



UNIVERSITÄT
LEIPZIG



University of Leipzig
Faculty of Physics and Earth Sciences
Institute of Meteorology

Characterisation of the properties of trade wind cumulus clouds with Sentinel-2 observations, including cloud cover, cloud height, cloud size distribution and their radiative effects

by

Oscar Ritter

A Thesis submitted for the degree of
Master of Science

First reviewer: Prof. Dr. Andreas Macke
Second reviewer: Prof. Dr. Johannes Quaas
Supervisor: Dr. Hartwig Deneke

Student-ID: 3751937
Mail: oscar.ritter@web.de

Submission date: 07.03.2022

Contents

1. Introduction	1
2. Definitions	4
2.1. Cloud fraction and cloud equivalent diameter	4
2.2. Cloud size distributions	4
2.3. Cloud fraction distributions	5
3. Dataset	7
3.1. Sentinel-2 Multi Spectral Imager (MSI) sensor	7
3.2. Description of the Sentinel-2 dataset	9
3.3. Shipbased Ceilometer data	9
4. Methods	11
4.1. Adaptation of a cloud mask	11
4.1.1. Theoretical background of cloud detection algorithms	11
4.1.2. Description of the algorithm	12
4.1.3. Threshold adjustment	13
4.2. Shadow detection algorithm	16
4.2.1. Description of the cloud shadow detection index (CSDI)	16
4.2.2. Definition of a CSDI threshold	16
4.2.3. Evaluation of the adjustet scene classification	17
4.3. Geometrical cloud base height estimation	20
4.3.1. Background of cloud height estimation using passive satellite remote sensing	20
4.3.2. Cloud base height retrieval algorithm	21
4.3.3. Comparison with ground-based ceilometer data	23
5. Analysis of Sentinel-2 MSI trade cumulus scenes	26
5.1. Cloud and shadow fraction	26
5.2. Cloud base height	28
5.3. Cloud size statistics	30
5.4. Resolution effects on cloud optical properties	32
6. Conclusion and Outlook	36
A. Appendix	39
A.1. Single cloud scenes	39
A.2. Overview images of analyzed scenes	40
B. References	43
C. List of acronyms	49
D. Acknowledgement	50
E. Declaration of Authorship	50

1. Introduction

Marine shallow cumulus clouds in the trade wind region, hereafter called trade cumulus, are found in large areas in the western parts of tropical oceans. These clouds form in large-scale environments characterized by a warm and moist layer as well as steady winds within the lowest kilometers of the atmosphere and below a temperature inversion, resulting from large-scale subsidence of dry, free-tropospheric air (Bony et al., 2017). Compared to stratocumulus decks, trade-cumulus is observed in regions with higher sea surface temperatures, a deeper boundary layer and smaller lower-tropospheric stability (Wood and Bretherton, 2006). The condensation and evaporation processes associated with the formation and dissipation of clouds are determined by the turbulent exchange of heat and moisture from the sea surface as well as the free troposphere (Bony et al., 2015). Due to the large albedo, trade cumulus strongly reflects incoming solar radiation, which cools the underlying atmosphere (Hartmann et al., 1992). Although the radiative processes depend primarily on small spatial and temporal scales, trade-cumulus clouds substantially influence the earth’s radiative budget as well as large-scale dynamics due to the widespread dispersion in the tropical oceans (Bony et al., 2017; Neggers et al., 2007).

The response of trade cumulus to a warming climate is critical for the estimation of global mean cloud feedbacks (Bony et al., 2017). In global climate models, parametrizations have to be used to represent small-scaled cloud-related processes. Unfortunately, not all processes can be quantified yet. The clouds respond more strongly to changes in the atmospheric environment in climate models than has been determined by observations (Vial et al., 2017; Nuijens et al., 2015). Therefore, tropical low clouds have been identified as the main source of the uncertainty between the model estimates of the global cloud feedback (Bony and Dufresne, 2005). Even though observational approaches have improved our understanding of the tropical low cloud feedback in recent years and the sign of the feedback is very likely positive, uncertainties about the strength of the feedback remain large. A better understanding of the governing processes is still necessary. (Klein et al., 2017).

To improve the understanding of key processes of trade cumulus through observations, the Elucidating the role of clouds-circulation coupling in climate (EUREC⁴A) field campaign was executed on 20-Jan-2020 to 20-Feb-2020 in the downstream winter-trades of the North Atlantic, eastward and south-eastward of Barbados. Its objective was to quantify the mass, energy and momentum balance in the subcloud layer, cloud microphysics as well as the link to the large scale environment and other components of the Earth system (Bony et al., 2017; Stevens et al., 2021). In-situ and remote sensing observations, including the High-flying aircraft (HALO), three low-flying in-situ aircrafts, four research vessels, long-range observations of the Barbados Cloud Observatory as well as 2500 drop- and radiosonde were taken to address the questions. (Stevens et al., 2021).

The campaign dataset was also planned to serve as a reference point for satellite retrievals, which offer a different perspective of clouds (Bony et al., 2017). These include the MSI on board the two polar-orbiting Sentinel-2 satellites, which were developed to provide optical observations over global terrestrial surfaces with a high spatial resolution up to 10 m (Drusch et al., 2012). During the campaign, the Sentinel-2 MSI observation coverage was extended to the ocean region east of Barbados. This provides one month high-resolution radiation data of tropical marine cloud scenes from this imager. So far, measurements have only been performed over coastal oceans and were used for the determination of the concentrations of water constituents (e.g. Pahlevan et al., 2017). In this work, atmospheric processes over marine surfaces are derived from Sentinel-2 MSI observations for the first time.

Spatially high resolution analyses of marine trade cumulus have previously been performed with other

satellite imagers, such as the imagers of various Landsat generations with horizontal resolutions between 80 m and 15 m as well as the Advanced Spaceborn Thermal Emission and Reflection Radiometer (ASTER) onboard the Terra satellite with a horizontal resolution up to 15 m, and a revisit frequency of 16 days (e.g. Wielicki and Welch, 1986; Cahalan and Joseph, 1989; Berendes et al., 1992; Koren et al., 2008; Zhao and Di Girolamo, 2007; Mieslinger et al., 2019). However, both Sentinel-2 satellites offer a 290 km wide observation path and thus the possibility of satellite based investigations from the micro- β to meso- β length scale according to Orlanski (1975) with a comparatively high revisit frequency of 5 days at the equator under the same viewing conditions. Therefore, the Sentinel-2 MSI observations of marine trade cumulus will help to characterize the individual macrophysical properties as well as the spatial organization.

The size of trade cumulus spans several meteorological length scales ranging from a few tens of meters to several kilometers in diameter. While smaller clouds can be observed in very large numbers, the number of larger clouds, on the other hand, is orders of magnitude lower. Several approaches (exponential, lognormal, power law) exist to describe empirical cloud size data by mathematical functions (Neggers et al., 2003). However, recent considerations commonly assume the size data follow a power law distribution (Mieslinger et al., 2019). The knowledge of a mathematical description of the size distribution along with cloud fraction as a variable for radiative flux forms an important basis for parameterizations of convective mass flux in global climate models (Mieslinger et al., 2019; Neggers et al., 2003). Cloud size distributions based on high-resolution satellite imagery showed that the small cumulus clouds provide the largest contribution to the cloud fraction (Zhao and Di Girolamo, 2007) as well as to the total cloud reflectance (Koren et al., 2008). In addition, it was shown that the estimated cloud fraction depends on the spatial resolution of the satellite sensor. The use of coarse resolution detectors generally leads to an overestimation of the cloud fraction, since most pixels contain reflectances of cloud and cloud-free region (Zhao and Di Girolamo, 2006; Koren et al., 2008). However, these publications only showed averaged values or analyzed a single scene and did not consider more closely at whether differences in change of the cloud fraction at coarser detector resolution between the cumulus scenes were apparent.

Observations of the Cloud base height (CBH) of trade cumulus are typically conducted using ground-based remote sensing techniques, such as ceilometers or Light Detection and Ranging (LIDAR) (e.g. Nuijens et al., 2014; Stevens et al., 2016). These techniques require an observation platform and provide information on the temporal evolution of height at a point, but not on the spatial distribution. Likewise, high-resolution satellite observations were used in only a few publications to characterize the spatial distribution of cloud base heights (e.g. Berendes et al., 1992; Böhm et al., 2019). The combination of high-resolution Sentinel-2 MSI observations and ground-based remote sensing measurements in the EUREC⁴A campaign offers the opportunity to develop a CBH estimation method for Sentinel-2 data.

Characterisation of cloud properties such as cloud cover, CBH and cloud size using satellite remote sensing requires a number of assumptions that allow the separation of cloudy areas from surface or atmospheric properties (Goodman and Henderson-Sellers, 1988). The method that separates into clear and cloudy areas depends on the goal of the cloud detection and on the characteristics of the sensor, since there is no single retrieval algorithm capable of performing under all circumstances at all locations of the earth (Rossow, 1989). Numerous experimental algorithms exist for the detection of cloud cover over ocean surfaces using high resolution satellite sensors, which separate clouds from the dark ocean surface by their differences in brightness in one (e.g. Wielicki and Welch, 1986; Zhao and Di Girolamo, 2007; Dey et al., 2008; Koren et al., 2008) or more (e.g. Werner et al., 2016) spectral channels based on threshold tests. On the other hand, for the detection of clouds at a single timestep with the Sentinel-2 MSI sensor, some detection algorithms (e.g. Zhu et al., 2015) of other high-resolution satellite sensors were adapted

in addition to the operational scene classification of pixels Sen2Cor (Louis et al., 2016). The cloud mask algorithm Function of mask (Fmask) (Zhu and Woodcock, 2012), originally developed for Landsat 7, has a high detection accuracy, but was developed for the separation of clouds over a wide variety of surfaces and is therefore relatively complex (Tarrio et al., 2020). This work will show, if the adaptation of a cloudmask for marine low-level clouds from Werner et al. (2016) leads to a similar performance even with a reduced number of decision tests.

The main objective of the thesis is the characterization of macrophysical properties of trade cumulus in terms of cloud fraction, shadow fraction, CBH and cloud size distribution based on high resolution Sentinel-2 MSI observations. For this purpose, the properties will be determined for 9 110x110 km² sized satellite scenes of small to medium trade cumulus and the differences between the scenes will be presented. In addition, the dependence of cloud properties on the spatial resolution of the satellite sensor will be investigated. As a basis, a cloud detection algorithm developed by Werner et al. (2016) for retrievals of low clouds with ASTER will be adapted to the Sentinel-2 data sets and the accuracy compared to other cloud masks will be evaluated. For CBH estimation, an algorithm following the method of Berendes et al. (1992) based on the geometric match between cloud and cloud shadow will be developed and compared with ship-based ceilometer measurements. The method also requires the successful detection of cloud shadows from the ocean surface. Therefore, a shadow detection algorithm by Amin et al. (2012) based on the Cloud shadow detection index (CSDI) will be adapted to the Sentinel-2 data sets.

The thesis is structured as follows: In section 2 mathematical definitions of the size variables are made. Section 3 introduces the Sentinel-2 MSI sensor and the dataset. The methods for detecting cloud and shadow areas as well as estimating cloud base height are presented and evaluated in section 4. Finally, the methods are applied to 9 trade cumulus scenes in section 5 and these are discussed in terms of cloud fraction, shadow fraction, cloud height, their cloud size distribution, and the influence of detector resolution on the determination of these cloud properties.

