Reply to referee L. Guanter

Dear Referee,

Below you will find detailed responses to your review of manuscript amt-2013-38, 'Retrieval of aerosol parameters from the oxygen A band in the presence of chlorophyll fluorescence'. The review helped to improve the manuscript and we would like to thank the referee for his effort and time.

Sincerely,

Bram Sanders (on behalf of the authors)

Note: page numbers in reviewer's comments refer to amtd manuscript, page numbers in responses refer to revised manuscript

--- --- ---

The manuscript deals with the retrieval of aerosol parameters and terrestrial chlorophyll fluorescence from spectrally-resolved measurements in the O2 A-band. Given that a number of spaceborne instruments measuring in O2A are currently operating, and that several others are planned for launch in the coming years, the topic of exploiting the O2 A-band for improved atmospheric and surface retrievals is of relevance to a wide community. In particular, the retrieval of fluorescence from space is evolving rapidly in the last couple of years, which is further supported by new studies in the line of the one presented in this manuscript.

My opinion is that the manuscript is timely and addresses an important field of research, and also that it is well written and presented. However, I think that the authors should address several critical points before the manuscript can be accepted for publication:

1) Representativeness of the simulations

As the authors discuss in the text, their precision estimates are significantly better than the ones provided by other authors (Frankenberg et al) dealing with a similar problem. In my opinion, this might be explained by an over-simplified simulation and retrieval set-up. For example, the effect of uncertainties in e.g. the temperature profile, surface pressure, aerosol optical properties, polarization or the direct/diffuse radiation ratio could change the precision estimates substantially through cross-correlation with the state vector parameters. Also, at least 2-3 parameters should be added to the state vector in order to account for the non-linear spectral shape of surface albedo (normally modeled by an n-order polynomial). Even in this simplified case, the authors mention "exceptions to the overall trends described in Sect. 4 exist" (p3199, L19).

The question is then to what extent the precision estimates achieved in this work would hold for a more realistic retrieval scenario including more state vector *elements and uncertainties in the forward model parameters. I consider that the authors should extend their simulation set-up so that their conclusions can really be considered representative of a real retrieval scenario. Simulations with only 4 free parameters and a flat and constant surface reflectance cannot recreate the complexity of the problem.*

Reply: We have now included in our retrieval simulations model errors in the single scattering albedo, surface pressure and temperature profile (see p.12-13). Furthermore, we have added calibration errors to the measurement error covariance matrix (next to the noise error). Finally, surface albedo and fluorescence emission in retrieval are described by a second-order polynomial (although we expect based on the literature that a linear wavelength dependence captures most of the spectral variation across the relatively small fit window). Effects of model errors on retrieval are investigated by including model parameters in the state vector with appropriate a priori errors. Thus, reported precision levels are estimates of the sum of retrieval errors due to measurement errors and smoothing errors but also model parameter errors. This approach takes into account that errors in retrieval parameters and forward model parameters may be correlated. We then take these precision values to represent realistic precision levels. Model errors in the aerosol phase function and the presence of more than one scattering layer are also discussed (p.16,l.17-28).

Concerning the reviewer's remark on "exceptions to the overall trends": We have observed in our O2 A band simulation studies that for specific scenarios, derivatives can suddenly become strongly linearly dependent and precision deteriorates. It is important to note that this is characteristic of O2 A band aerosol retrieval in itself and not so much the result of a simultaneous retrieval of aerosol and fluorescence parameters. If fluorescence is not fitted and is absent in the simulation, we observe the same thing. A small clarification has been added on p.18, l.4-6.

We prefer not to discuss the effect of ignoring polarization, as we feel that it is of minor importance for the main purpose of this manuscript. From retrieval simulations we know that the effect of ignoring polarization on retrieved aerosol pressure is limited.

