# *Interactive comment on* "Preliminary verification for application of a support vector machine based cloud detection method to GOSAT-2 CAI-2" *by Yu Oishi et al.*

**Anonymous Referee #1**

This work describes some validation results of a VSM cloud detection method for GOSAT-2 CAI-2. The results indicate the accuracy of CLAUDIA3-CAI over rainforests is improved compared with the older version (CLAUDIA1-CAI). However, there are several important points in the manuscript that are not clear:

1. The language throughout the paper needs to be improved, the topic of this study is mainly on the cloud detection, while in Introduction section, the discussion is mainly on the $CO_2$ product. I suggest to reorganize the Introduction section and provided more background knowledge on the motivation of this work. 2. In the abstract, the authors referred the algorithm is developed in another paper (Page 1, Line 19-20), There is no

detailed discussion on the difference between these two close related works. 3. The uncertainty of the CLAUDIA-CAI is not fully analyzed and discussed. Several cases are shown in Figure 9 and 10 to describe the performance of CLAUDIA-CAI over various land cover types, however, there is no quantitative results and discussion. There needs to be a better description in Section 3.1. 4. Figure 6 and 7 are very similar with each other in the upper panel. It would be more concise to combine these two figures.

5. Page 2 line 24-25: what is meant by "Accuracy 16 times higher than at present is required assuming that the MRV for REDD+ and JCM needs an accuracy of 10 %"?

6. Page 5 line 21-30: "GOSAT returns to a similar footprint after 44 orbits (44 CAI paths) in three days. The satellite ground path of one orbit is divided into 60 equidistant CAI frames. We used the GOSAT CAI L1B product, which general users could download from the GOSAT User Interface Gateway (GUIG, https://data.gosat.nies.go.jp), for various land cover types on the beginning of the month from 2012 to 2014 in the same as the previous study (Oishi et al., 2017) (Table 2), and for rainforests (Table 3). Currently GUIG has been changed to GOSAT Data Archive Service (GDAS, https://data2.gosat.nies.go.jp/index_en.html). The spatial resolution of these products (pixel size at nadir) is 500 m, the image size is 2048 × 1355 pixels (approximately 1000 × 680 km). The CLAUDIA algorithm requires a land/sea mask and a surface albedo data. The CAI L1B product includes the Shuttle Radar Topography Mission's 15" land/sea mask. For areas with latitudes higher than ±60°, the USGS Global Land 1-KM AVHRR Project mask is used. Surface albedo data at 1/30° resolution was generated from the CAI L1B data from 10 recurrent cycles by separating the land and water regions." This section describes the data involved in the validation of algorithm, while the spatial resolution is inconsistent, did you resample the data? 7. Page 7 table 2: the last line "1 April 2012–1 March" data lacks the information of the year. 8. Page 7 table 3: the GOSAT product applied to Borneo seems only half of that to Amazon, why? 9. Page 8 figure 6: "Solar Zenith Angle >=85°(Night)" Could the authors provide some related references? 10. Page 12 line 4-5: "clear despite cloudy" and "cloudy

despite clear" are vague, and these terms should be replaced with already explained alphabets. Furthermore, it is better to provide references to the two formulas. 11. Page 14 figure 9, Page 17 figure 11, Page 18 figure 12, Page 21 figure13-14: the name of vertical axis just like "Percent" could be added. 12. Page 19 table 5: the second CLAUDIA3-CAI 0.56 should be corrected to 0.5. 13. Page 23 line 9-10: The authors wrote the results that "The averaged accuracy of CLAUDIA3 used with GOSAT CAI data (CLAUDIA3-CAI) was approximately 89.5 % in tropical rainforests, which was greater than that of CLAUDIA1-CAI (85.9 %) for the test cases presented here." But how to calculate 89.5% and 85.9% two values? 14. Page 23 line 6: purpose—> the purpose 15. Page 24 line 8: form—> from