., Zhou, X., Mearns, L., Wang, G., Kiang, C. S., Saylor, R. D., Luo, C., Huang, Y., Steiner, A., and Giorgi, F.: Case study of the effects of atmospheric aerosols and regional haze on agriculture: An opportunity to enhance crop yields in China through emission controls?, Proceedings of the National Academy of Sciences, 96, 13 626–13 633, doi:10.1073/pnas.96.24.13626, http://www.pnas.org/content/96/24/13626.abstract, 1999.

Chimot, J., Vlemmix, T., Veefkind, J. P., de Haan, J. F., and Levelt, P. F.: Impact of aerosols on the OMI tropospheric NO_2 retrievals over industrialized regions: how accurate is the aerosol correction of cloud-free scenes via a simple cloud model?, Atmospheric Measurement Techniques, 9, 359–382, doi:10.5194/amt-9-359-2016, http://www.atmos-meas-tech.net/9/359/2016/, 2016.

de Graaf, M., Stammes, P., Torres, O., and Koelemeijer, R. B. A.: Absorbing Aerosol Index: Sensitivity analysis, application to GOME and comparison with TOMS, Journal of Geophysical Research: Atmospheres, 110, n/a–n/a, doi:10.1029/2004JD005178, http://dx.doi.org/10.1029/2004JD005178, d01201, 2005.

de Haan, J. F.: DISAMAR Algorithm Description and Background Information, Royal Netherlands Meteorological Institute, De Bilt, the Netherlands, 2011.

de Haan, J. F., Bosma, P. B., and Hovenier, J. W.: The adding method for multiple scattering calculations of polarized light, , 183, 371–391, http://adsabs.harvard.edu/abs/1987A%26A...183..371D, provided by the SAO/NASA Astrophysics Data System, 1987.

de Leeuw, G., Holzer-Popp, T., Bevan, S., Davies, W. H., Descloitres, J., Grainger, R. G., Griesfeller, J., Heckel, A., Kinne, S., Klüser, L., Kolmonen, P., Litvinov, P., Martynenko, D., North, P., Ovigneur, B., Pascal, N., Poulsen, C., Ramon, D., Schulz, M., Siddans, R., Sogacheva, L., Tanré, D., Thomas, G. E., Virtanen, T. H., von Hoyningen Huene, W., Vountas, M., and Pinnock, S.: Evaluation of seven European aerosol optical depth retrieval algorithms for climate analysis, Remote Sensing of Environment, 162, 295 – 315, doi:http://dx.doi.org/10.1016/j.rse.2013.04.023, http://www.sciencedirect.com/science/article/pii/S0034425713003507, 2013.

de Smedt, I., Müller, J.-F., Stavrakou, T., van der A, R., Eskes, H., and Van Roozendael, M.: Twelve years of global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors, Atmospheric Chemistry and Physics, 8, 4947–4963, doi:10.5194/acp-8-4947-2008, http://www.atmos-chem-phys.net/8/4947/2008/, 2008.

Dirksen, R. J., Folkert Boersma, K., de Laat, J., Stammes, P., van der Werf, G. R., Val Martin, M., and Kelder, H. M.: An aerosol boomerang: Rapid around-the-world transport of smoke from the December 2006 Australian forest fires observed from space, Journal of Geophysical Research: Atmospheres, 114, n/a–n/a, doi:10.1029/2009JD012360, http://dx.doi.org/10.1029/2009JD012360, d21201, 2009.

Dubovik, O., Holben, B., Eck, T. F., Smirnov, A., Kaufman, Y. J., King, M. D., Tanré, D., and Slutsker, I.: Variability of Absorption and Optical Properties of Key Aerosol Types Observed in Worldwide Locations, Journal of the Atmospheric Sciences, 59, 590–608, doi:10.1175/1520-0469(2002)059<0590:VOAAOP>2.0.CO;2, http://dx.doi.org/10.1175/1520-0469(2002)059<0590:VOAAOP>2.0.CO;2, 2002.

Dufresne, J.-L., Gautier, C., Ricchiazzi, P., and Fouquart, Y.: Longwave Scattering Effects of Mineral Aerosols, Journal of the Atmospheric Sciences, 59, 1959–1966, doi:10.1175/1520-0469(2002)059<1959:LSEOMA>2.0.CO;2, http://dx.doi.org/10.1175/1520-0469(2002)059<1959:LSEOMA>2.0.CO;2, 2002.

Eck, T. F., Holben, B. N., Dubovik, O., Smirnov, A., Goloub, P., Chen, H. B., Chatenet, B., Gomes, L., Zhang, X.-Y., Tsay, S.-C., Ji, Q., Giles, D., and Slutsker, I.: Columnar aerosol optical properties at AERONET sites in central eastern Asia and aerosol transport to the tropical mid-Pacific, Journal of Geophysical Research: Atmospheres, 110, n/a–n/a, doi:10.1029/2004JD005274, http://dx.doi.org/10.1029/2004JD005274, d06202, 2005.

Feingold, G., Cotton, W. R., Kreidenweis, S. M., Janel, and Davis, T.: The impact of giant cloud condensation nuclei on drizzle formation in stratocumulus: Implications for cloud radiative properties, J. Atmos. Sci, pp. 4100–4117, 1999.

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, n/a–n/a, doi:10.1029/2005JD006618, http://dx.doi.org/10.1029/2005JD006618, d14203, 2006.

Haykin, S.: Neural Networks: A Comprehensive Foundation, International edition, Prentice Hall, https://books.google.com.tr/books?id=M5abQgAACAAJ, 1999.

Hewson, W., Barkley, M. P., Gonzalez Abad, G., Bösch, H., Kurosu, T., Spurr, R., and Tilstra, L. G.: Development and characterisation of a state-of-the-art GOME-2 formaldehyde air-mass factor algorithm, Atmospheric Measurement Techniques, 8, 4055–4074, doi:10.5194/amt-8-4055-2015, http://www.atmos-meas-tech.net/8/4055/2015/, 2015.

Hovenier, J. W. and Hage, J. I.: Relations involving the spherical albedo and other photometric quantities of planets with thick atmospheres, , 214, 391–401, http://adsabs.harvard.edu/abs/1989A%26A...214..391H, provided by the SAO/NASA Astrophysics Data System, 1989.

IPCC: Solomon, S., Qin, D., Manning, M., Chen, Z., M., M., Averyt, K. B., Tignor, M., and Miller, H. L.: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

IPCC: The Core Writing Team Pachauri, R. K. and Meyer, L. A.: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, IPCC, Geneva, Switzerland, http://www.ipcc.ch/report/ar5/syr/, 2014.

Jethva, H., Torres, O., and Ahn, C.: Global assessment of OMI aerosol single-scattering albedo using ground-based AERONET inversion, Journal of Geophysical Research: Atmospheres, 119, 9020–9040, doi:10.1002/2014JD021672, http://dx.doi.org/10.1002/2014JD021672, 2014JD021672, 2014.

Kanaya, Y., Irie, H., Takashima, H., Iwabuchi, H., Akimoto, H., Sudo, K., Gu, M., Chong, J., Kim, Y. J., Lee, H., Li, A., Si, F., Xu, J., Xie, P.-H., Liu, W.-Q., Dzhola, A., Postylyakov, O., Ivanov, V., Grechko, E., Terpugova, S., and Panchenko, M.: Long-term MAX-DOAS network observations of NO_2 in Russia and Asia (MADRAS) during the period 2007 and 2012: instrumentation, elucidation of climatology, and comparisons with OMI satellite observations and global model simulations, Atmospheric Chemistry and Physics, 14, 7909–7927, doi:10.5194/acp-14-7909-2014, http://www.atmos-chem-phys.net/14/7909/2014/, 2014.

Kaufman, Y. J., Tanre, D., and Boucher, O.: A satellite view of aerosols in the climate system, Nature, 419, 215–223, doi:10.1038/nature01091, http://dx.doi.org/10.1038/nature01091, 2002.

Kleipool, Q. L., Dobber, M. R., de Haan, J. F., and Levelt, P. F.: Earth surface reflectance climatology from 3 years of OMI data, Journal of Geophysical Research: Atmospheres, 113, n/a–n/a, doi:10.1029/2008JD010290, http://dx.doi.org/10.1029/2008JD010290, d18308, 2008.

Koelemeijer, R. B. A., de Haan, J. F., and Stammes, P.: A database of spectral surface reflectivity in the range 335–772 nm derived from 5.5 years of GOME observations, Journal of Geophysical Research: Atmospheres, 108, n/a–n/a, doi:10.1029/2002JD002429, http://dx.doi.org/10.1029/2002JD002429, 4070, 2003.

Krotkov, N. A., McClure, B., Dickerson, R. R., Carn, S. A., Li, C., Bhartia, P. K., Yang, K., Krueger, A. J., Li, Z., Levelt, P. F., Chen, H., Wang, P., and Lu, D.: Validation of SO_2 retrievals from the Ozone Monitoring Instrument over NE China, Journal of Geophysical Research: Atmospheres, 113, n/a–n/a, doi:10.1029/2007JD008818, http://dx.doi.org/10.1029/2007JD008818, d16S40, 2008.

Kuhlmann, G., Lam, Y. F., Cheung, H. M., Hartl, A., Fung, J. C. H., Chan, P. W., and Wenig, M. O.: Development of a custom OMI NO_2 data product for evaluating biases in a regional chemistry transport model, Atmospheric Chemistry and Physics, 15, 5627–5644, doi:10.5194/acp-15-5627-2015, http://www.atmos-chem-phys.net/15/5627/2015/, 2015.

Lee, C., Martin, R. V., van Donkelaar, A., O'Byrne, G., Krotkov, N., Richter, A., Huey, L. G., and Holloway, J. S.: Retrieval of vertical columns of sulfur dioxide from SCIAMACHY and OMI: Air mass factor algorithm development, validation, and error analysis, Journal of Geophysical Research: Atmospheres, 114, n/a–n/a, doi:10.1029/2009JD012123, http://dx.doi.org/10.1029/2009JD012123, d22303, 2009.

Lee, J., Kim, J., Yang, P., and Hsu, N. C.: Improvement of aerosol optical depth retrieval from MODIS spectral reflectance over the global ocean using new aerosol models archived from AERONET inversion data and tri-axial ellipsoidal dust database, Atmospheric Chemistry and Physics, 12, 7087–7102, doi:10.5194/acp-12-7087-2012, http://www.atmos-chem-phys.net/12/7087/2012/, 2012.

Lee, K. H., Li, Z., Wong, M. S., Xin, J., Wang, Y., Hao, W.-M., and Zhao, F.: Aerosol single scattering albedo estimated across China from a combination of ground and satellite measurements, Journal of Geophysical Research (Atmospheres), 112, D22S15, doi:10.1029/2007JD009077, 2007.

Leitão, J., Richter, A., Vrekoussis, M., Kokhanovsky, A., Zhang, Q. J., Beekmann, M., and Burrows, J. P.: On the improvement of NO_2 satellite retrievals – aerosol impact on the airmass factors, Atmospheric Measurement Techniques, 3, 475–493, doi:10.5194/amt-3-475-2010, http://www.atmos-meas-tech.net/3/475/2010/, 2010.

Levelt, P. F., Hilsenrath, E., Leppelmeier, G. W., van den Oord, G. H. J., Bhartia, P. K., Tamminen, J., de Haan, J. F., and Veefkind, J. P.: Science Objectives of the Ozone Monitoring Instrument, IEEE Transactions on Geoscience and Remote Sensing, 44, 1199–1208, doi:10.1109/TGRS.2006.872336, 2006.

Levy, R. C., Remer, L. A., Mattoo, S., Vermote, E. F., and Kaufman, Y. J.: Second-generation operational algorithm: Retrieval of aerosol properties over land from inversion of Moderate Resolution Imaging Spectroradiometer spectral reflectance, Journal of Geophysical Research: Atmospheres, 112, n/a–n/a, doi:10.1029/2006JD007811, http://dx.doi.org/10.1029/2006JD007811, d13211, 2007.

Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., and Hsu, N. C.: The Collection 6 MODIS aerosol products over land and ocean, Atmospheric Measurement Techniques, 6, 2989–3034, doi:10.5194/amt-6-2989-2013, http://www.atmos-meas-tech.net/6/2989/2013/, 2013.

Lin, J.-T., Martin, R. V., Boersma, K. F., Sneep, M., Stammes, P., Spurr, R., Wang, P., Van Roozendael, M., Clémer, K., and Irie, H.: Retrieving tropospheric nitrogen dioxide from the Ozone Monitoring Instrument: effects of aerosols, surface reflectance anisotropy, and vertical profile of nitrogen dioxide, Atmospheric Chemistry and Physics, 14, 1441–1461, doi:10.5194/acp-14-1441-2014, http://www.atmos-chem-phys.net/14/1441/2014/, 2014.

