***Response to Interactive comment on* "A sensitivity study on the retrieval of aerosol vertical profiles using the oxygen A-band"**

We would like to thank the reviewer for the valuable comments. In the following response we will address each comment specifically. As a result of the comments we added additional calculations to our manuscript, which will be fully shown in the final manuscript and its supplement.

This is a useful paper which quantifies the amount of vertical profile information that can be extracted from satellite-based measurements in the oxygen-A band and studies the dependence of the information content on the spectral resolution and acquisition time of the sensors. For this purpose, the authors carried out numerical simulations with the radiative transfer model VLIDORT and used the optimal estimation theory to determine the amount of vertical information.

This paper is of great interest for researchers, who uses the principle of aerosol retrieval using O2 A-band. Indeed, the paper details the potentials but also the limitations of the oxygen-A band in regards to the technical specifications of future space-born sensors. The work has been carefully carried out and the results are presented exhaustively. However, for the researchers who are not particularly familiar with the information content analysis, the paper can be very difficult to read because of the lack of examples. I encourage the authors, more particularly in the third paragraph, to give some practical examples to link the formalism with the context of the retrieval of aerosol vertical profiles.

Consequently, I recommend that the manuscript be accepted after minor revisions.

Page 8 - You assumed a constant solar zenith angle of 45° and an instrument looking down at 30° off-nadir. Does the observation geometry influence the results? Did you perform computations with other view and sun zenith angles?

We performed new computations for two instrument viewing angle (nadir and 30° off-nadir) and four solar zenith angle SZA (30°, 45°, 60° and 75°). The viewing geometry seems to have minimal impact on the information content. Degrees of Freedom (DoF) at 30° off-nadir are slightly higher when compared to the nadir case, due to the longer absorption path length.

The variation of the DoF for different SZA, on the other hand, shows more variation of the information content. The results of this new sensitivity tests are implemented in the manuscript (page 16 line 21-26 and page 17 line 1-6) with relative comparison plot (Fig.4) and an additional table in the supplement (Table S1 in the supplement):

"In order to explore the dependence of DoF on observation geometry, we vary the solar zenith angle (30◦ , 45◦ , 60◦ and 75◦) and the viewing angle (nadir and 30◦ off-nadir) implementing the urban scenario for the aerosol

and assuming a ω = 0.95 as single scattering albedo. Figure 4 shows the improvement of the information content due to resolution and higher values of SZA, for both observations. A different solar illumination changes the scattering angle of the particles, allowing a better information retrieval. At 5 cm^{-1} resolution, DoF improve from 1.39 (SZA=30◦) to 1.87 (SZA=75◦) and from 4.49 to 5.22 (same solar angles) at 0.05 cm^{-1} , for the nadir case. The same change in resolution for the 30◦ off-nadir viewing angle allows a similar improvement in the DoF with a change from 1.57 (SZA=30◦) to 1.96 (SZA=75◦) at lower resolution and from 4.74 to 5.33 at 0.05 cm^{-1} resolution, for the same solar angles. As expected, the off-nadir viewing angle shows higher absolute DoF due to the longer light path."

Page 8 - For study cases with maritime aerosols and considering the chosen observation geometry, why did you not carry out the computations using the complete kernel-model BRDF/M instead of using a constant albedo? I am wondering if the use of the BRDF/M considering different wind speeds could impact significantly your results.

We agree with the reviewer that an information content sensitivity test over the ocean with different wind speeds using the BRDF/M option in VLIDORT would be very interesting. However, a more comprehensive investigation of the information content for a particular case, for example analyzing the impact of the roughness caused by different wind speeds in the marine case, was beyond the aim of this study.

We have added one more case (Vegetation) to provide results for an albedo of 0.3. This allows a better understanding of the albedo dependence of the retrievals. In general we see a difference of ~0.4 – 0.5 DoF between a case with 0.05 (marine) and 0.3 (vegetation) albedo. With the exception of a glint case, one can thus argue that a change in albedo due to ocean surface roughness will likely also lead to variations of this magnitude.

Page 9 and figure 1, page 33 - The term "marine-artic" is quite confusing. It is not a unique study cases but two different study cases. Instead of "marine-artic", you should write "marine or artic". It will help the reader to understand that it corresponds to two different study cases as shown on figure 4.

We addressed this comment by changing the profiles definition and the corresponding plot. Because we introduced the results for a vegetation case in the revised manuscript, we now call this the "Low extinction" aerosol profile, which is then used for the vegetation, marine and arctic scenarios (page 9 line 10), which only differ in their albedo.

Page 12 - Could you give some examples of high quality future instruments?