2. Definitions

Remote sensing of clouds using passive satellite methods is based on the detection of spectral radiances by the satellites instrument. The radiances are directed upward at the Top of atmosphere (TOA) as a result of emission and scattering of electromagnetic radiation by the atmosphere and the underlying surface. The derived properties should therefore be understood as quantities integrated over the entire path (Stephens and Kummerow, 2007). In the visible to shortwave-infrared wavelengths, the detected radiance results mostly from the reflection of incident solar radiation. Therefore, spectral TOA reflectances are used in the analysis, which characterize the ratio of the detected to the incident solar radiance. A technical definition of the Sentinel-2 MSI reflectance is performed in section 3.1.

In the context of this work, cloud size, cloud equivalent diameter, cloud fraction, and frequency distributions of cloud size and cloud fraction are derived from Sentinel-2 MSI reflectances. The size variables and their mathematical relationships are presented below.

2.1. Cloud fraction and cloud equivalent diameter

Cloud objects are considered as contiguous areas of pixels identified as cloudy by the scene classification (see chapter 4.1.1). The cloud area A is therefore calculated from the summation of all $10 \times 10 \text{ m}^2$ pixels of a cloud object and corresponds to the thickness-dependent projection area of the cumulus cloud along the observation vector onto the Earth's surface (Neggers et al., 2003). In addition to the area of a cloud, the cloud equivalent diameter D as a typical length scale measure is calculated from the cloud area by assuming a circular cloud:

$$D = \sqrt{\frac{4 \cdot A}{\pi}}. \quad (1)$$

The cloud fraction F is defined as the ratio of the total projected area of all cloud objects to the area A_{scene} of the considered cloud scene. The proportion of clouds of one diameter to the total cloud fraction $F(D)$ can therefore be calculated from the total number of clouds of one diameter $N(D)$ by

$$F(D) = \frac{N(D) \cdot \pi \cdot D^2}{4 \cdot A_{scene}}. \quad (2)$$

2.2. Cloud size distributions

For the description of trade wind cumulus, it is important to find a functional relation for cloud size density, defined as the probability density function of the number of clouds with a size variable such as cloud equivalent diameter D (Neggers et al., 2003). The total number of clouds N can be derived from the integral of the cloud size density $N(D)$ over all cloud equivalent diameters dD

$$N = \int_0^{\infty} N(D) dD \quad (3)$$

and is a conserved quantity independent from the chosen size variable. Since both the number and size of trade cumulus span several orders of magnitude, but the majority of clouds are found in small cloud sizes, several approaches exist to describe empirical cloud size data by mathematical functions. Plank (1969) and Wielicki and Welch (1986) described $N(D)$ with an exponential function, for example. However, in the latest publications of observed size distributions, a power function is used to describe $N(D)$ (e.g. Cahalan and Joseph, 1989; Neggers et al., 2003; Zhao and Di Girolamo, 2007; Koren et al., 2008; Mieslinger et al., 2019):

$$N(D) = a \cdot D^b. \quad (4)$$

The exponent of the power function b is called the scaling parameter and is typically in the range $-3 \leq b \leq -2$ for natural phenomena (Mieslinger et al., 2019). For a sufficiently large number of sampled cloud objects, this can be derived by minimizing the squared residuals in a linear regression from a log-log histogram of the cloud size distribution. The mean sizes of the histogram bins are used as reference points. The scaling parameter b results from the slope of the linear regression:

$$\log(N(D)) = \text{constant} + b \cdot \log(D). \quad (5)$$

The derivation of a functional description by a linear least squares fit can lead to strong deviation in the determined scaling parameters when using too small data sets, because the size classes are incompletely occupied for the large cloud sizes (Clauset et al., 2009). Therefore, to increase the number of cloud objects in the upper size classes, size distributions using exponentially increasing histogram bins (logarithmic binning) are also determined according to Mieslinger et al. (2019). The power functions approximated to the cloud size distributions with logarithmic binning $N(\log(D))$ are mathematically related to the size distributions with equidistant binning $N(D)$ in the following way:

$$N(\log(D)) = N(D) \cdot \frac{dD}{d\log D} = N(D) \cdot D \cdot \ln(10) = a \cdot \ln(10) \cdot D^{b+1} \quad (6)$$

Some previous publications derived cloud size distribution using the projected area A as the size variable and the logarithmic binning scheme (Koren et al., 2008). The approximated power function to the cloud size distribution $N(\log(A))$ can be calculated from the size distributions of the equivalent diameter $N(D)$ as follows:

$$N(\log(A)) = N(D) \cdot \frac{dD}{d\log(A)} = N(D) \cdot \frac{\frac{\pi}{4} \cdot D^2 \cdot \ln(10) \cdot 2}{\pi \cdot D} = a \cdot \frac{1}{2} \ln(10) \cdot D^{b+1} = a' \cdot A^{\frac{b+1}{2}} \quad (7)$$

The exponents of the power functions using different binning methods can thus be compared via:

$$b(N(D)) \propto b(\log(D)) - 1 \quad (8)$$

and

$$b(N(D)) \propto 2 \cdot b(N(\log(A))) - 1. \quad (9)$$

2.3. Cloud fraction distributions

This work will also focus on the contribution of clouds with an equivalent diameter D to the total cloud fraction F of the scene. For this purpose, the slope of the cloud fraction contribution $\frac{dF}{dD}$ is computed from equation 2, which turns to be

$$\frac{dF}{dD} = \frac{\pi}{4} \cdot a \cdot (b + 2) \cdot D^{b+2}. \quad (10)$$

Consequently, the probability density function of the cloud fraction contribution shows a monotonically increasing behavior for a scaling parameter $b \geq -2$ and a monotonically decreasing behavior for $b \leq -2$. A monotonically increasing behavior would correspond to a larger contribution of large clouds to the total cloud fraction, a monotonically decreasing behavior to a larger contribution of smaller clouds to the total

cloud fraction. However, it must be noted that this calculation is made under the assumption that the cloud areas within a bin are also continuously distributed like the cloud equivalent diameters and the distribution can be described with a power law (Zhao and Di Girolamo, 2007). Due to the finite resolution and irregular geometry of the cloud objects, the assumption might be violated in some bins.

3. Dataset

3.1. Sentinel-2 MSI sensor

The Sentinel-2 MSI is a high-resolution sensor with 13 spectral bands aboard Sentinel-2A and Sentinel-2B, two sun-synchronous satellites at 786 km altitude. The satellites were launched in 2015 and 2017, respectively, by the European Space Agency (ESA) as part of the Global Monitoring for Environment and Security (GMES) program (Szantoi and Strobl, 2019). The mission provides continuity to services relying multi-spectral high-resolution optical observations over global terrestrial surfaces (Drusch et al., 2012).

Band	$\lambda_{central}$ (nm)	$\Delta\lambda$ (nm)	Resolution (m)	L_{ref} ($\text{W}\cdot\text{m}^2\cdot\text{sr}^{-1}\cdot\mu\text{m}^{-1}$)	Band parallax ($^\circ$)	Temporal offset (s)
1	443	20	60	129	3.14	2.31
2	492	66	10	128	1.26	-
3	560	36	10	128	1.70	0.53
4	665	31	10	108	1.93	1.01
5	704	15	20	74.5	2.18	1.27
6	740	15	20	68	2.40	1.53
7	783	20	20	67	2.62	1.79
8	833	106	10	103	1.50	0.26
8a	865	21	20	52.5	2.92	2.05
9	945	20	60	9	3.18	2.59
10	1375	31	60	6	1.76	0.85
11	1610	90	20	4	2.30	1.47
12	2190	180	20	1.5	2.86	2.09

Table 1: Sentinel-2 MSI spectral band characteristics: central wavelength $\lambda_{central}$, bandwidth at full width half maximum $\Delta\lambda$, spatial resolution, reference radiance level L_{ref} , along-track parallax angle between odd and even detectors and temporal offset in relation to Band 2 (Drusch et al., 2012; Fletcher, 2012). The bands used in the analysis are marked in bold.

Both identical satellites are maintained in the same orbit with a phase delay of 180 degree and overpass the equator in a 10.30 am descending node (14 and 3/10 revolutions per day), to make a compromise between minimizing cloud cover and ensuring suitable sun illumination (Drusch et al., 2012). The Sentinel-2 mission provides systematic observations of land areas, islands larger than 100 km² as well as coastal and inland waters between 56°S and 84°N (Gascon et al., 2017). Within the framework of calibration sites or measurement campaigns, observations can also be requested in other areas, such as the area of the southern Sargasso sea east and south-east of Barbados (Gascon et al., 2017).

Table 1 shows an overview of the properties and resolutions of the 13 MSI spectral bands. Bands 1-4 are in the Visible (VIS), bands 5-9 bands are in the Near-infrared (NIR), and bands 10-12 are in the Shortwave infrared (SWIR) spectral range. Band 2 in the red, band 3 in the green, band 4 in the blue as well as band 8 in the NIR spectral range operate with a resolution of 10 m, the wider bands 11 and 12 in the SWIR as well as 5 further narrow bands in the Visible and near-infrared (VNIR) spectral range with a resolution of 20 m. Three additional bands provide observations with a resolution of 60 m in the edge of water-vapor absorption lines, mainly for the detection of cirrus clouds and the atmospheric correction (Gascon et al., 2017). However, Sentinel-2 MSI does not include a band in the thermal infrared, which is important for

cloud-detection and cloud-top-height estimation, since the emission of radiation is temperature dependent and optical thick cumulus stand out with lower values of the brightness temperature compared with the relatively warm ocean surface (Zhu et al., 2015).

The 12 detectors on the two focal planes of the VNIR and SWIR bands are employed in a staggered configuration, which covers a 20.6° field of view and enables a large swath width of 295 km (Drusch et al., 2012). The large field of view at both satellites allow a high revisit frequency of 5 days at the equator under the same viewing conditions. In comparison, the Terra ASTER observes in a more narrow field of view of 60 km, which leads to a larger revisit frequency of 16 days (Werner et al., 2016).