Lin, J.-T., Liu, M.-Y., Xin, J.-Y., Boersma, K. F., Spurr, R., Martin, R., and Zhang, Q.: Influence of aerosols and surface reflectance on satellite NO_2 retrieval: seasonal and spatial characteristics and implications for NO_x emission constraints, Atmospheric Chemistry and Physics, 15, 11 217–11 241, doi:10.5194/acp-15-11217-2015, http://www.atmos-chem-phys.net/15/11217/2015/, 2015.

Luger, G. and Stubblefield, W.: Artificial intelligence : Structures and strategies for complex problem solving, Benjamin/Cummings, http://www.worldcat.org/title/artificial-intelligence-structures-and-strategies-for-complex-problem-solving/oclc/26305570&referer=brief_results, 1998.

Lyapustin, A., Wang, Y., Xiong, X., Meister, G., Platnick, S., Levy, R., Franz, B., Korkin, S., Hilker, T., Tucker, J., Hall, F., Sellers, P., Wu, A., and Angal, A.: Scientific impact of MODIS C5 calibration degradation and C6+ improvements, Atmospheric Measurement Techniques, 7, 4353–4365, doi:10.5194/amt-7-4353-2014, http://www.atmos-meas-tech.net/7/4353/2014/, 2014.

Ma, J. Z., Beirle, S., Jin, J. L., Shaiganfar, R., Yan, P., and Wagner, T.: Tropospheric NO_2 vertical column densities over Beijing: results of the first three years of ground-based MAX-DOAS measurements (2008 and 2011) and satellite validation, Atmospheric Chemistry and Physics, 13, 1547–1567, doi:10.5194/acp-13-1547-2013, http://www.atmos-chem-phys.net/13/1547/2013/, 2013.

Omar, A. H., Winker, D. M., Vaughan, M. A., Hu, Y., Trepte, C. R., Ferrare, R. A., Lee, K.-P., Hostetler, C. A., Kittaka, C., Rogers, R. R., Kuehn, R. E., and Liu, Z.: The CALIPSO Automated Aerosol Classification and Lidar Ratio Selection Algorithm, Journal of Atmospheric and Oceanic Technology, 26, 1994–2014, doi:10.1175/2009JTECHA1231.1, http://dx.doi.org/10.1175/2009JTECHA1231.1, 2009.

Park, S. S., Kim, J., Lee, H., Torres, O., Lee, K.-M., and Lee, S. D.: Utilization of O_4 slant column density to derive aerosol layer height from a space-borne UV–visible hyperspectral sensor: sensitivity and case study, Atmospheric Chemistry and Physics, 16, 1987–2006, doi:10.5194/acp-16-1987-2016, http://www.atmos-chem-phys.net/16/1987/2016/, 2016.

Platt, U. and Stutz, J.: Differential Optical Absorption Spectroscopy (DOAS), Principles and Applications, doi:10.1007/978-3-540-75776-4, 2008.

Prospero, J. M.: Long-term measurements of the transport of African mineral dust to the southeastern United States: Implications for regional air quality, , 104, 15, doi:10.1029/1999JD900072, 1999.

Ramanathan, V., Crutzen, P. J., Kiehl, J. T., and Rosenfeld, D.: Aerosols, Climate, and the Hydrological Cycle, Science, 294, 2119–2124, doi:10.1126/science.1064034, http://science.sciencemag.org/content/294/5549/2119, 2001.

Rosenfeld, D.: Suppression of Rain and Snow by Urban and Industrial Air Pollution, Science, 287, 1793–1796, doi:10.1126/science.287.5459.1793, http://science.sciencemag.org/content/287/5459/1793, 2000.

Rosenfeld, D., Lahav, R., Khain, A., and Pinsky, M.: The Role of Sea Spray in Cleansing Air Pollution over Ocean via Cloud Processes, Science, 297, 1667–1670, doi:10.1126/science.1073869, http://science.sciencemag.org/content/297/5587/1667, 2002.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J.: Learning representations by back-propagating errors, Nature, 323, 533–536, doi:10.1038/323533a0, http://dx.doi.org/10.1038/323533a0, 1986.

Sanders, A. F. J., de Haan, J. F., Sneep, M., Apituley, A., Stammes, P., Vieitez, M. O., Tilstra, L. G., Tuinder, O. N. E., Koning, C. E., and Veefkind, J. P.: Evaluation of the operational Aerosol Layer Height retrieval algorithm for Sentinel-5 Precursor: application to O_2 A band observations from GOME-2A, Atmospheric Measurement Techniques, 8, 4947–4977, doi:10.5194/amt-8-4947-2015, http://www.atmos-meas-tech.net/8/4947/2015/, 2015.

Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., Strugnell, N. C., Zhang, X., Jin, Y., Muller, J.-P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., d'Entremont, R. P., Hu, B., Liang, S., Privette, J. L., and Roy, D.: First operational BRDF, albedo nadir reflectance products from {MODIS}, Remote Sensing of Environment, 83, 135 – 148, doi:http://dx.doi.org/10.1016/S0034-4257(02)00091-3, http://www.sciencedirect.com/science/article/pii/S0034425702000913, the Moderate Resolution Imaging Spectroradiometer (MODIS): a new generation of Land Surface Monitoring, 2002.

Schaul, T., Bayer, J., Wierstra, D., Sun, Y., Felder, M., Sehnke, F., Rückstieß, T., and Schmidhuber, J.: PyBrain, Journal of Machine Learning Research, 11, 746–746, doi:10.1.1.193.3547, http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.193.3547, 2010.

Shaiganfar, R., Beirle, S., Sharma, M., Chauhan, A., Singh, R. P., and Wagner, T.: Estimation of NO_x emissions from Delhi using Car MAX-DOAS observations and comparison with OMI satellite data, Atmospheric Chemistry and Physics, 11, 10 871–10 887, doi:10.5194/acp-11-10871-2011, http://www.atmos-chem-phys.net/11/10871/2011/, 2011.

Sneep, M., de Haan, J. F., Stammes, P., Wang, P., Vanbauce, C., Joiner, J., Vasilkov, A. P., and Levelt, P. F.: Three-way comparison between OMI and PARASOL cloud pressure products, Journal of Geophysical Research: Atmospheres, 113, n/a–n/a, doi:10.1029/2007JD008694, http://dx.doi.org/10.1029/2007JD008694, d15S23, 2008.

Sontag, E. D.: Feedback stabilization using two-hidden-layer nets, IEEE Transactions on Neural Networks, 3, 981–990, doi:10.1109/72.165599, 1992.

Spada, F., Krol, M. C., and Stammes, P.: McSCIA: application of the Equivalence Theorem in a Monte Carlo radiative transfer model for spherical shell atmospheres, Atmospheric Chemistry and Physics, 6, 4823–4842, doi:10.5194/acp-6-4823-2006, http://www.atmos-chem-phys.net/6/4823/2006/, 2006.

Stammes, P.: Spectral radiance modelling in the UV-visible range, in Proceedings IRS-2000: Current Problems in Atmospheric Radiation, Studies in geophysical optics and remote sensing, A deepak publishing, 101 Research drive, PO BOX 7390, Hampton, VA 23666 USA, http://apps.webofknowledge.com/InboundService.do?product=WOS&SID=P1RNzp4xEQjDzvcN8Fu&UT=

WOS%3A000176920100100&SrcApp=Wiley_Online_Library&DestFail=http%3A%2F%2Fwww.webofknowledge.com&action=retrieve&Init=Yes&SrcAuth=LinksAMR&Func=Frame&customersID=LinksAMR&IsProductCode=Yes&mode=FullRecord, 2001.

Torres, O., Bhartia, P. K., Herman, J. R., Ahmad, Z., and Gleason, J.: Derivation of aerosol properties from satellite measurements of backscattered ultraviolet radiation: Theoretical basis, Journal of Geophysical Research: Atmospheres, 103, 17 099–17 110, doi:10.1029/98JD00900, http://dx.doi.org/10.1029/98JD00900, 1998.

Torres, O., Bhartia, P. K., Herman, J. R., Sinyuk, A., Ginoux, P., and Holben, B.: A Long-Term Record of Aerosol Optical Depth from TOMS Observations and Comparison to AERONET Measurements., Journal of Atmospheric Sciences, 59, 398–413, doi:10.1175/1520-0469(2002)059<0398:ALTROA>2.0.CO;2, 2002.

Torres, O., Tanskanen, A., Veihelmann, B., Ahn, C., Braak, R., Bhartia, P. K., Veefkind, P., and Levelt, P.: Aerosols and surface UV products from Ozone Monitoring Instrument observations: An overview, Journal of Geophysical Research: Atmospheres, 112, n/a–n/a, doi:10.1029/2007JD008809, http://dx.doi.org/10.1029/2007JD008809, d24S47, 2007.

Torres, O., Ahn, C., and Chen, Z.: Improvements to the OMI near-UV aerosol algorithm using A-train CALIOP and AIRS observations, Atmospheric Measurement Techniques, 6, 3257–3270, doi:10.5194/amt-6-3257-2013, http://www.atmos-meas-tech.net/6/3257/2013/, 2013.

Twomey, S. A., Piepgrass, M., and Wolfe, T. L.: An assessment of the impact of pollution on global cloud albedo, Tellus B, 36B, 356–366, doi:10.1111/j.1600-0889.1984.tb00254.x, http://dx.doi.org/10.1111/j.1600-0889.1984.tb00254.x, 1984.

Veefkind, J. P., Boersma, K. F., Wang, J., Kurosu, T. P., Krotkov, N., Chance, K., and Levelt, P. F.: Global satellite analysis of the relation between aerosols and short-lived trace gases, Atmospheric Chemistry and Physics, 11, 1255–1267, doi:10.5194/acp-11-1255-2011, http://www.atmos-chem-phys.net/11/1255/2011/, 2011.

Veefkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H. J., de Haan, J. F., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P. F.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote sensing of environment., 120, 70—83, doi:10.1016/j.rse.2011.09.027, http://dx.doi.org/10.1016/j.rse.2011.09.027, 2012.

Veefkind, J. P., de Haan, J. F., Sneep, M., and Levelt, P. F.: Improvements of the OMI O_2-O_2 Operational Cloud Algorithm and Comparisons with Ground-Based Radar-Lidar Observations, Atmospheric Measurement Techniques Discussions, 2016, 1–28, doi:10.5194/amt-2016-48, http://www.atmos-meas-tech-discuss.net/amt-2016-48/, 2016.

Veihelmann, B., Levelt, P. F., Stammes, P., and Veefkind, J. P.: Simulation study of the aerosol information content in OMI spectral reflectance measurements, Atmospheric Chemistry and Physics, 7, 3115–3127, doi:10.5194/acp-7-3115-2007, http://www.atmos-chem-phys.net/7/3115/2007/, 2007.

Vlemmix, T., Piters, A. J. M., Stammes, P., Wang, P., and Levelt, P. F.: Retrieval of tropospheric NO_2 using the MAX-DOAS method combined with relative intensity measurements for aerosol correction, Atmospheric Measurement Techniques, 3, 1287–1305, doi:10.5194/amt-3-1287-2010, http://www.atmos-meas-tech.net/3/1287/2010/, 2010.

Wagner, T., Dix, B., Friedeburg, C. v., Frieß, U., Sanghavi, S., Sinreich, R., and Platt, U.: MAX-DOAS O_4 measurements: A new technique to derive information on atmospheric aerosols—Principles and information content, Journal of Geophysical Research: Atmospheres, 109, n/a–n/a, doi:10.1029/2004JD004904, http://dx.doi.org/10.1029/2004JD004904, d22205, 2004.

Wagner, T., Burrows, J. P., Deutschmann, T., Dix, B., von Friedeburg, C., Frieß, U., Hendrick, F., Heue, K.-P., Irie, H., Iwabuchi, H., Kanaya, Y., Keller, J., McLinden, C. A., Oetjen, H., Palazzi, E., Petritoli, A., Platt, U., Postylyakov, O., Pukite, J., Richter, A., van Roozendael, M., Rozanov, A., Rozanov, V., Sinreich, R., Sanghavi, S., and Wittrock, F.: Comparison of box-air-mass-factors and radiances for Multiple-

Table 1. Ensemble of parameters and values associated with the simulated learning data set (see Sect. 3.2). Aerosols are simulated with a Henyey-Greenstein scattering phase function (Hovenier and Hage, 1989)

Parameter	List of values
Solar Zenith Angle (θ_0) [°]	9.267, 21.167, 32.892, 44.217, 54.940, 64.814
Viewing Zenith Angle (θ) [°]	0.0, 9.267, 21.167, 32.892, 44.217
Relative azimuth angle ($\phi - \phi_0$) [°]	0., 30., 60., 90., 120., 150., 180.
Surface pressure (Ps) [hPa]	1013., 963.
Surface albedo (A)	0.025, 0.05, 0.075, 0.1
τ	0.0, 0.05, 0.1, 0.2, 0.4, 0.5, 0.6, 0.9, 1.25, 2.0, 3.0
Aerosol layer pressure (ALP) [hPa]	975., 925., 850., 750., 700., 650., 550., 350., 150.
Aerosol Single Scattering Albedo (ω_0)	0.9, 0.95
Ångström coefficient (α)	1.5
asymmetry parameter (g)	0.7

Axis Differential Optical Absorption Spectroscopy (MAX-DOAS) geometries calculated from different UV/visible radiative transfer models, Atmospheric Chemistry and Physics, 7, 1809–1833, doi:10.5194/acp-7-1809-2007, http://www.atmos-chem-phys.net/7/1809/2007/, 2007.

