METHODOLOGY   **Open Access**

Theoretical basis of the algorithms and early phase results of the GCOM-C (Shikisai) SGLI cloud products

Takashi Y. Nakajima[1*] ⓘ, Haruma Ishida[2], Takashi M. Nagao[3], Masahiro Hori[4], Husi Letu[5], Riko Higuchi[6], Naoya Tamaru[6], Naritoshi Imoto[6] and Akihiro Yamazaki[7]

Abstract

This paper discusses the cloud/clear discrimination algorithm (CLAUDIA) and the cloud microphysical properties algorithm (CAPCOM), which are used by the Second-generation GLobal Imager (SGLI) aboard the GCOM-C satellite, launched in December 2017. Also described are the preliminary results of cloud products' validation. CLAUDIA was validated by comparing cloud fractions derived from satellite data against data from whole-sky images captured by ground-based fisheye cameras. User's accuracy and producer's accuracy were mostly high at around 90%, and the resulting overall accuracy was also high, ranging from 83 to 100% (average of all sites was 90.5%). CLAUDIA has proven to be sufficiently accurate to apply a cloud mask to measurements and meets the requirements for releasing data for SGLI cloud flag products (the minimum for a successful GCOM-C mission). CAPCOM was evaluated by comparing cloud properties obtained by SGLI products against data from MODIS collection 6 products (MOD06). This was done for both ocean and land in the low to middle latitudes (60° N–60° S) from August 22, 2018 to September 14, 2018. The comparison showed good correlation coefficients for cloud optical thickness, effective particle radius, and cloud-top temperature for water clouds: 0.88 (0.83), 0.92 (0.88), and 0.94 (0.92) for ocean (land), respectively. CAPCOM data for ice cloud optical thickness correlated well with data from MODIS products: 0.86 (ocean) and 0.82 (land).

Keywords: GCOM-C, Shikisai, SGLI, Climate change, Cloud properties, Remote sensing, Cloud evolution process, Radiation budget, CLAUDIA, CAPCOM

Introduction

The Intergovernmental Panel on Climate Change (IPCC) Assessment Report-5 (AR-5) predicted that the atmospheric temperature will increase in the 2100's by 0.3 ℃ – 1.7 ℃ in the RCP2.6 scenario and 2.6 ℃ – 4.8 ℃ in the RCP8.5 scenario, with uncertainties of 0.4 ℃ and 2.2 ℃, respectively. RCP stands for representative concentration pathway, and 2.6 and 8.5 denote radiative forcings of 2.6 and 8.5 W/m^2 in 2100 (IPCC 2013). The large uncertainties in temperature predictions reflect our lack of knowledge of cloud properties and related processes such as those of indirect aerosols and this uncertainty will likely drive further study of such properties. The IPCC also reported that the largest uncertainties in global atmospheric temperature prediction are due to the effects of atmospheric aerosol forcing. Direct effects of aerosols in the atmosphere are both warming and cooling, cooling being dominant. But warming is predominant for absorption aerosols such as black carbon. The indirect effect of aerosols is known as the cloud-adjustment effect.

The injection of aerosols into the atmosphere changes cloud systems in two ways: by the Twomey effect (Twomey 1977, Twomey et al. 1984), which increases cloud albedo; and by the Albrecht effect (Albrecht 1989), which increases the cloud lifetime by increasing the liquid water path (LWP) in the cloud and thus suppressing precipitation. Both effects cool the surface temperature by changing the radiation budget. Overall, radiative forcing of cloud-adjustment effects is assumed to be − 1.33 to − 0.03 W/m^2. The uncertainty is still large (IPCC 2013), so further investigation is needed. This explains why aerosol and cloud

* Correspondence: nkjm@yoyogi.ycc.u-tokai.ac.jp
[1]Research and Information Center, Tokai University, 4-1-1, Kitakaname, Hiratsuka, Kanagawa 259-1292, Japan
Full list of author information is available at the end of the article

properties have attracted attention as have the transformation of aerosols to rain in clouds and the energy budget.

There are many methods for determining cloud properties: ground-based observatories, aircraft, and satellites for instance. They all have various trade-offs to consider. Ground observatories and aircraft observations are poor for understanding the synoptic scale of atmospheric phenomena, such as aerosols and clouds. Satellite observation is the best method for observing aerosols and clouds on the synoptic to global scale.

Satellite observations provide data via remote sensing rather than direct measurement, adding further uncertainty in the retrieved physical parameters. But data of higher accuracy can be obtained through better calibrations and validation. One of the best tools is the geostationary meteorological satellite, which can take frequent observations, although three or more satellites are required to cover the globe. Another way to make full-scale observations is by using polar orbiting satellites. A single one of these satellites can cover the entire globe. The clarity of observations is an advantage too, although 2 or 3 days are required for full coverage.

One such polar orbiting sensor is the Second-generation GLobal Imager (SGLI) aboard the Global Change Observation Mission-Climate (GCOM-C, "Shikisai" in Japanese), which was launched on December 23, 2017 (Fig. 1). It was conceived, designed, and developed specifically for this purpose. The SGLI has 19 channels covering the ultraviolet to thermal infrared spectrum and is a refinement of the sensors used in the GLobal Imager (GLI) aboard the Advanced Earth Observing Satellite II (ADEOS II, "Midori-II" in Japanese), which was launched in 2002 (Nakajima et al. 1998). Table 1 summarizes the channel specifications of the SGLI. The SGLI has some advantages over visible-to-infrared imagers on other polar orbiters and geostationary meteorological satellites. First, the SGLI's channels have a 250-m resolution

which is optimal for ground and ocean imagery. Second, it has two polarization and bidirectional channels in the visible (673.5 nm) and near-infrared (868.5 nm) regions, which are effective in making measurements of aerosols over land. The 380-nm channel is also useful in monitoring aerosols over land because aerosol reflectance is low at these wavelengths (Nakajima et al. 2005).

This paper focuses on cloud observations using the SGLI. The SGLI atmospheric group has developed two algorithms for obtaining information about clouds. One discriminates between clouds and clear skies and is called the CLoud and Aerosol Unbiased Decision Intellectual Algorithm, or CLAUDIA. The other determines the microphysical properties of a cloudy area and is called the Comprehensive Analysis Program for Cloud Optical Measurement, CAPCOM. Table 1 summarizes the SGLI channels used by these algorithms (the two right-most columns). CLAUDIA uses the VN8, VN11, SW1, SW2, SW3, T1, and T2 channels; CAPCOM uses the SW1, SW3, SW4, and T1 channels. Here, VN, SW, and T indicate visible-to-near-infrared, shortwave, and thermal wavelengths, respectively.

The next section will explain the theoretical basis of the algorithms used by CLAUDIA and CAPCOM. The "Results and Discussion" section reports preliminary validation data for SGLI cloud products. Cloudy/clear discrimination was validated by data obtained from ground-based whole-sky cameras. For cloud property validation, the paper compares the SGLI output with that of similar sensors, such as the MODerate resolution Imaging Spectroradiometer (MODIS). The final section summarizes this paper.

Methods/Experimental

Cloud/clear discrimination algorithm, CLAUDIA

SGLI data analysis needs to discriminate between cloudy and clear pixels before it can determine cloud and aerosol properties, so correct discrimination is required for valid,

Fig. 1 GCOM-C satellite launch on the H-2A rocket from Tanegashima Space Center, Japan, on December 23, 2017

Table 1 GCOM-C SGLI channel specifications and channel use by CLAUDIA and CAPCOM. Lmax denotes maximum observed radiance for each channel

	Channel	Center Wavelength VNR, SWI: nm TIR: μm	Band width	Lmax VNR-SWI: W/m2/sr/μm TIR:Kelvin	Spatial Resolution (m)	CLAUDIA	CAPCOM
VNR: Non -Polarization	VN1	380	10	210	250		
	VN2	412	10	250			
	VN3	443	10	400			
	VN4	490	10	120			
	VN5	530	20	350			
	VN6	565	20	90			
	VN7	673.5	20	62			
	VN8	673.5	20	210		✓	
	VN9	763	12	350			
	VN10	868.5	20	30			
	VN11	868.5	20	300		✓	
VNR: Polarization	P1	673.5	20	250	1000		
	P2	868.5	20	300			
SWI	SW1	1050	20	248	1000	✓	✓
	SW2	1380	20	103		✓	
	SW3	1630	200	50	250	✓	✓
	SW4	2210	50	20	1000		✓
TIR	T1	10.8	0.7	340	250	✓	✓
	T2	12.0	0.7	340		✓	

efficient retrievals. The SGLI has several channels that are effective in cloud/clear discrimination via an algorithm based on radiance intensities from the visible to infrared. CLAUDIA, a cloud/clear discrimination algorithm, was developed by Ishida and Nakajima (2009) for visible-to-infrared imagers and introduced the clear confidence level (CCL), as does the algorithm for the MODIS cloud-masking product, MOD35 (Ackerman et al. 1998), but it is designed to perform unbiased discrimination. Nakajima et al. (2011) reported the effectiveness of CLAUDIA in detecting clouds compared to the specifications of sensors used in MODIS, the SGLI, the multi-spectral imager (MSI) on the Earth Cloud, the Aerosol and Radiation Explorer (EarthCARE) satellite, the Advanced Very High Resolution Radiometer (AVHRR, model 2), and the Cloud and Aerosol Imager (CAI) on the Greenhouse Gases Observing Satellite (GOSAT) satellite. This subsection describes the principles behind the design of CLAUDIA.

The cloud/clear discrimination algorithm
The cloudy/clear discrimination algorithm is based on the radiative properties of clouds and the contrast between imagery of clouds and the ground surface (i.e., a clear sky). It utilizes (a) reflectance of solar radiance, (b) dependence of reflectance on wavelength, (c)

reflectance of solar radiance at the absorption wavelength, (d) emission in the thermal infrared region, and (e) dependence of thermal emission on wavelength. It uses one visible (VN8), four near-infrared (VN11, SW1, SW2, SW3), and two thermal infrared (T1, T2) SGLI channels (Table 1). The details of the radiative properties and how to use them in SGLI cloudy/clear discrimination will be explained.

(a) Reflectance of solar radiance

In general, clouds with a certain optical thickness have a higher reflectance of solar radiance than earth surfaces; therefore, reflectance can be an efficient indicator for differentiating between cloudy and clear pixels. In particular, absorption by atmospheric molecules in the visible region is small, so a high reflectance is always caused by clouds. Channel VN11 (868.5 nm) is in the near-infrared region but molecular absorption is also low, and the effect of Rayleigh scattering is minimal in the near-infrared channels. Thus reflectance data from channel VN11 is used for cloud discrimination over bodies of water such as oceans, rivers, and lakes. On the other hand, the reflectance data from channel VN8 (673.5 nm) is more

suitable for land reflectance than in the near-infrared, where reflection from vegetation is large and reduces the contrast between land and clouds. The reflectance data from channel SW1 (1050 nm) is also used. Pixels that are in the swath of SW1 but not VNR are tested via SW1 only.

The reflectance, R_f, is calculated based on the radiance detected by the sensor and database of solar insolation as follows:

$$R_f = \frac{\pi I}{\mu_0 \left(\dfrac{F_0}{d^2}\right)} \tag{1}$$

where I is the radiance measured in a channel, μ_0 is the cosine of the solar zenith angle, F_0 is the solar irradiance at the top of the atmosphere, and d is the distance between the sun and the earth in astronomical units. Calculating solar irradiance requires the response function for each channel. Cloudy/clear discrimination can be done accurately from using the minimum albedo, the minimum surface reflectance data for each channel over 1 month; these values are taken as the surface albedo. If these data are not available, the minimum albedo can be calculated from MODIS data instead. The reflectance per used channel is then compared to the minimum albedo. However, it is difficult to discriminate between clouds and surfaces with higher reflectance (such as snow-covered areas or desert) using reflectance alone.

(b) Dependence of reflectance on wavelength

The reflectance of clouds changes little with wavelength in the visible region, as clouds are usually white. In contrast, the reflectance of several types of land surfaces does change with wavelength, so the difference (or ratio) of reflectances at different wavelengths can provide a good indication of a cloudy sky. In particular, clouds over vegetation can be distinguished by the difference in reflectance between the visible and infrared regions. The normalized difference vegetation index (NDVI) is generally used as an indicator of the level of vegetation and is given by:

$$NDVI = \frac{R_f(868.5 \text{ nm}) - R_f(673.5 \text{ nm})}{R_f(868.5 \text{ nm}) + R_f(673.5 \text{ nm})} \tag{3}$$

The value of NDVI is high for areas with vegetation, whereas it is close to 0 for clouds.

Over sandy deserts or areas with sparse vegetation, the difference between the reflectance in the infrared and visible regions is small and the reflectance is large, similar to that for clouds. In the near-infrared, the reflectance of deserts tends to increase monotonically with the wavelength, whereas the reflectance of clouds does not depend on wavelength. The ratio between the reflectance of VN11 (868.5 nm) and SW3 (1630 nm) is given by:

$$R_f(868.5 \text{ nm})/R_f(1630 \text{ nm}) \tag{4}$$

This can be used for cloud discrimination over deserts, but it is not so effective for light surfaces such as ice and snow or icy areas. It is difficult to distinguish such surfaces and clouds based on the dependence of reflectance on wavelength.

(c) Reflectance of solar radiance at the absorption wavelength

The reflectance of solar radiance at the absorption wavelength can be used to detect high clouds, such as cirrus. Channel SW2 (1380 nm) is located within near a water vapor absorption band. With a clear sky, the intensity of the 1380 nm wavelength reaching a satellite is low because the atmosphere readily absorbs this wavelength, obscuring the influence of difference between the surface and low clouds. In contrast, if there is an object at high altitude in the atmosphere, the radiance is reflected by the object and the absorption is smaller, increasing the intensity detected by the satellite. For this reason, channel SW2 is particularly useful for detecting high clouds, even if they are optically thin. However, the radiance measured by SW2 can sometimes be large for clear skies over highlands because of lower absorption by the atmosphere.

(d) Emission in the thermal infrared region

Because clouds are usually colder than land, radiance in the thermal infrared region can be used for cloud/clear discrimination, especially for high clouds. However, the radiance of channel T1 (10.8 μm) changes greatly with the surface temperature, which makes it hard to determine the threshold for detecting clouds. This is why the authors used the radiance received through T1 only in the restoral test as follows: if the brightness temperature of channel T1 is sufficiently larger than the general (or climatological) surface temperature, the area is assigned as clear sky regardless of the results of other tests.

(e) Dependence of thermal emission on wavelength

Water vapor absorption at the wavelength of channel T2 (12.0 μm) is slightly larger than that of channel T1 (10.8 μm). Channel T2 also provides a larger imaginary part of the refractive index for ice than does channel T1.

As a result, the brightness temperatures of channels T1 and T2 (called a "split window") of thin cirrus clouds are generally greater than that of clear sky and optically thin clouds (Inoue 1987). The authors therefore use the split window to detect thin cirrus clouds.