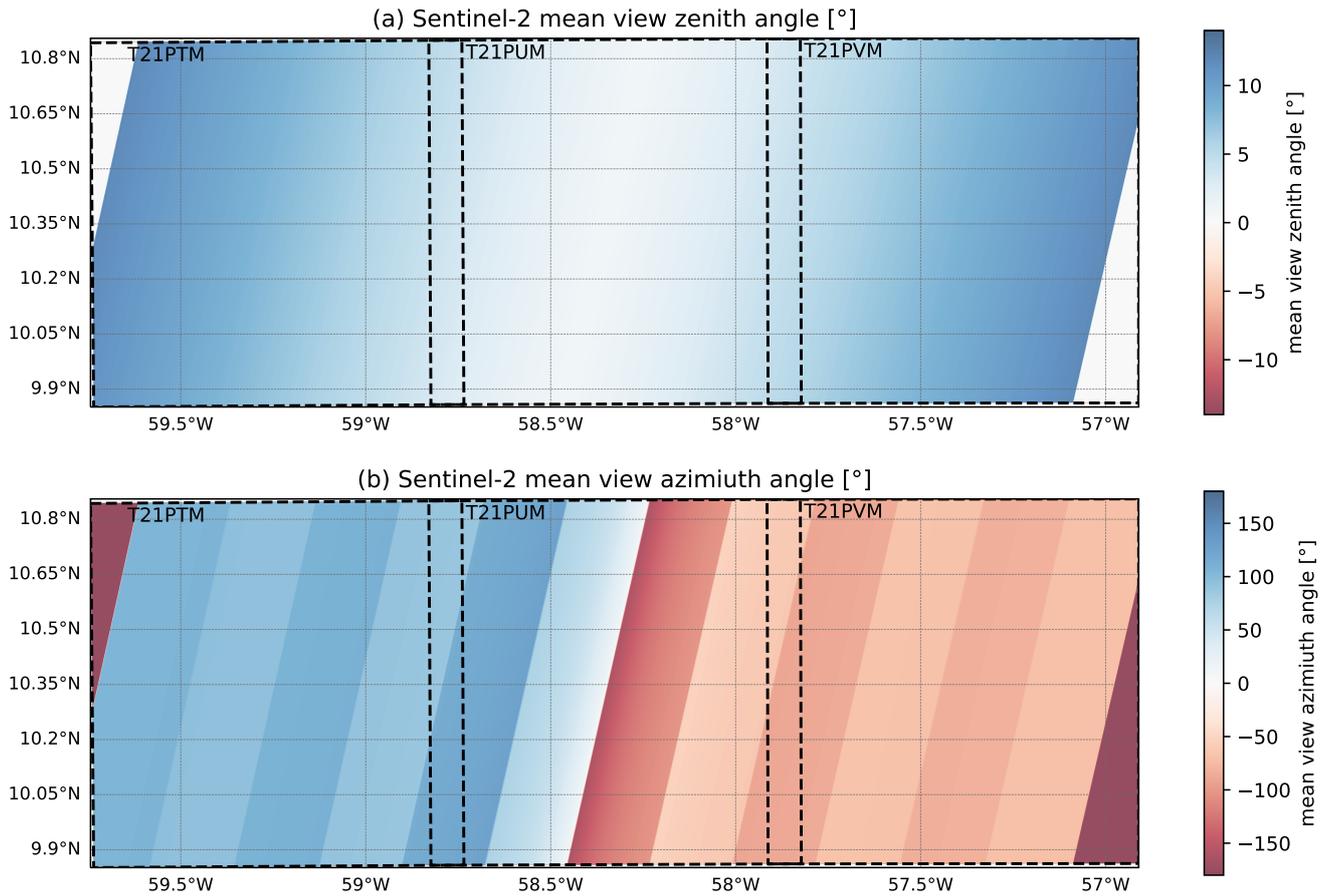


Figure 1: Satellite observation angles (a) θ_{sat} and (b) ϕ_{sat} in a cross section of the observation path on 31-Jan-2020. The angles were averaged from all bands involved in the CBH estimation (Bands 2, 3, 4, 8 and 12). The black lines indicate the size and location of the data tiles T21PTM, T21PUM and T21PVM

The arrangement of the detectors leads to differences in the observation geometry. Figure 1 shows the zenith and azimuth angles of observation averaged over the 5 bands used in this analysis for a cross-section of the observation path on 31-Jan-2020. Each detector covers a 25.9 km wide observation area, overlapping with the neighboring detector by 1.9 km (Gascon et al., 2017). As a consequence of the two-row staggered arrangement, the azimuth angle of the observation differs between odd and even detectors. The variability in the viewing angle lead to different radiometric properties of e.g. the ocean surface, due to the individual angular relationship between the observation angle and the solar angle (Gascon et al., 2017). The arrangement of the bands on the detectors also differs as a consequence of the mirror concept

(Gascon et al., 2017). As the recording of single bands is done with a push-broom concept, a time shift in the observation of an area with two different bands results by up to 2.31s. Since the satellite itself has moved during this time, this leads to different observation angles between two bands in the same detector field of view (Gascon et al., 2017; Frantz et al., 2018). While this parallax between two bands can be accounted with a geometric correction for stationary objects, the displacement is still visible for non-stationary objects with an unknown altitude (e.g. clouds) (Frantz et al., 2018).

The spectral, equalized digital counts of the raw image $d_\lambda(i, j)$ are converted for each pixel i, j into the TOA spectral reflectances $R_\lambda(i, j)$ by

$$R_\lambda(i, j) = \frac{\pi \cdot d_\lambda(i, j) \cdot r_{SE}^2}{A_k \cdot F_0(\lambda) \cdot \cos \theta_0(i, j)} \quad (11)$$

using the sun-earth-distance r_{SE}^2 , an absolute calibration coefficient A_k of the instrument, the incoming spectral solar irradiance $F_0(\lambda)$ and the solar zenith angle $\theta_s(i, j)$. The conversion is part of a 5-step processing of the compressed raw data into the published geolocated TOA dataset with sub-pixel multi-spectral and multi-date registration in Level 1C (Drusch et al., 2012). During processing, a radiative correction is applied to the computed reflectances that removes dark and blind pixels, interpolates defective pixels, and performs a pixel response non-uniformity correction. In addition, a linear correction is applied to the SWIR pixels to reduce electronic cross-talk effects (Gascon et al., 2017). In the last step, a geometric refining is applied to increase the accuracy in the geolocation and the reflectances are transformed to a cartographic UTM/WGS84 projection. The datasets are provided as tiles of 110x110 km² size each, overlapping with the neighboring tile by 5 km in all directions.

3.2. Description of the Sentinel-2 dataset

For the selection of the analyzed scenes, Red-green-blue image (RGB) quicklooks of the observations in a range of 10°- 20°N and 50°- 60°W were first created for the period between 20-Jan-2020 and 20-Feb-2020 and sorted according to the presence of trade cumulus and the absence of major clusters or multilayer cloud cover. The Level 1C TOA datasets of the selected overflights were downloaded via the Copernicus Open Data API Hub¹. For the high resolution datasets RGB images of the individual scenes were created and examined for the absence of thin cirrus cloud cover again.

For the determination of property averages, but also to highlight differences in the characteristics between trade-cumulus scenes, 9 tiles with a size of 110x110 km² were selected from all datasets. The tiles are located in three cross-sections of the observation path on 31-Jan-2020, 02-Feb-2020, and 05-Feb-2020 and were selected based on the presence of shallow to moderate trade cumulus throughout the cross-section. For these tiles, the Bottom of atmosphere (BOA) datasets from Level 2A were also downloaded, which include atmospherically corrected reflectances as well as a scene classification generated with the Sen2Cor processor (Richter et al., 2011). Table 2 shows the position and properties of the sensor and solar geometry for the selected scenes.

3.3. Shipbased Ceilometer data

Within the scope of the thesis, the CBH over ocean surfaces will be derived from Sentinel-2 observations. For validation, a comparison with CBH measurements from 2 Jenoptik CHM15k ceilometers on board the research vessels Meteor and L'Atalante will be performed. Ceilometers are active remote sensing devices

¹The Copernicus Sentinel-2 data of the year 2020 used in this thesis is available at <https://scihub.copernicus.eu/apihub/odata/v1/>

Date	Tile	Latitude[°N] Longitude[°W]	viewing zenith [°] viewing azimuth [°]	sun zenith [°] sun azimuth [°]
2020-01-31	T21PTM	9.85 - 10.85	1.6 - 11.8	36.6 - 38.0
		57.81 - 58.83	88.0 - 124.0	137.0 - 138.9
2020-01-31	T21PUM	9.86 - 10.85	1.0 - 5.6	36.0 - 37.4
		58.73 - 59.74	-163.1 - -128.8	138.1 - 139.9
2020-01-31	T21PVM	9.86 - 10.86	3.0 - 11.9	35.5 - 36.8
		56.91 - 57.91	-90.6 - -58.5	139.1 - 141.1
2020-02-02	T21QYU	16.17 - 17.18	0.8 - 7.2	40.1 - 41.5
		54.08 - 55.13	-164.4 - 132.2	143.3 - 145.0
2020-02-02	T21QZU	16.15 - 17.17	1.4 - 11.3	35.9 - 37.3
		53.14 - 54.19	-130.7 - -55.5	138.8 - 140.7
2020-02-05	T21PTP	11.66 - 12.66	6.1 - 11.8	37.1 - 38.4
		58.74 - 59.77	95.2 - 112.1	136.7 - 138.5
2020-02-05	T21PUP	11.67 - 12.66	0.8 - 8.7	36.5 - 37.9
		57.82 - 58.84	-164.2 - 131.8	137.7 - 139.6
2020-02-05	T21PVP	11.67 - 12.66	0.8 - 9.9	35.9 - 37.3
		56.91 - 57.92	-164.2 - 59.0	138.8 - 140.7
2020-02-05	T21PWP	11.67 - 12.66	7.2 - 11.9	35.4 - 36.7
		55.99 - 57.00	84.2 - -69.6	134.2 - 136.2

Table 2: Overview of the position and satellite and solar geometry of the data tiles used in the analysis.

that operate on the LIDAR principle. A vertically emitted laser pulse is scattered by particles in the atmosphere and a backscatter profile is generated from the power scattered to the receiver (Emeis et al., 2004). The height of the first 3 cloud bases are derived from strong inhomogeneities in the backscatter profile with a range resolution of 5 m and the uncertainty of the height derivation is given as ± 5 m (Heinemann, 2015). The first base height will be used as CBH.

Despite the numerous research vessels that provided ground-based remote sensing data in the observation area as part of the EUREC⁴A campaign, only two scenes on 10-Feb-2020 with a good overlap between the observation path and vessel measurements and the presence of small to medium-sized trade cumulus could be selected for the evaluation. Therefore, in addition to the characterization, the Sentinel-2 datasets in tile T21PVR (13.48 - 14.47°N, 56.90 - 57.92°W) and tile T21PTM (9.85 - 10.85°N, 57.81 - 58.83°W) on 10-Feb-2020 were used for the evaluation. The associated ceilometer data sets were obtained from the EUREC⁴A AERIS data server ².

²<https://observations.ipsl.fr/aeris/eurec4a/25-Feb-2020>

4. Methods

4.1. Adaptation of a cloud mask

The analysis of the radiative properties of clouds requires classification of image pixels into those in which the radiative flux detected by the satellite sensor originates exclusively from optically thick clouds, exclusively from the cloud-free surface, and those in which both the surface and clouds contribute to the radiative flux. Therefore, an algorithm for classifying pixels into confidently cloudy, probably cloudy, probably clear, and confidently clear cloud classes is presented below by Werner et al. (2016) and adapted to the Sentinel-2 MSI sensor.

4.1.1. Theoretical background of cloud detection algorithms

Algorithms to retrieve clouds from satellite images can be divided into fixed threshold tests, statistical methods based on brightness distributions of multiple bands or spatial coherence, and radiative transfer methods from which cloud optical thickness is directly derived (Goodman and Henderson-Sellers, 1988). In this work, cloudy regions are detected using fixed threshold test, but also via the ratio of reflectances in two spectral regions, the ratio tests. The adjustment of the threshold values is done via the analysis of distinct clusters in frequency distributions of the reflectances of a scene.

Fixed detection thresholds are determined for observables - such as the bidirectional reflectance - on which cloudy and clear pixels show a strong contrast to each other (Rossow, 1989). The contrast can be chosen in spatial as well as in temporal dimension and the assignment of pixel properties is often done by using a fixed decision tree that includes tests in multiple spectral regions. Since the MSI images on both Sentinel-2 satellites are characterized by a broad observational swath and a high revisit frequency, both single-scene and multi-temporal algorithms have been developed.

The operational processing of TOA reflectances stored in Level-1C to BOA reflectances in Level-2A includes a scene classification (Sen2Cor) step, that assigns pixels to three different cloud classes and a cloud shadow class, respectively. It is based on single band thresholds as well as ratio tests (Richter et al., 2011). Some tests are also based on differences in the reflectance of two bands. In addition, two threshold values are defined for each reflectance tests and the cloud probability changes linearly from 0 to 1 between both thresholds (Richter et al., 2011). The scene classification is operationally provided at a resolution of 20 m as part of the Level-2A dataset.