Wagner, T., Deutschmann, T., and Platt, U.: Determination of aerosol properties from MAX-DOAS observations of the Ring effect, Atmospheric Measurement Techniques, 2, 495–512, doi:10.5194/amt-2-495-2009, http://www.atmos-meas-tech.net/2/495/2009/, 2009.

Xiao, Q., Zhang, H., Choi, M., Li, S., Kondragunta, S., Kim, J., Holben, B., Levy, R. C., and Liu, Y.: Evaluation of VIIRS, GOCI, and MODIS Collection 6 AOD retrievals against ground sunphotometer observations over East Asia, Atmospheric Chemistry and Physics, 16, 1255–1269, doi:10.5194/acp-16-1255-2016, http://www.atmos-chem-phys.net/16/1255/2016/, 2016.

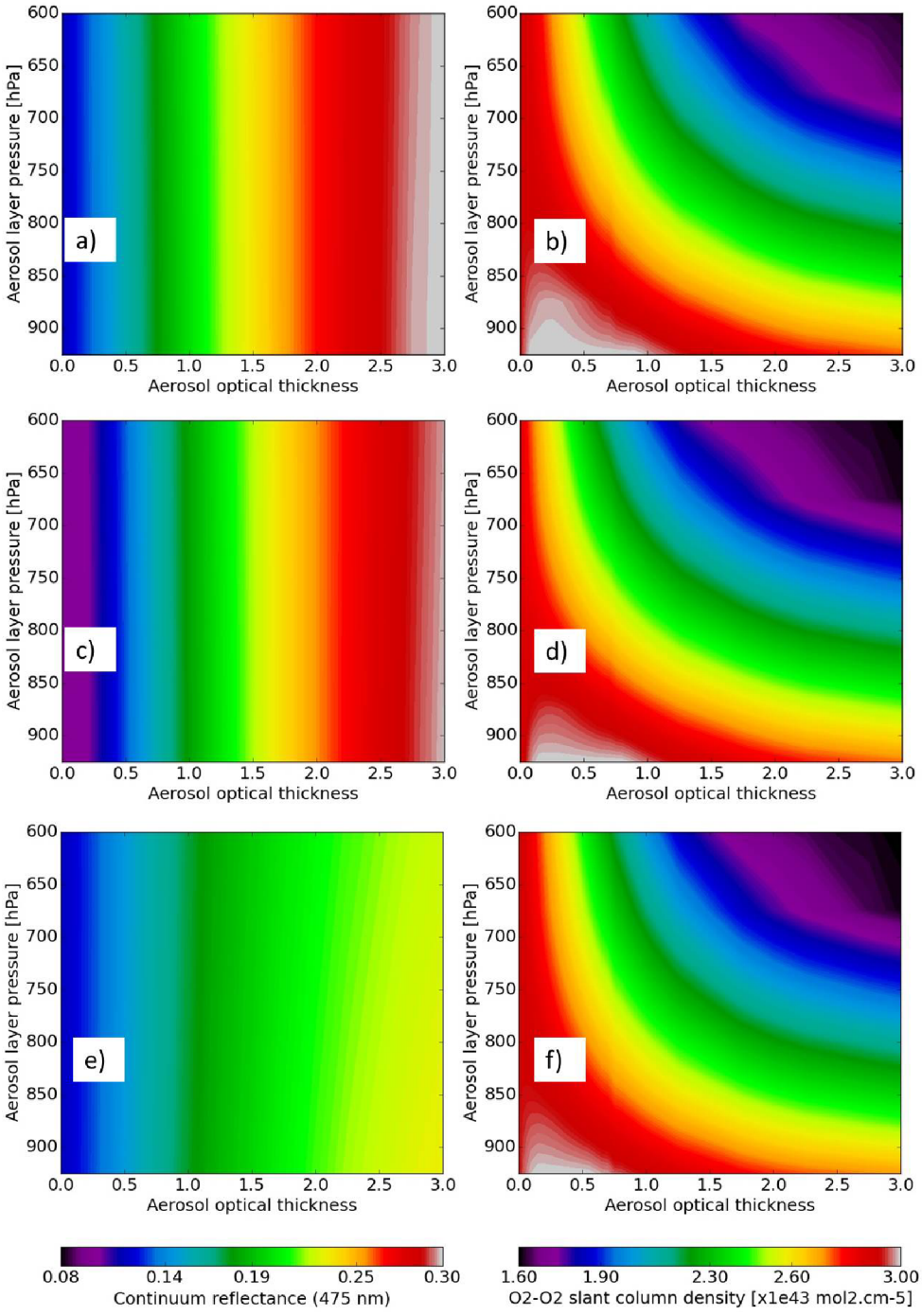


Figure 1. ~~τ~~ Continuum reflectance $R_c(475nm)$ and ~~aerosol layer pressure (ALP) as a function of~~ O_2-O_2 slant column density $N^s_{O_2-O_2}$ ~~, continuum reflectance at 475 nm~~ R_c ~~as~~ as a function of $\tau(550\text{nm})$ and ~~true τ values~~ aerosol layer pressure for the following conditions: climatology mid-latitude summer temperature, NO_2, O_3 and H_2O profiles, ~~surface albedo=0.05,~~ $\theta_0 = 32°$, $\theta = 32°$, surface pressure = ~~1013~~1010 hPa and fine ~~scattering~~ aerosol particles ($\alpha = 1.5$, ~~ω_0=0.95,~~ $g = 0.7$). ~~The ellipses represent the R_c and $N^s_{O_2-O_2}$ results derived from the simulations. The background colors then result from~~ the radial basis interpolation of the DOAS fit results depicted by the ellipses~~:~~ **(a)** ~~τ~~ and ~~aerosol pressure as a function of DOAS fit parameters,~~ **(b)** ~~τ as a function of ALP~~ surface albedo = 0.07 and ~~continuum reflectance at 475 nm~~aerosol $\omega_0 = 0.95$, , **(c)** ~~ALP as a function of true τ~~ and ~~slant column density~~**(d)** surface albedo = 0.03 and aerosol $\omega_0 =$

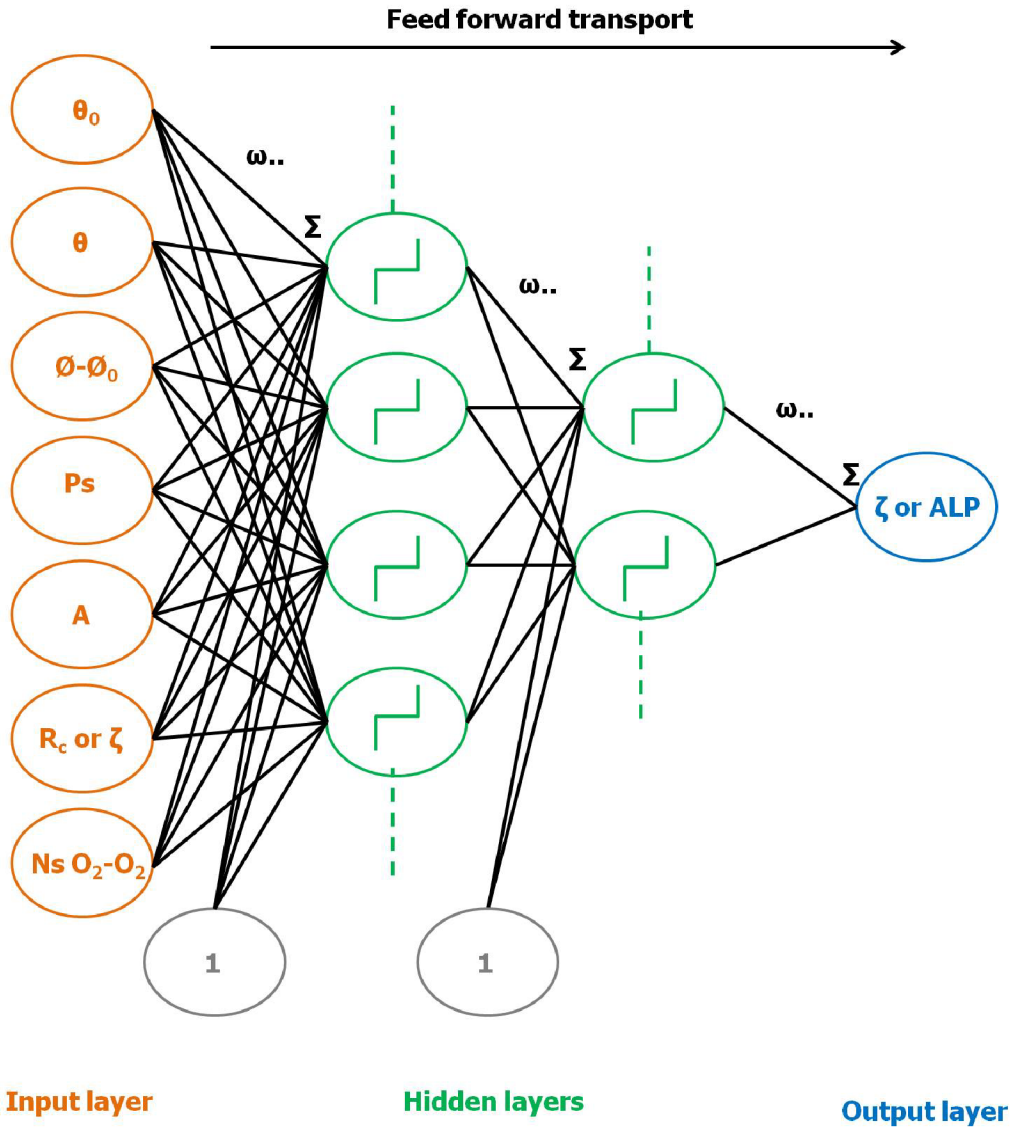


Figure 2. Diagram of Multilayer Perceptron (MLP) Neural Network (NN) architecture designed for Aerosol Layer Pressure (ALP) and aerosol optical thickness τ retrieval algorithms from the OMI O_2-O_2 spectral band at 477 nm. The input parameters are based on the list given in Table 1. The different considered approaches for the MLP design and their applications are more detailed in Sect. 3. Each circle represent a specific processor (named neuron) including either an input / output variable (in the input / output layer) or the activation function (i.e. sigmoid function in the hidden layer). The synaptic weights $\omega..$ ensure the connections of neurons between two consecutive layers. A weighted sum $\sum$ is performed before the transport through the activation function. Note the presence of the bias neurons, prior to the activation functions in the hidden layers. For simplicity, bias neurons are commonly visualized as values added to each neuron in the input and hidden layers of a network, but in practice are treated in exactly the same manner as other weights: all biases are simply weights associated with vectors that lead from a single node whose location is outside of the main network and whose activation is always 1. While the synaptic weights essentially change the steepness of the activation functions, the bias neurons allow to modify the origin of these functions from 0 to positive or negative values.