Clear confidence level

Cloud/clear discrimination is done by comparing the radiance from the SGLI to the threshold value, which defines the boundary between clouds and clear sky. Several tests have two ends, discriminating as clear or cloud when the observed value is within a certain range. This method is called the threshold test. The cloudy/clear discrimination algorithm consists of many threshold tests to improve accuracy because a threshold test that is effective for a certain cloud or surface type may not be appropriate for another type. However, it is difficult to define a threshold value accurately in all cases because the optical thickness of a cloud changes gradually, and hence the boundary between cloud and clear sky is not well defined. It is inevitable that some ambiguity arises. Several cloudy/clear discrimination algorithms interpret ambiguous areas as clear (cloud conservative) or cloud (clear conservative) under a fail-safe determination. The cloudy/clear discrimination algorithm for the SGLI has to be unbiased because its results are used to analyze both cloud and aerosol properties. An algorithm developed by Ishida and Nakajima (2009) for neutral cloudy/clear discrimination leads to unbiased results and is used in the SGLI.
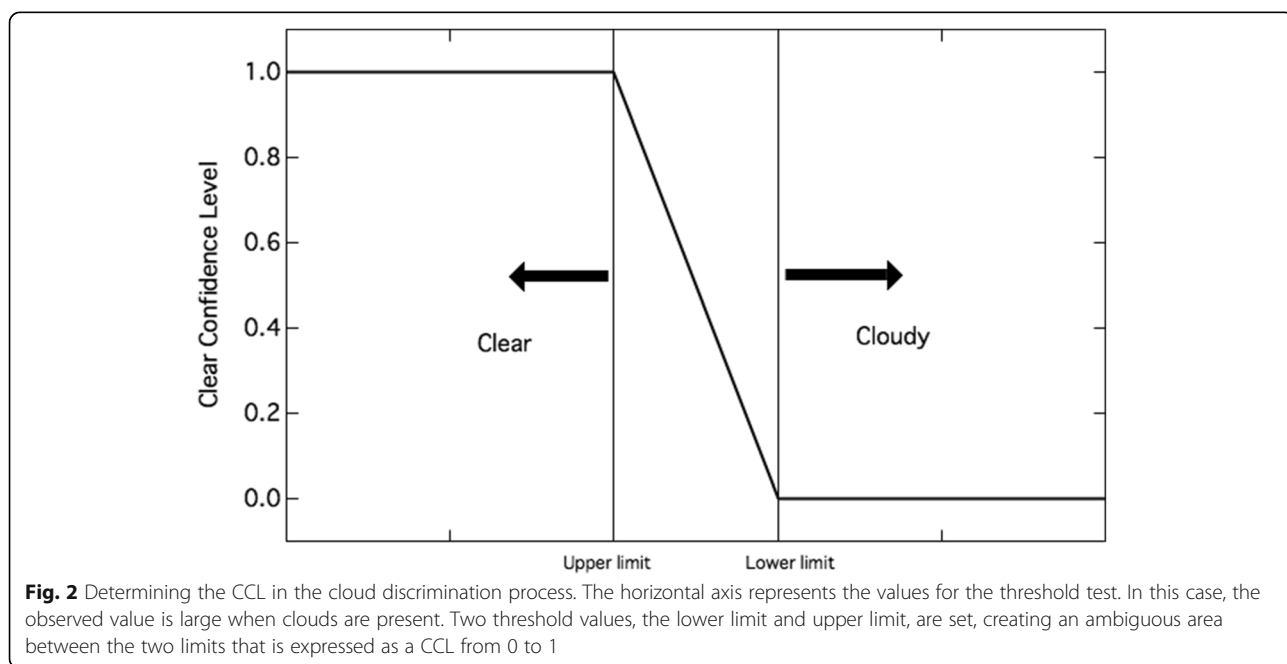
To achieve neutral cloud/clear discrimination, Ishida and Nakajima (2009) introduced the CCL and categorization by threshold tests. Two preset threshold values, an upper limit and a lower limit, are defined for the threshold test rather than one value, as is typical (Fig. 2). The CCL is determined by comparing the observed radiance against the two thresholds. A confidence level of 0 means that the area is cloudy and a level of 1 means that the area is clear sky. A CCL between 0 and 1 indicates an ambiguous area. If it makes sense to set a range around a certain value as the criterion for determining cloudy or clear skies, upper and lower boundaries for the ranges must each be identified by setting two thresholds, giving a total of four thresholds (Fig. 3).
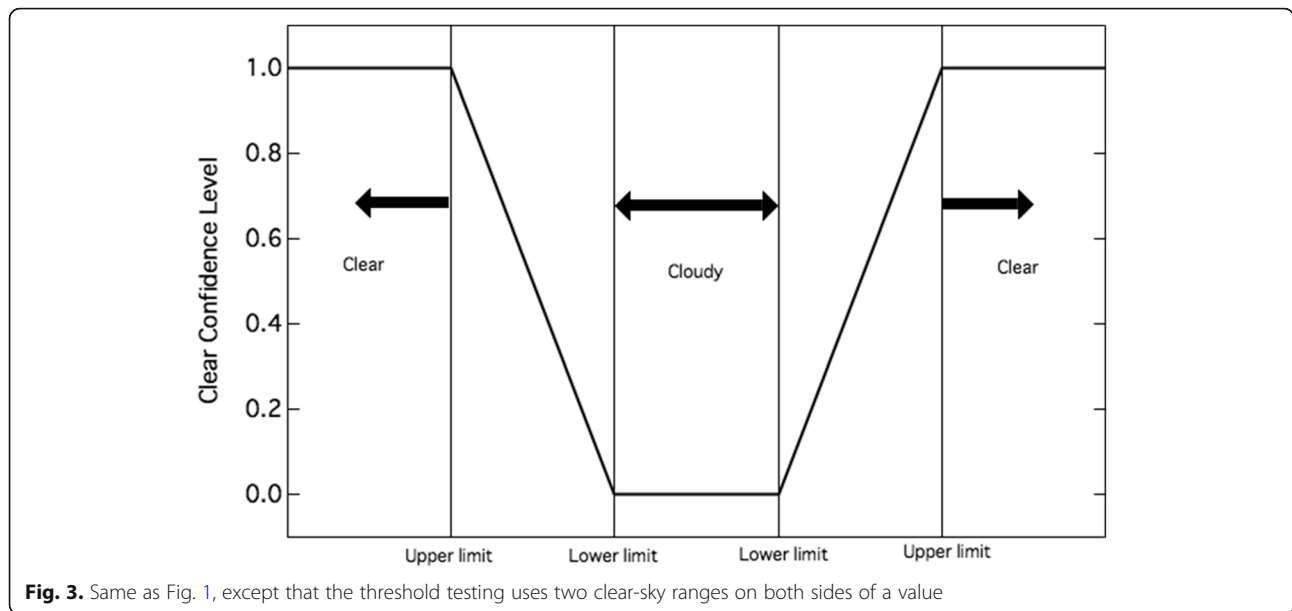
An overall CCL must be determined, integrating the CCL of all the threshold tests. For realizing neutral discrimination of cloudy or clear skies, threshold tests are categorized into two groups: those that tend to identify clear skies incorrectly as cloudy (group 1), and those that tend to identify cloudy areas incorrectly as clear (group 2).

The representative CCL for group 1, G_1, is determined to be cloud conservative:

$$G_1 = 1 - \sqrt[n]{(1-F_1)\cdot(1-F_2)\cdots(1-F_k)\cdots(1-F_n)} \qquad (5)$$

where F_k is the CCL of the kth threshold test in group 1. Equation (5) means that if any of the threshold tests result in a CCL of 1, $G_1 = 1$ (clear) regardless of the other threshold tests of group 1, whereas $G_1 = 0$ (cloudy) only when all the F_k values are 0.

Fig. 2 Determining the CCL in the cloud discrimination process. The horizontal axis represents the values for the threshold test. In this case, the observed value is large when clouds are present. Two threshold values, the lower limit and upper limit, are set, creating an ambiguous area between the two limits that is expressed as a CCL from 0 to 1

Fig. 3. Same as Fig. 1, except that the threshold testing uses two clear-sky ranges on both sides of a value

In contrast, the representative CCL for group 2, G_2, is determined to be clear conservative:

$$G_2 = \sqrt[n]{F_1 \cdot F_2 \cdots F_k \cdots F_n} \qquad (6)$$

which implies that if any of the threshold tests result in a CCL of 0, $G_2 = 0$ (cloudy) regardless of the other threshold tests of group 2, whereas $G_2 = 1$ (clear) only when all the F_k values are 1. Finally, the overall CCL, Q (which corresponds to Equation 2 of Ackerman et al. 1998), is determined from the representative values for the two groups as

$$Q = \sqrt{G_1 \cdot G_2} \qquad (7)$$

Q is the output product of the cloudy/clear discrimination algorithm; that is, $Q = 1$ (0) indicates a clear (cloudy) sky, and $0 < Q < 1$ indicates ambiguity. The cloudy/clear discrimination algorithms are designed to be used in parallel, not in cascade; that is, all applicable threshold tests are performed independently, without considering the results from the other tests, with the exception of the one restoral test with T1 mentioned above.

The SGLI cloud mask product contains additional information and has greater utility. Horizontal inhomogeneity is sometimes a useful indication of broken clouds, especially over water. The relative standard deviation of reflectance at 673.5 nm (land) or 868.5 nm (water) within 3×3 pixels around a central pixel is used as an index of the inhomogeneity.

Heavy aerosols and especially dust tend to be incorrectly assigned as cloudy, since aerosols are usually difficult to detect. Dust in the atmosphere tends to have a larger reflectance at shorter wavelengths (such as blue) but a smaller split window value than clear sky, and to be more horizontally homogeneous than clouds. Thus the authors have designed a simple test for suspected dust based on the optical characteristics of dust at solar and thermal infrared wavelengths. Cirrus probability is diagnosed from R_f (1380 nm), which is sensitive to high clouds as explained above. The phase of cloud particles (water, ice, or mixed) is roughly distinguished by the cloud-top temperature, which is inferred from T1 (10.8 μm) and the split window, with thresholds that are empirically determined based on the CloudSat CLDCLASS-LIDAR and MODIS-AUX products.

### Data flow in the algorithm

This section explains the data flow in the cloudy/clear discrimination algorithm. Figure 4 is a flow chart of the algorithm.

(a) Input information

Input information used for cloudy/clear discrimination is categorized as follows: (Ia) measurement data; (Ib) scan geometry information; (Ic) algorithm-specific information; (Id) geophysical information; and (Ie) ancillary information (Table 2).

The (Ia) data includes the calibrated radiance data for each pixel. The (Ib) information includes the solar zenith angle, satellite zenith angle, relative azimuth angle, latitude and longitude, and date and time of observation. Threshold values for threshold tests are categorized as (Ic) information. The (Id) information is required for calculating

Fig. 4 Flow chart for the cloud/clear discrimination algorithm
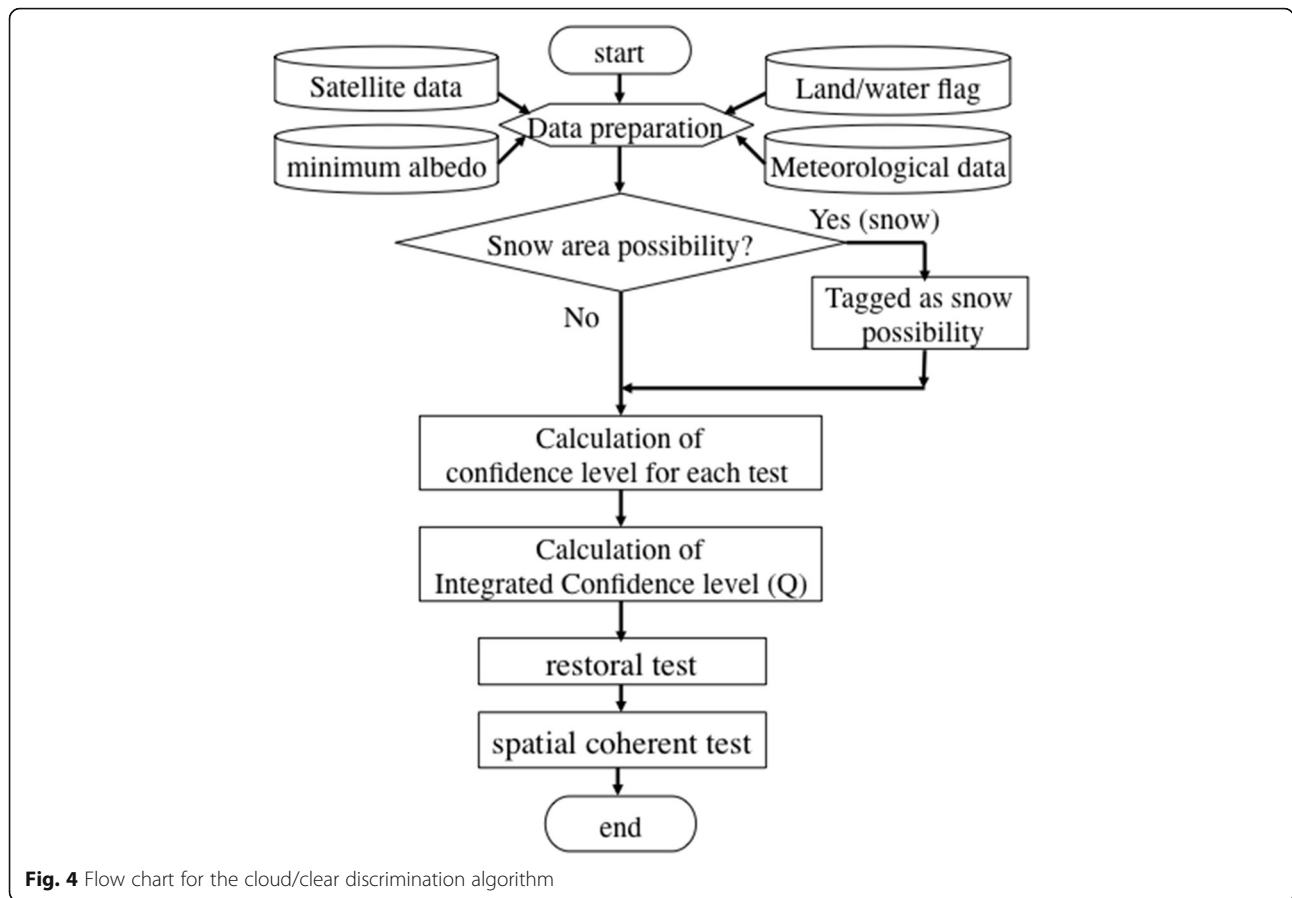
reflectance, such as the solar irradiance data, considering the SGLI instrument function. The (Ie) information consists of ancillary information, such as the database of surface minimum albedo for channels VN8 and SW1, and land/water flag data, which are derived from the Shuttle Radar Topography Mission water body data.

(b) Preprocessing

Before the algorithm performs cloudy/clear discrimination, it determines whether it is night or day, water or

Table 2 Input information for the cloud discrimination algorithm

Input information	Description
(Ia) Measurement data	Radiance
(Ib) Scan geometry information	Solar zenith angle, satellite zenith angle, relative azimuth angle, latitude and longitude, date and time of observation
(Ic) Algorithm-specific information	Threshold values, etc.
(Id) Geophysical information	Solar irradiance data, etc.
(Ie) Ancillary information	Surface minimum albedo in channels VN8 and SW1 Land/Water flag

land, or areas are sunglint from the input data. The sequence of applied threshold tests is arranged according to conditions. The determinations of night or day is based on the solar zenith angle: when the angle is 85° or more, it is assumed to be night. Note that this paper focuses on remote sensing in the daytime, although the GCOM-C product operated by JAXA also does cloudy/clear discrimination at night using thermal infrared bands only. Water, which includes lakes and rivers as well as the ocean, is distinguished from land by comparing the latitude and longitude of the pixel position against the 1 km land/water flag data. When the latitude is 66.6° or more north or south, the area is regarded as a polar region. If the cone angle between the solar incident and satellite direction is lower than 35° over the water surface, this area is identified as sunglint. In addition, the possibility of snow cover is evaluated based on the normalized difference snow index (NDSI):

$$\text{NDSI} = \frac{R_f(530\,\text{nm}) - R_f(1630\,\text{nm})}{R_f(530\,\text{nm}) + R_f(1630\,\text{nm})} \tag{8}$$

Over land with an altitude above 2 km and a latitude greater (smaller) than 30° (− 30°), if NDSI > 0.4, R_f (1630 nm) < 0.02, R_f (530 nm) > 0.15, R_f (868.5

nm) < 0.02, and brightness temperature Tb (10.8 μm) < 280 K, then the "Snow/Ice Probability Flag" in the bit field is set to "yes." For water areas with latitudes greater (or smaller) than 40° (– 40°), the threshold setting for this flag is the same as for land, except for R_f (1630 nm) < 0.01. Note that this flag is for quality assurance and does not affect cloudy/clear discrimination.