Another mature scene classification algorithm is the Function of mask (Fmask), a single-date, object based method for cloud and cloud shadow detection in Landsat 4-7 data (Zhu and Woodcock, 2012). It was adopted by Zhu et al. (2015) to Sentinel-2 data. The detection algorithm considers both a fixed threshold decision tree based on cloud physical properties and a cloud probability layer based on the spatial variability of reflectances in a variety of steps. A prediction of the potential location of the cloud shadow is used to link detected shadow pixels to the cloud object. Frantz et al. (2018) improved the separation of low clouds using the displacement of elevated objects as a consequence of the viewing angle parallax. In contrast to the other cloud detection algorithms, the Fmask produces only one cloud and cloud shadow category at a horizontal resolution of 20 m.

Examples of multi-temporal cloud masks include the MACCS-ATCOR³ Joint Algorithm (MAJA) of Hagolle et al. (2017) in which clouds are detected based on their spectral differences from the last cloud-free composite image and the Multi-temporal mask (Tmask) of Zhu and Woodcock (2014), which compares

³MACCS, the Multi-sensor Atmospheric Correction and Cloud screening algorithm Center d'Etudes Spatiales de la Biosphere (CESBIO) and ATCOR, the Atmospheric and Topographic Correction software from German Aerospace Center (DLR)

all pixels classified as cloud-free to a modeled ground reflectance based on cloud detection by the Fmask and identifies deviations as cloud or cloud shadow. Since the campaign period was limited to one month, only a few overlaps of the satellite path are available for each observation point. Furthermore, the radiative properties of the ocean surface differ due to the presence of sun glint in some scenes. As a consequence, only single-scene cloud masks are considered below.

The presented cloud detection algorithms were developed to successfully perform scene classification over numerous surface types with different reflectance properties. However, a globally performing detection algorithm is not necessary at this time, since the treatment of clouds in this work is limited exclusively to selected cumulus scenes over ocean surfaces. For this type of scenes, the variability of background radiative properties is small, and depends on the ocean wave slope distribution, Aerosol optical depth (AOD), and how the viewing angle is aligned to the scattering angle of the incident solar radiation (Mieslinger et al., 2021). Therefore, a cloud detection algorithm introduced by Werner et al. (2016) is described below. It was developed to study effects of sensor resolution on remote sensing products of marine boundary layer clouds using Terra-ASTER reflectance datasets.

4.1.2. Description of the algorithm

The algorithm uses a hybrid approach based on a four-test decision tree with fixed thresholds and requires a selection of scenes with presence of low-level water clouds and the absence of high-level cirrus clouds. According to the Moderate Resolution Imaging Spectrometer (MODIS) cloud masking scheme introduced by Platnick et al. (2003) pixels will be classified as confidently cloudy, probably cloudy, probably clear or confidently clear. A binary decision between clear and cloudy pixels classifies confidently cloudy and probably cloudy pixels as a cloud, confidently clear and probably clear pixels as clear areas, respectively. The four decision tests and the decision tree of the cloudmask is shown in Figure 2:

- i) Near-infrared (NIR) reflectances $R_{0.83}$ (band 8) have to exceed a threshold of 0.065 to become flagged as confidently cloudy. $R_{0.83} < 0.03$ are decided as confidently clear.
- ii) Pixels with Short-wave-infrared (SWIR) band 12 reflectances $R_{2.19} > 0.02$ are kept as confidently cloudy, with $R_{2.19} > 0.015$ as probably cloudy and with $R_{2.19} > 0.01$ as probably clear. This decision separates the absorbing ocean surface from bright clouds.
- iii) A band 8 and band 4 reflectance ratio $r_1 = \frac{R_{0.83}}{R_{0.66}}$ tests the constant spectral behavior of clouds in the visible and NIR range. Therefore, pixels with $0.80 < r_1 < 1.75$ keep the confidently cloudy flag, $r_1 < 0.70$ leads to a confidently clear flag.
- iv) Another band 3 and band 4 reflectance ratio $r_2 = \frac{R_{0.56}}{R_{0.66}}$ is based on a strong decrease in Rayleigh scattering with increasing wavelength to distinguish thin cumulus clouds from ocean surface. Pixels $r_2 < 1.20$ keep their classification as confidently cloudy, $r_2 < 1.35$ as probably cloudy and $r_2 < 1.45$ as probably clear.

Adapting the detection algorithm to Sentinel-2 observations can be done, because the central wavelengths and bandwidths of the bands used in the ASTER sensor differ only slightly from the spectral ranges in the Sentinel-2 MSI. However, even small changes in the spectral response function used may require adjustments to the limits, as discussed in the following section. In addition, the original cloud mask includes a fifth test based on radiances in the thermal infrared, which is applied to correct false cloudy pixels under strong sun glint or due to increased horizontal photon transport. Since the Sentinel-2 MSI

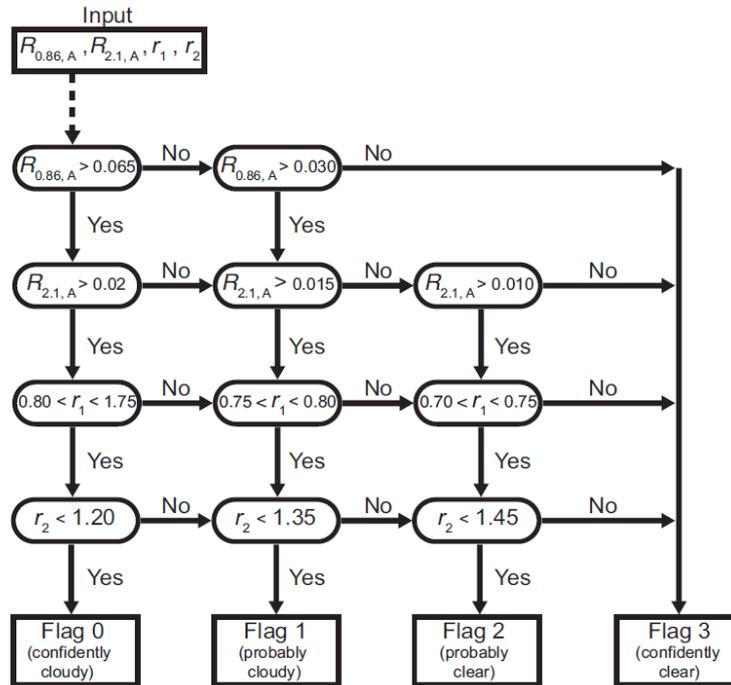


Figure 2: Outline of the cloud detection algorithm by Werner et al. (2016). It illustrates the decision tree including the four cloudiness tests based on the reflectances $R_{0.83} = R_{0.86,A}$ and $R_{2.19} = R_{2.1,A}$ as well as the ratios r_1 and r_2 . Note that the ASTER band wavelengths R_A slightly differ to the Sentinel-2 wavelengths shown in the description.

does not include a thermal channel, this step is not performed and a selection of scenes by the absence of strong sunglint is necessary.

4.1.3. Threshold adjustment

As it is shown in Yang and Di Girolamo (2008), thresholds for distinguishing between clear and cloudy pixels can be derived from the Probability density function (PDF) of the observed quantities. In an ideal situation, a gap between a maximum of clear pixels and a maximum of cloudy pixels would define the threshold. However, overlapping frequency distributions are observed as a result of different optical cloud thickness, 3D radiation effects in broken cloud fields with different cloud top heights, as well as due to different ground reflectances (Wielicki and Welch, 1986; Koren and Joseph, 2000). Clear pixels brightened by scattering can therefore have higher reflectances than shaded or optical thin cloud edge pixels. Since the selection of a perfect threshold is no longer possible due to this overlap in the PDFs, the objective of the analysis must be considered to improve the accuracy of the threshold-based detection. Werner et al. (2016) targeted a cloud-conservative selection in order to subsequently develop successful retrievals of cloud properties. A cloud-conservative detection algorithm sets the thresholds, to avoid that clear pixels are misclassified as cloudy. In contrast, some clear-conservative detection algorithms are also included in the analysis, which set the thresholds, to avoid that cloudy pixels are misclassified as clear (Yang and Di Girolamo, 2008). Since the objective is to derive cloud fraction and cloud base height from the detected cloudy pixels, the adjustment of the thresholds based on visual inspection aims at reducing errors in the determination of the cloud fraction, but also avoid misclassification of clear pixels as cloudy.

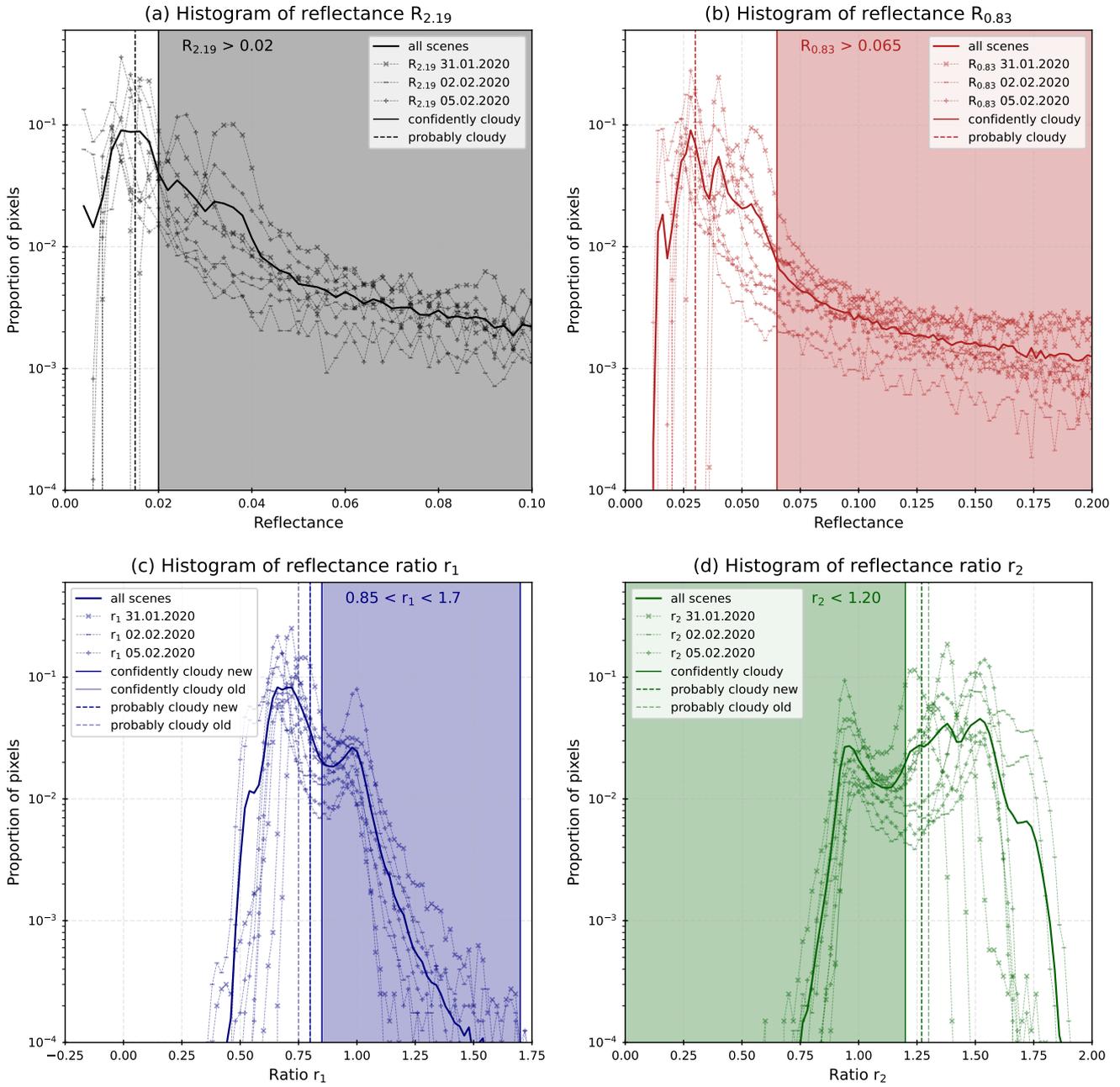


Figure 3: Probability density functions (PDF) of the Werner et al. (2016) cloud detection threshold quantities in $2 \times 2 \text{ km}^2$ single-cloud subsets (thin lines) and PDF of all subsets (thick line). The vertical lines visualize the original (thin) and adjusted (thick) thresholds.