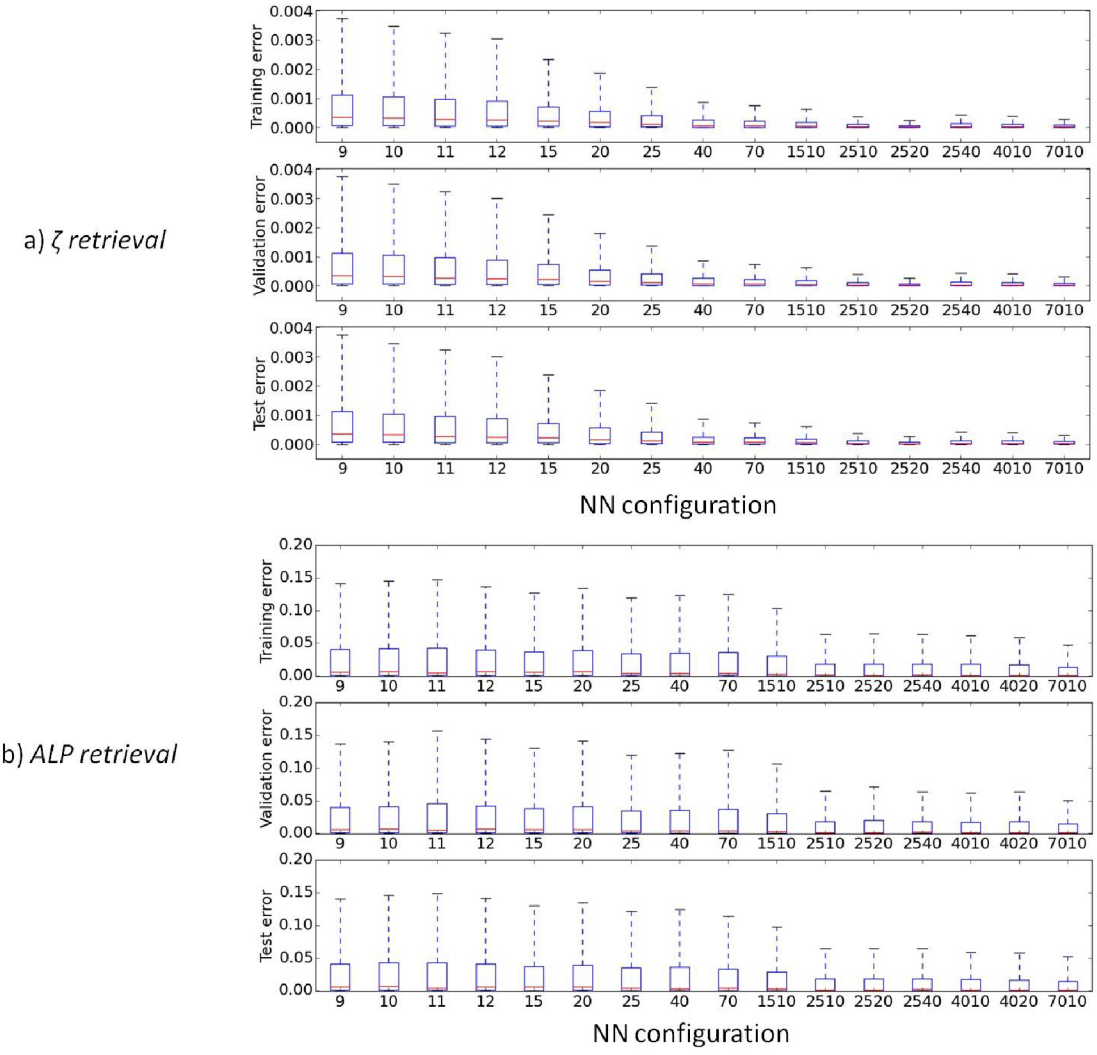
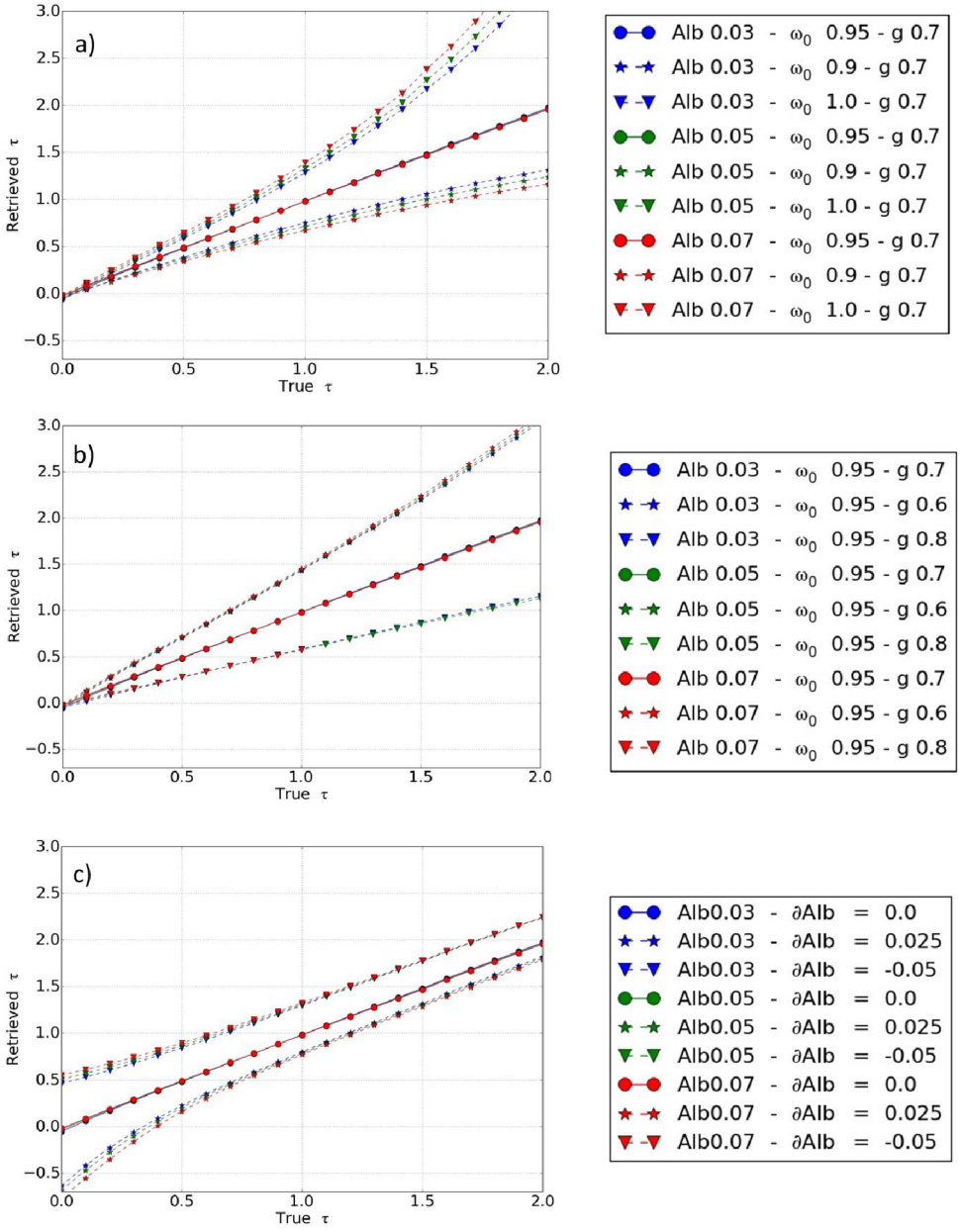


Figure 3. Box-whisker plots of the square of errors (see Eq. 7) obtained for different Neural Network (NN) configurations, at the end of their training, over the supervised data set (training-validation-test). The NNs XX have one hidden layer where XX indicate the number of neurons. The NNs YYXX have two hidden layers where YY and XX are the number of neurons in the 1st and 2nd hidden layer respectively: **(a)** NNs for τ retrieval, **(b)** NNs for ALP retrieval. Note that errors are computed over normalized output and true ~~τ~~ $\tau(550\text{nm})$ and ALP values (between -1 and 1) due to the definition of the sigmoid functions (see Sect. 3.1).



~~Box-whisker plots~~, **(c)** Sensitivity of ~~aerosol retrieval biases induced by~~ τ(550 nm) retrievals to a surface albedo ~~and SCD $N^S_{O_2-O_2}$ uncertainties. The assumed conditions are: fine scattering aerosols~~ bias (~~α~~$\partial Alb =$ ~~1.5~~0.0, -0.025, 0.05) with $\omega_0 = 0.95$, $g = 0.7$~~),~~ ~~climatology mid-latitude summer temperature, , and profiles, surface pressure=1010, surface albedo =0.03-0.05-0.07and combination of $\theta_0 - \theta =$: **(a)** Retrieved τ(550) bias due to surface albedo uncertainty, **(b)** Retrieved ALP bias due to surface albedo uncertainty, **(c)** Retrieved τ bias due to SCD uncertainty"~~.

~~Box-whisker plots~~, **(c)** Sensitivity of ~~aerosol retrieval biases induced by~~ τ(550 nm) retrievals to a surface albedo ~~and SCD $N^S_{O_2-O_2}$ uncertainties. The assumed conditions are: fine scattering aerosols~~ bias (~~α~~$\partial Alb =$ ~~1.5~~0.0, -0.025, 0.05) with $\omega_0 = 0.95$, $g = 0.7$~~),~~ ~~climatology mid-latitude summer temperature, , and profiles, surface pressure=1010, surface albedo =0.03-0.05-0.07and combination of $\theta_0 - \theta =$: **(a)** Retrieved τ(550) bias due to surface albedo uncertainty, **(b)** Retrieved ALP bias due to surface albedo uncertainty, **(c)** Retrieved τ bias due to SCD uncertainty"~~.

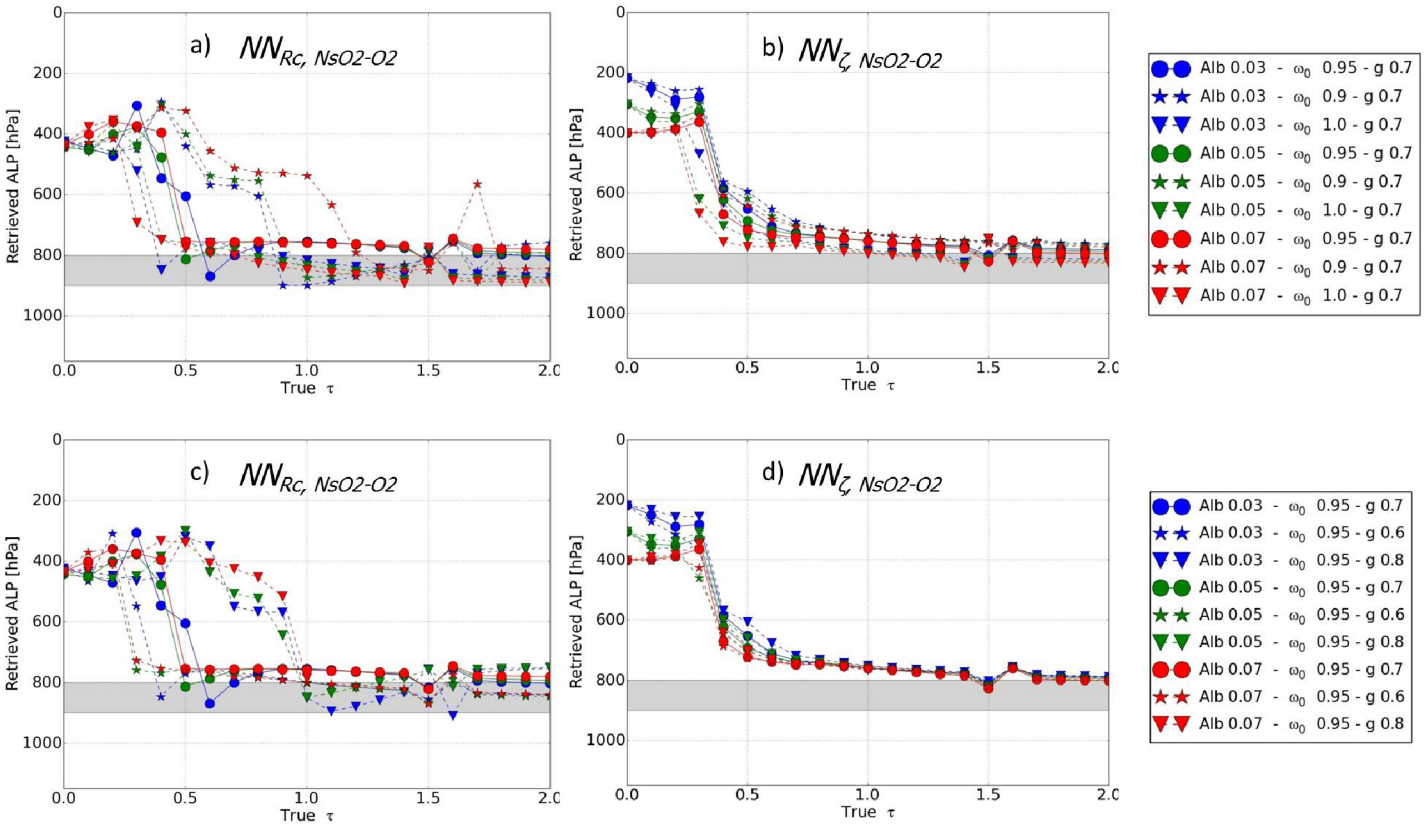


Figure 5. Simulated ALP retrievals, based on noise-free synthetic spectra with aerosols, as a function of true τ. All the retrievals are achieved with the two NN configurations ($NN_{R_c,N^s_{O_2-O_2}}$ and $NN_{\tau,N^s_{O_2-O_2}}$) (see Sect. 3.1) trained with $\omega_0 = 0.95$. The assumed geophysical conditions are: temperature, H_2O, O_3, and NO_2 from climatology mid latitude summer, $\theta_0 = 25°$ and $\theta = 25°$, $Ps = 1010\,\mathrm{hPa}$. The ~~referent~~ reference aerosol scenario assumes fine scattering particles ($\alpha = 1.5$, $\omega_0 = 0.95$, $g = 0.7$) located between 800 and 900 hPa: **(a)** and **(b)** Sensitivity of ALP retrievals to the aerosol single scattering albedo ($\omega_0 = 0.95$ or 0.9) in the synthetic spectra, **(c)** **(d)** Sensitivity of ALP retrievals to the aerosol asymmetry parameter ($g = 0.6$, 0.7 or 0.8) in the synthetic spectra.