(c) Determining the CCL for each threshold test

Table 3 summarizes the threshold values for each threshold test applied in the cloudy/clear discrimination. The thresholds of reflectance R_f (673.5 nm) and R_f (1050 nm) for land depend on the surface minimum albedo (denoted by $R_{f\ \text{surface}}$ in Table 3(b)) of the pixel position. If a pixel over water is judged to be in a sunglint area, the threshold for the reflectance is increased according to the cone angle. The relationship between the cone angle and the increase in the threshold, denoted by $R_{f\ \text{sun-glint}}$ in Table 3(a), is

Table 3 Thresholds for cloud screening with SGLI. a: Over water, b: Over land, and c: Over polar regions. R_f and T_b are reflectance and brightness temperature, respectively

	Threshold test		Lower limit	Upper limit
a. Water regions				
1	R_f(868.5 nm)		0.195 + $R_{f\ \text{sun-glint}}$	0.045+ $R_{f\ \text{sun-glint}}$
1	NDVI	Smaller End	-0.10	-0.22
		Larger End	0.22	0.46
1	R_f (1050 nm)		0.195 + $R_{f\ \text{sun-glint}}$	0.045+ $R_{f\ \text{sun-glint}}$
2	Tb(10.8 μm)-Tb(12.0 μm)		3.0 [K]	2.6 [K]
2	R_f (1380 nm)		0.015	0.005
-	Tb(10.8 μm) [*1]		297.5 [K]	-
b. Land regions				
1	R_f (673.5 nm)		0.195+ $R_{f\ \text{surface}}$	0.045+ $R_{f\ \text{surface}}$
1	NDVI	Smaller End	-0.1	-0.22
		Larger End	0.22	0.46
1	R_f (868.5 nm)/ R_f (1630 nm)		1.05	1.00
1	R_f (1050 nm)		0.195+ $R_{f\ \text{surface}}$	0.045+ $R_{f\ \text{surface}}$
2	Tb(10.8 μm)-Tb(12.0 μm)		3.0 [K]	2.6 [K]
2	R_f (1380 nm)		0.040	0.030
-	Tb(10.8 μm) [*1]		297.5 [K]	-
c. Polar regions				
1	R_f (673.5 nm)		0.14+ $R_{f\ \text{surface}}$	0.06+ $R_{f\ \text{surface}}$
1	NDVI	Smaller End	-0.1	-0.22
		Larger End	0.22	0.46
2	R_f (1380 nm)		0.060	0.030
-	Tb(10.8 μm) [*1]		297.5 [K]	-

[*1]Restoral test

given in Table 4. For angles between those in the table, the increase is determined by linear interpolation. Finally, the overall CCL (Q) is determined by integrating the CCLs of the threshold tests.

(d) Outputting the cloud/clear discrimination results

Output from the cloudy/clear discrimination algorithm is expressed in 16 bits, which include the overall CCL and other miscellaneous data. Table 5 shows the bit fields of an output file. The overall CCL is expressed in a bit field with three bits as shown in Table 6, and the cone angle is expressed by two bits as shown in Table 7. After deriving the overall CCL, additional information is derived, which (including the possibility of snow cover) does not influence the result of cloudy/clear discrimination. The horizontal inhomogeneity flag is set to "yes" if the relative standard deviation is greater than 0.25 for land or 0.1 for water. Referring to NOAA/NESDIS/STAR (2010) and Mouri et al. (2016), the heavy aerosol flag over land is set to "yes" if the horizontal inhomogeneity flag is "yes," the snow/ice probability flag is "no," and

$$R_f\ (1380\ \text{nm}) < 0.035 \tag{9a}$$

$$R_f\ (412\ \text{nm})/R_f\ (443\ \text{nm}) < 1.16 \tag{9b}$$

$$R_f\ (412\ \text{nm}) > 0.2 \tag{9c}$$

$$R_f\ (1630\ \text{nm}) < 0.3 \tag{9d}$$

$$R_f\ (868.5\ \text{nm}) < 0.3 \tag{9e}$$

$$Tb\ (10.8\ \mu\text{m})-Tb(12.0\ \mu\text{m}) < 0\ \text{K} \tag{9f}$$

Equation (9a) excludes cirrus clouds. The threshold tests for water are the same as for land, but the threshold settings are:

$$R_f\ (1380\ \text{nm}) < 0.01 \tag{10a}$$

$$R_f\ (412\ \text{nm})/R_f\ (443\ \text{nm}) < 1.25 \tag{10b}$$

$$0.15 < R_f\ (412\ \text{nm}) < 0.25 \tag{10c}$$

$$R_f\ (1630\ \text{nm}) < 0.25 \tag{10d}$$

$$R_f\ (868.5\ \text{nm}) < 0.3 \tag{10e}$$

Table 4 Increase in the threshold values for reflectance tests on sunglint areas

Cone angle (°)	Increase in the threshold
35	0
25	0.013
15	0.075

Table 5 Bit fields of the output files for the cloud discrimination algorithm

Bit field	Description	Bit
0	Cloud mask algorithm flag	0 Not determined, 1 Determined
3, 2, 1	Clear confidence level flag	Table 6
4	Day/Night flag	0 Night, 1 day
5	Land/Water flag	0 Water, 1 Land
6	Snow/Ice probability flag	0 Yes, 1 No
8, 7	Sun-glint cone angle flag	Table 7
9	Heavy aerosol flag	0 Yes, 1 No
10	Cirrus probability flag	0 Yes, 1 No
11	Horizontal inhomogeneity flag	0 Yes, 1 No
13, 12	Cloud phase flag	00 Uncertain, 01 Liquid 10 Ice, 11 Mixed
14	Cloud shadow flag (TBD)	0 Yes, 1 No
15	VN data availability flag	0 Not available, 1 Available

$$Tb\,(10.8\,\mu m) - Tb(12.0\,\mu m) < 0.5\,K \tag{10f}$$

$$Tb\,(10.8\,\mu m) > 268.15\,K \tag{10g}$$

The cirrus probability flag is set to "yes" if R_f (1380 nm) > 0.035. The cloud phase flag is set according to cloud thermodynamic phase discrimination as mentioned in the next subsection. In addition, the bit field includes the VN data availability flag, which is set to "not available" if the VN bands (380–868.5 nm) are out of swath, because the swath of VN (1150 km) is narrower than those of SWIR and TR (1400 km).

Algorithm for the cloud thermodynamic phase

The SGLI cloud flag product includes cloud thermodynamic phase information for day and night derived from SGLI thermal infrared channels (Table 5). One of the best choices for cloud-phase discrimination is a combination of

Table 6 Expressions of overall CCLs in bits

Overall clear confidence level	Bit
0.00	000
0.00–0.17	001
0.17–0.33	010
0.33–0.50	011
0.50–0.67	100
0.67–0.83	101
0.83–1.00	110
1.00	111

Table 7 Expressions of cone angles in bits

Cone angle (°)	Bit
0–15	00
15–25	01
25–35	10
35–45	11

the 8.5- and 10.8-μm channels used in the MODIS L2 cloud product because of the larger difference between imaginary indices of refraction for water and ice over 8 to 12 μm (Baum et al. 2000). However, the SGLI does not have an 8.5-μm channel, so a combination of the 10.8- (T1) and 12-μm (T2) channels is used. Figure 5 shows the most frequent cloud phases (represented in color) as a function of the brightness temperature (Tb) of the 10.8-μm channel and the difference between the brightness temperatures of the 10.8- and 12-μm channels extracted by a combined use of the CloudSat CLDCLASS-LIDAR and MODIS-AUX products. They and their documents are distributed by the CloudSat Data Processing Center (http://www.cloudsat.cira.colostate.edu). Figure 5 has the following cloud-phase classes that we extract from the CloudSat CLDCLASS-LIDAR product: single-layer liquid-phase cloud (LIQ); single- or multi-layer ice- or mixed-phase cloud (ICE); and multi-layer cloud (MUL), in which an ice-phase cloud lies above a liquid-phase cloud. Both ICE and MUL have ice-phase clouds at the cloud top. For simplicity and robustness, the SGLI cloud phase algorithm discriminates cloud phases using the two boundaries shown by dotted lines in Fig. 5: The cloud-phase flag is set to "ICE" if Tb (10.8 μm) – Tb (12.0 μm) > 0.08 Tb (10.8 μm) – 21 and Tb (10.8 μm) < 265 K, whereas this flag is set to "LIQ" if Tb (10.8 μm) – Tb (12.0 μm) < 0.08 Tb (10.8 μm) – 21, and the other flags are labeled as "MUL." Although the response functions of the SGLI T1 and T2 channels are slightly different from those of MODIS bands 31 and 32, we apply the thresholds with no corrections. Figure 6 shows images of (a) the color composite of the SW4, SW3, and SW1 channels, and (b) the cloud-phase flag. The images were created from the GCOM-C/SGLI observations over the Indian Ocean on March 23, 2018, which captured Tropical Cyclone Marcus. The labels in Fig. 6 match the labels in Fig. 5. Clear pixels identified by the CLAUDIA are classified automatically into the class "Uncertain (UNC)." And, cloudy pixels are classified into the classes "LIQ," "ICE," and "MUL" as shown in Fig. 5.

CAPCOM: algorithm for retrieving cloud properties

CAPCOM is an algorithm for retrieving cloud properties from the visible to infrared imager radiances. This algorithm has been used by the Global Imager (GLI) on ADEOS-II, CAI on GOSAT, AHI on Himawari-8, and
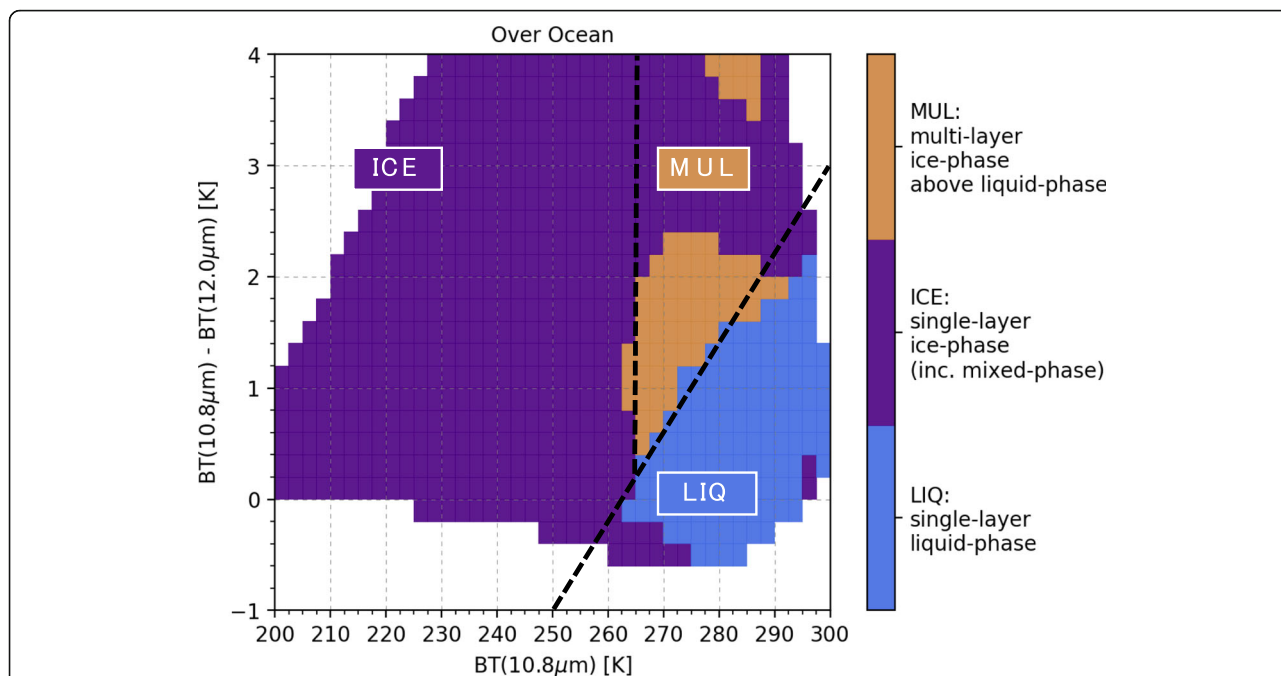
Fig. 5 The most frequent cloud phases as a function of the brightness temperature (Tb) of the 10.8-μm channel and the difference between the brightness temperatures of the 10.8- and 12-μm channels extracted by a combined use of CloudSat CLDCLASS-LIDAR and MODIS-AUX products. The color represents most frequent cloud phases: the cloud-phase flag is set to "ICE" if Tb (10.8 μm) − Tb (12.0 μm) > 0.08 Tb (10.8 μm) − 21 and Tb (10.8 μm) < 265 K, whereas this flag is set to "LIQ" if Tb (10.8 μm) − Tb (12.0 μm) < 0.08 Tb (10.8 μm) − 21, and the other flags are labeled as "MUL"

SGLI on GCOM-C, and will be used for MSI on the Earth-CARE satellite, which will be launched in 2021. CAPCOM was developed by Nakajima and Nakajima (1995) based on the principle shown in the Nakajima-King diagram (Fig. 7). It was applied to two AVHRR scene images for retrieving warm-water cloud properties. Kawamoto et al. (2001) applied a water vapor correction function to CAP-COM to allow the algorithm to retrieve cloud properties globally. Scattering models of the spherical Mie particles and the non-spherical Voronoi particles (Ishimoto et al. 2010) are used to develop the look-up table (LUT) for CAPCOM to retrieve water-cloud and ice-cloud properties (letu et al. 2012, 2016, 2018), respectively. This subsection describes the operating principles of CAPCOM.

CAPCOM uses the LUT iteration method to retrieve the target geophysical parameters from satellite-derived radiance data. CAPCOM uses a non-absorption channel (SW1), an absorption channel (SW4 primary, SW3 secondary), and a thermal channel (T1) to derive cloud optical thickness (τ_c), cloud effective particle radius (r_e), and cloud-top temperature (T_c). Newton-Raphson methods were used to obtain τ_c, r_e, and T_c, simultaneously. In addition to these radiance data, some ancillary input data, such as the vertical profile of the temperature, pressure, water vapor, and ground

albedo, are also used to calculate related geophysical parameters, and cloud top height (Z_c) and cloud-top pressure (P_c) are retrieved by comparing cloud-top temperature (T_c) and a temperature profile from ancillary data (objective analysis data).

