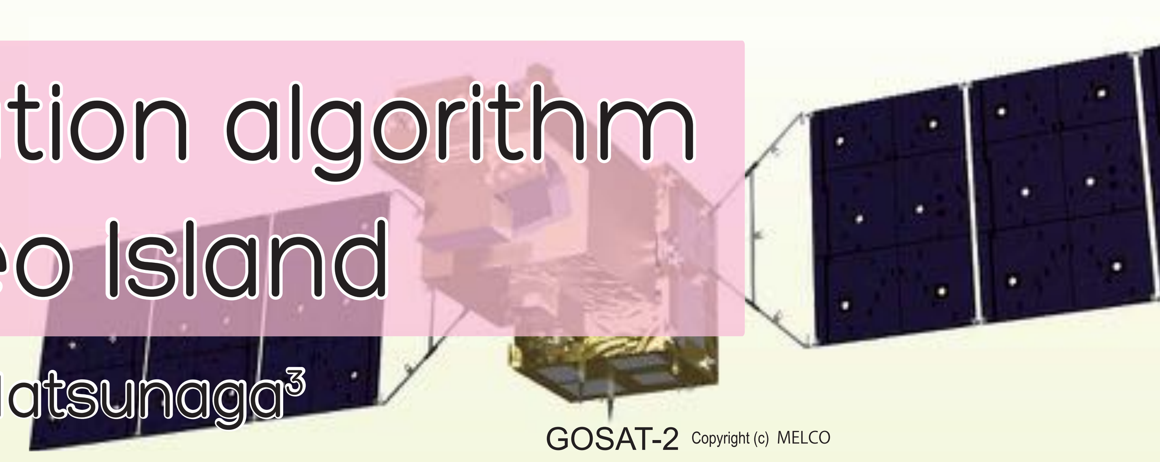
Primary verification of new cloud discrimination algorithm used with GOSAT TANSO-CAI in Borneo Island



Yu Oishi[1*], Haruma Ishida[2], Takashi Y. Nakajima[1], and Tsuneo Matsunaga[3]

[1] Tokai University, [2] Meteorological Research Institute, Japan, [3] National Institute for Environmental Studies, Japan

*Yu Oishi, E-mail address: oishi.yu@tokai-u.jp

1. INTRODUCTION

About GOSAT

Greenhouse Gases Observing SATellite (GOSAT) was launched in 2009 to determine global atmospheric CO_2 and CH_4 concentrations. GOSAT is equipped with two Earth-observing instruments: the Thermal And Near-infrared Sensor for carbon Observation-Fourier Transform Spectrometer (TANSO-FTS) and TANSO-Cloud and Aerosol Imager (CAI). The presence of clouds in the instantaneous field-of-view (IFOV) of the FTS leads to incorrect estimates of the concentrations[1)].

Cloud Discrimination

Thus, an important role of CAI is to perform cloud discrimination to identify and reject cloud-contaminated FTS data. Conversely, overestimating clouds reduces the amount of FTS data that can be used to estimate greenhouse gases concentrations. This is a serious problem in the region of tropical rainforest regions, such as Borneo, where the amount of useable FTS data is small because of cloud cover[2)].

New Cloud Discrimination Algorithm for GOSAT-2

Preparations are continuing for the launch of GOSAT-2 in fiscal year 2017[3)]. To improve the accuracy of the estimates of greenhouse gases concentrations, we need to refine the existing CAI cloud discrimination algorithm (CLAUDIA1). A new cloud discrimination algorithm using support vector machines (CLAUDIA3) was developed[4)].

This Study

Visual inspections can use the locally optimized thresholds, although CLAUDIA1 and CLAUDIA3 use common thresholds all over the world. Thus, the accuracy of visual inspections is better than that of these algorithms in the limited regions without areas such as ice and snow, where it is difficult to distinguish cloud and ground surfaces[2)]. In this study we evaluated the accuracy of CLAUDIA3 by comparing it with CLAUDIA1 and visual inspections of the same CAI images in Borneo.

2. USED DATA and ALGORITHMS

Study Area and Data

The orbits of the GOSAT platform repeat themselves every 44 revolutions (44 paths) around the Earth. Each path is divided into 60 frames. We used 1) CAI L1B products on Path 7, Frame 30-31 in Borneo, 2010, 2) Surface Albedo data generated from CAI L3 global reflectance distribution products.