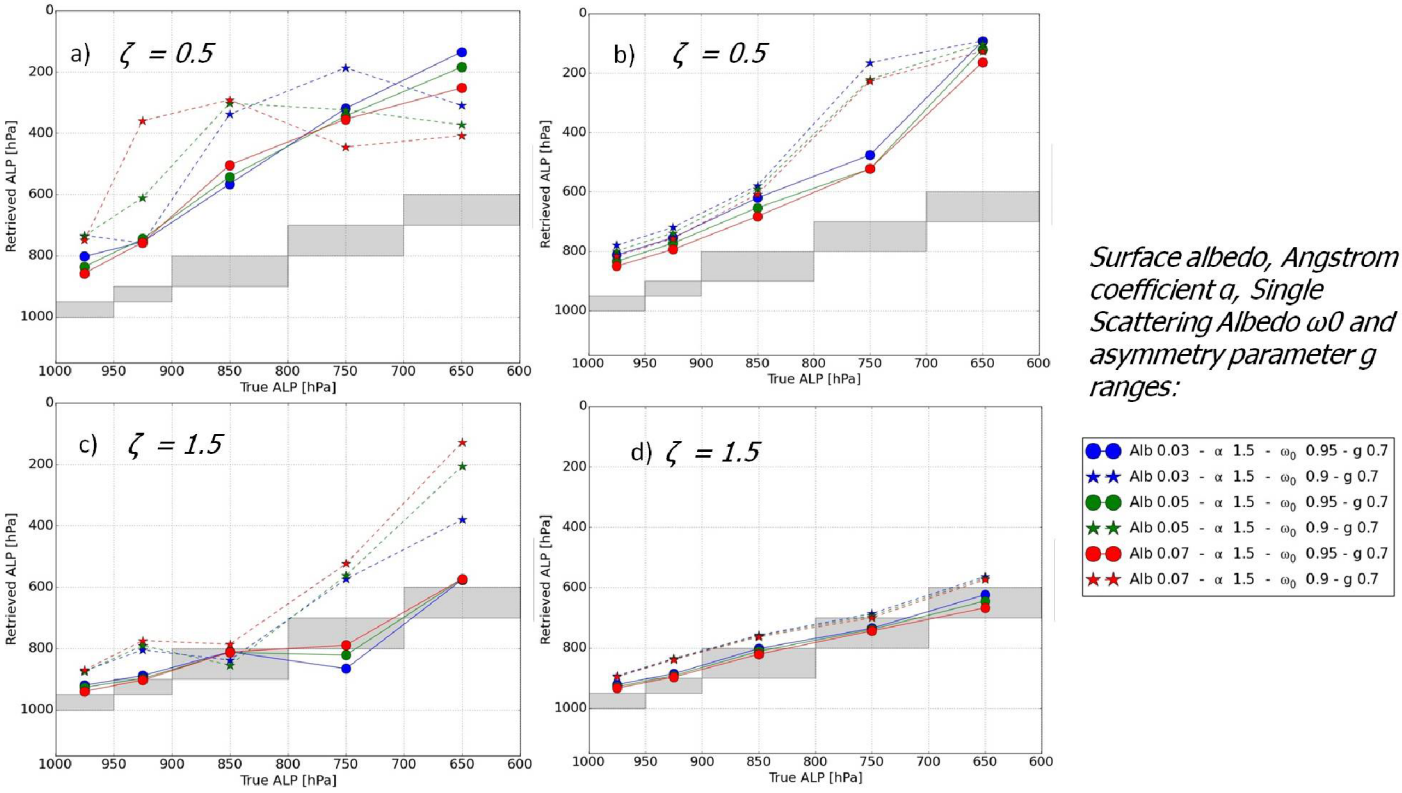


Figure 6. Simulated ALP retrievals vs. true ALP for 2 τ values (0.5 and 1.5) and the two NN configurations (see Sect. 3.1) and for the following conditions: temperature, H_2O, O_3, and NO_2 from climatology mid latitude summer, $\theta_0 = 25°$, $\theta = 25°$, surface pressure = 1010 hPa and fine scattering aerosol particles ($\alpha = 1.5$, $\omega_0 = 0.95$, $g = 0.7$): **(a)** $NN_{R_c,N^s_{O_2-O_2}}$ and ~~τ~~$\tau(550nm)$ = 0.5, **(b)** $NN_{\tau,N^s_{O_2-O_2}}$ with true ~~τ~~$\tau(550nm)$ value as input and ~~τ~~$\tau(550nm)$ = 0.5, **(c)** $NN_{R_c,N^s_{O_2-O_2}}$ and ~~τ~~$\tau(550nm)$ = 1.5, **(d)** $NN_{\tau,N^s_{O_2-O_2}}$ with true ~~τ~~$\tau(550nm)$ value as input and ~~τ~~$\tau(550nm)$ = 1.5

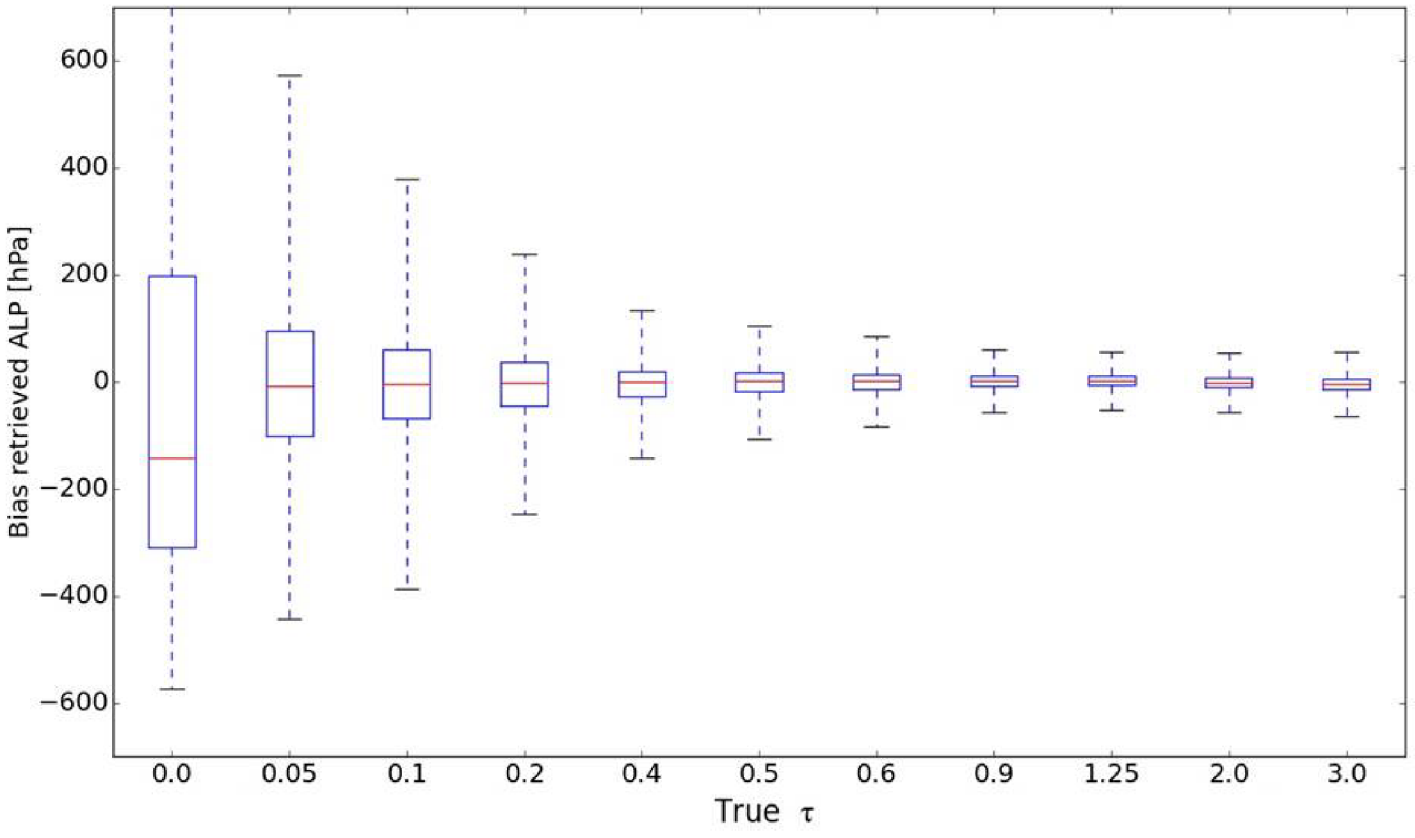


Figure 7. Box-whisker plots of retrieved Aerosol Layer Pressure (ALP) biases as a function of true ~~τ~~ $\tau(550\text{nm})$ from $NN_{R_c,N^s_{O_2-O_2}}$ configuration over the the whole learning data set

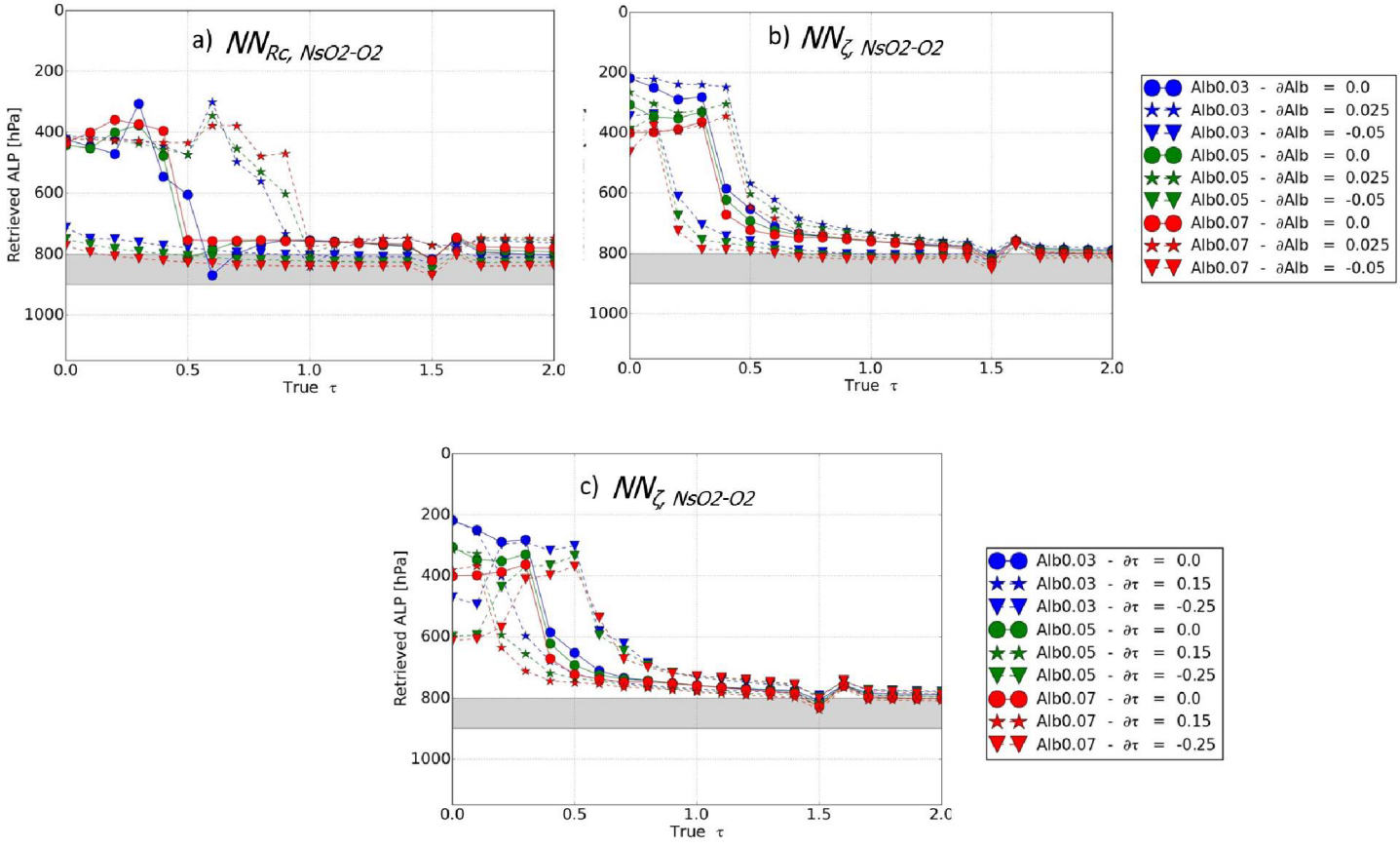


Figure 8. Simulated ALP retrievals, based on noise-free synthetic spectra with aerosols, as a function of true τ. The retrievals are achieved with the NN configurations ($NN_{R_c,N^s_{O_2-O_2}}$ and / or $NN_{\tau,N^s_{O_2-O_2}}$) (see Sect. 3.1) trained with $\omega_0 = 0.95$. The assumed geophysical conditions are: temperature, H_2O, O_3, and NO_2 from climatology mid latitude summer, $\theta_0 = 25°$ and $\theta = 25°$, $Ps = 1010$ hPa. The reference aerosol scenario assumes fine scattering particles ($\alpha = 1.5$, $\omega_0 = 0.95$, $g = 0.7$) located between 800 and 900 hPa: **(a)** and **(b)** Sensitivity of ALP retrievals to a surface albedo bias ($\partial Alb = 0.0, -0.025, 0.05$) with $NN_{R_c,N^s_{O_2-O_2}}$ and $NN_{\tau,N^s_{O_2-O_2}}$. **(c)** Sensitivity of ALP retrievals to a $\tau(550\text{nm})$ bias ($\partial\tau(550\text{nm}) = 0.0, -0.025, 0.05$) with $NN_{\tau,N^s_{O_2-O_2}}$.