Principles of the algorithm

CAPCOM uses the following LUTs: cloud-reflected radiance in channels SW1 and SW4; transmissivities and reflectivity in channels SW1 and SW4; and transmissivity in channel T1. Figure 8 illustrates the flow of the analysis. The Newton-Raphson method is used to iterate the main loop in the program. For water-cloud analysis, the cloud geometrical thickness Dc was obtained using the liquid water path (LWP) and liquid water content (LWC). For the ice-cloud analysis, Dc = 1 km was assumed since LWP = 2/3 $\tau_c r_e$ is not correct for ice clouds. For water-cloud analysis, stratocumulus clouds are assumed for liquid water content with a value of 1.28×10^{-6} g/cm^{-3}. The Dc value was used for water-cloud analysis since the LUT grid system uses Dc. Some related parameters, such as cloud liquid water path and cloud top height (Z_c), are also calculated in CAPCOM with temperature-slicing data ($T(z)$) from the objective analysis dataset.
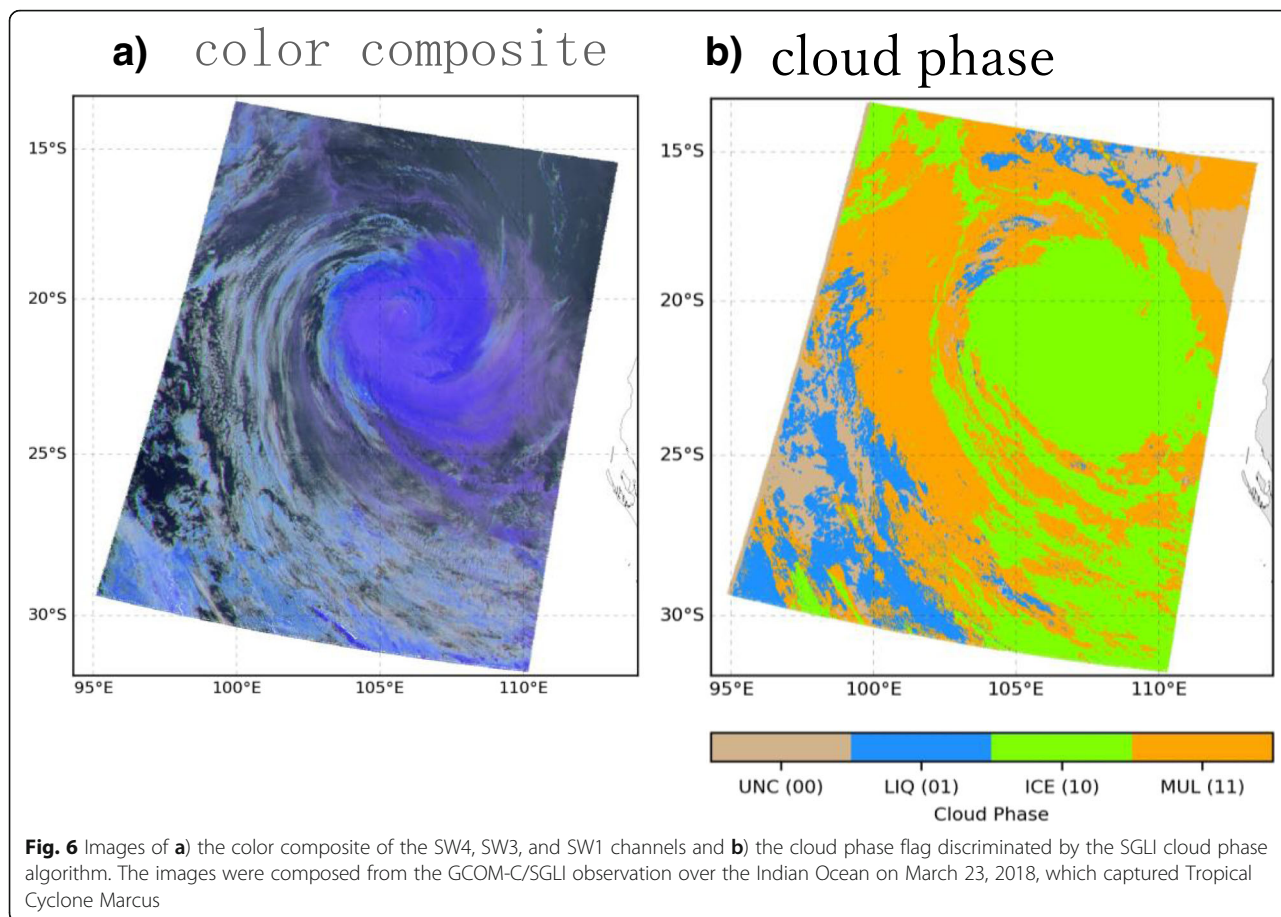
Fig. 6 Images of **a**) the color composite of the SW4, SW3, and SW1 channels and **b**) the cloud phase flag discriminated by the SGLI cloud phase algorithm. The images were composed from the GCOM-C/SGLI observation over the Indian Ocean on March 23, 2018, which captured Tropical Cyclone Marcus

Physical and mathematical aspects of CAPCOM

The solar reflectance method uses a non-absorbing channel (such as 673.5, 868.5, or 1050 nm) and water-absorbing shortwave infrared wavelengths (such as 1630 and 2210 nm) for the simultaneous retrieval of the cloud optical thickness, τ_c, at a wavelength of 500 nm and the effective particle radius, r_e. This paper mainly discusses the solar reflectance method using the SGLI channels SW1 (1050 nm), SW4 (2210 nm), and T1 (10.8 μm). Despite the inclusion of 10.8 μm in this analysis, this will be referred to as the "solar reflectance method" since the primary two bands are 1050 nm and 2210 nm.

The effective particle radius of the water clouds is defined by:

$$r_e \equiv \frac{\int_0^\infty r^3 n(r) dr}{\int_0^\infty r^2 n(r) dr} \tag{11}$$

where $n(r)$ is the number size distribution as a function of particle radius r. The effective particle radius of the ice clouds is defined by:
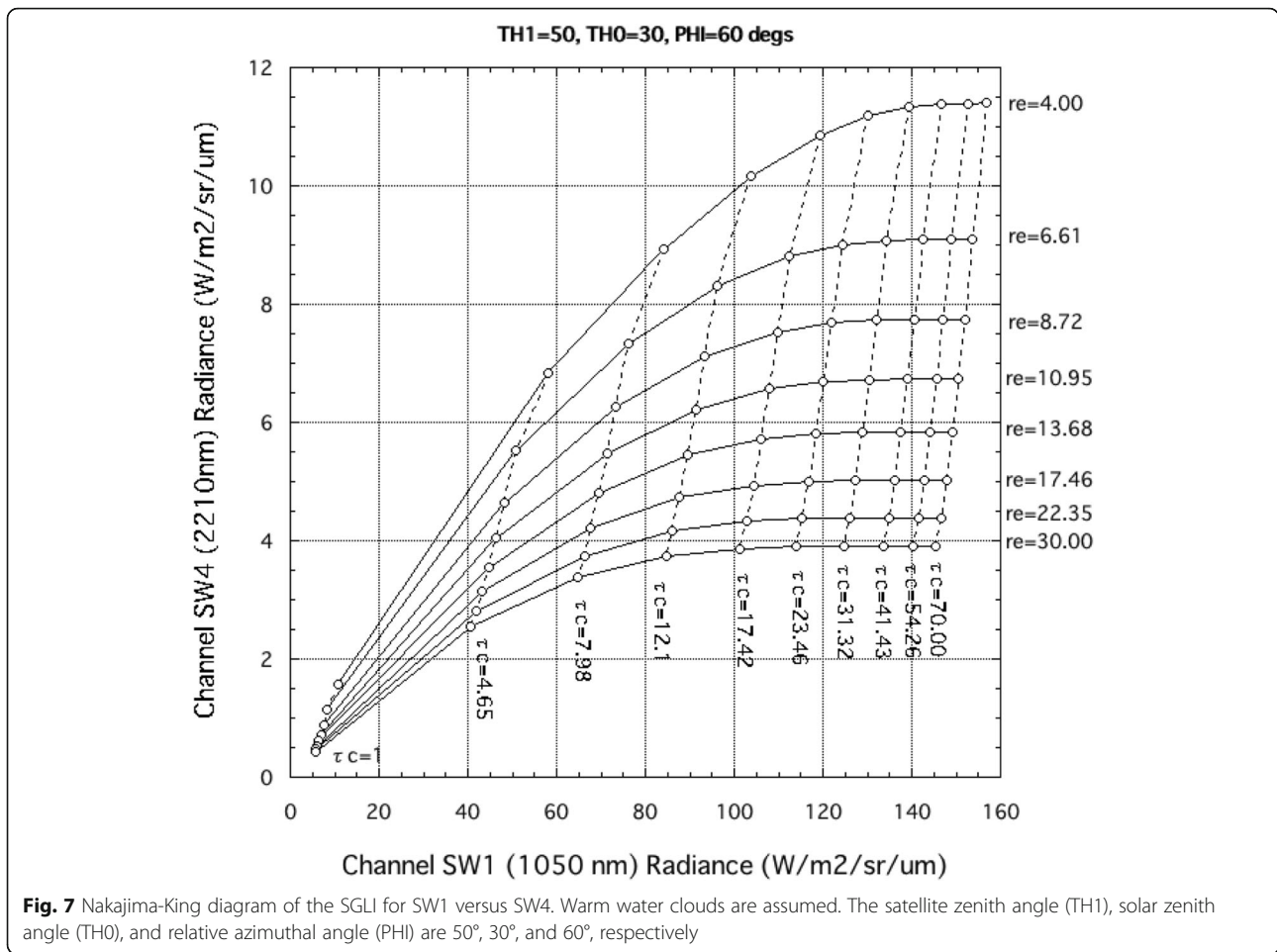
$$r_e \equiv \frac{3 \int_0^\infty V(L) n(L) dL}{4 \int_0^\infty A(L) n(L) dL} \tag{12}$$

where L is the maximum dimension of the Voronoi particle. $V(L)$ and $A(L)$ are the volume and projected area of the Voronoi single particle, respectively. For the satellite signal simulation, we use an accurate, efficient radiative transfer scheme (Nakajima and Tanaka 1986, 1988) extended to include the thermal radiative transfer (Stamnes et al. 1988).

We retrieve τ_c at a wavelength of 500 nm and r_e from SGLI channels SW1 and SW4 because these channels primarily depend on the cloud optical thickness and the effective particle radius, respectively. Although the concept of retrieval is simple, there are difficulties in determining the cloud properties from the measured SGLI spectral radiance. The solar radiation reflected by the ground must be removed from the observed radiance.

Formulations of the radiative components

According to the radiative transfer theory for plane paral=l layers with an underlying Lambert surface,

Fig. 7 Nakajima-King diagram of the SGLI for SW1 versus SW4. Warm water clouds are assumed. The satellite zenith angle (TH1), solar zenith angle (TH0), and relative azimuthal angle (PHI) are 50°, 30°, and 60°, respectively

any unexpected radiation component, such as the solar radiation reflected by the ground surface, is removed from the satellite-received radiance, I_{obs}, to decouple the radiation component reflected by the cloud layer, I. The unexpected radiation component is expressed in the second term of Equation (Letu et al., 2016):

$$I(Z_c, D_c, \tau_c, r_e; \mu, \mu_0, \phi) = I_{obs}(Z_c, D_c, \tau_c, r_e; \mu, \mu_0, \phi)$$
$$-t(Z_c, D_c, \tau_c, r_e; \mu)\frac{A_g}{1-\bar{r}(Z_c, D_c, \tau_c, r_e)A_g}t(Z_c, D_c, \tau_c, r_e; \mu_0)\frac{\mu_0 F_0}{\pi}$$
(13)

where F_0 is the extraterrestrial solar flux, μ_0 and μ are the cosines of the solar and satellite zenith angles, and ϕ is the azimuthal angle of the satellite relative to the sun. Z_c and D_c are the top height and the geometrical thickness of the cloud, respectively. A_g is the ground albedo, and t and $\bar{r}$ are the transmissivity and spherical albedo of the cloud. These formulations are exact for monochromatic radiance. We introduce a further process of averaging of the variables in the

formulations with respect to the wavelength. For example, t is averaged with a subchannel response function of SGLI as:
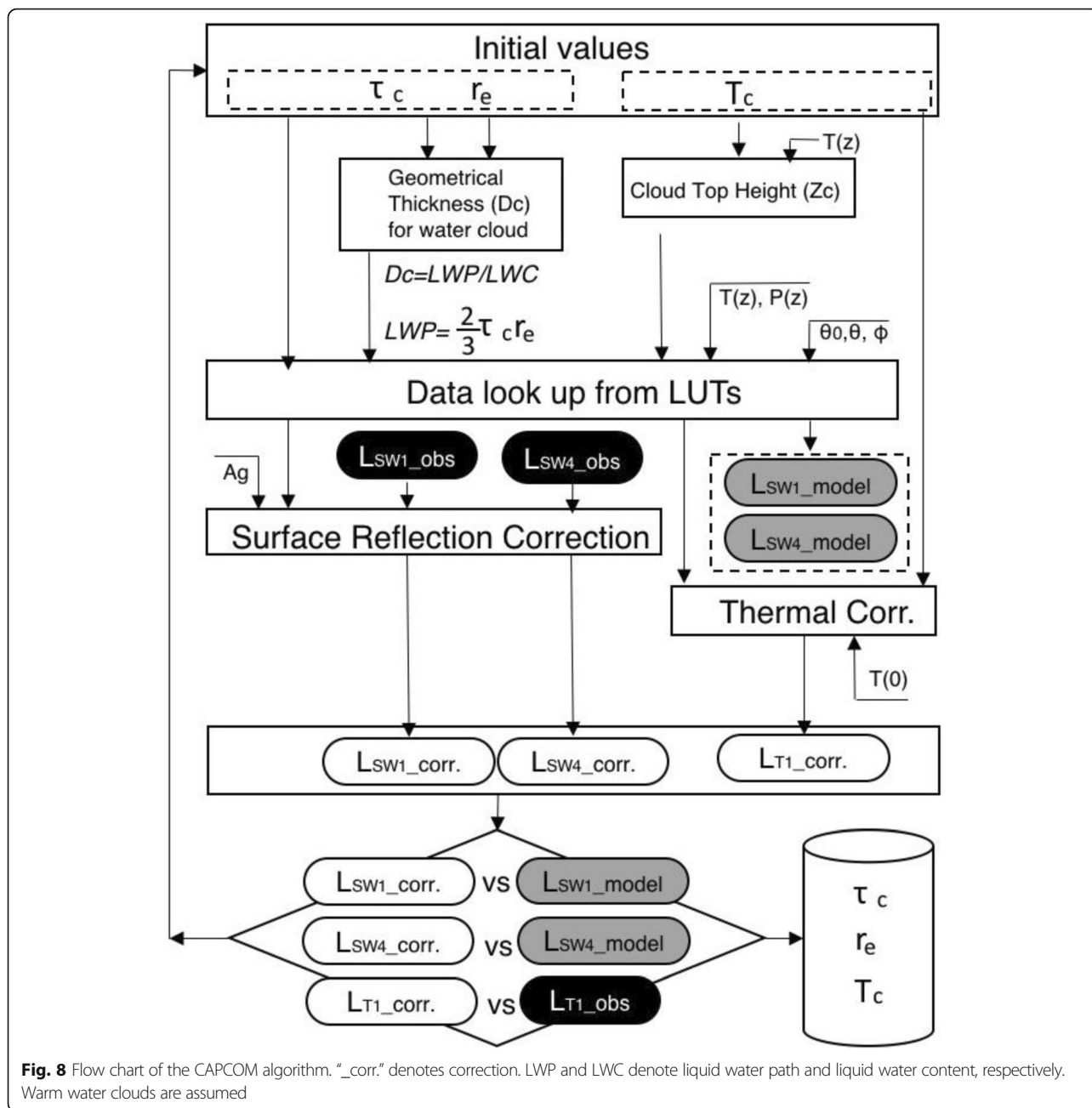
$$t = \sum_{n=1}^{N}\phi_n\left\{\sum_{k=1}^{M}\xi_{n,k}\times t_{n,k}\right\}/\sum_{n=1}^{N}\phi_n$$
(14)

where ϕ_n is the response function of the nth subchannel wavelength for each SGLI channel, $\xi_{n,k}$ is the weight of the kth k-distribution, and $t_{n,k}$ is the transmissivity for the kth k-distribution at the nth wavelength.