At the very end of the conclusion section, we already indicated the NASA Panchromatic Fourier Transform Spectrometer (PanFTS) , as one possible next generation instrument that could benefit from our study. This instrument will cover roughly the 0.26-15 [microns] spectral range at very high spectral resolution.

Page 12, Eq. (7) - I may miss something but I do not understand why you divided Δν. According to Eq. (4) and Figure 2, Lc(ν) is in radiance units (W.m-2.sr-1 = J.s-1.m-2.sr-1). The numerator units are Joules, whereas the denominator units are J. photon-1.cm-1. It results that the N units are photon.cm and not photon.

The reviewer is correct; there is a mistake in the formula. The radiance term in the numerator in the Eq.(7) is expressed as the directional spectral intensity in frequency units [W sr-1 m-2 cm]. We corrected Eq. (7) accordingly, in order to have a dimensionless quantity for the number of photons. We also added a sentence at page 11 line 1 to clarify the units issue in Eq.(7).

"...(the term radiance refers to the directional spectral intensity in frequency units [W m−2 sr−1 cm]) ..."

Page 12, paragraph "Information content analysis" - In the third paragraph, I think that it quite difficult to identify the different steps of the methods from a practical point of view.

You use the wavelength-dependence radiances through the K matrix, don't you? What is the retrieved vector?

You wrote "According to the Bayesian formalism, the sensitivity of the retrieval of a set of parameters, to the measurements (...)". Could you please give some examples of the retrieved parameters? It will help the reader.

Thanks for pointing this out. We have expanded the first part of Section 3 to provide a better explanation of our method and to put it into the context of normal optimal estimation remote sensing retrievals. The following text was added/revised (page 12, line 21-25, page 13 and page 14 line 1-10).

*"An atmospheric state vector **x**, containing the parameters of interest, is related to a measurements vector **y**, through a forward model **F(x,b)**, which is a function of **x** and other model parameters not retrieved (vector **b**):*

$$\mathbf{y} = \mathbf{F(x,b)} + \boldsymbol{\varepsilon}$$

*where **ε** is the measurements error vector (e.g., instrument noise).*
*For atmospheric remote sensing retrievals, **F(x,b)** is normally a radiative transfer model. The solution to the inverse problem is then to use the forward model and the information from the measurement vector **y** to construct an estimate of the atmospheric state. For an inverse problem, the a priori state vector **xa** is often used as an initial estimate. **xa** is ideally chosen to be close to the true state **x**, based on existing knowledge of the atmosphere.*
*Because of the nonlinear nature of the dependence of the model on the true state vector (and the real measurements), the solution for a full retrieval of the retrieved state vector, **x** ˆ, has to be found using an iterative process, which ends when an optimal agreement between the model **F(x,b)** and the real measurements vector **y** is reached. Such an approach, however, requires a good knowledge of the properties of the parameters involved in the measurements (aerosol micro-physical and optical properties, spectroscopic parameters, model uncertainties, surface properties) and the relative errors for every measurement (instrument noise).*
*Here we are interested in a theoretical and more general quantification of the aerosol retrieval information content for different scenarios, with no reference to any specific mission, instrument or location, we perform a one-step retrieval instead, assuming the a priori parameters to be known and a linear dependence of the forward model **F(x,b)** on the state vector **x**. The forward model is directly applied to our first guess (linear regime), assumed to be the true state, for which we define an a priori uncertainty. We then derive the corresponding information content, calculated using the optimal estimation method. We remark that this methodology is not applicable for retrieval with real data, where the high non-linearity of the physical process and the imperfect knowledge of the a priori parameters requires the use of an iterative approach.*
According to the Bayesian formalism, the linearisation of the forward model about the state vector can be expressed as:

y *=* ***Kx*** *+* ***ε***

where ***K*** *is the functional derivative matrix, also called Jacobian, which represents the change in the measurement for a unit change in the retrieved parameter.*
For this study, the vertical aerosol extinction profile represents the retrieved parameter, ***x****, and the elements of* ***K*** *represent the derivative of the radiance with respect to the aerosol extinction coefficients, for every single layer and every single wavenumber. VLIDORT provides analytic derivatives of the Stokes vector field with respect to any atmospheric or surface property (Spurr, 2006). The Jacobians required in Eq. (9) can be calculated from these weighting functions by the application of the chain rule. For every scenario, we fix the model parameters, calculating the radiance and the Jacobian with a single run of the code, before using these quantities for the estimation analysis.*

Page 15 - You investigated the effect of different Sa on the total DoF for urban aerosol scenario and nadir looking geometry. However, page 8, you indicated that you assume a constant instrument looking down at 30° off-nadir. Why did you change the value of the view zenith angle?