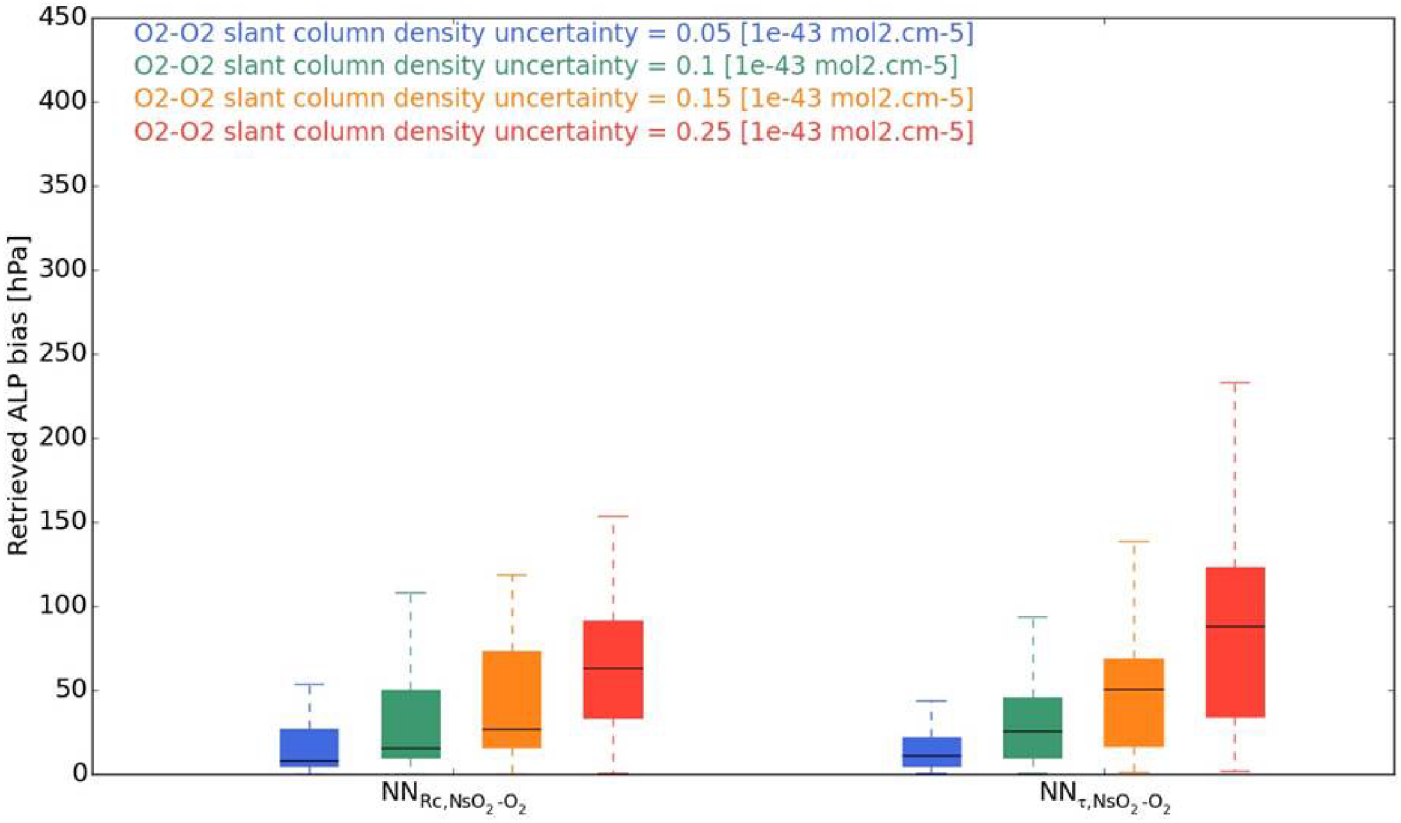


Figure 9. Box-whisker plots of ALP retrieval biases induced by O_2-O_2 SCD $N^s_{O_2-O_2}$ uncertainties. The assumed conditions are: fine scattering aerosols ($\alpha =$ 1.5, $\omega_0 =$ 0.95, $g =$ 0.7), climatology mid-latitude summer temperature, NO_2, O_3 and H_2O profiles, surface pressure = 1010 hPa, surface albedo =[0.03-0.05-0.07] and combination of $\theta_0 - \theta =$ [25°-25°, 50°-25°, 25°-45°].

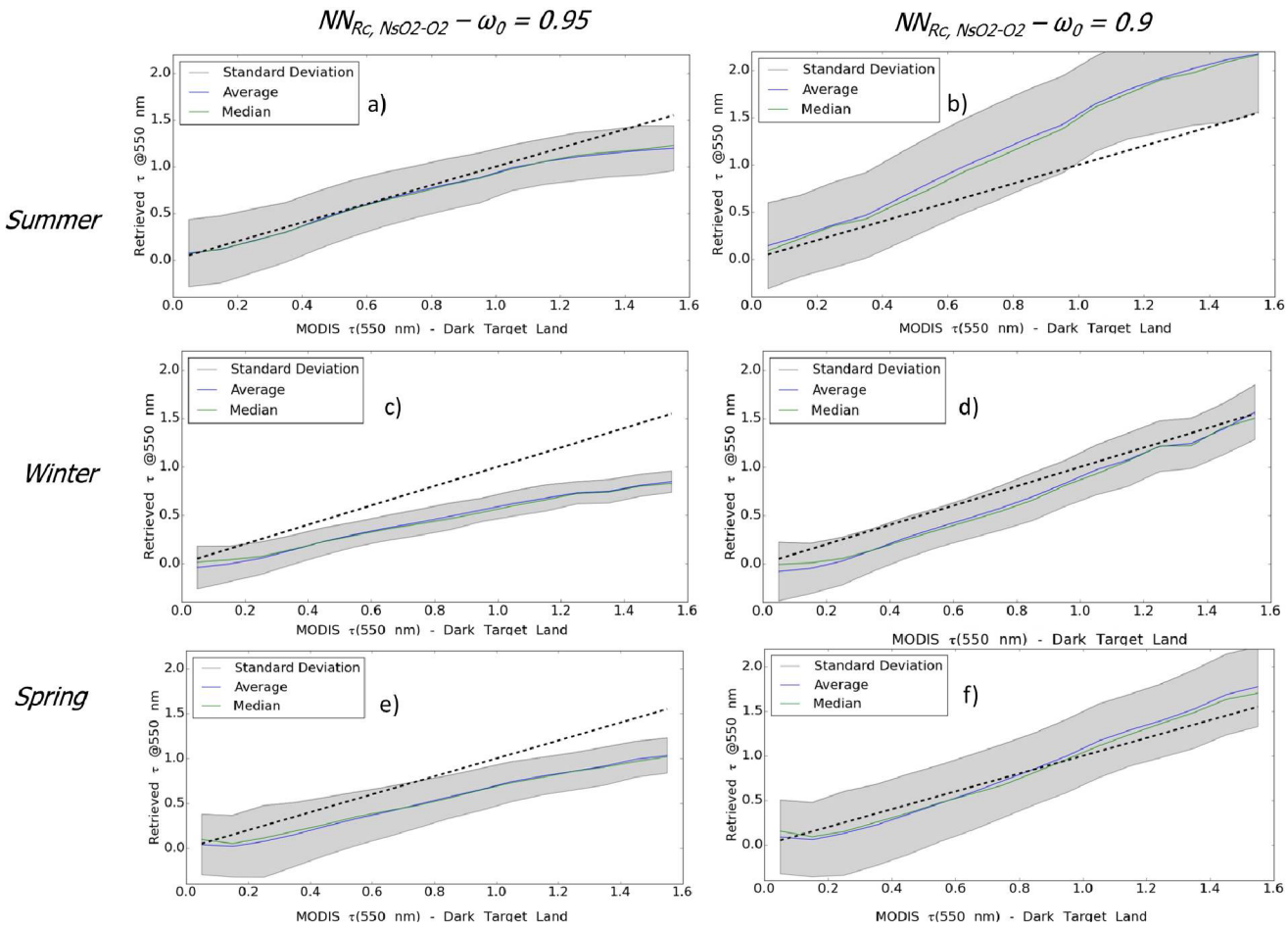


Figure 10. Collocated MODIS Aqua and retrieved OMI $\tau(550\ \mathrm{nm})$ based on the OMLER surface albedo, over East China for cloud-free scenes and summer, winter and spring seasons. Statistics are computed over 3 years ~~(2005, 2006, 2007)~~2005-2007: **(a)**, **(c)** and **(d)** assumed aerosol model with $\omega_0 = 0.95$, **(b)**, **(d)** and **(f)** assumed aerosol model with $\omega_0 = 0.9$.

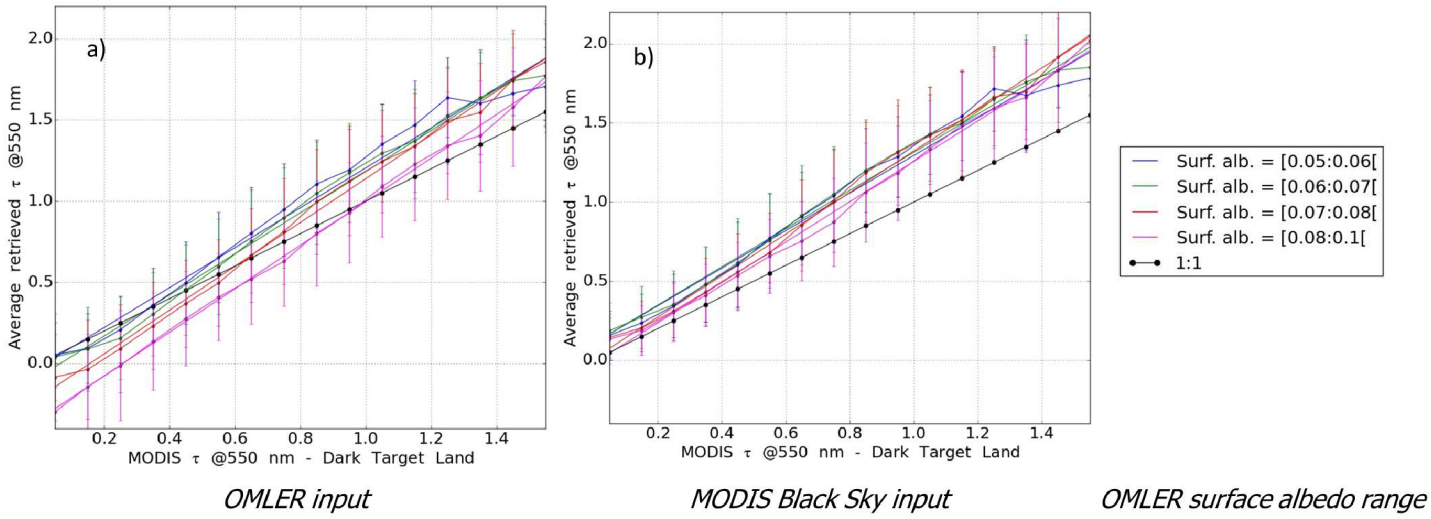


Figure 11. Collocated retrieved OMI retrieval and MODIS Aqua ~~$\tau(550)$~~ $\tau(550\text{nm})$ from Dark Target algorithm over land, over ~~East China~~ North-East Asia for cloud-free scenes in autumn season. Retrievals are depicted as a function of OMLER surface albedo ranges (Kleipool et al., 2008) ~~and for 2 seasons (winter and autumn~~see Sect. 5.1). The dotted black line is the reference 1:1. Statistics are computed over 3 years ~~(2005, 2006,2007)~~ 2005-2007: (**a**) ~~and (b)~~ OMI retrievals ~~in winter~~ based on OMLER~~and MODIS black sky surface albedo (see Sect. 5.1) respectively~~, ~~(c) and (d)~~ (**b**) OMI retrievals ~~in autumn~~ based on ~~OMLER and~~ MODIS black sky surface albedo~~(see Sect. 5.1) respectively.~~