2) Accuracy vs precision

Related to the previous point, non-expert readers might be confused by the small errors reported in Fig. 4. Apart from the potential over-simplification of the retrieval approach described before, systematic errors are not considered in the error budget. Even though the authors state clearly that those figures are only precision estimates from the propagation of instrumental noise, they also refer to comparisons of those precision errors with "scientific user requirements" which at least in the case of fluorescence will always include systematic errors. Actually, we know that biases can become very important when fitting wide spectral windows (>5-10nm) due to cross-correlation of fluorescence with the state vector parameters describing surface albedo. In fact, the formulation of surface reflectance in the forward model is critical for such wide fitting *windows, and more sophisticated approaches than polynomials in wavelength are necessary (this is not the case for narrow spectral windows containing only Fraunhofer lines).*

In my opinion, providing precision errors for a fluorescence retrieval method dealing with a relatively broad fitting window, as it is the case here, is misleading. In this case, precision errors may be significantly smaller than biases. The authors should consider to perform realistic end-to-end simulations in which both accuracy and precision are properly evaluated.

Reply: As explained above, we have now included the effect of model errors in single scattering albedo, surface pressure and temperature profile in precision estimates, and we discuss effects of model errors in phase function and errors due to the presence of more than one scattering layer. We also take calibration errors into account. We believe that we have captured the most important retrieval uncertainties and we assume reported precision estimates to be representative for realistic retrievals. We explicitly state this assumption on p.15,l.1-3. Furthermore, when comparing precision levels with values from the literature, we carefully state that 'errors […] may be put into perspective by comparing them against the benchmark numbers provided in Table 2.' (p.16,l.30-p.17,l.1)

Concerning accuracy errors vs precision errors: The effect of a forward model parameter error on the retrieved state is $\mathbf{G}_y\mathbf{K}_b(\mathbf{b}-\hat{\mathbf{b}})$ (Rodgers, 2000, Sect.3.2). The covariance of this accuracy error can be included in the a posteriori error by adding the model parameters to the state vector (with a priori errors equal to the variance of **b**). This is what we do. Hence, reported precision estimates include errors due to model parameter errors, as also explained above. Note that the focus of this analysis is really to describe single-retrieval precision: Fluorescence retrievals may be compiled into spatio-temporally averaged maps (thereby reducing the random error component), but this is not desirable for retrieved aerosol parameters (e.g. when for aviation safety purposes).

We have no indications that more sophisticated approaches to describing the surface reflectivity across the O2 A band than low-order polynomials are needed for aerosol retrieval.

We have redone simulations for the full state vector in Fig. 1 with an extended fit window (750 – 775 nm, i.e. length of the window has more than doubled). As before, surface albedo and fluorescence are described by a second-order polynomial. We observe that retrieval precision does not change except for precision of retrieved fluorescence, which improves (more FH-lines). In addition, we also find no indications that correlations between surface albedo and fluorescence are systematically higher.

3) O2 vs Fraunhofer lines

The first part of Section 5 presents a critical analysis of other works (Frankenberg et al 2011, 2012) also dealing with the retrieval of fluorescence from O2A *measurements. In particular, the authors put a lot of emphasis on the discussion of the information content provided by Fraunhofer lines for fluorescence retrievals, and state e.g. in the abstract "we also show that most of the fluorescence signal is provided by in-filling of the O2A band and to a lesser extent by filling-in of Fraunhofer lines".*

On the one hand, I think that this O2A/Fraunhofer discussion is unnecessary for this work, especially with such a direct language as the one used in Section 5.

Reply: Agree. The intended focus of the manuscript is retrieval of aerosol parameters, which we haven't made clear enough. The discussion of FH/O2 contributions is perhaps not that relevant for the purpose of this manuscript. We have therefore replaced these simulations with simulations investigating the dependence of retrieval precision on the a priori error in the fluorescence emission (Section 4.3). This is more relevant for O2 A aerosol retrieval: a Fs-retrieval using FH-lines might be implemented as a pre-processing step to provide an a priori value and constrain retrieval of aerosol parameters.