Existing Algorithm (CLAUDIA1)

CLAUDIA1 comprises the calculation of clear-sky confidence level (CCL) for every threshold test and their comprehensive integration. Integrated-CCL of 0 means that the pixel is cloudy and 1 means that the pixel is cloud-free. Ambiguous pixels between cloudy and cloud-free are described by numerical values from 0 to 1[5)]. The threshold below which the integrated-CCL counts the pixel as cloud-free for GOSAT FTS L2 is 0.33, otherwise the pixel regarded as cloudy[6)].

New Algorithm (CLAUDIA3)

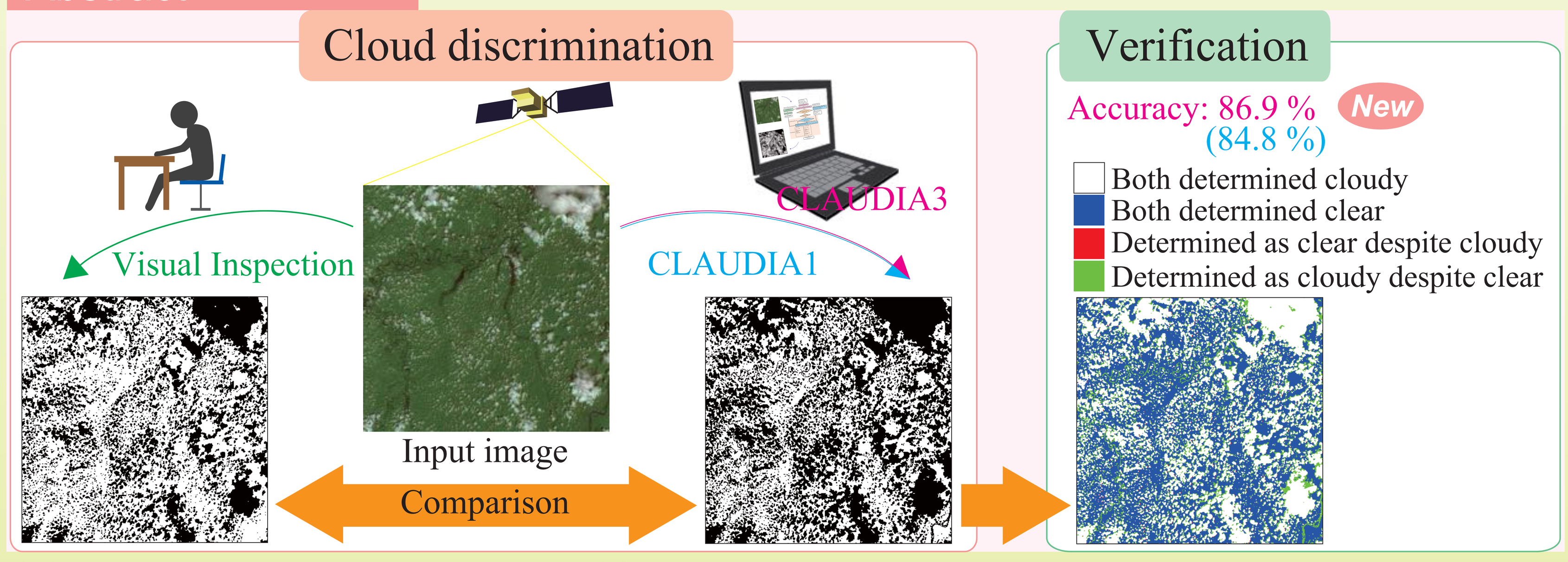
CLAUDIA1 performs cloud discrimination by using thresholds set based on experience. CLAUDIA3 uses SVM to decide the thresholds objectively by using multivariate analysis. CLAUDIA3 applies the Kernel Trick[7)] to soft-margin SVM[8)]. The kernel uses a second-order polynomial;

$$K(\boldsymbol{x}_i, \boldsymbol{x}) = \frac{(\boldsymbol{x}_i \bullet \boldsymbol{x} + 1)^2}{2},$$