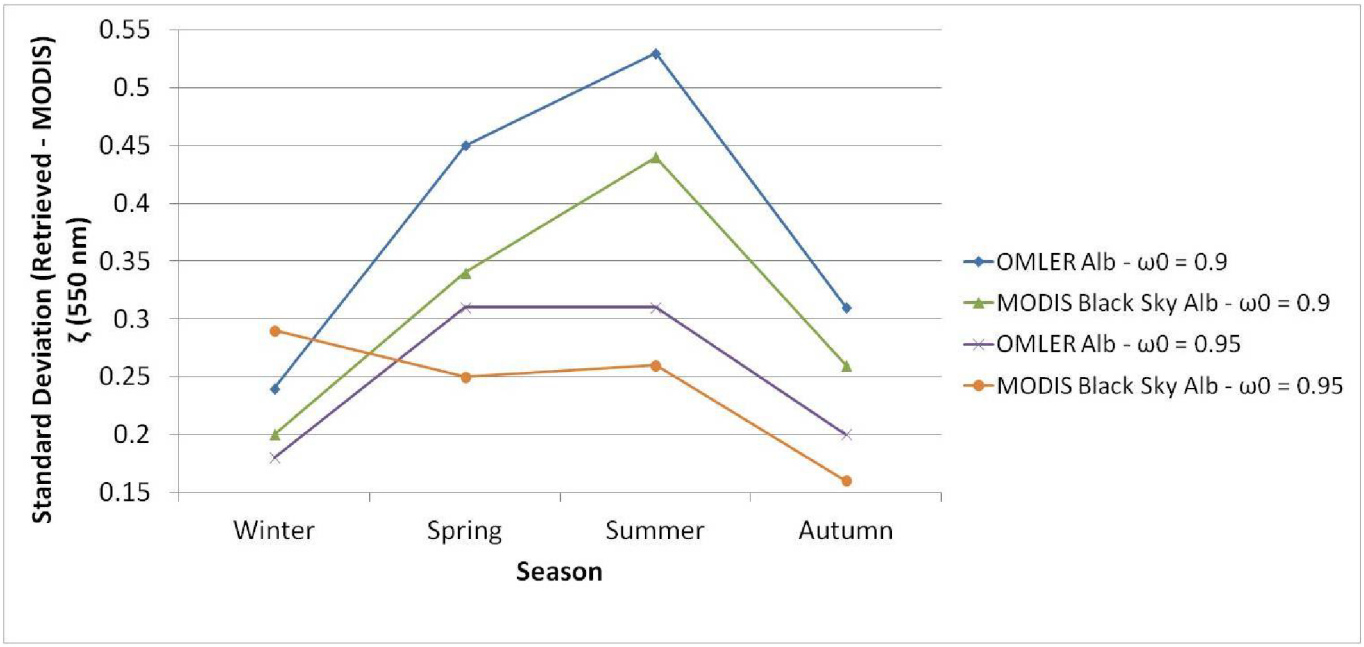


Figure 12. Standard deviation of the differences between OMI ~~τ~~ $\tau(550\text{nm})$ retrievals and MODIS ~~$\tau(550$~~) $\tau(550\text{nm})$ from Dark Target Land algorithm over land for all the individual cloud-free observations over North-East Asia. The retrievals are obtained over 3 years ~~(2005~~2005-2007, ~~2006,2007),~~ and for the 4 seasons. Aerosol single scattering albedo $\omega_0 = 0.95$ and $\omega_0 = 0.9$ are assumed. OMLER and MODIS black sky surface albedo are alternatively considered.

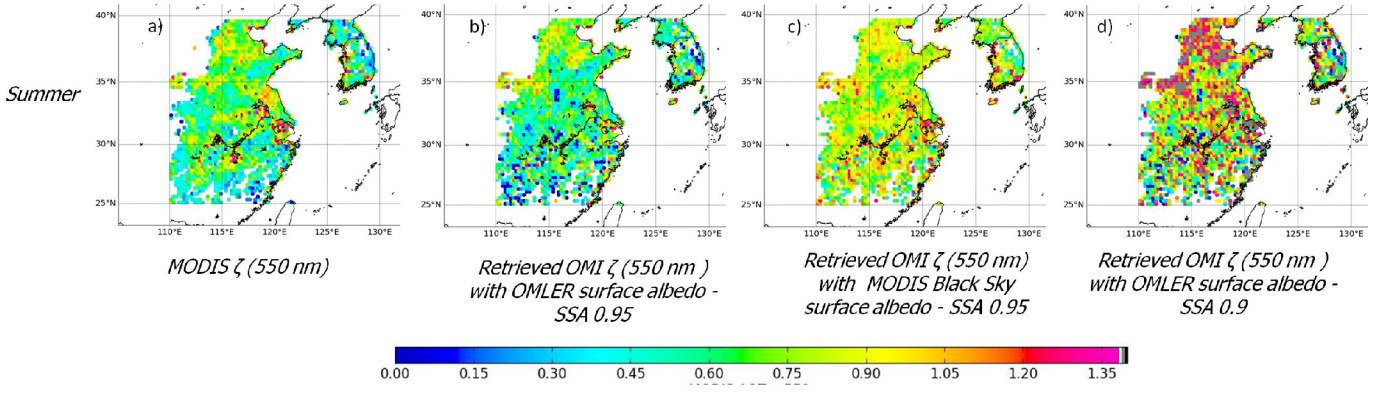


Figure 13. Spatial averages of $\tau(550\text{ nm})$ values, over ~~East China~~ North-East Asia for cloud-free scene. Statistics are computed over 3 years ~~(2005, 2006,2007)~~ 2005-2007 for summer.

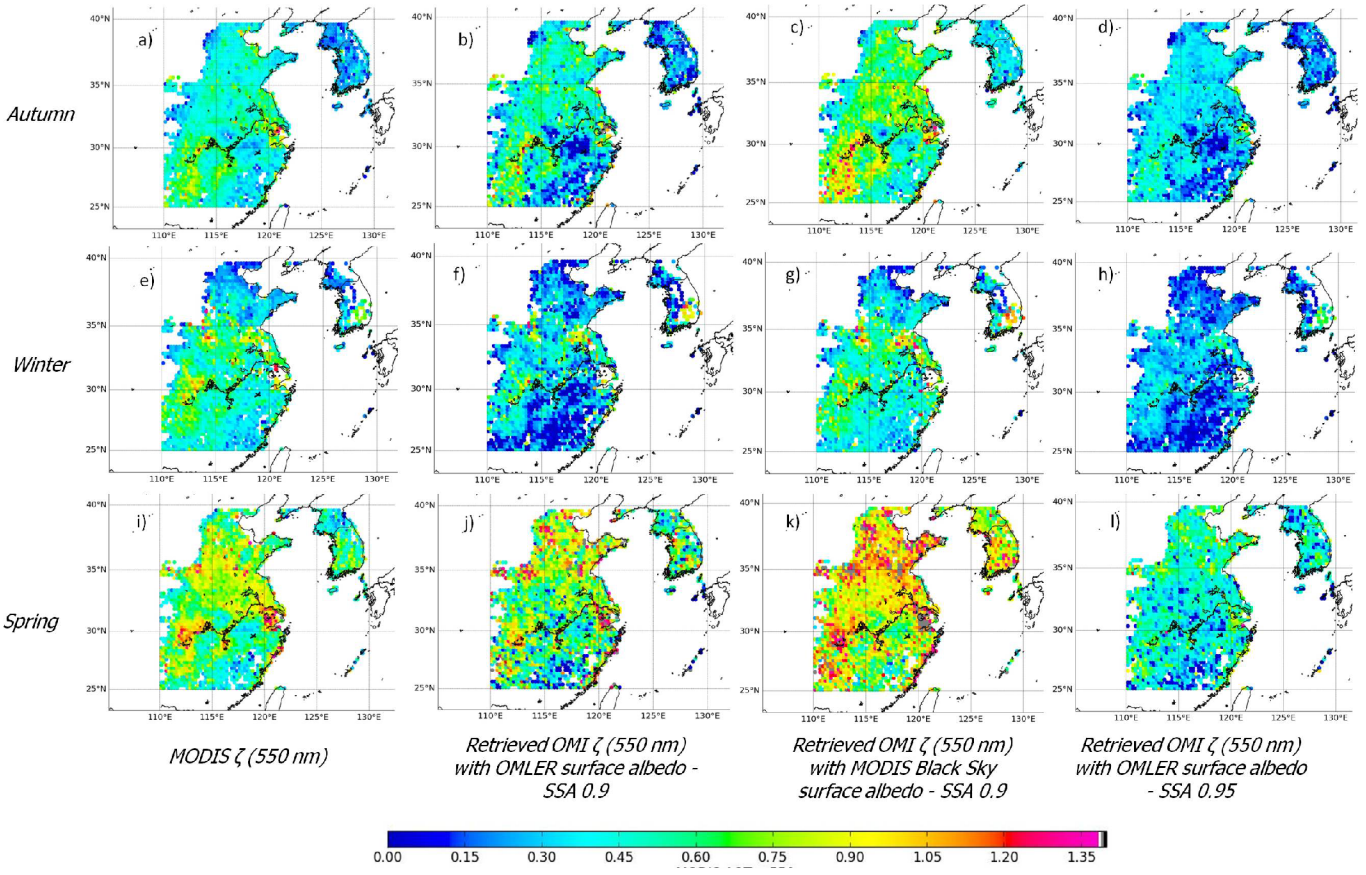


Figure 14. Spatial averages of τ(550 nm) values, over ~~East China~~ North-East Asia for cloud-free scene. Statistics are computed over 3 years ~~(2005, 2006,2007)~~ 2005-2007 for the 3 seasons, autumn, winter and spring.