Thresholds	confidently cloudy		probably cloudy	
	original	adjusted	original	adjusted
$R_{2.16}$	0.02	0.02	0.15	0.15
$R_{0.83}$	0.065	0.065	0.030	0.040
ratio r_1	0.80 - 1.70	0.85 - 1.70	0.75 - 0.8	0.8 - 0.85
ratio r_2	< 1.20	< 1.15	< 1.35	< 1.25

Table 3: Selected thresholds used in the four-step decision tree to classify a pixel as confidently cloudy respectively probably cloudy. Comparison between original thresholds in Werner et al. (2016) and adjusted thresholds in this work.

Figure 3(a) and (b) show the PDF of the SWIR and the NIR reflectance of 9 manually chosen scene subsets with a size of $2 \times 2 \text{ km}^2$. All subsets contain at minimum one shallow cumulus cloud. The cloud shadow area and the surrounding clear ocean surface are displayed in Figure 13 in the appendix of this thesis. The histograms show a continuous PDF with at least one broader maximum. While in 5 scenes this maximum can be observed in the range $0.15 \leq R_{2.19} \leq 0.2$ and $0.25 \leq R_{0.83} \leq 0.3$, respectively, two scenes from 31-Jan-2020 and two scenes from 05-Feb-2020 indicate a broader maximum in the range $R_{2.19} > 0.2$ respectively $R_{0.83} > 0.35$. The reflectance value of the clear ocean surface can be derived from these maxima, whereas the value is higher in the latter scenes due to the stronger reflection of the incident solar radiation in the viewing direction. Most scenes show a further, narrow maximum in each case immediately above the smallest reflectance value. This results from the contribution of the shaded surface pixels to the frequency distribution. The reflectance of the cloud shadows in two scenes on 31-Jan-2020 is thus in the reflectance range of the sunlit ocean surface in 5 scenes.

As a consequence of the different background reflectances, the PDF calculated from all scenes has several local maxima, which makes it difficult to select a global threshold from the reflectances. Therefore, in some previous analyses of high-resolution satellite imagery, the thresholds were determined manually for individual scenes (Zhao and Di Girolamo, 2007; Koren et al., 2008). Since this is time-consuming for larger statistics and, in addition, the Sentinel-2 MSI sensor geometry results in different surface reflectances in one dataset, a scene-dependent threshold selection is not applied in this work. Only the limit for probably cloudy pixels in band 8 has been increased to a threshold of 0.4 to avoid misclassification of the ocean surface as probably cloudy in some scenes.

In contrast to the single-band reflectances, the PDFs of the reflectance ratios $r_1 = \frac{R_{0.83}}{R_{0.66}}$ and $r_2 = \frac{R_{0.56}}{R_{0.66}}$ indicate similar progression with two distinct maxima for all scenes. This is mainly due to the strong decrease of rayleigh scattering and increase of absorption of water bodies with increasing wavelength in the VNIR range. Therefore, clear oceans pixels show values of 0.55-0.75 for r_1 , and values of 1.25-1.6 for r_2 , while the homogenous scattering of cloud droplets in all VNIR wavelengths leads to values around 1 for both ratios. Based on the distributions, a slight adjustment of the thresholds of both ratios in the direction of the frequency maximum of the cloudy pixels was possible. This should reduce the misclassification between bright surface pixels and dark cloud pixels even without a further test in the terrestrial wavelength range. The adjusted thresholds in comparison with the original cloud detection thresholds are displayed in table 3.

4.2. Shadow detection algorithm

The calculation of cloud height based on the geometric relationship between cloud, cloud shadow and the satellite detector requires the successful differentiation of ground pixels into shaded and those that are not shaded by a cloud. Since the operational Sen2Cor processor is not able to detect cloud shadows over water bodies and classifies shadows as water pixels, a shadow detection algorithm must be adapted for the characteristics of the Sentinel-2 MSI sensor (Richter et al., 2011).

4.2.1. Description of the cloud shadow detection index (CSDI)

There have been few publications on the identification of cloud shadows over water surfaces in medium- to high-resolution satellite imagery. The MODIS cloud detection algorithm includes a cloud shadow test for clear pixels based on two reflectance thresholds and a reflectance ratio test in the VNIR wavelength range (Ackermann et al., 2006). The Fmask algorithm determines potential cloud shadow layers over various surfaces using the intensity differences between shaded and surrounding sunlit areas in a single NIR band (Zhu and Woodcock, 2012). A similar approach was taken by Amin et al. (2012) for satellite imagery from the Hyperspectral Imager for the coastal ocean (HICO) aboard the International Space Station (ISS). Reflectances from numerous channels in the blue-green region of the visible spectrum were amplified by integration over all wavelengths and shadows were subsequently separated using a contrast threshold. Since reflectances for shaded pixels over water surfaces differ between scenes depending on atmospheric conditions and observation geometry, the cloud shadow detection based on a fixed reflectance threshold is not possible. In addition, shadow detection was developed by Amin et al. (2012) specifically for fast and direct identification of cloud shadows over water surfaces. Therefore, this algorithm is used in the following.

In contrast to the HICO, only 3 high-resolution spectral bands are in the wavelength range between 450 nm and 670 nm. Thus, for the amplification of the contrast between shaded and adjacent sunlit regions, it is not possible to determine an integrated brightness value over all wavelengths, but only a mean reflectance value (MV):

$$MV = \frac{1}{3} \cdot (R_{0.49} + R_{0.56} + R_{0.65}) \quad (12)$$

with the reflectance values $R_{0.49}$ of band 2, $R_{0.56}$ of band 3 and $R_{0.65}$ of band 4. To separate shaded pixels by using a constant threshold, the MV of a pixel MV_i is normalized by the spatial mean of the MV of all clear pixels within a Spatial adaptive sliding box (ASB), called the cloud shadow detection index (CSDI):

$$CSDI = \frac{MV_i}{\overline{MV_{ASB}}} \quad (13)$$

The calculation of the spatial mean $\overline{MV_{ASB}}$ requires the removal of all pixels identified as probably cloudy or confidently cloudy by the cloud mask. Since the ASB should contain both shaded and sunlit pixels or only sunlit pixels, a size of $5 \times 5 \text{ km}^2$ (500×500 pixels at 10 m resolution) was chosen. Therefore, shadows are not detected within 2.5 km from the edges of the data sets.

4.2.2. Definition of a CSDI threshold

Figure 4(a) shows the PDF of the calculated CSDI for all subsets. For the majority of pixels in the subsets, a CSDI value between 0.95 and 1.1 was calculated. Another, smaller frequency maximum is shown by

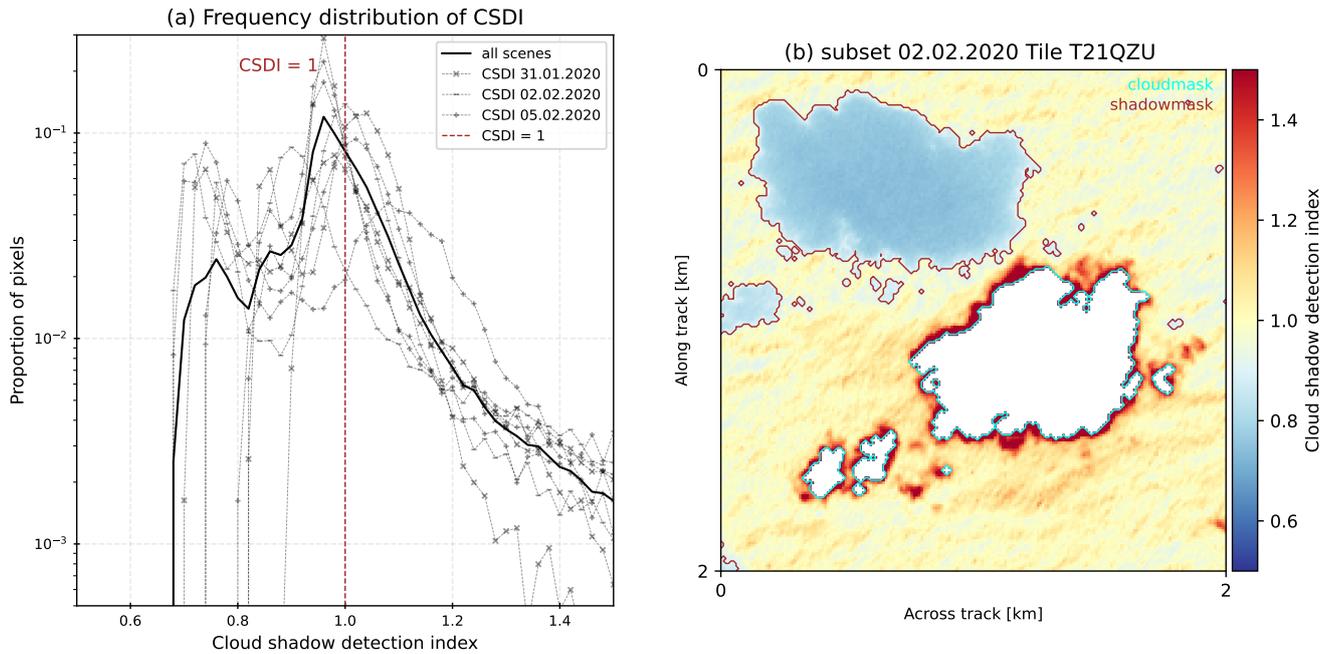


Figure 4: (a) PDFs of calculated cloud shadow detection index (CSDI) in $2 \times 2 \text{ km}^2$ single-cloud subsets (thin lines) and PDF of all subsets (thick line), (b) Example subset (02-Feb-2020, tile T21QZU) with calculated CSDI and scene classification.

almost all scenes at $0.7 \leq \text{CSDI} \leq 0.85$. The location of the second, shadow induced maximum depends on the proportion of shaded pixels and the contrast between pure diffuse and reflected plus diffuse radiation.

Contrary to the original idea of contrast-based detection, not all sunlit areas show a $\text{CSDI} \geq 1$. This may be caused by two reasons: First, as can be seen in the Figure 4(b) example, the ocean wave slopes leads to inhomogeneous reflectance properties of the sunlit ocean surface, violating the condition of homogeneous surface properties within the ASB. The heterogeneity is not preventable for detector resolutions that are finer than the mean wavelength of the ocean waves. Second, not all cloudy pixels at the edge of detected clouds are included in the cloud mask, and thus large reflectance values are included in the calculation of the mean reflectance of clear pixels. A buffer around detected cloud objects for shadow analysis could prevent this misinterpretation.

To avoid false detection of sunlit ocean pixels as shaded, a fixed CSDI threshold as in the original shadow detection algorithm does not lead to a successful shadow detection in all scenes. The threshold for classifying a pixel as a shadow pixel is therefore adjusted individually for each tile. For this purpose, the PDF of the CSDI is calculated in the detection algorithm and the threshold is set to a CSDI of 0.04 below the maximum of the PDF. All pixels whose CSDI is below this threshold but greater than 0 are classified here as cloud shadow pixels.