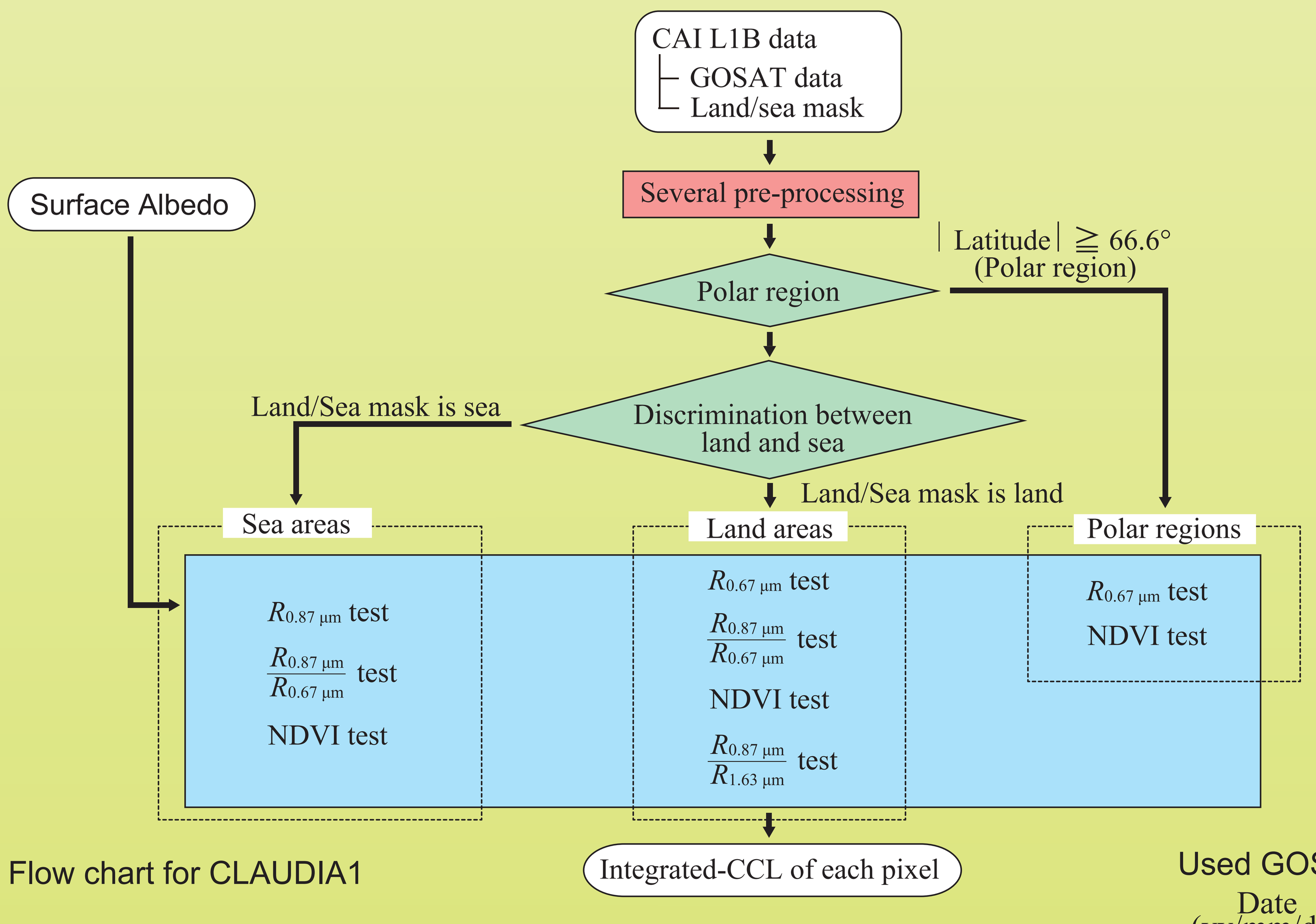
where K is the kernel function, $\boldsymbol{x}_i$ is the support vectors, and $\boldsymbol{x}$ is input data.

For CLAUDIA3, the integrated-CCL of 0.5 corresponds to the separating hyperplane of clear support vectors and cloudy support vectors.