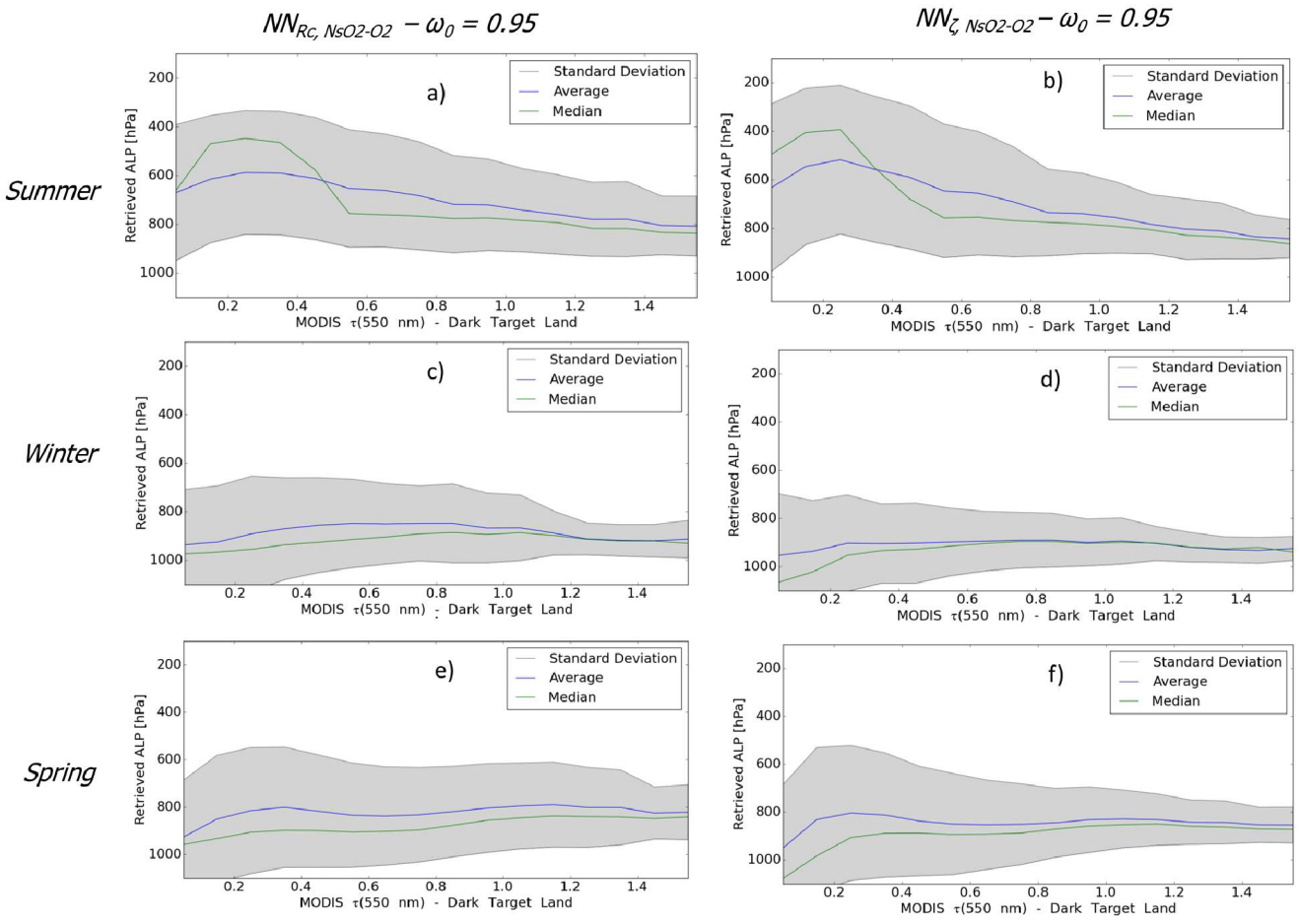
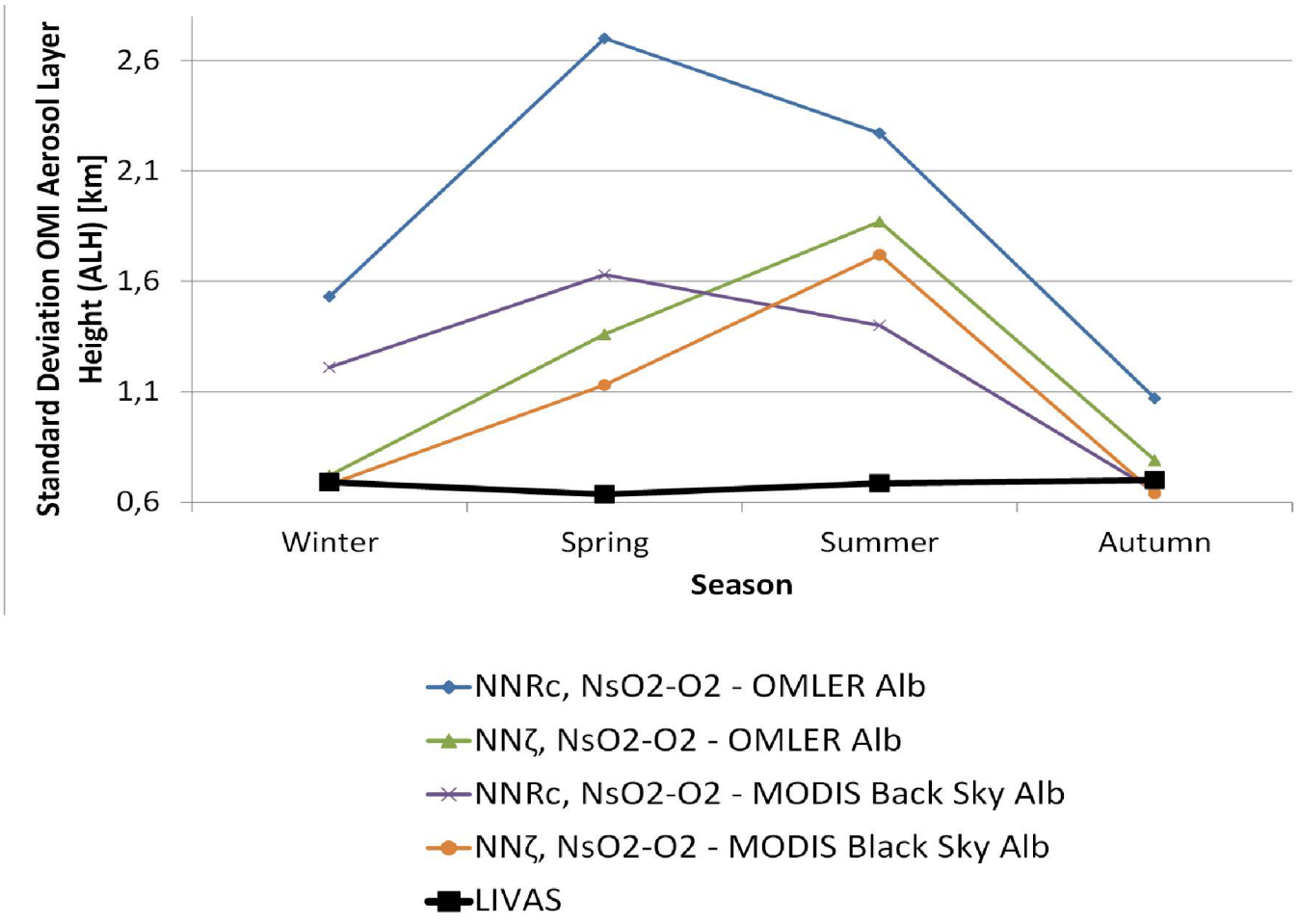


Figure 15. Collocated retrieved OMI ALP (assumed aerosol model with $\omega_0 = 0.95$ and based on the OMLER surface albedo) and MODIS Aqua $\tau(550\ \mathrm{nm})$, over North-East Asia for cloud-free scenes and summer, winter and spring seasons. Statistics are computed over 3 years ~~(2005, 2006, 2007)~~2005-2007: **(a)**, **(b)** and **(c)** OMI retrievals are from the $NN_{R_c,N^s_{O_2-O_2}}$ configuration, **(d)**, **(e)** and **(f)** OMI retrievals are from the $NN_{\tau,N^s_{O_2-O_2}}$ configuration with MODIS Aqua $\tau(550\ \mathrm{nm})$, from Dark Target Land algorithm, as input (see Sect. 5.1).



~~Comparison of the average of the OMI aerosol layer height (ALH) retrievals obtained for the higher collocated MODIS τ values, with the LIVAS ALH climatology databasecf. The retrievals are achieved over East China for cloud-free scene, over 3(2005, 2006,2007), and for the 4 seasons. OMI and MODIS black sky surface albedo are considered alternatively. The 2 NN configurations are considered (see~~ Sect. ~~5.16.1).Cases with and without the temperature correction applied on the are illustrated: **(a)** and **(b)** OMI ALH from the $NN_{R_c,N^s_{O_2-O_2}}$ configuration with and without temperature correction respectively, **(c)** and **(d)** OMI ALH from the $NN_{\tau,N^s_{O_2-O_2}}$ configuration with MODIS τ(550), from Dark Target Land algorithm, as input, with and without temperature correction respectively.~~

~~Comparison of the average of the OMI aerosol layer height (ALH) retrievals obtained for the higher collocated MODIS τ values, with the LIVAS ALH climatology databasecf. The retrievals are achieved over East China for cloud-free scene, over 3(2005, 2006,2007), and for the 4 seasons. OMI and MODIS black sky surface albedo are considered alternatively. The 2 NN configurations are considered (see~~ Sect. ~~5.16.1).Cases with and without the temperature correction applied on the are illustrated: **(a)** and **(b)** OMI ALH from the $NN_{R_c,N^s_{O_2-O_2}}$ configuration with and without temperature correction respectively, **(c)** and **(d)** OMI ALH from the $NN_{\tau,N^s_{O_2-O_2}}$ configuration with MODIS τ(550), from Dark Target Land algorithm, as input, with and without temperature correction respectively.~~

Figure 16. Standard deviation of the OMI ~~aerosol layer height (~~ALH ~~)~~ retrievals obtained for the cloud-free scenes with MODIS $\tau(550\text{nm}) \geq 1.0$ over North-East Asia. The retrievals are ~~achieved over East China~~one for cloud-free scene, over 3 years ~~(2005, 2006,2007), and for the 4 seasons~~2005-2007. Aerosol single scattering albedo $\omega_0 = 0.95$ is assumed. OMLER and MODIS ~~black sky~~ Black Sky surface albedo, and the NN configurations ($NN_{R_c,N^s_{O_2-O_2}}$ and $NN_{\tau,N^s_{O_2-O_2}}$ with MODIS τ(550 nm) as input) are alternatively considered ~~.~~

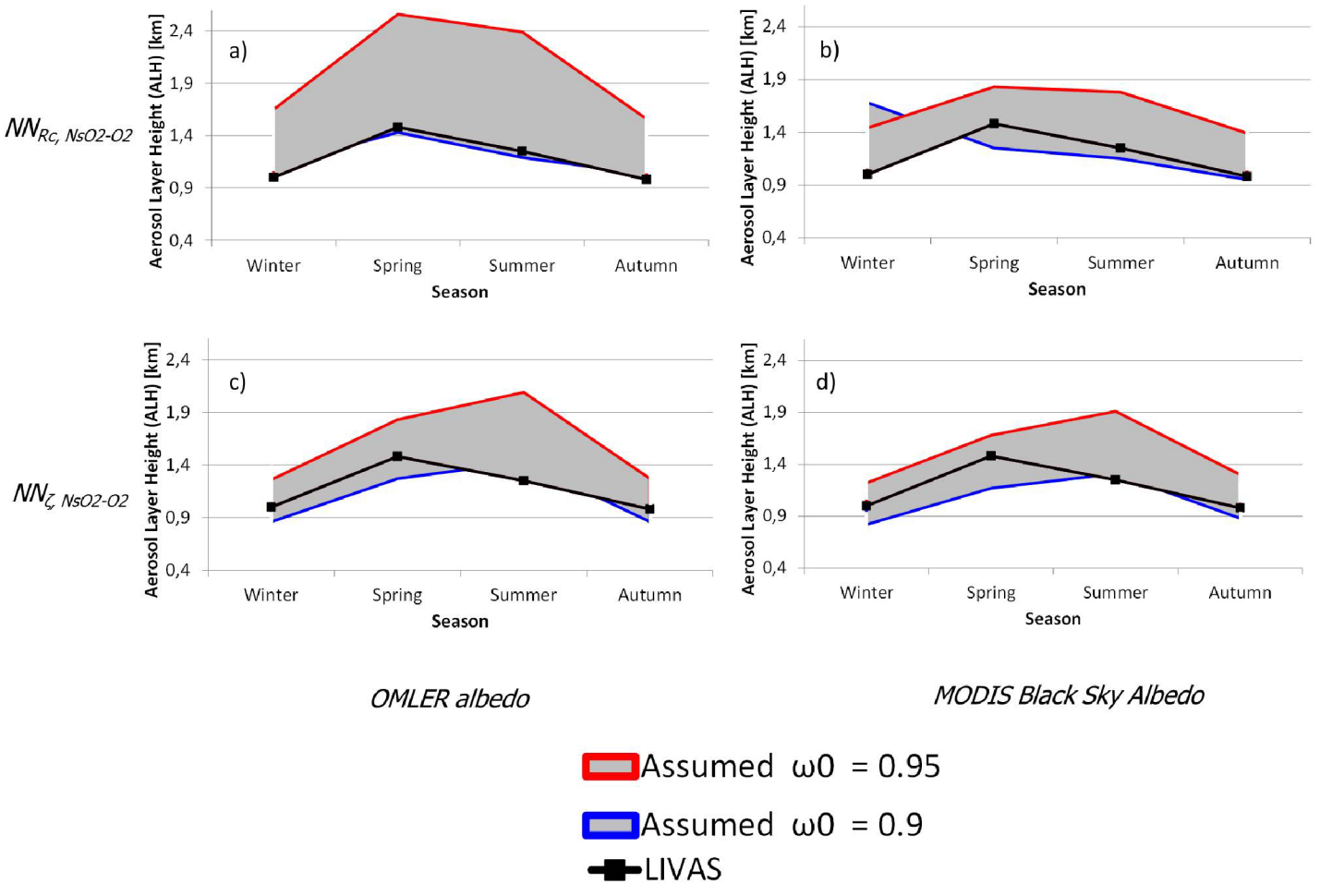


Figure 17. Comparison of the average of the OMI ALH retrievals obtained over scenes with collocated MODIS $\tau \geq 1.0$, with the LIVAS ALH climatology database. The retrievals are achieved over North-East Asia for cloud-free scenes, over 3 years 2005-2007, and for the 4 seasons. OMI and MODIS Black Sky surface albedos are alternatively considered. The 2 NN OMI ALH algorithms are used (see Sect. 5.1): $NN_{R_c,N^s_{O_2-O_2}}$ based on OMI R_c(475 nm) and $NN_{\tau,N^s_{O_2-O_2}}$ based on MODIS τ(550 nm). A temperature correction is applied to the OMI $N^s_{O_2-O_2}$ prior to the retrievals (cf. Sect 6.1): **(a)** $NN_{R_c,N^s_{O_2-O_2}}$ algorithm and OMLER surface albedo, **(b)** $NN_{R_c,N^s_{O_2-O_2}}$ algorithm and MODIS Black Sky surface albedo, **(c)** $NN_{\tau,N^s_{O_2-O_2}}$ algorithm, and OMLER surface albedo, **(d)** $NN_{\tau,N^s_{O_2-O_2}}$ algorithm and MODIS Black Sky surface albedo.