Results and Discussion
Comparison with ground-based whole-sky camera images
Accuracy evaluation strategy
The output of the cloudy/clear discrimination algorithm (CLAUDIA) was compared with whole-sky images (WSIs) acquired with sky cameras located at several ground stations in Japan (Table 8). One is a sky camera system consisting of a single-lens reflex digital camera (D7000, NIKON). This is the same system used by Letu et al. (2014) and Kuji et al. (2018), and details can be found in their paper. The other system is a Skyview

Fig. 8 Flow chart of the CAPCOM algorithm. "_corr." denotes correction. LWP and LWC denote liquid water path and liquid water content, respectively. Warm water clouds are assumed

(PSV-100, PREDE), which was also used by Yabuki et al. (2014). Both systems capture color WSIs in JPEG format. Cloud fractions were first determined from the WSIs based on the image analysis method described in the next subsection, and were then compared with the cloud fractions determined by CLAUDIA to evaluate the accuracies of SGLI cloud flag products based on the validation guidelines for the GCOM-C mission (Hori et al. 2018). SGLI cloud flag products were also compared to evaluate the performance of CLAUDIA with products derived from other satellite sensors, such as MODIS.

Analysis of WSIs

WSIs acquired with the two different sky camera systems are similar true-color RGB images in JPEG format with different pixel resolutions (640 × 480 for the Skyview and 4928 × 3264 for the NIKON system). Although the image sizes are different, the pixel resolution of the Skyview image is large enough to extract the cloud fraction (percentage) within the fisheye image circle. The approach described below was used to discriminate cloudy and clear pixels based on an image analysis technique similar to that used by

Table 8 Sites and types of sky cameras used for evaluating the CLAUDIA outputs and the amount of training data used for calculating cloud probabilities in SI-BI space for each site

Location (Lat/Lon)	Sky-camera type	The number of training data		
		Clear scenes	Cloudy scenes	Total
Ny-Alesund (78.93 N/11.86 E)	Skyview	32	77	109
Sapporo (43.08 N/141.34 E)	D7000	72	70	142
Tsukuba (TKSC) (36.07 N/140.13 E)	D7000	96	113	209
Tsukuba (MRI) (36.06 N/140.13 E)	Skyview	128	131	259
Kumamoto (32.84 N/130.87 E)	D7000	31	50	81
Miyakojima (24.74 N/125.33 E)	Skyview	33	31	64
Syowa St. (69 S/39.59 E)	Skyview	65	80	145

Fig. 9 Method of analysis for the logistic regression approach. **a** Two-dimensional histograms for clear training data and **b** cloudy training data. **c** An example logistic regression curve derived for the clear and cloudy cases within the bin class of BI shown in red dotted boxes in **a**, **b**, and **d** calculated clear probability in the BI-SI space
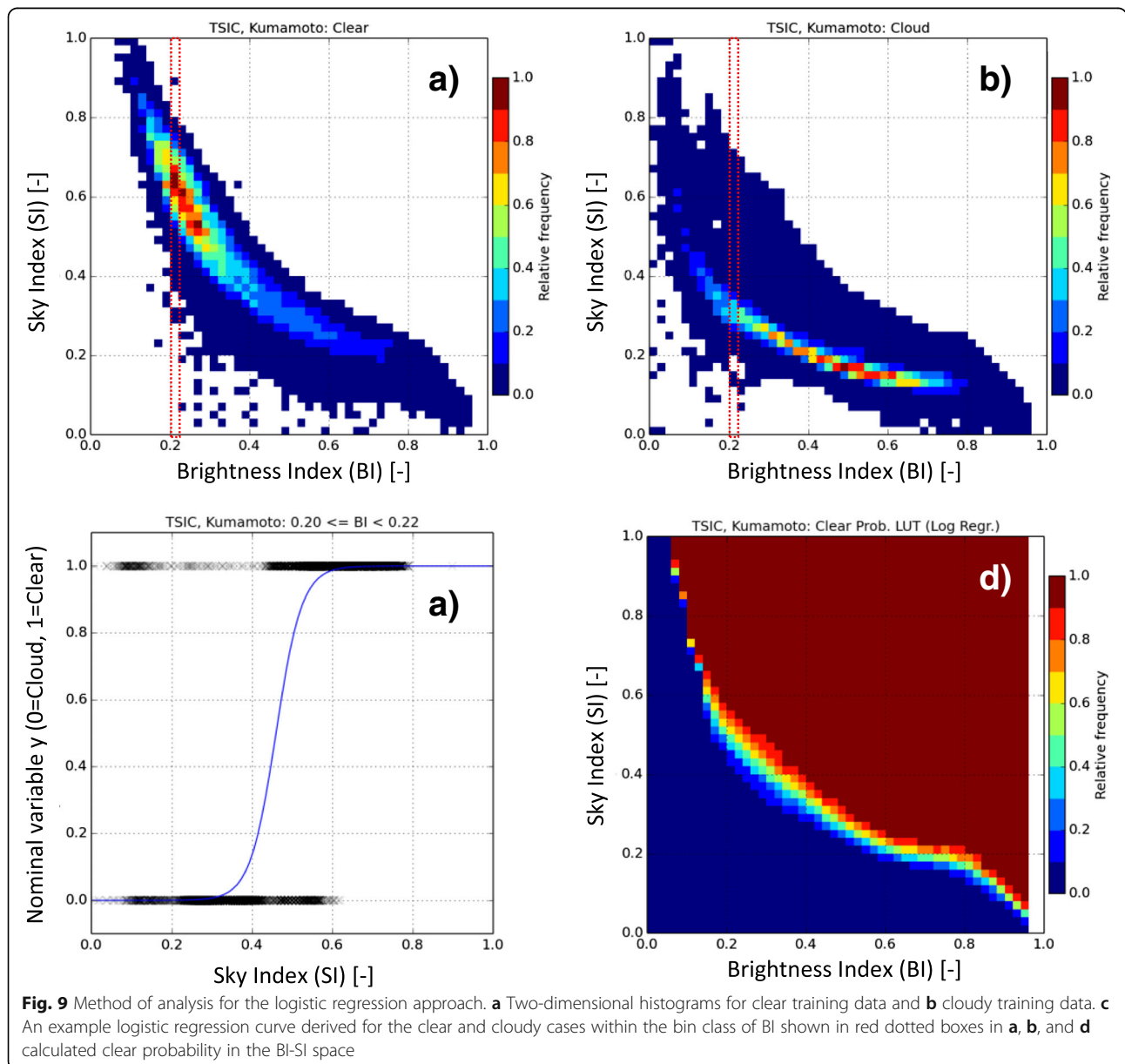
Yoshimura and Yamashita ([2013](#)). They developed a threshold method using the sky index (SI) and the brightness index (BI) as calculated using the following equations:

$$SI = (Blue–Red)/(Blue + Red) \qquad (15)$$

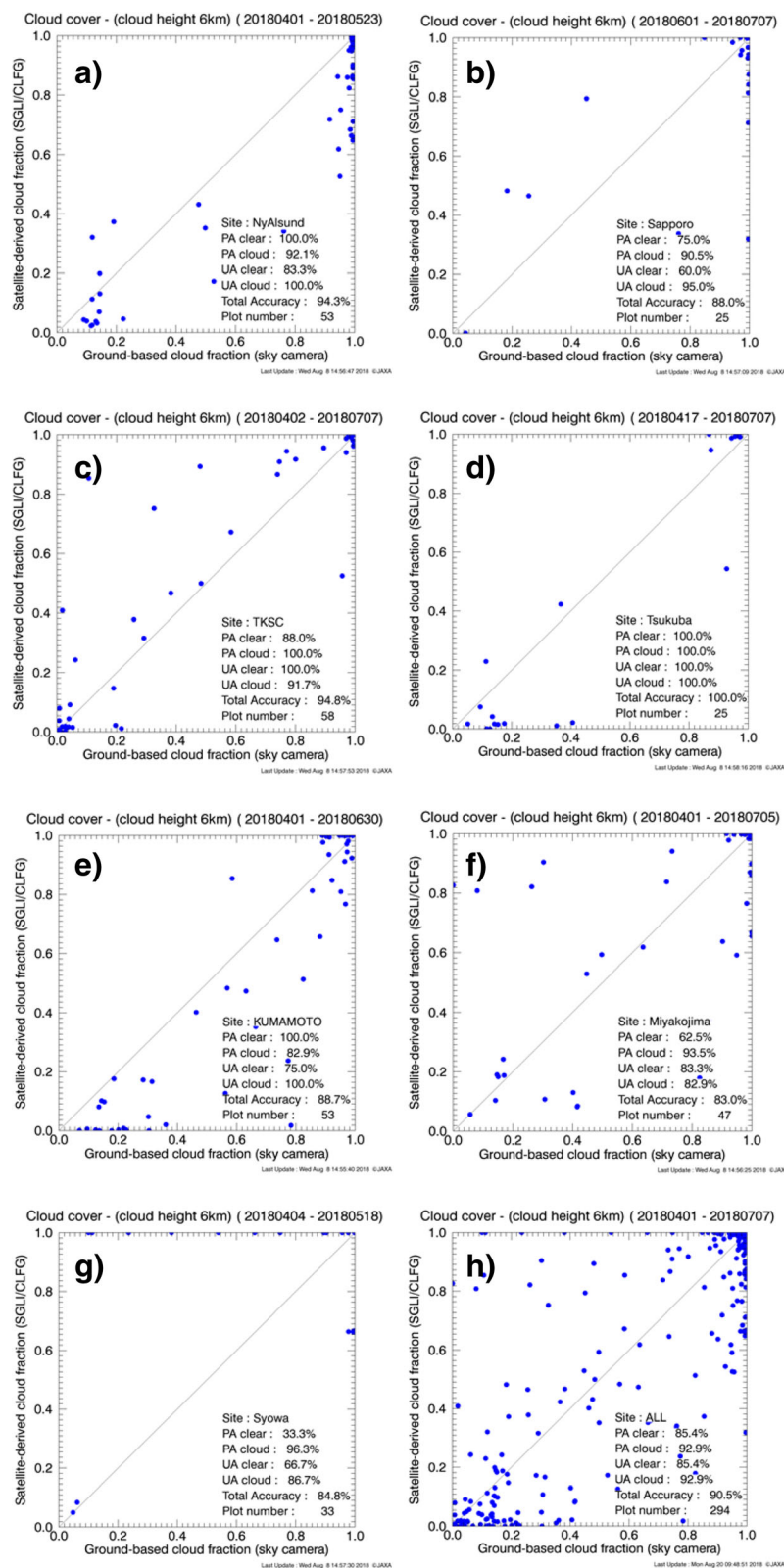$$BI = (Red + Green + Blue)/(255 \times 3) \qquad (16)$$

Here, Red, Green, and Blue are the digital numbers of the individual colors stored in the JPEG file. Using these two indices, Yoshimura and Yamashita ([2013](#)) identified cloudy and clear pixels in WSIs with an exponential curve to define the threshold in the SI-BI space. A similar approach is used here, although instead of using exponential curve estimates of the cloudy and clear probability, they were made in the same two-dimensional SI-BI space based on the logistic regression approach (Fig. [9](#)). Cloudy/clear probabilities in the SI-BI space were calculated for each bin, which is defined along the BI axis, using training data carefully selected for clear and uniformly cloudy cases by visual inspection (Table [8](#)), and the cloud/clear probabilities were stored in LUTs for all the sky camera sites. For example, SI data determined for clear and cloudy cases within a bin class of BI ranging from 0.20 to 0.22 (boxes drawn in red in Fig. [9](#)a, b) were extracted and used as nominal variables (y), that is, $y = 0$ for the cloudy case and $y = 1$ for clear cases. Then, the nominal variables are curve-fitted with a sigmoid function as a function of SI to generate a smoothed clear probability curve as shown in Fig. [9](#)c. Figure [9](#)d shows the clear probabilities generated for all the bins of BI from 0.0 to 1.0. Figure [10](#) shows examples of the cloud/clear discrimination results for the Kumamoto site based on the logistic regression approach, which

identified clear and cloudy pixels corresponding to clear or cloudy skies, and also for mixtures of clear and cloudy conditions. Once the fractions of cloudy pixels within the field of view (image circle located 60° from the zenith) of WSIs were obtained, they were compared to the cloud fractions derived from SGLI cloud flag products (CLAUDIA outputs) at the locations of the sky camera sites. In this comparison, we defined the common fields of view (radii) of both SGLI and sky cameras to be 1.75, 10.4, and 17.3 km by assuming that the heights of cloud layers are at 1, 6, and 10 km, respectively which correspond to the low, middle, and high clouds under the International Satellite Cloud Climatology Project cloud classification (Rossow and Schiffer [1999](#)).

Results

Figure [11](#) compares cloud fractions derived from ground-based sky cameras with those from SGLI observations for all sites for a cloud height of 6 km. Similar results were obtained for the other cloud heights and thus are not shown here. Most of the data points are concentrated around the lower left (clear sky) and upper right (cloudy sky) corners, although some are scattered around the lower right and upper left corners. The scattered points around the lower right corner are due to the spatial heterogeneity of clouds, including altitude biases, which causes biases in the field of view of the sky camera and SGLI. In contrast, the points scattered around the upper left corner are due to the mistaken detection of thin cirrus clouds within WSIs in the SI-BI space-based analysis. In addition, for Syowa Station (Fig. [11](#)g), the current version of CLAUDIA fails to discriminate between clear and cloudy conditions correctly when the solar zenith is

Fig. 10 Examples of raw sky camera images (upper images) and analyzed sky camera images (lower images) for **a** completely clear sky, **b** completely cloudy sky, and **c** a mixed clear/cloudy sky
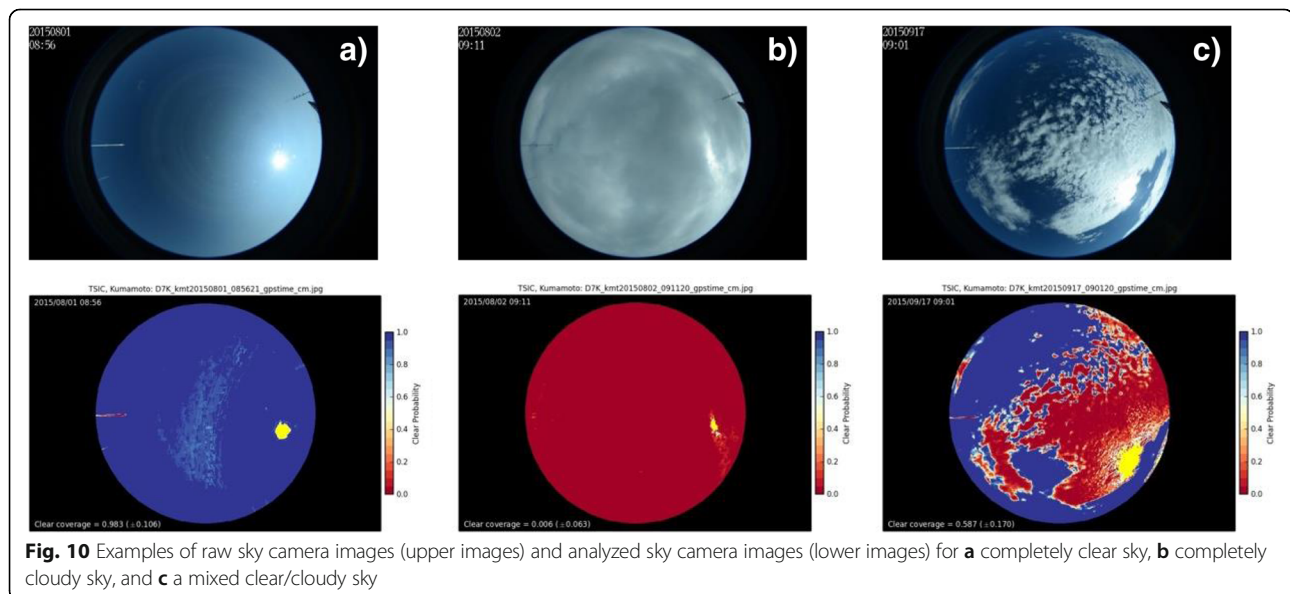
Fig. 11 Comparison of cloud fractions based on ground-based sky camera observations with those derived from SGLI data using the CLAUDIA algorithm. Observations taken from **a**) Ny-Alesund, **b**) Sapporo, **c**) Tsukuba (TKSC), **d**) Tsukuba (MRI), **e**) Kumamoto, **f**) Miyakojima, **g**) Syowa Station, and **h**) all sites