Abstract



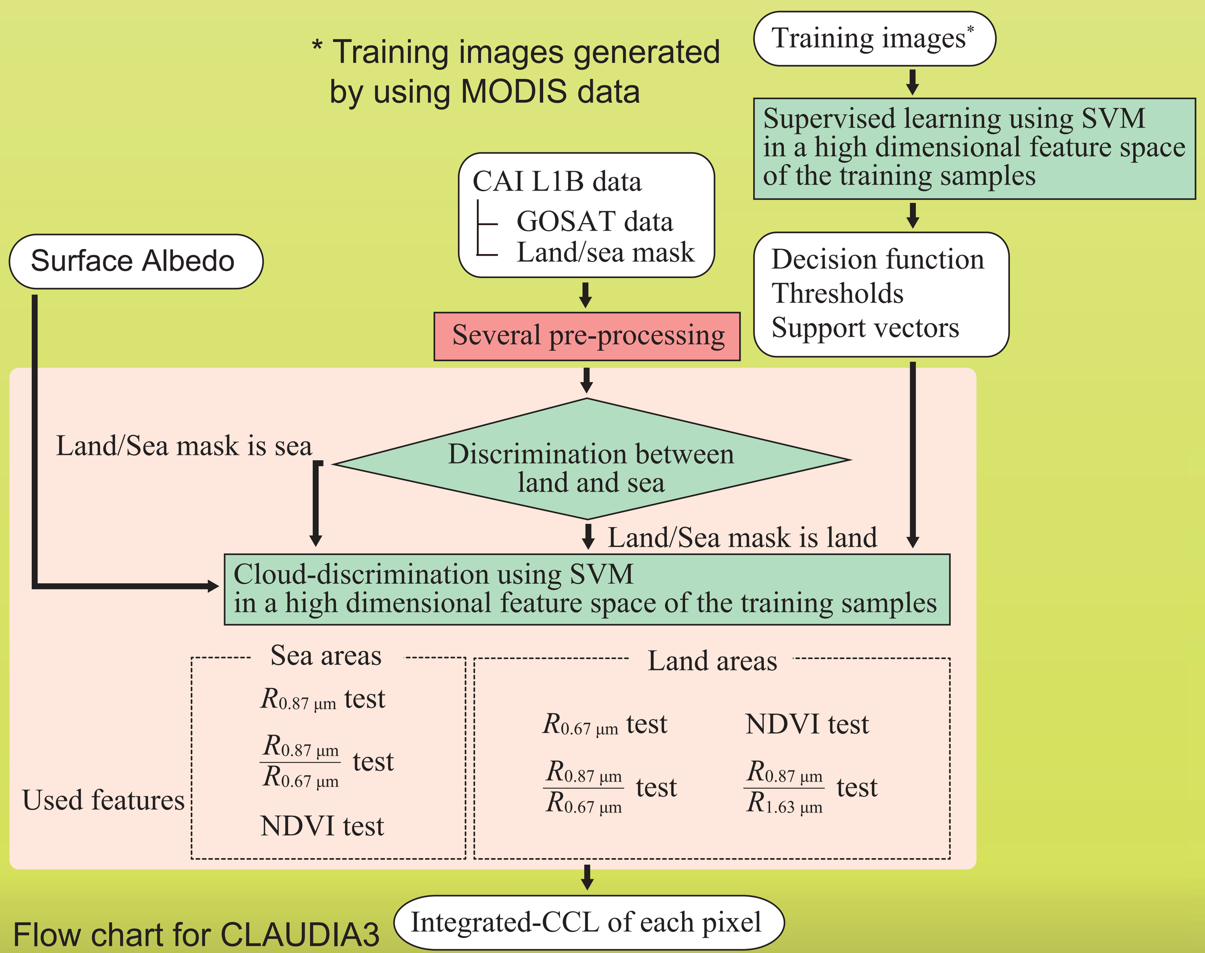
1) Cut 400 x 400 pixels around the center of CAI L1B images
2) Perform visual inspection of the pixels cut from the CAI L1B images
3) Perform cloud discrimination by using CLAUDIA1 and CLAUDIA3
   For CLAUDIA1, we produced output images setting the integrated-CCL threshold to 0.33.
   For CLAUDIA3, we produced output images setting the integrated-CCL threshold to 0.5.
4) Compare output with visual inspection
   We colored the images by comparing the visual inspection images with the output images pixel-by-pixel.