4.2.3. Evaluation of the adjusted scene classification

Both cloud and shadow detection algorithms (further: scene classification) were computed for three $10 \times 10 \text{ km}^2$ sized trade cumulus scenes to compare the accuracy with the Sentinel-2 Fmask classification by Zhu et al. (2015) and scene classification from the operational Sen2Cor processor. Scenes were visually selected considering all observation days used in the analysis, different optical thicknesses and cloud sizes, and different reflectances of the sunlit pixels. As evident in the RGBs of the scenes shown in Figure 5, the

left scene from 31-Jan-2020 demonstrate flat and thin cumulus clouds over typical background reflectance. In contrast, the two scenes from 02-Feb-2020 and 05-Feb-2020 differ primarily in their cloud fraction and the reflectances of the sunlit ocean surface. The right scene was additionally selected from the eastern edge of the observation path, and thus the observation was made using a relatively large detector zenith angle.

In the figures below, the scene classification forms a layer superimposed on the gray values of a reflectance averaged from all channels (band 12, band 8, band 4, and band 3) included in the computation. The reflectance was limited to a maximum of 0.1, especially to analyze the exact location of cloud edges and shaded cloud areas. Pixels identified as confidently cloudy (Sen2Cor: cloud, high probability) were masked white, pixels identified as probably cloudy (Sen2Cor: cloud, medium probability) were masked yellow. Since the Fmask has only one cloud class, all pixels identified as cloud are shown with white color. Clear sunlit pixels are shown in blue, while cloud shadows on the ocean surface are indicated in brown.

In a visual evaluation with the RGB or mean reflectance, the adjusted scene classification shows good performance. Bright pixels within optically thick cloud cover were identified as confidently cloudy, while most of the less bright pixels at the cloud edge were identified as probably cloudy. Misinterpretation of cloudy pixels as clear is evident in shaded areas of deep cumulus clouds, especially in the middle scene. Combining adjacent cloud pixels into a cloud object should reduce this misinterpretation. In contrast, the detection of optically very thin and small-scale clouds succeeds well. In cases with large background reflectances, sporadic misinterpretations of clear pixels as cloudy occurs in the edge region of clouds influenced by the horizontal photon transport.

In comparison, the results of the Fmask algorithm are similar to those. Since the Fmask is calculated at a lower resolution and clouds are already merged into cloud objects during the analysis, the number of pixels interpreted as clear within cloudy areas is smaller. However, the classification of shaded cloud areas on 05-Feb-2020 succeeds better with the adjusted scene classification. The calculated cloud fraction in the scenes differ only slightly between the two detection algorithms, and are slightly higher in the Fmask.

In contrast, a comparison of the Sen2Cor scene classification with the mean reflectance reveals a misclassification of optically thin and shaded cloud regions as clear pixels. In the scene of 31.01. this leads to a 51% lower cloud fraction than calculated by the adjusted scene classification. This may be caused by two reasons: First, water bodies are excluded using a ratio test between a VIS channel in the blue wavelength range and a SWIR channel with a linear decreasing cloud probability between thresholds of 2 and 4 (Richter et al., 2011). Shaded cloud areas show ratio values larger than 2, due to a stronger contribution from Rayleigh scattering compared to illuminated cloud areas. Secondly the reflectance test based on band 4 ($R_{0.65}$) uses a high upper threshold, which reduces the cloud probability of optical thin clouds too strong.

Previous analyses of the Fmask and Sen2Cor classification on Sentinel-2 scenes showed higher producers accuracy of the Fmask (78%, 85%) for cloud detection compared to the Sen2Cor (75%, 66%), highlighting the underestimation of cloudy areas as clear pixels in the Sen2Cor mask (Tarrío et al., 2020; Zekoll et al., 2021). The cloud producers accuracy is the ratio of pixels, which are classified as cloud by the algorithm compared to the pixels, that are subjectively visible as cloudy pixels. However, these evaluations were performed over land surfaces and with a default buffer radius of 60 m in the Fmask, which also basically assigned cloud edges to the cloud object. For the selected ocean scenes, there are larger differences between the two scene classifications.

A visual evaluation of cloud shadow detection accuracy in the adjusted scene classification also show good results. In particular, larger cloud shadows of optically thick cumulus clouds are detected precisely,

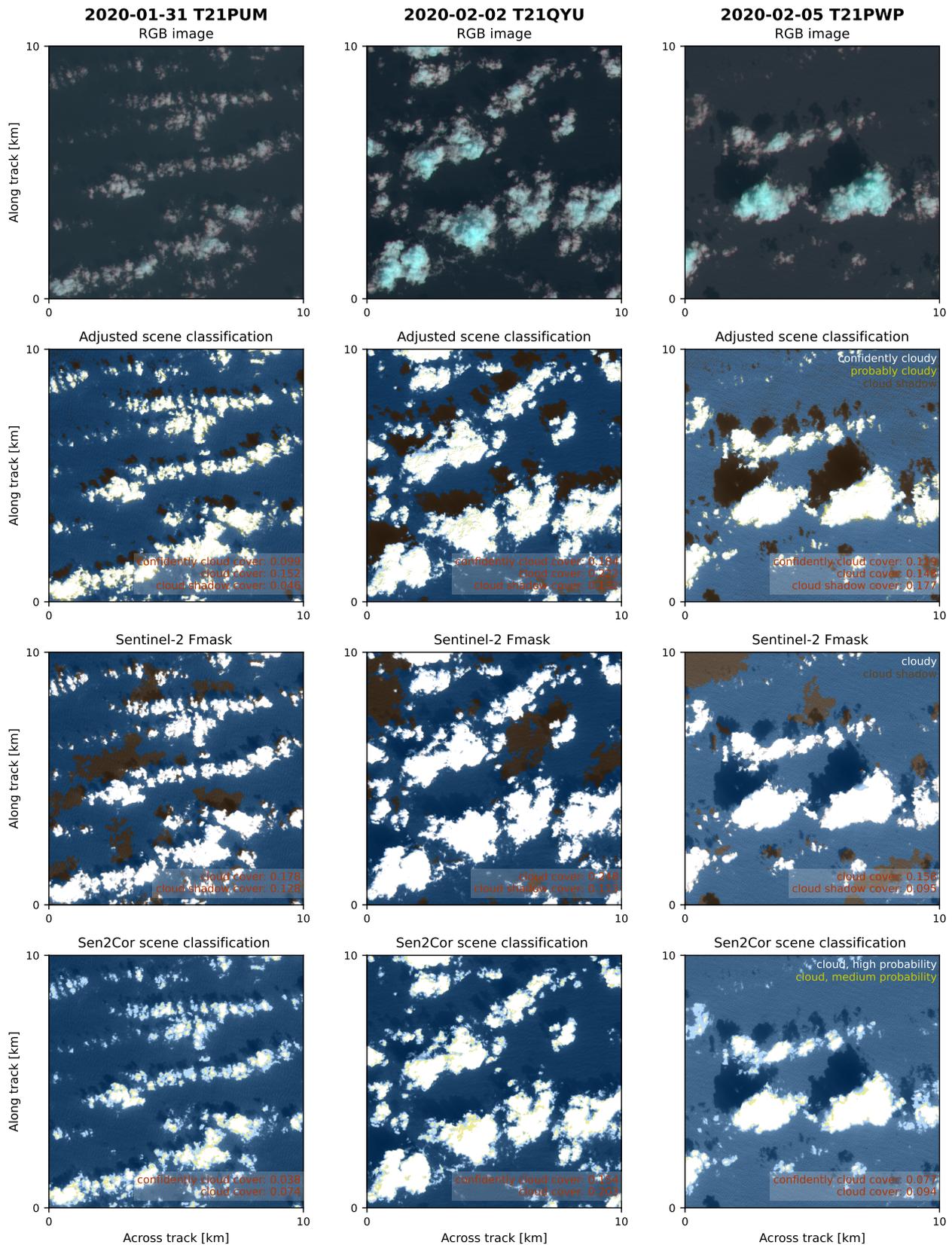


Figure 5: Comparison of the adjusted cloud and shadow detection (scene classification), the Sentinel 2 Fmask and the operational Sen2Cor scene classification for three $10 \times 10 \text{ km}^2$ trade cumulus scenes. The RGB image of the scene is shown above, the images below show a mean reflectance overlaid by the scene classification of each algorithm (white: confidently cloudy, yellow: probably cloudy, blue: clear pixel, brown: cloud shadow). The calculated scene fractions of clouds and shadows are shown in the bottom right.

but the method derived directly from the reflectance contrast allows also the detection of very small cloud shadows, as can be seen in the scene of 05-Feb-2020. A false detection of cloud shadows as clear pixels can occur with optically thinner clouds, whose shadow areas are still partially reached by direct solar radiation. Furthermore, some misclassifications of ocean wave shadows as cloud shadows are visible in the scene of 05-Feb-2020. Previous analyses also indicate a low overall accuracy of 50% for the detection of cloud shadows using the Fmask algorithm.

In summary, successful detection of clouds and cloud shadows over ocean surfaces is possible with the adjusted scene classification and it can feature similar or even better results compared to established algorithms. Considering the selection of scenes according to the presence of low cloud cover and the absence of high cirrus, this method can be used for further analysis and the cloud base height estimation.

4.3. Geometrical cloud base height estimation

In the following, a method for estimating CBH based on the geometric relationship between detected cloud and shadow objects is presented. The method is compared with shipbased ceilometer measurements of the first cloud base height (CBH₁) on the 10-Feb-2020.

4.3.1. Background of cloud height estimation using passive satellite remote sensing

The derivation of cloud heights by means of passive satellite remote sensing is difficult, since, in contrast to active remote sensing, there is no information about the propagation time of radiation between the scattering object and the sensor. For this reason, several approaches have been developed to determine both CBH and cloud top height of low clouds from measurements of medium- to high-resolution satellite sensors.

Cloud top height estimations of shallow cumulus clouds are often performed using the brightness temperature in the infrared spectral region. This is done by converting local minima of brightness temperature in completely cloudy pixels to a cloud top height using local atmospheric profiles (e.g. Zhao and Di Girolamo, 2007) or zonally averaged, "apparent lapse rates" of brightness temperature (e.g. Baum et al., 2012; Mieslinger et al., 2019). Other approaches to derive cloud top height are based on stereoscopic techniques, that analyze disparities caused by cloud height and wind drift in detected cloud features from at least two observation angles (e.g. Hasler, 1981; Seiz et al., 2006).

The estimation of CBH using satellite remote sensing received less attention, what makes it interesting to focus on this cloud property. Böhm et al. (2019) showed that the estimation of CBH of convective clouds with homogeneous condensation level is possible via the spatial minimum value of cloud top height derived by the stereoscopic method. However, this method already requires a successful cloud top height product. Berendes et al. (1992) calculated CBH in high-resolution Landsat-MSS shallow cumulus scenes from the geometric relationship between the edge of a cloud object and an associated cloud shadow object, as well as the solar and observation angles. The pairing of two objects is performed using a Generalized Hough Transformation and the assumption that the cloud height is constant within 100 m around the edges of the clouds. Deriving the height from the cloud-shadow geometry is also used in some object-based cloud detection algorithms such as Fmask by Zhu and Woodcock (2012) or a cloud mask for Chinese medium-high resolution satellite sensors by Zhong et al. (2017) to predict the shadow position.