Table 9 Summary of cloud detection accuracies

Accuracies[a]	Ny-Alesund	Sapporo	Tsukuba (TKSC)	Tsukuba (MRI)	Kumamoto	Miyakojima	Syowa St.	All
N	53	25	58	25	53	47	33	294
UA$_{cloud}$	100.0	95.0	91.7	100.0	100.0	82.9	86.7	92.9
PA$_{cloud}$	92.1	90.5	100.0	100.0	82.9	93.5	96.3	92.2
OA	94.3	88.0	94.8	100.0	88.7	83.0	84.8	90.5

[a]For a cloud height of 6 km
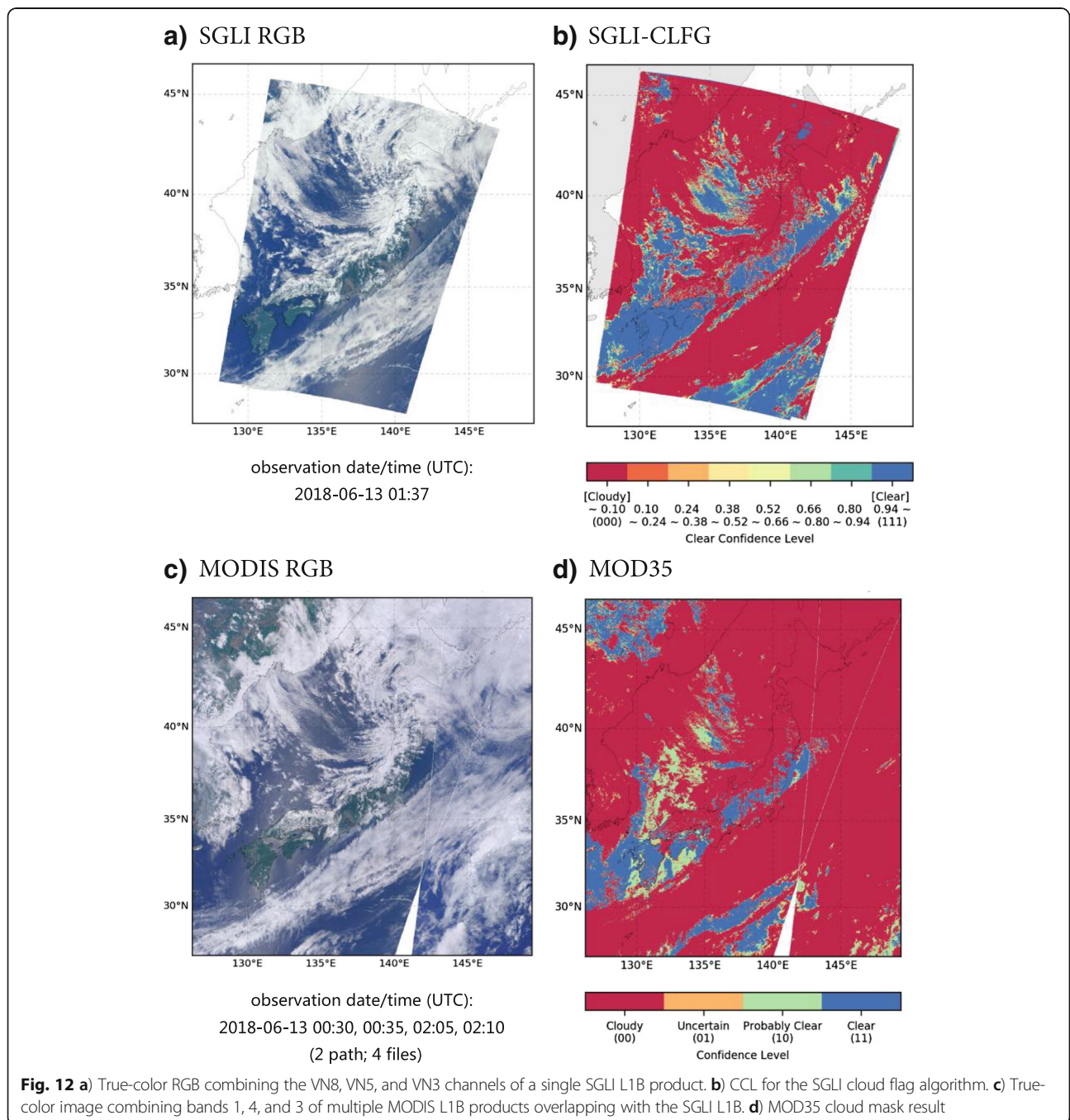
Fig. 12 a) True-color RGB combining the VN8, VN5, and VN3 channels of a single SGLI L1B product. **b**) CCL for the SGLI cloud flag algorithm. **c**) True-color image combining bands 1, 4, and 3 of multiple MODIS L1B products overlapping with the SGLI L1B. **d**) MOD35 cloud mask result

large (80° to 90°). Except for these scattered points, the satellite-derived cloud fractions are consistent with those derived from the sky camera. The user's accuracies (UA) and producer's accuracies (PA) were calculated based on a binary classification in which the data plots with a cloud fraction higher (less) than 50% are assumed to be cloudy (clear) for both the sky camera and the satellite. The calculated UA and PA for cloud detection and the overall accuracies (OA) are summarized in Table 9. UA and PA are mostly high around 90% and the resulting OA is also

high, ranging from 83 to 100% (the average of all sites is 90.5%). Therefore, the SGLI cloud detection algorithm (CLAUDIA) has sufficient accuracy for cloud masking, which meets the threshold for releasing data from the SGLI cloud flag product, that is, the minimum success criterion of the GCOM-C mission (Hori et al. 2018). To understand the causes of the scattered points that reduce the UAs and PAs, the accuracies of both sky camera and SGLI-derived cloud fractions need to be refined further considering, for example, the frequencies of cloud appearance

a) SGLI RGB

observation date/time (UTC): 2018-03-29 01:47

b) SGLI-CLFG

[Cloudy]   [Clear]
~ 0.10   0.10   0.24   0.38   0.52   0.66   0.80   0.94 ~
(000)   ~ 0.24   ~ 0.38   ~ 0.52   ~ 0.66   ~ 0.80   ~ 0.94   (111)
Clear Confidence Level

c) MODIS RGB

observation date/time (UTC): 2018-03-29 01:40, 01:45, 03:20, 03:25 (2 path; 4 files)

d) MOD35

Cloudy (00)   Uncertain (01)   Probably Clear (10)   Clear (11)
Confidence Level
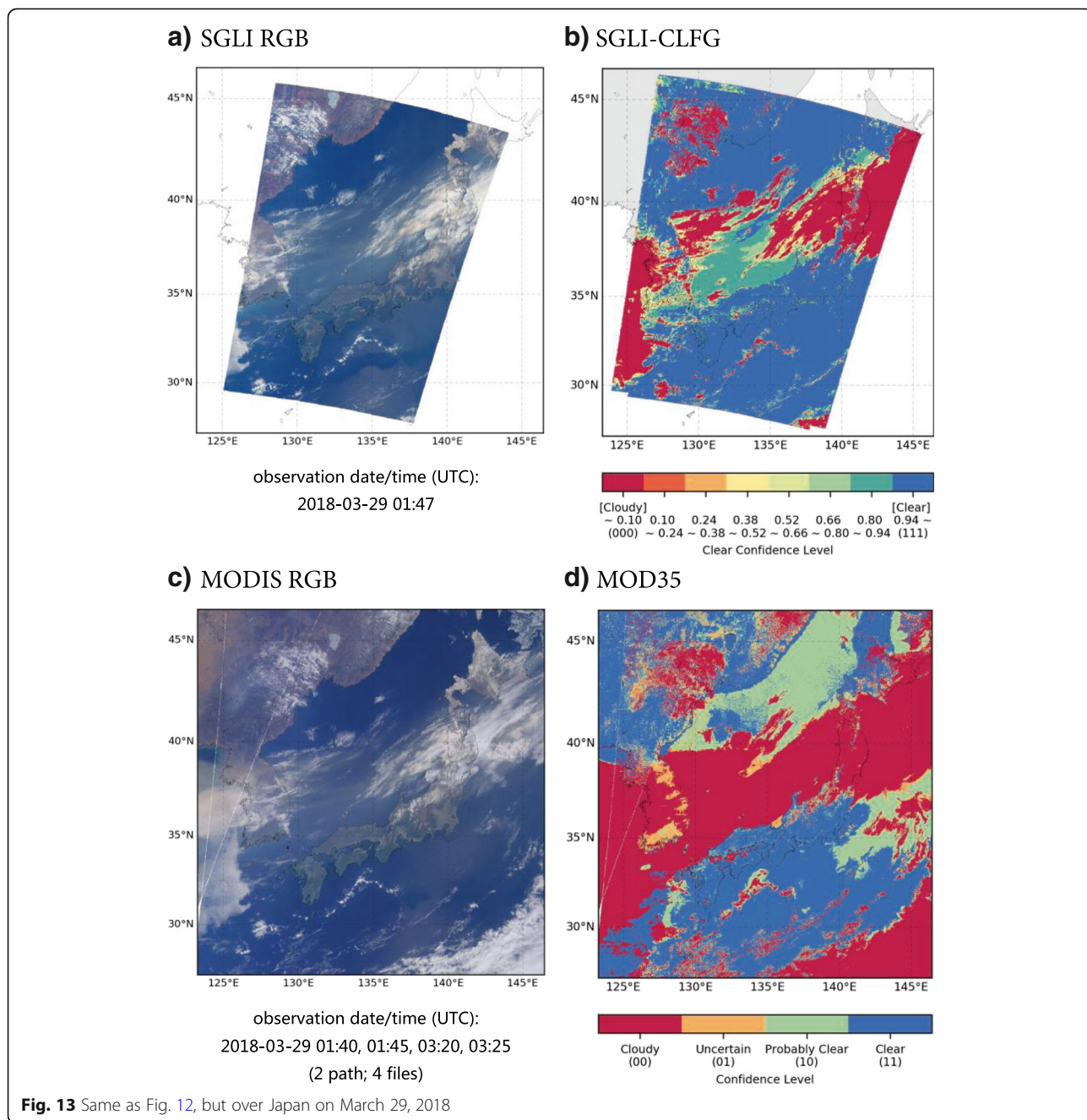
Fig. 13 Same as Fig. 12, but over Japan on March 29, 2018

heights at individual sites measured by micropulse lidar (Shiobara et al. 2003).
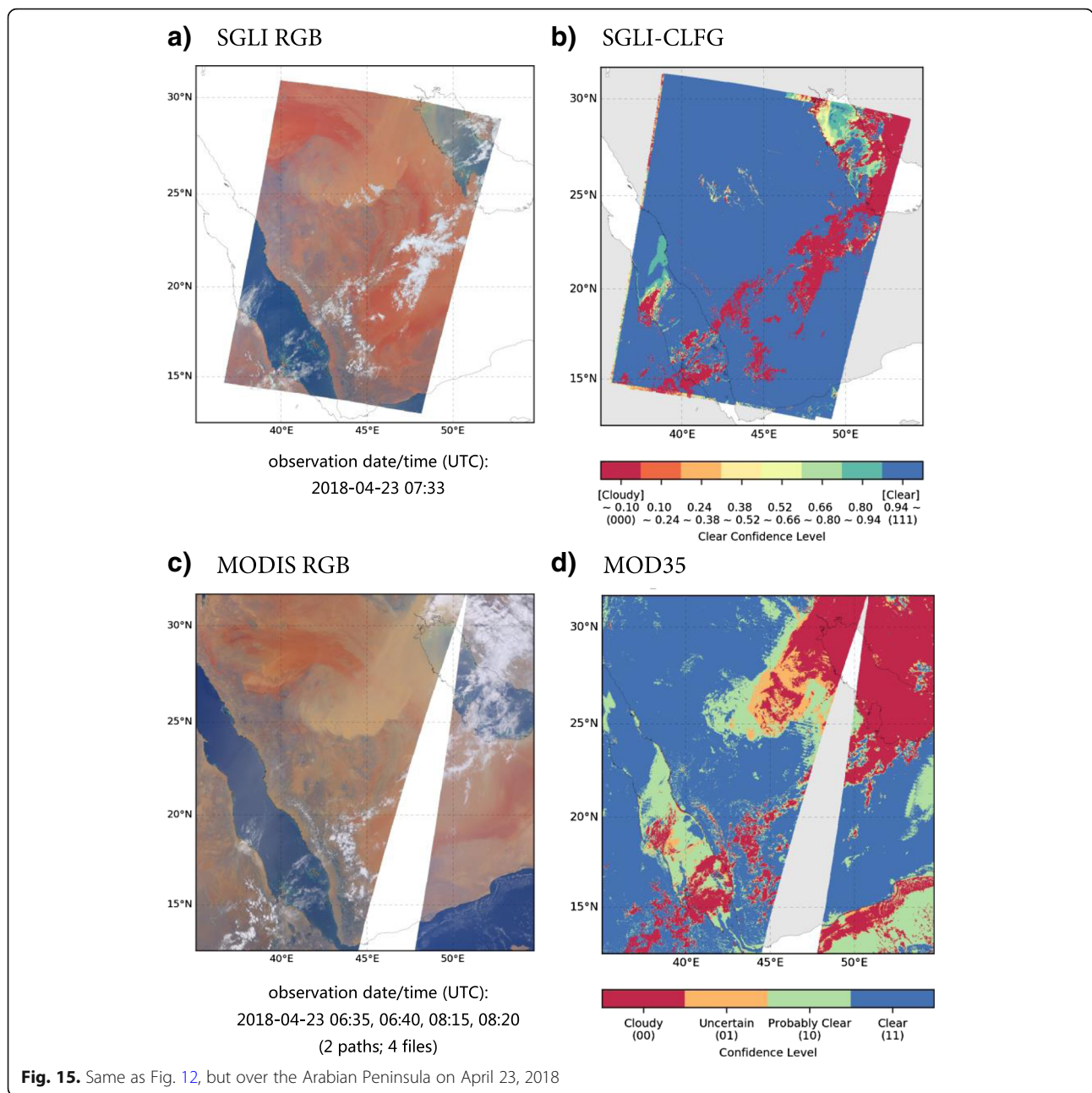
Comparison with other sensors
Cloud/clear discrimination
The authors evaluated SGLI cloud cloud/clear discrimination (referred to as the SGLI cloud flag [SGLI-CLFG]) by comparing it with the SGLI visible composite image and the MODIS cloud-mask product (MOD35). The orbits of GCOM-C and Terra are similar: both have a sun-synchronous sub-recurrent orbit with a 10:30 local sun time at the descending node. However, their swath and repeat cycles are different: the GCOM-C/SGLI VN (IR) instrument has a swath of 1150 km (1400 km) and a repeat cycle of 34 days. It can observe the whole globe approximately in every 2 days. The Terra/MODIS instrument has a swath of 2330 km and a repeat cycle of 16 days. It can cover the entire Earth's surface every 1 to 2 days. To clarify the characteristics of SGLI-CLFG (the Cloud Flag), five scenes were selected (Figs. 12, 13, 14, 15, and 16) with different cloud properties and ground surface characteristics. Each scene has four panels: (a) the SGLI true-color image combining the VN8, VN5,