Flow chart for CLAUDIA1

Used GOSAT CAI L1B product

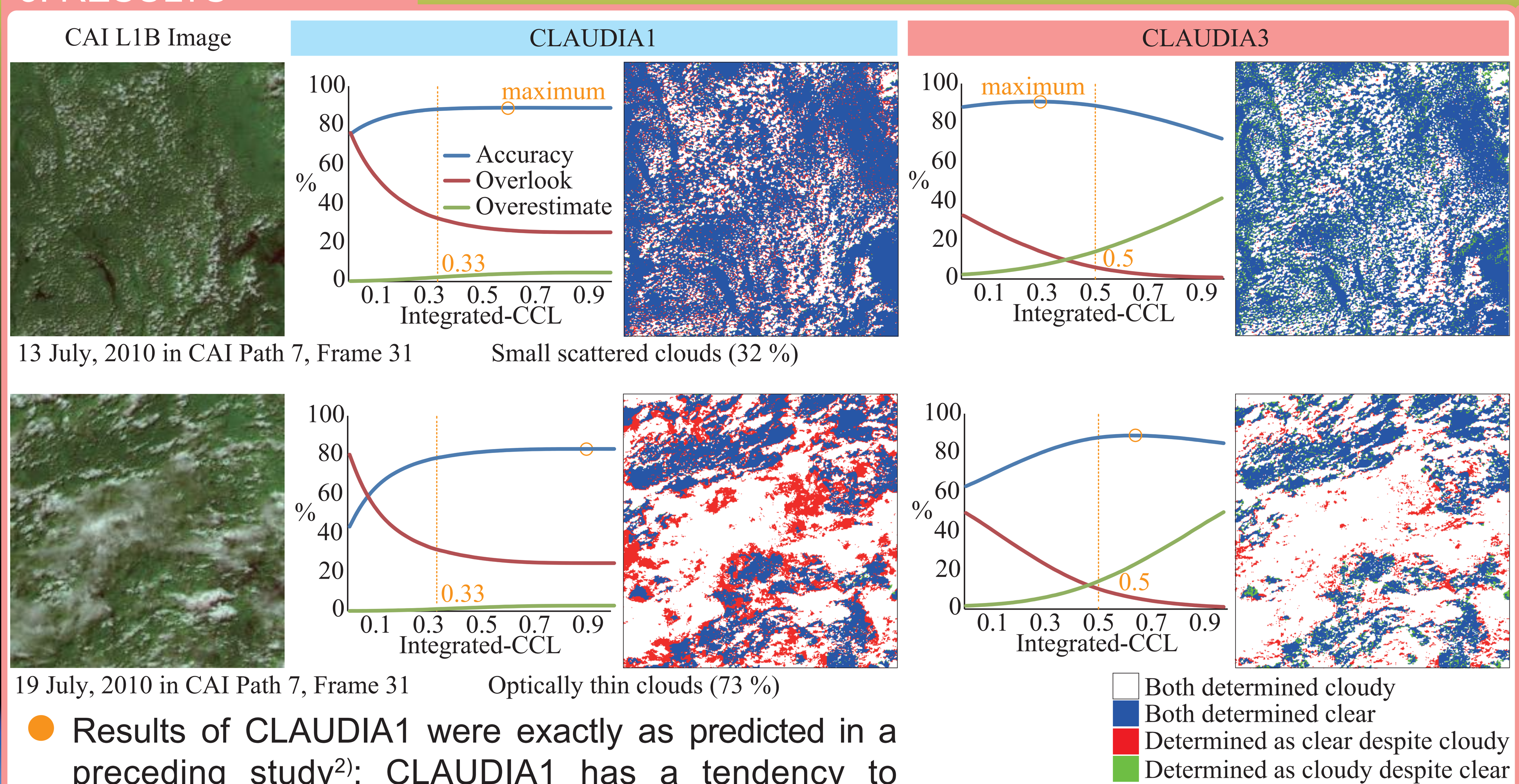
Date (yy/mm/dd)	Location (CAI Path_Frame)
10/04/02	7_30
10/01/02	7_31
10/04/01	7_31
10/07/01	7_31
10/07/07	7_31
10/07/13	7_31
10/07/19	7_31
10/07/28	7_31
10/09/02	7_31
10/11/01	7_31