Since Sentinel-2 MSI does not have a channel in the infrared spectral range, an estimate using the brightness temperature is not possible. A stereoscopic method would be conceivable due to the detector arrangement, but requires knowledge of the wind speed. Therefore, an estimation based on geometry

between cloud and shadow objects is developed in the following:

4.3.2. Cloud base height retrieval algorithm

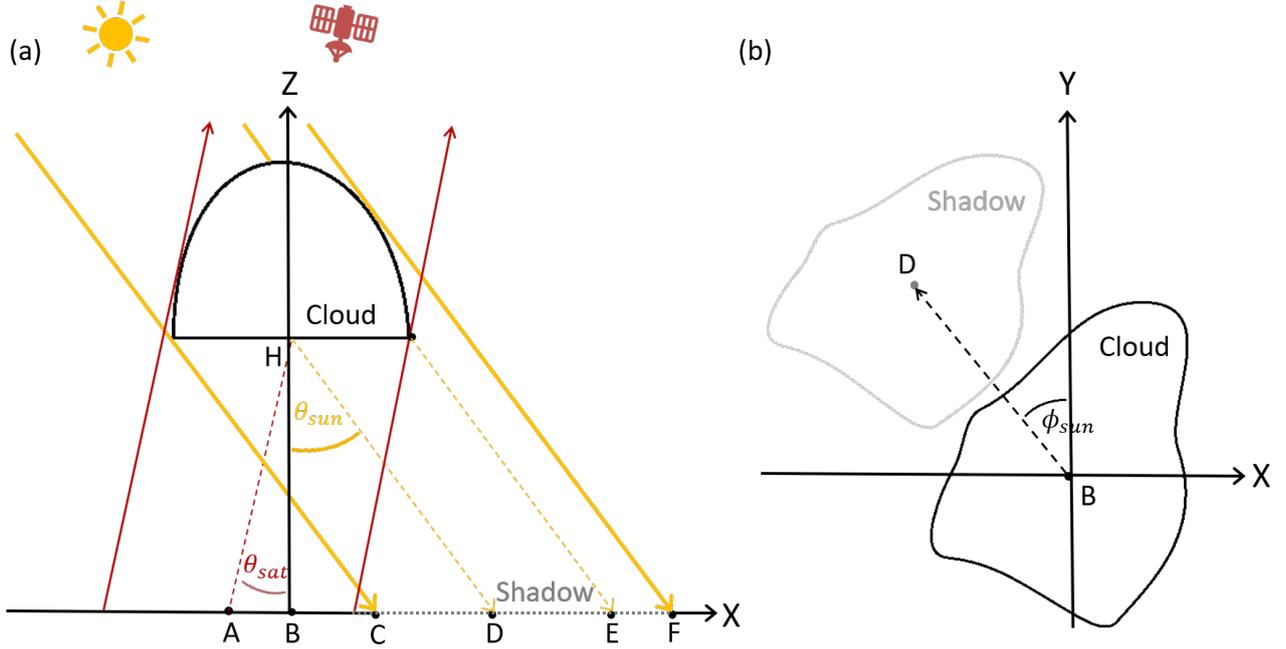


Figure 6: Schematic representation of the CBH estimation with reference points (A-H) in the X-Z cross section (a) and X-Y base section (b).

Figure 6 displays a schematic drawing of the CBH geometry and its estimation based on a satellite image as well as the knowledge of sun- and satellite geometry. The basic idea of the approach is that the height of a cloud object can be determined from the horizontal distance $d = \overline{BD}$ between the center of the cloud object B and the center of an associated shadow object D . The geometry is first considered in a vertical cross-section. Assuming that a satellite sensor would look to nadir, the CBH $h_{base} = \overline{BH}$ could be calculated with the solar zenith angle θ_{sun} as

$$h_{base} = \frac{d}{\tan(\theta_{sun})}. \quad (14)$$

Since in the edge region of the observation path the observation angles θ_{sat} deviate from the zenith by up to 12° , this leads to a displacement of the elevated cloud object by the distance $d'_{sat} = \overline{AB}$:

$$d'_{sat} = h_{base} \cdot \tan(\theta_{sat}). \quad (15)$$

It must be taken into account in the cloud height calculation that the distance between the observed center A of the cloud and the center shadow D differs from the real distance as $d_{obs} = \overline{AD} = d - d'_{sat}$, and the height estimation changes to

$$h_{base} = \frac{d_{obs}}{\tan(\theta_{sun}) - \tan(\theta_{sat})}. \quad (16)$$

In the three-dimensional perspective, the azimuth angle of the sun ϕ_{sun} as well as the satellite sensors ϕ_{sat} must be considered additionally. The latter results from the staggered arrangement of the sensors.

Therefore the components d_x and d_y of the real displacement vector $d_x = d \cdot \sin(\phi_{sun})$, $d_y = d \cdot \cos(\phi_{sun})$ between the cloud object and the shadow object can be calculated as

$$d_x = h_{base} \cdot [\tan(\theta_{sun}) \cdot \sin(180^\circ - \phi_{sun}) + \tan(\theta_{sat}) \cdot \sin(180^\circ - \phi_{sat})] \quad (17)$$

$$d_y = h_{base} \cdot [\tan(\theta_{sun}) \cdot \cos(180^\circ - \phi_{sun}) + \tan(\theta_{sat}) \cdot \cos(180^\circ - \phi_{sat})]. \quad (18)$$

The estimation of CBH is based on the assumption of a flat cloud. As can be seen in the schematic drawing in Figure 6, the vertical extension of cumulus clouds can lead not only to shading of cloud areas at larger solar zenith angles, but also to a surface shadow enlarged by the distance $\Delta d = \overline{EF}$. Berendes et al. (1992) estimated the resulting error in the cloud height calculation Δh assuming a semicircular cloud with radius r to be

$$\Delta h_{base} = \frac{r \cdot (1 - \cos(\theta_{sun}))}{\sin(\theta_{sun})}. \quad (19)$$

However, trade wind cumulus shows various shapes, which deviate strongly from a semicircular assumption. Since a vertical extension can also lead to different shapes of cloud and its shadow, the matching method used here deviates from the Generalized Hough Transformation in Berendes et al. (1992). Instead the selection method by Zhu and Woodcock (2012) based on the area overlap between cloud and shadow object is chosen in the following.

The cloud-shadow matching is based on scene classification derived by the cloud and shadow masking algorithms. Both pixels labeled as confidently cloudy or probably cloudy are considered in the binary decision between clear and cloudy pixels. The required solar and observation angles are provided in the datasets with a resolution of $5 \times 5 \text{ km}^2$. To achieve higher accuracy in the calculation of the vector between cloud and shadow, the angles are interpolated to the spatial sensor resolution of 10 m. The procedure of the object matching is as follows:

- (1) The scene classification is subjected to a binary morphology operation of opening and closing to reduce objects and holes caused by noise.
- (2) Pixels adjacent to other pixels of the same classification (using 8-way connectedness) are considered as one cloud object respectively shadow object.
- (3) All objects having an area less than 100 pixels ($10,000 \text{ m}^2$) are removed to avoid misdetections near the edges of clouds and in shaded regions of ocean waves.
- (4) For each cloud object, the azimuth angle of the motion vector is calculated from the solar and observation angles at the center of the object. The cloud object is shifted along the vector calculated from equations 17 and 18 in a predetermined distance interval. Since the scenes are selected by the presence of low cumulus clouds, the distance is limited to 200 pixels (corresponding to a height of about 2800 m-3000 m). For each moving step, the overlapping area with shadow objects is calculated and normalized with the area of the cloud object (match similarity).
- (5) If the match similarity between shifted cloud object and a shadow object reaches a maximum with a value $\frac{A_{overlap}}{A_{cloud}} \geq 0.3$, the CBH is calculated from the moving distance at the point of maximum similarity using equation 14. The threshold value was adopted from Zhu and Woodcock (2012) to reduce mismatches. To avoid misinterpretation of smaller fluctuations as maxima, the match

similarity functions are smoothed using a Gaussian filter. If multiple maxima are detected within the moving distance, the CBH is calculated from the first maximum, that exceeds the threshold.

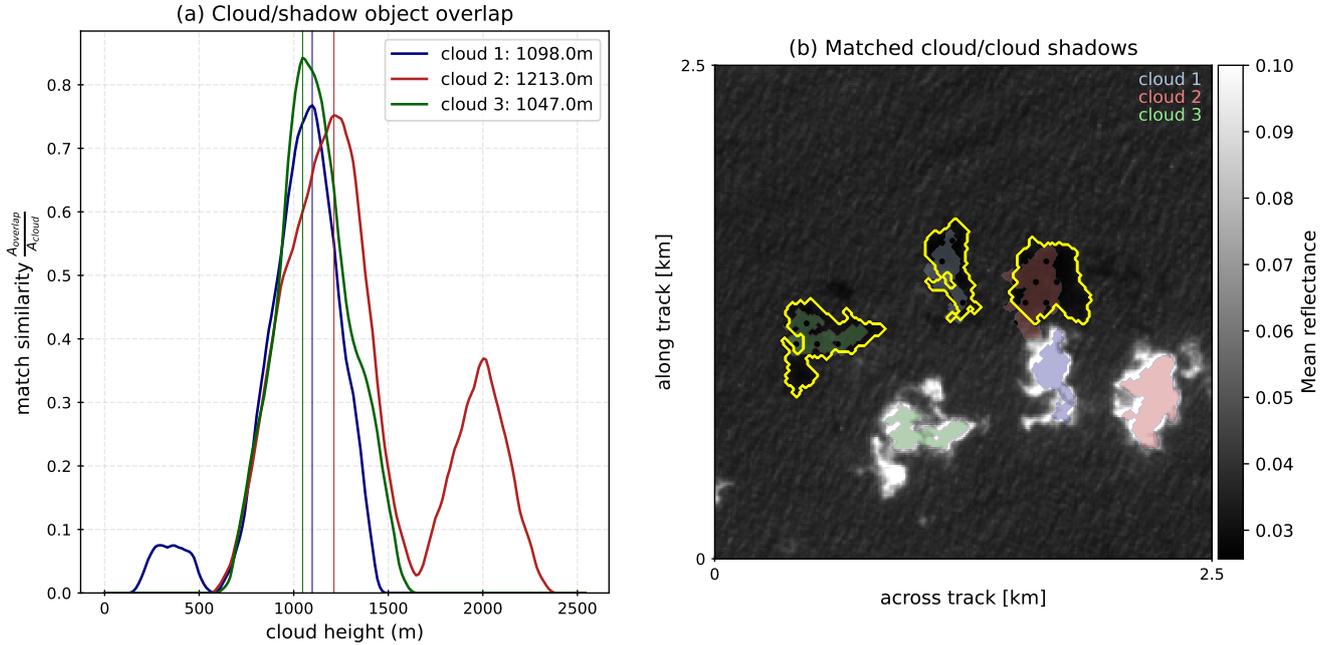


Figure 7: Cloud-shadow matching with 3 example cloud objects, (a) area overlap between the moved cloud objects and shadow objects as a function of CBH. The maximum is indicated by a vertical line. (b) Mean reflectance, cloud objects (filled contours), shadow objects (yellow lines) and moved cloud objects at maximum overlap (filled contours with dots).

Figure 7 illustrates the CBH calculation based on an example scene containing 3 shallow cumulus clouds. The detected cloud objects (filled contours) could be successfully matched to the identified shadow objects (yellow contours), and show as maximum in the match similarity function at a CBH between 1047m and 1200 m. However, despite interpolation of detector and sun angles, the matched cloud objects (filled contours with dots) in this scene show a slight shift to the shadow objects. This could be a consequence of averaging the viewing angles over all bands involved in the scene classification. Also noticeable are the significant differences in the shape of associated objects, which would make a shape-based assignment without smoothing difficult, especially for complex object structures.