Fig. 14 Same as Fig. 12, but over the Amazon forest on May 21, 2018

a) SGLI RGB

observation date/time (UTC):
2018-04-23 07:33

b) SGLI-CLFG

[Cloudy]   [Clear]
~ 0.10 0.10   0.24   0.38   0.52   0.66   0.80   0.94 ~
(000) ~ 0.24 ~ 0.38 ~ 0.52 ~ 0.66 ~ 0.80 ~ 0.94   (111)
Clear Confidence Level

c) MODIS RGB

observation date/time (UTC):
2018-04-23 06:35, 06:40, 08:15, 08:20
(2 paths; 4 files)

d) MOD35

Cloudy   Uncertain   Probably Clear   Clear
(00)   (01)   (10)   (11)
Confidence Level

Fig. 15. Same as Fig. 12, but over the Arabian Peninsula on April 23, 2018

and VN3 channels of a single SGLI level-1b product for the scene; (b) the CCL image of the SGLI-CLFG derived by applying CLAUDIA to the SGLI level-1b product; (c) the MODIS true-color image combining bands 1, 3, and 4 of multiple MODIS level-1b products overlapping with the SGLI level 1b scene; and (d) the confidence level image of MOD35. The CCL of the SGLI-CLFG consists of eight gradations (3 bits) obtained by discretizing the real value of 0.0 to 1.0 derived from Q in Equation (7), whereas the confidence level of MOD35 is one of four categories (cloud, uncertain, probably clear, clear). Figure 12

shows images over Japan captured on June 14, 2018, which include various clouds over both ocean and land with different top heights, optical thicknesses, and horizontal scales. Due to the differences in the orbits of the GCOM-C and Terra satellites, sunglint overlaps with the Japanese archipelago in the SGLI image, and appears around the western part of the MODIS image. The confidence levels of SGLI-CLFG and MOD35 are generally similar, although the SGLI-CLFG is continuous even in the sunglint area. This result implies that the sunglint correction shown in "Results and Discussion" section and Table 4 works

a) SGLI RGB

observation date/time (UTC):
2018-04-02 01:37

b) SGLI-CLFG

[Cloudy] [Clear]
~ 0.10   0.10   0.24   0.38   0.52   0.66   0.80   0.94 ~
(000) ~ 0.24 ~ 0.38 ~ 0.52 ~ 0.66 ~ 0.80 ~ 0.94 (111)
Clear Confidence Level

c) MODIS RGB

observation date/time (UTC):
2018-04-02 01:15, 01:20, 02:50, 02:55
(2 path; 4 files)

d) MOD35

Cloudy   Uncertain   Probably Clear   Clear
(00)   (01)   (10)   (11)
Confidence Level
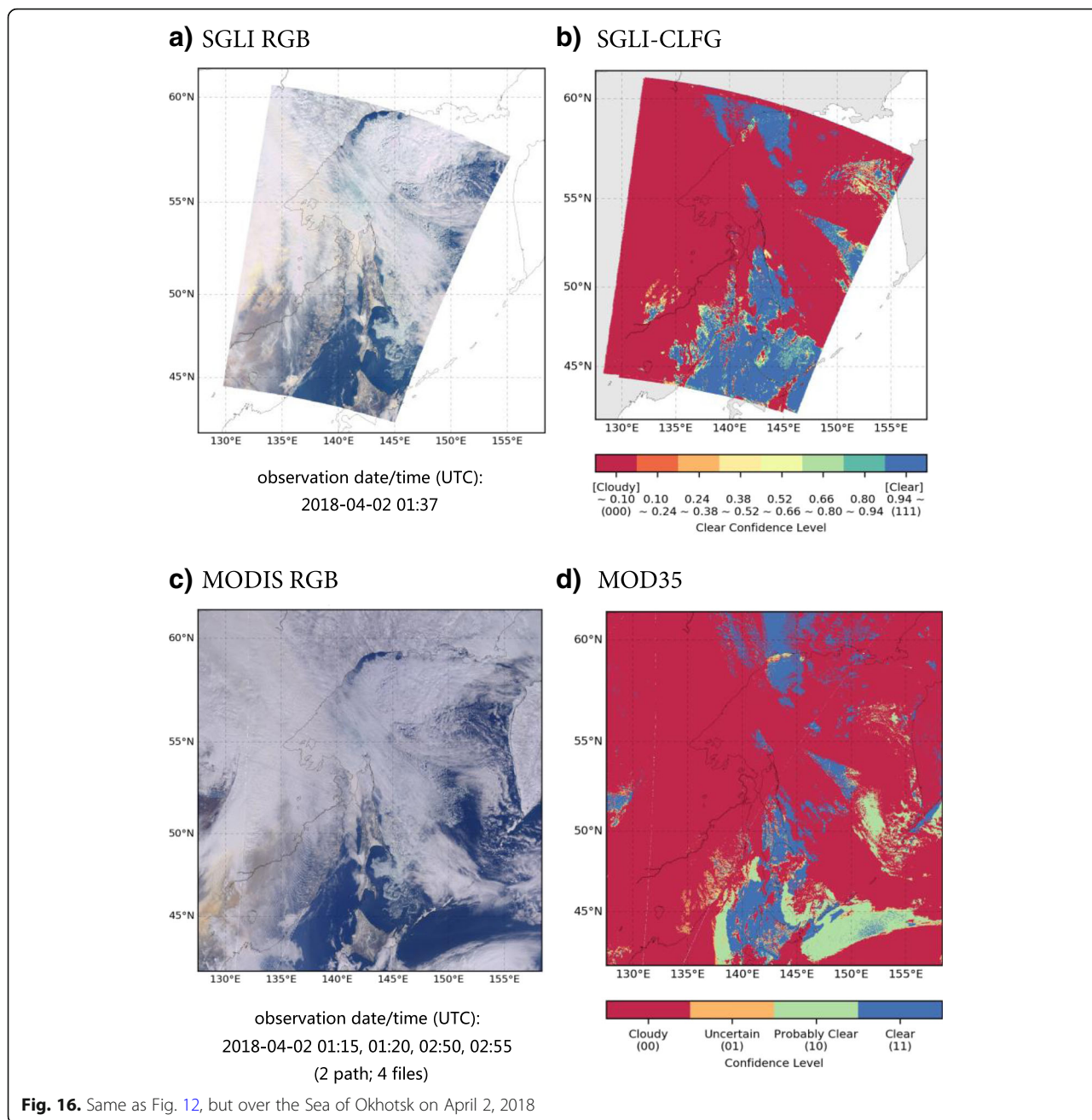
Fig. 16. Same as Fig. 12, but over the Sea of Okhotsk on April 2, 2018

well. The MOD35 tends to classify clear pixels strongly influenced by sunglint as probably clear. The other characteristic of the SGLI-CLFG is that the CCL depends on the apparent optical thickness of the cloud edge and small cumulus clouds because CLAUDIA was designed to allow cloud distinction ambiguity. Figure 13 shows the images capturing the belt of yellow sand flowing from Asia, traversing the Japan Sea, and mainland Japan on March 29, 2018. The CCL values of the SGLI-CLFG in the yellow sand belt are almost above 0.66, which means probably clear. However, MOD35 misclassified the same pixels as

cloudy. Figure 14 shows the images capturing the Amazon forest on May 21, 2018, which include many clouds of small horizontal extent, comparable to the instantaneous field of view (IFOV) of the SGLI and MODIS. The figures clearly show the difference between SGLI-CLFG and MOD35. In Fig. 14d, cloudy pixels are spread without gaps because MOD35 tends to be clear conservative. On the other hand, in Fig. 14b, the CCL of the SGLI-CLFG varies depending on the apparent optical thickness of cloud edges and small cumulus clouds because CLAUDIA is designed for neutral cloud/clear discrimination. Figure 15

shows images of the Arabian Peninsula captured on April 23, 2018, which include a dust storm in the center of the peninsula. Compared with the SGLI true-color image, the identification of SGLI-CLFG is reasonable. In addition, the CCLs of the SGLI-CLFG in the dust storm pixels the MOD 35 misidentified as cloudy are from 0.94 to 1. Finally, Fig. 16 shows an image around the Sea of Okhotsk on April 2, 2018. For pixels covered with snow and/or sea ice, it is generally difficult to distinguish clear sky and clouds because they have a spectrum similar to clouds, namely, white with high reflectance. Nevertheless, the CCL of the SGLI-CLFG is reasonable compared with the true-color image and MOD35.

Cloud properties

The cloud properties retrieved from the SGLI measurements with CAPCOM (described in "Results and Discussion" section) were evaluated by comparing them to the cloud properties with the MODIS Cloud Product (MOD06) Collection 6.1. Variables that were validated were cloud optical thickness, cloud particle effective radius, and cloud-top temperature for water and ice clouds. The data used were from over oceans and land in the middle and low latitudes (60° N–60° S) from August 22, 2018 to September 14, 2018. To reduce variability induced by cloud horizontal inhomogeneity and the gap between the IFOVs of the SGLI and MODIS, a spatial averaging

(within a 1° × 1° grid) and temporal averaging (during the above time period) were applied to retrievals. Only data with $\tau_c > 5$ were used to evaluate r_e and T_c of water- and ice-phase clouds to avoid estimation bias due to contamination from clear-sky radiance. In addition, only data with $T_c > 265$ K are used in the evaluation of τ_c, r_e, and T_c of water-phase clouds to avoid contamination from ice or mixed-phase cloud. Note that the thresholds are applied to the pixel-level retrievals.

(a) Water-phase clouds

Figure 17 shows scatter plots of (a, d) optical thickness, (b, d) effective radius, and (c, e) cloud-top temperature for water cloud from the SGLI retrievals compared with the MOD06 retrievals. Panels (a, b, c) are for ocean and (d, e, f) are for land. First, the results shown in Fig. 17a, d indicate that the SGLI τ_c values over ocean (land) are consistent with the MOD06 τ_c values with a sufficiently small bias of 0.7 (0.5), a reasonable root mean square error (RMSE) of 2.3 (3.8), and a high correlation coefficient (R) of 0.88 (0.83). The results shown in Fig. 17b, e indicate that the SGLI r_e values over ocean (land) are also consistent with the MOD06 r_e values with a bias of − 0.0 (− 0.5) μm, RMSE of 1.5 (1.8) μm, and R of 0.92 (0.88). Finally, the results shown in Fig. 17c (17f) indicate that the SGLI T_c values over ocean (land) are also consistent with the MOD06 T_c values with a bias of 0.4 (0.8) K,

Fig. 17 Scatter plots of **a**, **d** optical thickness, **b**, **e** effective radius, and **c**, **f** top temperature for water clouds from the SGLI retrievals compared with the MOD06 retrievals. Panels (**a**–**c**) are for ocean and (**d**–**f**) are for land
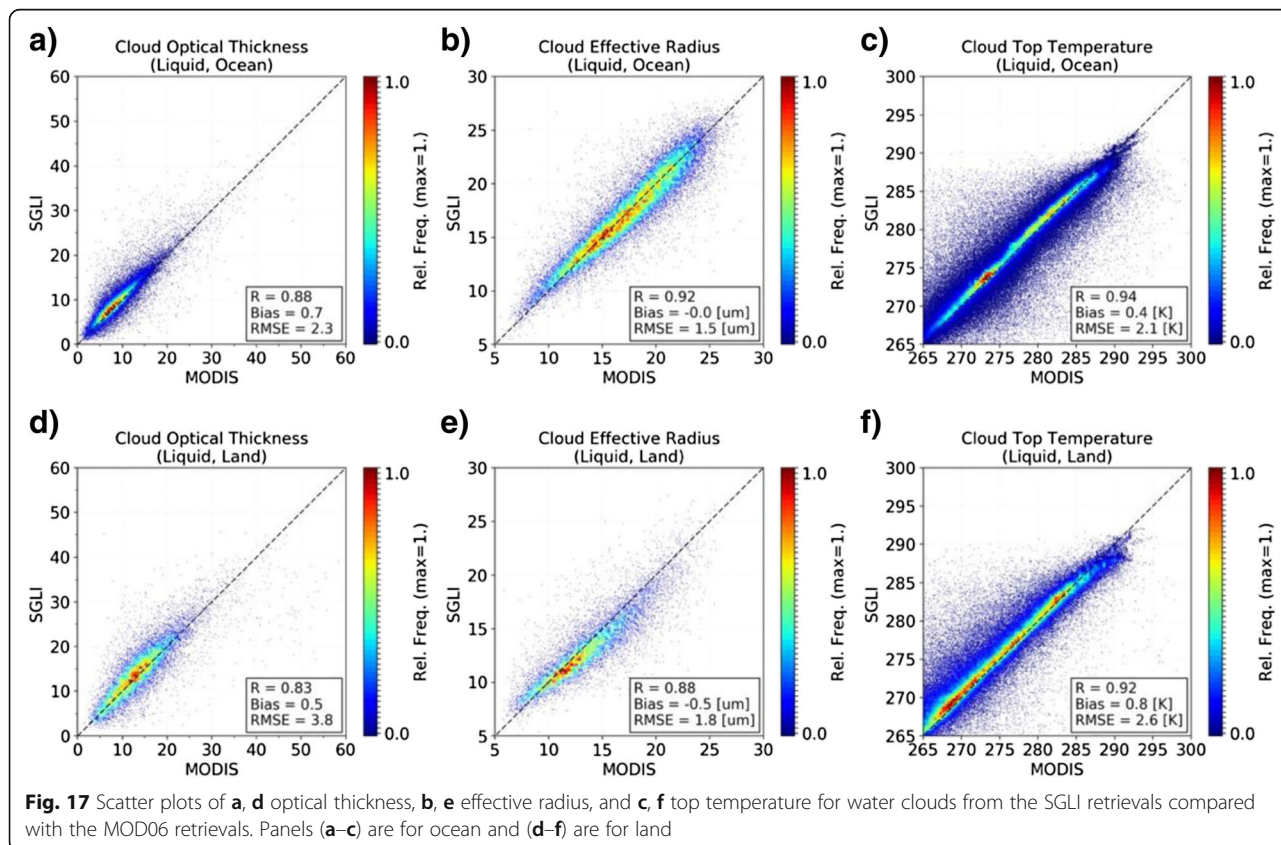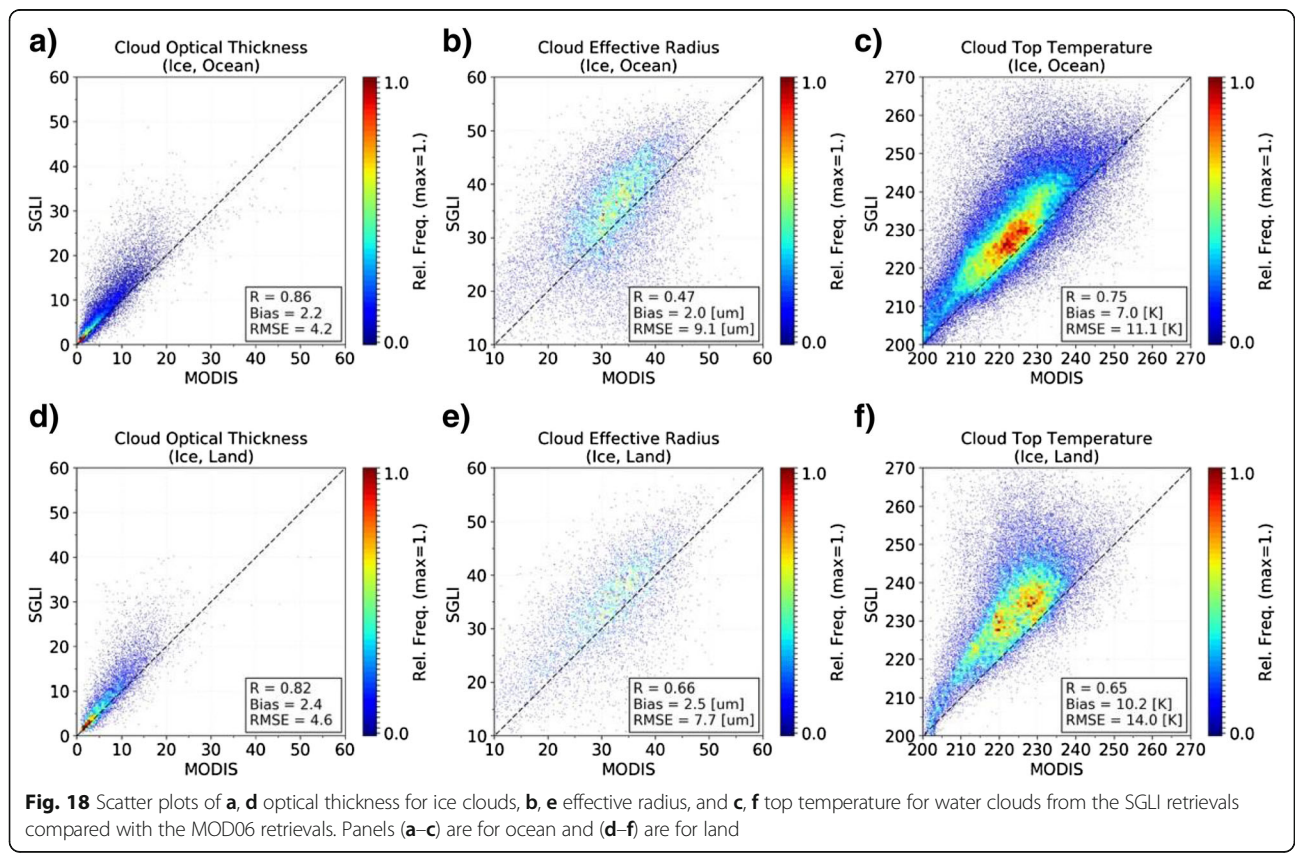
RMSE of 2.1 (2.6) K, and R of 0.94 (0.92). The finding that the SGLI-derived cloud properties are consistent with those of MOD06 despite the differences in sensor hardware, algorithm implementation, and wavelengths indicates the validity of SGLI's radiometric and geometric calibration and the high maturity of the water cloud retrieval algorithm.

