Interactive comment on “A simple and versatile cloud-screening method for MAX-DOAS retrievals” by C. Gielen et al.

Anonymous Referee #1

Received and published: 22 July 2014

General Comments

The manuscript entitled ‘a simple and versatile cloud-screening method for MAX-DOAS retrievals’ by Gielen et al. presents a method for the detection of clouds using the observed radiance and/or \( \text{O}_4 \) slant column density. Such a cloud flag can be used to filter out cloud-contaminated observations, but is also a valuable information by itself. Therefore, the topic of this manuscript fits well within the scope of AMT.

In general, the applied methods appear to be valid and the paper is well written. However, as outlined in the specific comments there are several aspects that lack conciseness and require some more detailed explanation.

It is mentioned several times (e.g., at the beginning of the abstract and in the introduction) that the method proposed here is based on zenith sky measurements only. This statement is a bit misleading as it raises the expectation that the method can be applied to zenith-sky instruments without any restrictions. However, the ‘multiple-scattering’ flag uses \( \text{O}_4 \) dSCDs which are based on measurements at 30° elevation angle and require either a MAX-DOAS instrument or a significant modification of the method proposed here. It should be mentioned that the method can be applied to zenith-sky measurements only with certain restrictions.

Apart from the measurements in Brussels, the improvement of the aerosol retrieval using the different cloud flags is not very convincing. For Xianghe, there is hardly any improvement of the agreement between AOD form MAX-DOAS and form sun photometer when applying a cloud filter. A statement on the level of improvement for the Jungfraujoch data cannot be made at all since, due to the small AOD at this site, there is virtually no correlation between MAX-DOAS and sun photometer. Here perhaps histograms of the difference between MAX-DOAS and sun photometer data would be more useful.

In summary, I feel that the conclusions should be formulated in a more balanced way regarding the capability of the algorithm to improve the retrieval of aerosol properties from MAX-DOAS measurements.

It is mentioned that the CI-based broken-cloud flag would be more sensitive to aerosols than the multiple-scattering flag, which would render the latter more suitable for the discrimination between clouds and high aerosol load. However, both flags are based on the detection of a high temporal variability of the measurements, in case of the broken-cloud flag based on CI and in case of the multiple-scattering flag based on \( \text{O}_4 \). Please explain why the sensitivity to aerosols should be higher for the CI-based flag than for the \( \text{O}_4 \)-based flag.

Specific Comments

5884.2 and 5885.11: These sentences suggest that only zenith sky spectra are used for cloud detection, but later on you also use measurements at 30° elevation.
Section 3: Very different wavelength pairs are used for the three sites. I understand that this originates from the different wavelength ranges of the instrument. However, wouldn’t it be useful to homogenise the definition of the CI for different instruments in the framework of this study for the sake of better comparability?

5889.10: The CI for Jungfraujoch is defined as $I_{405}/I_{560}$, but you state above (5887.25) that the instrument only covers a wavelength range up to 550 nm. How can this be?

5889.23ff: Here it should be explained in more detail how the simulations with DAK were performed, including a description of the aerosol profile shape, cloud thickness and cloud and aerosol optical properties. I could imagine that the cloud base height has a significant impact on the modelled quantities, in particular on $O_4$. It would therefore be important to vary also the cloud base height in the model simulations.

5891.1: I do not understand the concept of ‘broken or scattered clouds in the line of sight’ as this is a property of the entire sky and not only of a particular viewing direction.

5892.3: The definition of the ‘bad’ region is quite vague. At which value exactly do you cut off the peak? Do you use a particular percentile?

5892.20 Please specify what $x$ means in this equation. If it is time, please replace $x$ by $t$. Please avoid having the same symbol $f$ both as function name on the left side and as parameter on the right side.

5892.21: it is not clear what you mean with $|CI - \text{model}| > C$. I suggest to replace this equation by $|CI(t) - f(t)| > C$. Why is $C$ different for the different sites and, more importantly, how exactly did you determine the different values for $C$? For the multiple-scattering flag, you used the relative change of the signal. Would this also be useful for a more general definition of the broken-cloud flag?

5893.7: Not only clouds but also aerosols can have these effects on the $O_4$ dSCD. How can you distinguish between changes in light path due to aerosols and due to clouds?

5894.1: It should be specified that ‘outlier’ means flagged as affected by multiple scattering. Again, I suggest to replace the equation by $|O_4(t) - f(t)| / f(t)| > C$.

5897.17: I do not understand what you mean with ‘RMS < 50%’: 50% compared to what? The RMS is an absolute value quantifying the (error-weighted) difference between measurement and retrieval. In an ideal retrieval, it should be close the number of elements in the measurement vector (see, e.g., Rodgers, 2000).

5897.2: The statement that ‘A removal of data with evidence for the presence of clouds, be it either based on the sky and broken-cloud flag or the multiple-scattering flag, results in a much better agreement with the AOD measurements and retrievals’ only applies to the measurements in Brussels, as the correlation analysis later on shows. Instead, it appears from the data shown in Fig. 9 that measurements flagged as ‘bad sky’ have systematically higher AOD values than the other data, and are most of the time also flagged as cloudy by the sun photometer algorithm. It would be useful to investigate in more detail to what extent cloud flags from DOAS and from sun photometer coincide.

5900.1: Do you have an explanation for the poorer correlation between MAX-DOAS AOD and Sun photometer AOD in Brussels compared to Xianghe?

5900.20: I do not agree with your conclusions. For Brussels, there is clearly an improvement of the agreement between MAX-DOAS and sun photometer after a cloud screening has been applied. For Jungfraujoch, the correlation is very poor no matter if a cloud screening has been applied or not. And for Xianghe, the correlation is already very good for all data, and does not improve after a cloud screening has been applied. In particular, I cannot support the conclusion that ‘for Xianghe, additional information from $O_4$ DSCDs is invaluable for a correct cloud identification’, since this screening method leads to only very small changes in the correlation coefficient, and moreover leads to a slope of 1.16 for the measurements at 477 nm which is worse than when CI-based cloud flags are applied (slope of 1.01).

5900.26: Again, you method is not only based on zenith observations but, in case of
the O$_4$-based flag, also on observations at 30° SZA.

5902.4: Given the results of your correlation analysis in Section 6.2, I do not agree with your statement that 'When we apply the cloud filter to our aerosol retrievals we find an improvement in the agreement with other co-located measurements, such as from cimel and Brewer instruments, both in correlation and slope'. This is only true for the measurements in Brussels.

Technical Corrections

5890.15: Junchfraujoch -> Jungfraujoch

5890.25: 3 -> three

5891.8: get rid of -> remove

5893.13: effect -> affect

5896.1: depends -> depend

5897.18: points -> point

5898.27: 'data for which no co-located AOD measurements are available'

Figure 1: Please mark the different parts of the figure (a, b, ...) to which you refer to on P. 5889. What are the units of the O$_4$ dSCD?

Figure 6: In the figure caption, what do you mean with 'calculated CI values'? Do you mean measured CI values?


C1826