After the nadir geometry used for the Sa and SSA sensitivity test, we decided to implement this new observation geometry to see if the DoF in the altitude tests can be increased with larger viewing angles, (still in a reasonable range of typical satellite viewing geometry). We found that the 30° off-nadir geometry provides higher total DoF (for a fixed solar zenith angle), due to the longer absorption path.

As mentioned in the response to comment 1, we have now expanded the calculations and respective discussion to a wider set of solar and viewing angle. The results of these calculations confirm that the aerosol information content is smaller in the nadir case at smaller illumination angles. We added new text in the manuscript at pag. 16 line 21-29, pag. 17 line 1-6.

"In order to explore the dependence of DoF on observation geometry, we vary the SZA (30◦ , 45◦ , 60◦ and 75◦) and the viewing angle (nadir and 30◦ off-nadir) implementing the urban scenario for the aerosol and assuming ω = 0.95 as single scattering albedo. Figure 4 shows the improvement of the information content due to resolution and higher values of SZA, for both observations. A different solar illumination changes the scattering angle of the particles, allowing a better information retrieval. At 5 cm−1 resolution, DoF improve from 1.39 (SZA=30◦) to 1.87 (SZA=75◦) and from 4.49 to 5.22 (same SZA) at 0.05 cm−1 , for the nadir case. The same change in resolution for the 30◦ off-nadir viewing angle allows a similar improvement in the DoF with a change from 1.57 (SZA=30◦) to 1.96 (SZA=75◦) at lower resolution and from 4.74 to 5.33 at 0.05 cm−1 resolution, for the same solar angles. As expected, the off-nadir viewing angle shows higher absolute DoF due to the longer lightpath."

Page 15, line 22 - What do you mean by "extra information"? Please, could you clarify?

The term "extra" was intended as "maximum" (information available). We agree that it is misleading and removed it from the text (page 17 line 17).

Page 16 - "The DoF improvement with spectral resolution of 2.7 (Urban) and (...) is similar for both scenario". This sentence is not clear. You should write: "of a factor 2.7 Urban) and 2.8 (Highly polluted) "

The reviewer is right. We replace the sentence (page 18 line 13). The revised sentence states now:

" The DoF improvement with spectral resolution is of a factor of 2.8 for the urban case and of 2.7 for the high polluted case."

Page 16 – Why is the DoF change surprisingly small in spite of the large change in aerosol extinction?

Is there any physical explanations linked with scattering or absorption processes of aerosols or the surface?

Our results do not give a direct explanation for this behavior, but we can speculate, based on our other results, that the small change in DoF is due to a reduced influence of surface albedo at higher extinction, i.e. less light reaching the surface. In addition, we also believe that multiple scattering at the very high extinctions could have played a role.

However, as reviewer #1 pointed out, the extinction in our highly polluted case was likely too large and in the revised version of our manuscript we have reduced the extinction in the highly polluted case to 0.2 km^{-1} and the urban case to 0.1 km^{-1}. The difference in DoF between these, more realistic, scenarios is now ~10% at high resolution and more than 15% at lower resolution.

We rephrased the sentence adding the relative new improvement factors for the urban and highly polluted case in the text, at page 18 line 13-16.

" The DoF improvement with spectral resolution is of a factor of 2.8 for the urban case and of 2.7 for the high polluted case. In general, the high polluted case has a 10–15 % higher DoF due to its higher BL aerosol extinction."

Page 17 - "It evident that (...) polluted scenario". Could you explain why?

As stated in the previous comment, we introduced a new urban polluted case with an extinction of 0.1 km^{-1} in a BLH of 1 km height. The elevated layer case shows a similar DoF improvement for the change in resolution, when compared with this new profile.

At 5 cm^{-1} resolution, the total DoF for the elevated layer is 1.4, which improves to 4.6 at 0.05 cm^{-1}. A two-order of magnitude change in the resolution improves the DoF of a factor of 3.3 for the elevated layer scenario. The polluted scenario has 2.01 DoF at 5 cm^{-1} and 5.38 at 0.05 cm^{-1}, improving the information content of a factor of 2.7. Despite the absolute values of the DoF are higher for the highly polluted case, the improving factor is lower, at every resolution.

We can argue again that the surface albedo could have played a role in the elevated layer case, with the opposite effect of the high extinction case. The influence of the light coming from the surface, could have been enhanced, because of the lower aerosol optical depth. Furthermore, light coming from the ground, is affected by a lower background aerosol extinction (BLH=0.4 km) near the surface.

Page 18 – Altitude ranges [5-15km] and [15-50km] do not correspond to altitude ranges on Figure 5b-f:

[5-10km] and [10-50km.]

The reviewer is correct. The plot label refers to an old test were the altitude ranges were different.

We replace the labels in Fig. 3 and Fig. 7.