Flow chart for CLAUDIA3

Study area

3. RESULTS

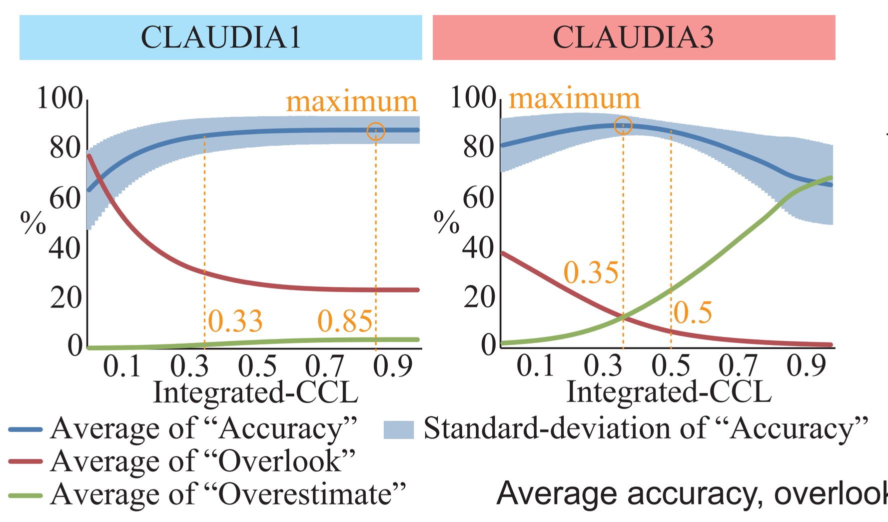


$$\text{Accuracy} = \frac{\text{Both cloudy} + \text{Both clear}}{\text{Total number of pixels}}$$

$$\text{Overlook} = \frac{\text{Clear despite cloudy}}{\text{Both cloudy} + \text{Clear despite cloudy}}$$

$$\text{Overestimate} = \frac{\text{Cloudy despite clear}}{\text{Both clear} + \text{Cloudy despite clear}}$$

- Results of CLAUDIA1 were exactly as predicted in a preceding study[2)]: CLAUDIA1 has a tendency to overlooks optically thin clouds and edges of clouds.
- CLAUDIA3 misjudged clear narrow muddy rivers as cloudy in the same manner as CLAUDIA1.
  - insufficient training data for muddy rivers
  - insufficient resolution of the surface albedo data
- CLAUDIA3 may be able to detect optically thin clouds that cannot be detected by visual inspection.
  - comparison with MODIS cloud mask product and satellite LiDAR data, such as CALIOP



	CLAUDIA1		CLAUDIA3	
Integrated-CCL threshold	0.33	0.85	0.50	0.35
Accuracy	84.8 %	87.5 %	86.9 %	88.8 %
Overlook	31.3 %	23.4 %	6.3 %	12.8 %
Overestimate	1.2 %	3.4 %	21.6 %	10.5 %

Average accuracy, overlook, and overestimate for all data for Borneo.

The most suitable integrated-CCL thresholds are 0.85 for CLAUDIA1 and 0.35 for CLAUDIA3 in Borneo.

Acknowledgments

This research is supported by the GOSAT-2 Project at the National Institute for Environmental Studies (NIES), Japan (2015, 2016). The authors would like to thank the GOSAT Project, GOSAT-2 Project, and Dr. T. Endo for their helpful comments; Mr. T. Hirose for his assistance with visual inspection.

References

1) O. Uchino et. al.: Influence of aerosols and thin cirrus clouds on the GOSAT-observed CO_2: A Case study over Tsukuba, Atmos. Chem. Phys., 12, pp. 3393-3404, 2012.
2) Y. Oishi et. al.: Evaluation of the accuracy of GOSAT TANSO-CAI L2 cloud flag product by visual inspection in the Amazon and of the impact of changes in the IFOV sizes of TANSO-FTS, Jour. Rem. Sen. Soc. Japan, 34, 3, pp. 153-165, 2014.
3) NIES GOSAT-2 Project (2016). GOSAT-2 Project at the National Institute for Environmental Studies, about GOSAT-2. www.gosat-2.nies.go.jp
4) H. Ishida et. al.: Development of a cloud-discrimination algorithm based on multivariate-analysis for a spaceborne multispectral imager, Proc. Met. Soc. Japan 2015 Aut. Meeting.
5) H. Ishida, T. Y. Nakajima: Development of an unbiased cloud detection algorithm for a spaceborne multispectral imager, JGR, 114, D07206, 2009.
6) Y. Yoshida et. al. (2010). ATBD for CO2 and CH4 columun amounts retrieval from GOSAT TANSO-FTS SWIR. https://data.gosat.nies.go.jp/GosatWebDds/productorder/distribution/user/ATBD_FTSSWIRL2_V1.1_en.pdf
7) B. Boser, I. Guyon, and V. Vapnik: A training algorithm for optimal margin classifiers. COLT '92 Proc. 5th workshop on computational learning theory.
8) C. Cotes and V. Vapnik: Support-vector networks. Machine Learning, 20, pp. 273-297, 1995.