(b) Ice-phase clouds

Figure 18 shows a comparison of the optical thickness, effective radius, and cloud-top temperature for ice clouds from MODIS and SGLI retrievals over ocean (top panels) and land (bottom panels). Figure 18a, d confirms that SGLI τ_c values over ocean (land) are consistent with the MOD06 τ_c values. The values of τ_c from SGLI are slightly larger than those from MODIS, and the biases of τ_c for the two satellite retrievals are 2.2 and 2.4 in the selected ocean and land areas, respectively. The RMSE and R values of the retrievals for the ocean (land) are 4.2 (4.6) and 0.86 (0.82), respectively. Figure 18b, e shows the comparison of the r_e values between the MODIS and SGLI retrievals over the ocean (land). The r_e values from the SGLI data are larger than those from the MODIS retrievals over both the ocean and the land. The bias, RMSE, and the R for the ocean (land) are 2.0 (2.5)

μm, 9.1 (7.7) μm, and 0.47 (0.66). These results can be explained by single-scattering properties (e.g., asymmetry parameter, single scattering albedo, extinction efficiency) of the ice particle habits used in ice-cloud retrievals of MODIS and SGLI products, as well as the different particle size distribution functions for the two satellite products (Letu et al., 2018). The Voronoi model and eight solid column aggregates with severely roughened surfaces ("8-clm-agg") were used in the ice-cloud retrievals for SGLI and MODIS data, respectively. Thus, the single-scattering properties of the Voronoi model are different from those of the 8-clm-agg model, which may lead to a difference in ice-cloud retrieval. In the SGLI ice-cloud retrievals, the log-normal size distribution function was used for ice-cloud properties, which is different from the gamma size distribution function used in the retrieval algorithm of the MODIS C6 cloud product. Figure 18c and f shows the T_c values from the MODIS and SGLI retrievals over the ocean (land). The T_c values from the SGLI retrievals are consistent with results from the MODIS retrievals, with R of 0.75 (0.65). However, the SGLI T_c values are larger than those from the MODIS retrievals, with a bias of 7.0 (10.2) K and an RMSE of 11.1 (14.0) K. In the MODIS-C6 product, the CO_2 slicing method is used to obtain T_c (Platnick et al. 2017). In contrast, CAPCOM in this

Fig. 18 Scatter plots of **a**, **d** optical thickness for ice clouds, **b**, **e** effective radius, and **c**, **f** top temperature for water clouds from the SGLI retrievals compared with the MOD06 retrievals. Panels (**a**–**c**) are for ocean and (**d**–**f**) are for land

study uses the single-channel thermal band data and cloud top properties from the reanalyzed data to retrieve T_c from the SGLI data (Nakajima and Nakajima 1995). Nevertheless, the T_c values from the SGLI and MODIS retrievals for water clouds agreed well (Fig. 17c, f), although the ice-cloud T_c values of the two products differed with the ice-scattering model used. One reason for the T_c difference between SGLI and MODIS is that the effective emission layers of clouds change with the wavelength used to derive T_c (e.g., the 10.8 μm brightness temperature will have more information about the inner cloud rather than the infrared CO_2 band). Another reason is that the CO_2 slicing method is more sensitive to high thin clouds than is the method of the single-channel thermal band used in this study. Thus, differences in T_c are considered to be caused by the differences in the retrieval method.

Conclusions

This paper described the theoretical basis of the algorithms for the GCOM-C SGLI sensor. The SGLI atmospheric group developed the cloudy/clear discrimination algorithm CLAUDIA and the cloud property retrieval algorithm CAPCOM. The CLAUDIA-output CCL for each observation pixel consists of 3 bits (eight gradations). This design allows an ambiguity of cloud coverage for a target pixel to be expressed. The results for CLAUDIA were evaluated by using the WSIs captured by ground-based fisheye camera systems. The evaluation stations were Ny-Alesund (78.93° N/11.86° E), Sapporo (43.08° N/141.34° E), Japan Aerospace Exploration Agency (JAXA) Tsukuba Space Center (TKSC) (36.07° N/140.13° E), Meteorological Research institute at Tsukuba (MRI) (36.06° N/140.13° E), Kumamoto (32.84° N/130.87° E), Miyakojima (24.74° N/125.33° E), and Syowa Station (69° S/39.59° E). Results showed that the user's accuracy and producer's accuracy were mostly high at around 90% and the overall accuracy was also high, ranging from 83 to 100% (average of all sites was 90.5%). Therefore, CLAUDIA had sufficient cloud-masking accuracy and met the threshold for the release of data for the SGLI cloud flag product, which is defined to be the minimum criterion for success of the GCOM-C mission.

CAPCOM retrieves cloud optical thickness, effective particle radius, and cloud top temperature for water- and ice-phase clouds by using the 1050-nm, 2210-nm, and 10.8-μm channels of the SGLI. The outputs were validated by comparing the retrieved values with the values of the MODIS standard product MOD06 over ocean and over land at middle and low latitudes (60° N −60° S) from August 22, 2018 to September 14, 2018. The comparison showed that the correlations of water cloud optical thickness, effective particle radius, and cloud top temperature were good, with correlation coefficients of 0.88 (0.83), 0.92

(0.88), and 0.94 (0.92) for ocean (land). The ice-cloud optical thickness also showed good correlations between the SGLI and MODIS products with correlation coefficients of 0.86 (0.82) over the ocean (land). SGLI gave higher effective particle radii and cloud-top temperatures for ice clouds than did MODIS owing to the effects of the ice-scattering model and size distribution function. These results meet the data release criterion defined for the GCOM-C mission. Standard GCOM-C products have been available to the public since data accuracy was validated in December 2018.

Author details
[1]Research and Information Center, Tokai University, 4-1-1, Kitakaname, Hiratsuka, Kanagawa 259-1292, Japan. [2]Meteorological Satellite Center, Japan Meteorological Agency, 3-235, Nakakiyoto, Kiyose, Tokyo 204-0012, Japan. [3]Atmosphere and Ocean Research Institute, The University of Tokyo, 5-1-5, Kashiwanoha, Kashiwa, Chiba 277-8568, Japan. [4]Earth Observation Research Center, Japan Aerospace Exploration Agency, 2-1-1, Sengen, Tsukuba, Ibaraki 305-8505, Japan. [5]State Key Laboratory of Remote Sensing Science, Institution of Remote Sensing and Digital Earth, Chinese Academy of Science, DaTun Road, No. 20 (North), Beijing 100101, China. [6]Remote Sensing Technology Center of Japan, JAXA Tsukuba Space Center Small-sized Satellite Test Building, 2-1-1, Sengen, Tsukuba, Ibaraki 305-8505, Japan. [7]Meteorological Research Institute, Japan Meteorological Agency, 1-1 Nagamine, Tsukuba, Ibaraki 305-0052, Japan.

References
Ackerman SA, Strabala KI, Menzel WP, Frey RA, Moeller CC, Gumley LE (1998) Discriminating clear-sky from clouds with MODIS. J. Geophys. Res. 103(D24):32,141–132,157

Albrecht BA (1989) Aerosols, cloud microphysics, and fractional cloudiness. Science 245:1227–1230

Baum BA, Soulen PF, Strabala KI, King MD, Ackerman SA, Menzel WP, Yang P (2000) Remote sensing of cloud properties using MODIS airborne simulator imagery during SUCCESS 2. Cloud thermodynamic phase. Journal of Geophysical Research-Atmospheres 105(D9):11781–11792

Hori M, Murakami H, Miyazaki R, Honda Y, Nasahara K, Kajiwara K, Nakajima TY, Irie H, Toratani M, Hirawake T, Aoki T (2018) GCOM-C data validation plan for land, atmosphere, ocean, and cryosphere. Trans. JSASS Aerospace Tech. Japan 16:218–223 https://doi.org/10.2322/tastj.16.218

Inoue T (1987) A cloud type classification with NOAA 7 split-window measurements. J. Geophys. Res. 92:3991–4000

IPCC (2013) Summary for policymakers. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York

Ishida H, Nakajima TY (2009) Development of an unbiased cloud detection algorithm for a spaceborne multispectral imager. Journal of Geophysical Research-Atmospheres 114. https://doi.org/10.1029/2008JD010710

Ishimoto H, Zaizen Y, Uchiyama A, Masuda K, Mano Y (2010) Shape modeling of mineral dust particles for light-scattering calculations using the spatial Poisson–Voronoi tessellation. J. Quant. Spectrosc. Radiat. Transfer 111:2434–2443

Kawamoto K, Nakajima T, Nakajima TY (2001) A global determination of cloud microphysics with AVHRR remote sensing. Journal of Climate 14:2054–2068

Kuji M, Murasaki A, Hori M, Shiobara M (2018) Cloud fractions estimated from Shipboard Whole-Sky Camera and Ceilometer Observations between East Asia and Antarctica. J. Meteorol. Soc. Japan 96:201–214 https://doi.org/10.2151/jmsj.2018-025

Letu H, Ishimoto H, Jerome R, Nakajima TY, Labonnote LC, Baran A, Nagao TM, Sekiguchi M (2016) Investigation of ice particle habits to be used for ice cloud remote sensing for the GCOM-C satellite mission. Atmospheric Chemistry and Physics 16(18):12287–12303

Letu H, Nagao TM, Nakajima TY, Matsumae Y (2014) Method for validating cloud mask obtained from satellite measurements using ground-based sky camera. Applied optics 53:7523–7533 https://doi.org/10.1364/AO.53.007523

Letu H, Nagao TM, Nakajima TY, Riedi J, Ishimoto H, Baran AJ, Shang H, Sekiguchi M, Kikuchi M (2018) Ice cloud properties from Himawari-8/AHI next-generation geostationary satellite: Capability of the AHI to monitor the DC cloud generation process. IEEE Transactions on Geoscience and Remote Sensing:1–11. https://doi.org/10.1109/TGRS.2018.2882803

Letu H, Nakajima TY, Matsui TN (2012) Development of an ice crystal scattering database for the global change observation mission/second generation global imager satellite mission: investigating the refractive index grid system and potential retrieval error. Applied Optics 51:6172–6178

Mouri K, Izumi T, Suzue H, Yoshida R (2016) Algorithm theoretical basis document for cloud type/phase product. Meteorological Satellite Center Technical Note 61:19–31

Nakajima T, Tanaka M (1986) Matrix formulation for the transfer of solar radiation in a plane-parallel scattering atmosphere. J Quant Spectrosc Radiat Transfer 35:13–21

Nakajima T, Tanaka M (1988) Algorithms for radiative intensity calculations in moderately thick atmospheres using a truncation approximation. J Quant Spectrosc Radiat Transfer 40:51–69

Nakajima TY, Nakajima T (1995) Wide-area determination of cloud microphysical properties from NOAA AVHRR measurements for FIRE and ASTEX regions. Journal of the Atmospheric Sciences 52:4043–4059

Nakajima TY, Nakajima T, Nakajima M, Fukushima H, Kuji M, Uchiyama A, Kishino M (1998) Optimization of the Advanced Earth Observing Satellite II Global Imager channels by use of radiative transfer calculations. Applied Optics 37:3149–3163

Nakajima TY, Sumi A, Imaoka K, Kikuchi N (2005) On the sciences obtained from ADEOS-II mission. SPIE, Sensors, Systems, and Next-Generation Satellites IX. In: Meynart R, Neeck SP, Shimoda H (eds) Proceedings of SPIE Vol. 5978. SPIE, Bellingham, p 597802

Nakajima TY, Tsuchiya T, Ishida H, Shimoda H (2011) Cloud detection performance of spaceborne visible-to-infrared multispectral imagers. Applied Optics 50:2601–2616

NOAA/NESDIS/STAR (2010) ABI Aerosol Detection Product. https://www.star.nesdis.noaa.gov/goesr/docs/ATBD/ADP.pdf. Accessed 12 Apr 2019.

Platnick S, Meyer KG, King MD, Wind G, Amarasinghe N, Marchant B, Arnold GT, Zhang Z, Hubanks PA, Holz RE, Yang P, Ridgway WL, Riedi J (2017) The MODIS cloud optical and microphysical products: Collection 6 updates and examples from Terra and Aqua. IEEE Transactions on Geoscience and Remote Sensing 55(1):502–525

Rossow WB, Schiffer RA (1999) Advances in understanding clouds from ISCCP. Bull. Amer. Meteorol. Soc. 80:2261–2288

Shiobara M, Yabuki M, Kobayashi H (2003) A polar cloud analysis based on Micro-pulse Lidar measurements at Ny-Alesund, Svalbard and Syowa, Antarctica. Physics and Chemistry of the Earth 28:1205–1212

Stamnes K, Tsay SC, Wiscombe W, Jayaweera K (1988) Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media. Appl Opt 27:2502–2509

Twomey S (1977) The influence of pollution on the shortwave albedo of clouds. J Atmos Sci 34:1149–1152

Twomey S, Piepgrass M, Wolfe TL (1984) An assessment of the impact of pollution on global cloud albedo. Tellus 36:356–366

Yabuki M, Shiobara M, Nishinaka K, Kuji M (2014) Development of a cloud detection method from whole sky color images. Polar Science 8:315–326

Yoshimura M, Yamashita M (2013) Contribution of ground-based cloud observation to satellite-based cloud discrimination. J. Environ. Sci. Eng. A 8:487–493

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.