Review of

"Using ocean-glint scattered sunlight as a diagnostic tool for satellite remote sensing of greenhouse gases"

by

A. Butz, S. Guerlet, O. P. Hasekamp, A. Kuze, and H. Suto.

Overview

This manuscript presents an analysis of roughly 3 years of GOSAT ocean-glint observations. A method is presented to identify very clear soundings that are free from light-path modifications in the O_2A band. This uses the so-called "upper-edge" method, which calculates a ratio between a retrieved O_2 column under a non-scattering assumption, to that of the predicted O_2 column based on meteorological reanalysis from ECMWF. The authors demonstrate that using the method allows them to see some changes in the GOSAT instrument over time (at least in the O_2A band), as well as evaluate the inter-consistency of CO_2 retrievals from different spectral windows.

The paper is well-written and should be published in AMT after addressing the following list of minor questions and issues.

Specific Comments & Questions

Line 155. How does selection of the R-branch of the O_2A band enhance sensitivity? Is this due to instrument problems that are worse in the P-branch? It seems that if you only use the R-branch, you make your results more temperature-dependent (because the temperature jacobians are opposite in the P vs. R branches). Also, please state if you fit anything related to temperature, or if you simply fix the T-profile to the prior meteorology.

Line 195: How sensitive are your results (in terms of selecting the upper edge ensemble) to the fitting or not of the O_2A band offset?

Line 190-215: It seems that the retrieval you describe is completely independent among the different windows. Ie., there are no parameters that have non-zero jacobians in more than one fit window. If this is the case, please state it in the manuscript, just to add clarity to what you've done.

Line 311-352: This is extremely interesting! For GOSAT data, can you state what additional amount or fraction of data are screened when you apply the 0.05 albedo criterion? It would be interesting to know how many of these double-layer cases there appear to be.

Line 365/Fig 4: Figure 4 implies that the distribution of the O_2 ratio only depends on time, but not on space. Did you examine if there is any kind of spatial dependence (if you

aggregate over reasonably large regions)?   Please explicitly state in the manuscript if/that you assume this, and if you've seen any evidence of any kind of spatial/latitudinal dependence.

Line 380: If I understand the method correctly, you take what is between the 95th and 99th percentile of the O_2 ratio you derive, for each 10-12 day period. This implies that you ALWAYS select EXACTLY 4% of the soundings in one time period. Have I got this correct? It implies that if one period is much cloudier than another, it doesn't matter; you will always select exactly 4% of the ocean glint soundings (at least with the upper-edge criterion alone; I realize the h2o water line screen and the lambertian albedo screen will further remove some soundings). If so, it would be helpful to state this in the paper, and please compare it to the throughput rate of some of the more traditional approaches over ocean. It seems that 4% is rather strict, but it is hard to say.

Section 5 general comment: What are the typical mean squared fit residuals of the fit in each window, expressed as a %? It would be interesting for the reader to see this; large variations in this between the bands might be further suggestive of spectroscopy shortcomings in particular windows. In fact, it might be information to see a plot of mean fractional fit residuals in each fit window. If you find such a plot informative, please consider adding it to the paper.

Section 5 comment 2: I notice that the predicted (posterior) error of (W6/W2)*100, at 0.26%, is significantly larger than the observed scatter in that quantity (0.62%). This is the only window for which this is true. I further notice that W6/W2 experiences large outliers in 2009-mid2010, which seem to go away. Please comment. Are these outliers primarily driven by W6 alone? This result implies that W6 retrievals appear dependent on something that is time-dependent, to which W2 and W3 are not dependent.

Technical/Grammatical Comments

Line 158: furtheron → further on
Line 215: constraint → constrained
Line 334: is → are
Line 484: is → are
Line 512: detected difference → detected mean difference
Line 525: later → latter