On the other hand, I was surprised by the findings in Fig.5 showing the apparent lack of impact of the Fraunhofer lines in the retrieval of fluorescence. This contradicts our own results using end-to-end retrieval simulations. I am attaching some figures from our own analysis. In short, I ran end-to-end simulations with the forward simulation data set and the statistical retrieval approach described in Joiner et al AMTD, 2013 (doi:10.5194/amtd-6-3883-2013). The end-to-end simulations were performed with and without a solar spectrum as the authors did to produce their Fig.5. Forward simulations comprise >200,000 case including different observation and illumination angles, atmospheric conditions (different values of surface pressure, T profile, aerosol optical thickness, model and height), and surface reflectance and fluorescence (from combinations of leaf area index and chlorophyll content). A constant SNR of 2000 is assumed. The mean and standard deviation of this test data set for the cases with and without solar irradiance are displayed in the Figs.1-2 of this review. Our end-to-end simulation results for the entire test data set are shown as a scatter plot in Fig.3 of this review. Diamond symbols and error bars show the mean and the standard deviation, respectively, derived from all the retrievals performed for the same surface state (reflectance and fluorescence spectra) under different atmospheric conditions and observation/illumination angles. Despite the almost identical fit residual obtained with and without the Fraunhofer lines (Fig.4 of this review), a very different retrieval performance is found, which contradicts the findings presented in this manuscript.

Of course the retrieval precision depends on the particular forward model configuration, state vector definition and associated assumptions. In this sense, the larger number of parameters inverted in our forward model (see Joiner et al) than in the one proposed in this manuscript makes our retrieval to be potentially more sensitive to instrumental noise. But nevertheless the improvements achieved with the the Fraunhofer lines seems concluding enough to challenge the authors' statement that most of the information is provided by the oxygen lines. I could provide the authors with the data base we generated to develop and test our own *fluorescence retrieval algorithms so that they can test some of the assumptions they are making in their approach.*

Reply: The results presented by the reviewer are interesting. We thank the reviewer for taking the effort to carry out these simulations. We have double-checked our own simulations and we find them trustworthy. At this point, we can only speculate then as to the origin of the differences. One aspect is that a wider fit window (747-780 nm) is used compared to our retrieval simulations (758-770 nm) and therefore more Fraunhofer lines are present. This may shift the balance into the direction of more information coming from these FH-lines. A careful analysis of the differences between the two retrieval approaches is needed, which we feel is beyond the scope of the present review.

The simulations with the flat solar spectrum have now been removed (see previous reply).

Other minor comments:

Title: it may be due to a personal bias, but my impression is that the manuscript is more focused on fluorescence than on the aerosol retrieval part, which is not clear in the title.

Reply: This is an important point. The intended focus of the manuscript is really the effect of fluorescence on aerosol retrieval (hence the title). We have not made this clear enough, as we apparently have given the impression that the manuscript was more directed towards fluorescence retrieval itself. We have modified the manuscript throughout and now focus more explicitly on aerosol retrieval. See also other remarks concerning this point.

Critical surface albedo (p.3199, L24): I think this concept was developed for multispectral data (MODIS-like). I am not sure that it applies to high spectral resolution data in which each spectrum samples very different atmospheric optical thickness for a relatively constant surface albedo).

Reply: Agree. The critical surface albedo is that particular ground albedo for which the continuum reflectance does not depend on AOT. For high spectral resolution data, one should in principle be able to distinguish between different AOTs from absorption in the O2 A band. However, one may still expect retrieval from high spectral resolution data to be more difficult for a critical surface albedo case than for a non-critical surface albedo case. From our retrieval simulations, it indeed seems that critical surface albedo cases typically have relatively flat chi-square functions or suffer from multiple minima. But we must also remark that sometimes retrievals that cannot be labeled as critical surface albedo cases are problematic as well. This is the reason why we wrote: "These singular cases often occur for optically thin layers over land and may be related (but not limited) to situations of a so-called critical surface albedo (e.g. Seidel and Popp, 2012)." We prefer to leave the sentence as it is. If the reviewer thinks this sentence needs more explanation, we are happy to modify it.