Review on "Assessing the Impact of Clouds on UV-visible Total Column Ozone Measurements in the High Arctic" by Zhao et al.

This work applies an algorithm to classify ZS-DOAS measurements at twilight according to cloud condition to understand the effect of clouds on the retrieval of total ozone column. This algorithm takes into account the colour index (CI), and the variation of CI and O_4 measurements along the day to differentiate between three cloud scenarios: cloudy, clear and intermediate.

The algorithm is applied to ZS-DOAS measurements for two different DOAS instruments located in a high latitude observatory at 80ºN.

The effect of clouds in DOAS TCO is investigated by comparison to Brewer TCO data series at the same observatory and to MERRA-2 data.

I think the long data series at Eureka observatory is highly valuable due to the difficulty to keep instruments measuring with high quality in remote locations for so long.

I think this work can be very interesting for the interpretation of DOAS TCO measurements, and to assess the way that clouds, even low clouds can affect to DOAS measurements in the visible region.

In my opinion this work should be published after addressing some mayor, minor and technical issues.

Major issues:

(1) To assign the label of cloudy, clear or intermediate, the variation of O_4 along the day is taken into account. I think this criterion can be stricter for GBS instrument than for SAOZ instrument due to their differences in FOV. As can be seen in figure 4 (although please, see technical comments about this figure), it seems that there are more "clear" data in the case of SAOZ than in the case of GBS. I was wondering if this fact could be due to the O_4 criterion. In figure 3, it is quite surprising that for year 2011 clear, cloud and intermediate cases are quite close for both instruments but this situation changes considerably for 2013 and it is clearly different for 2017. But both instruments are located in the same observatory, how is possible that the number of clear/cloudy days in 2017 can be that different? Maybe the O_4 criterion is too permissive for SAOZ and too strict for GBS? This could also have an effect in the difference on the bias for both instruments when compared to Brewer. If the algorithm is not properly working for SAOZ, some clear days can, in fact, be affected by clouds and that would explain the better agreement between SAOZ$_{CS}$ and Brewer than GBS$_{CS}$ and Brewer.

(2) To be sure that the effect you observe in the bias when CS is applied to GBS TCO is only due to the presence of clouds, have you take into account that most of cloudy days happen out of the summer? What SZA do you use to calculate TCO at summer? Could the observed bias to Brewer have some to do with the major weight of summer days when you eliminate the cloudy days?

Minor issues:

(1) Due to the high latitude of the observatory it is not possible to have DOAS measurements along the entire year. Please, in the description of the instrument include what is the annual period of measurements. From figure 4 and from data along the text it seems that the period is late winter to late autumn? It would be nice to know the months when DOAS and Brewer can measure.

(2) Section 4.1. Why the current agreement to Brewer and GBS is better in this work than in the previous work by Adams et al.?

(3) Taking into account the current results, it seems that in the case of Hendrick et al., not all the observed discrepancies between DOAS and Brewer were due to the temperature dependence of XS used in Brewer analysis or in this work the Brewer analysis takes into account this dependence? Do you observe also the same seasonal

difference (taking into account that you cannot observe the entire spring nor fall at 80ºN) that Hendrick et al. in the bias against the Brewer?

(4) Section 4.1.1, please indicate at any part of the text that the weather classification used here and in figure 5 is made by using meteorological data. If not, it is a little confusing.

Technical issues

(1) Figure 4. Please, unify ticks in the horizontal axis. The lower graph is different from the previous ones and this makes very difficult to see properly the measurement periods. Grid in the middle of each year would be also very helpful.

Colours in the legend are not coincident with the ones in the graphs.

As GBS_{CS} or $SAOZ_{CS}$ are over imposed to GBS and SAOZ respectively, it seems that there are more data for the CS filtered data than without any filter. This is a little bit confusing at first, I am not sure that it can be addressed, maybe using hollow symbols for CS case?

If possible it would be nice a greater graph.

(2) Sometimes the DOAS instrument GBS is called UT-GBS, please unify nomenclature along the text.