4.3.3. Comparison with ground-based ceilometer data

In the following, the accuracy of CBH estimation will be evaluated by comparing it to established remote sensing methods. As shown in section , only a small selection of scenes with overlap between observation path and ship measurements in the campaign period is available. The comparison is therefore limited to 2 trade cumulus scenes recorded on 10-Feb-2020 in the area 13.9°N 57.2°W (tile T21PVR) and 10.5°N, 58.9°W (tile T21PTM), respectively. Considering a cloud scene as homogeneous as possible, and a sufficient number of detected cloud objects, a scene size of 20x20 km² was selected. The two scenes differ in cloud fraction of 6.6% and 11.4%, respectively, but have a similar number of approximately 175 detected cloud objects of at least 10.000 m² in size.

For comparison, the CBH₁ from ceilometer measurements on board the research vessels Meteor and

L'Atalante is used. The distance traveled by the ships within the hour of the overflight (14UTC - 15UTC, yellow line) and the coordinates at the time of the scene recording (14:29UTC, yellow dot) are shown in Figure 8(a). Within the observation interval, 26 cloud base heights in the upper scene and 48 cloud base heights in the lower scene were recorded by the ceilometer.

The estimated CBHs are shown as color overlays in Figure 5(a) and as frequency distributions for both scenes and measurement methods in Figure 5(b). The height of most of the detected cloud objects in tile T21PVR was determined to be between 1000 m and 1300 m using the presented matching method, with a slightly increasing trend from north to south within the section. However, some cloud objects were assigned unrealistic large CBH up to 2200 m, since no higher-level clouds can be determined by visual observation. These objects are often located in the direction of the sun's azimuth angle to neighboring cloud objects. This causes that a part of the cloud shadow is obscured by the neighboring cloud. As a result, the match similarity reaches the threshold firstly at the next cloud shadow of the neighboring cloud. This results in a too great matching distance. For larger cloud objects whose shadow is obscured, the detected shadow area is too small in proportion and the match similarity remains below the threshold of 0.3. Therefore, no cloud height is assigned to these clouds.

In the lower tile T21PUM, these connections between cloud and incorrectly assigned shadow objects occur somewhat more frequently as a result of the higher cloud fraction. In addition, there are more smaller cloud objects in the edge area of larger clouds, whose CBH calculation was not carried out because the object size was too small. Assuming a homogeneous CBH within a small image section, this exclusion of small objects, however, enables a faster calculation.

The distribution of CBH calculated by the Sentinel-2 algorithm shown in Figure 8(b) shows a small scattering around the mean value of 1099 m for Tile T21PVR, while in Tile T21PTM the frequency maximum around the mean value of 1173 m is much broader. In both scenes, the CBH_1 detected by the ceilometer are in a very narrow range with a mean value of 757 m and 814 m, which underlines the assumption of a largely homogeneous CBH within the scene. Consequently, too high CBHs are calculated by the algorithm for the majority of the cloud objects. The cause of the overestimation is probably the extension of the shadow area along the displacement vector for vertically extend clouds. As shown in equation 19, this discrepancy can be corrected for the distance of the sunward edges of the cloud and shadow object assuming a semicircular cloud, or used to estimate the cloud top height using known CBHs. Furthermore, a correction of the overestimation could be possible by simplifying the object geometry to elliptical objects and comparing the eccentricity of both objects. However, both correction methods will be successful only for low cloud fraction and higher zenith angles, and may exacerbate the deviation of the calculated cloud height in case of overlap between cloud and shadow objects. Successful base height estimation at higher occultation levels requires the introduction of new estimation methods. Potential could be offered by the height-dependent shift of the object in the horizontal when viewed with two different detectors, as used by Frantz et al. (2018) for cloud detection. However, since the wind also leads to a displacement of cloud objects due to the time shift between two observing channels, knowledge of an average wind vector and the introduction of a single-band cloudmask is required for accurate cloud height determination of the cloud edge. Therefore, the evaluation of the cloud height in the context of this work will be performed without further correction and the observed overestimation has to be taken into account.

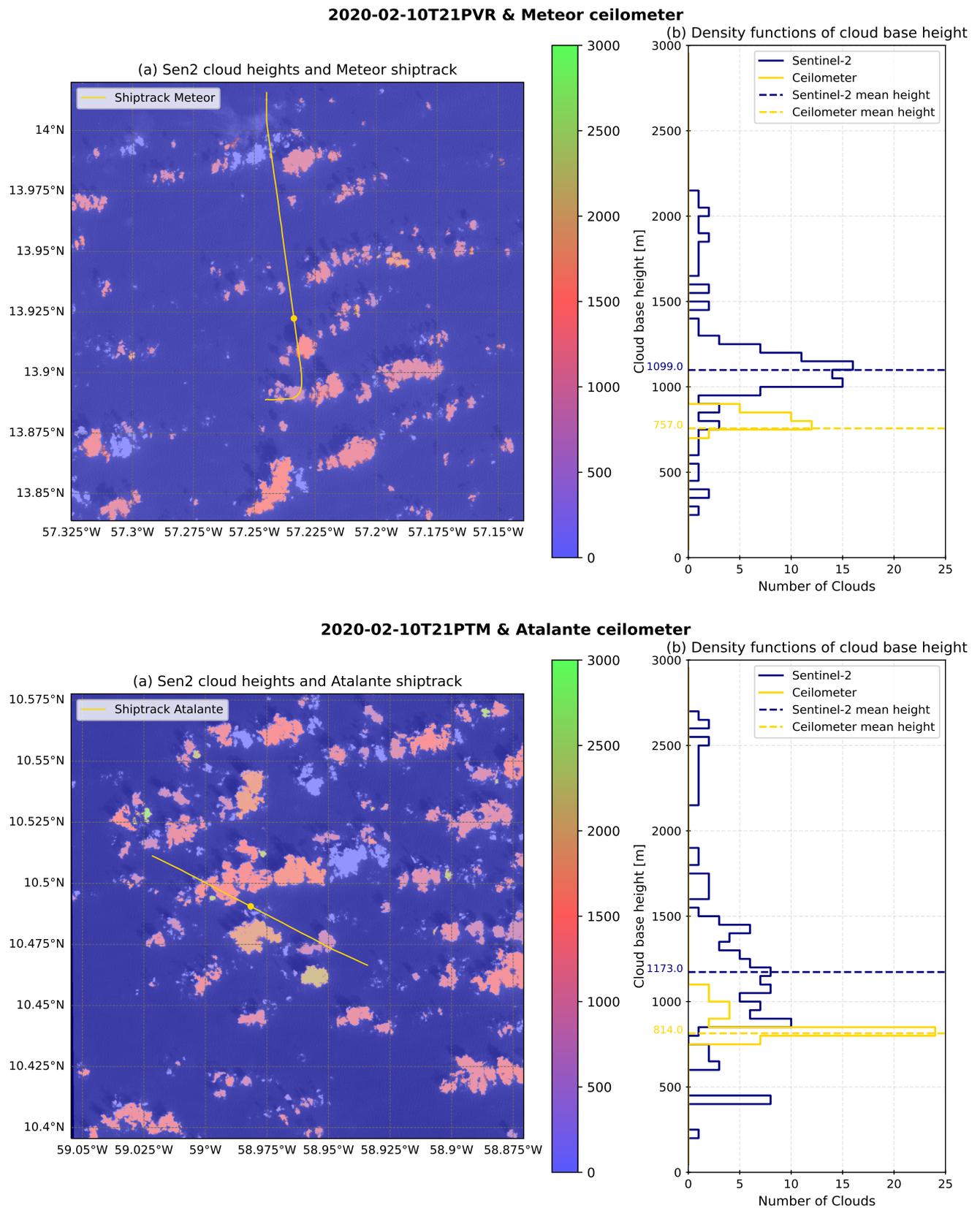


Figure 8: (a) Mean reflectance and estimated CBHs of two 20x20 km² scenes at 10th February 2020 as well as track (yellow line) and position (yellow dot) of the research vessels within the observation hour and time. (b) Distributions of estimated CBH from Ceilometer (CBH₁) and Sentinel-2 measurements.

5. Analysis of Sentinel-2 MSI trade cumulus scenes

In the following, cloud fraction, shadow fraction, and cloud base heights (CBH) calculated from Sentinel-2 MSI observations, as well as cloud size distributions are presented for 9 selected trade cumulus scenes. The scenes were each taken from a cross-section of the 31-Jan-2020, 02-Feb-2020, and 05-Feb-2020 observation paths. In the cross section from 02-Feb-2020, one scene (tile T21QXU) was excluded from the calculation due to the presence of a stratiform cloud layer. The selected sections show very fine to medium sized cumulus clouds, which can be assigned to the organization type "Sugar" according to the classification of mesoscale organization by Stevens et al. (2020) due to its low degree of self-organization. However, in the southern part of the 02-Feb-2020 cross section, a transition to the "Flower" organization type can be observed. The RGB images, scene classification, and calculated cloud heights for all scenes are shown in Figures 14, 15, and 16 in the appendix of the thesis.

5.1. Cloud and shadow fraction

The significance of a cloud cover value depends not only on the chosen determination method, but also on the size of the area for which it is determined. Therefore, in order to show not only the mean cloud fraction, but also the differences within a scene, each tile was divided into 16 subscenes with a size of about $27.5 \times 27.5 \text{ km}^2$ and both the amount of cloud objects and shadow objects in the total area of the subscene were calculated for the highest sensor resolution of 10 m. The size of the subscenes is thus close to the horizontal grid size of the current ECMWF reanalysis dataset ERA5 of 31 km, (e.g. Hersbach et al., 2020).

Figure 9(a) displays the statistics of the calculated cloud fraction and figure 9(b) the statistics of the shadow fraction visible from the satellite for all selected scenes. In each case, the boxplots show the median of the values as a red line, the mean as a red triangle, lower and upper quartiles as boundaries of the box as well as the 05th and 95th percentiles as boundaries of the whiskers. To avoid misinterpretation, the subsets of scenes, which contains the edge of the observation path, were excluded from the fraction calculation. In some tiles, therefore, less than 16 subsets were included in the analysis.

Cloud fraction varies significantly both within a tile and between scenes. The lowest value of 0.01 was determined for a nearly cloud-free subset on 31-Jan-2020 (tile T21PTM), the highest value of 0.29 for a subset on 02-Feb-2020 (tile T21QZU), where two larger cloud clusters could be observed. In addition, for tile T21PTM, which is largely interspersed with optically thin and very small-scale cumulus, the lowest mean fraction value of 0.037 was calculated; the highest mean fractions were found in tiles T21QZU and T21PWP (only 4 subsets at the edge of the observation footprint) due to the presence of some larger cumulus clusters and cloud streets.

For the remaining tiles, cloud fraction was mostly found in a range between 0.04 and 0.11, which also corresponds to the uncertainty range of the lower and upper quartiles of all subsets. The mean cloud fraction in this thesis is 0.083. This value is close to the cloud fraction of 0.086 and 0.087 determined by Zhao and Di Girolamo (2007) for 152 and Mieslinger et al. (2019) for 1158 Terra-ASTER trade cumulus scenes with 15 m resolution, respectively. Furthermore, 70% of the scenes in this analysis shows cloud fractions below 0.1, which is consistent with Zhao and Di Girolamo (2007). This indicates that the nature of the shallow cumulus scenes analyzed in this study corresponds to previous investigations using a different high resolution satellite sensor. It is important to note that the cloud detection algorithm was used to estimate CBH with a cloud conservative detection method. The aforementioned cloud fraction evaluations are also based on cloud conservative estimates and a comparison is thus possible. However,

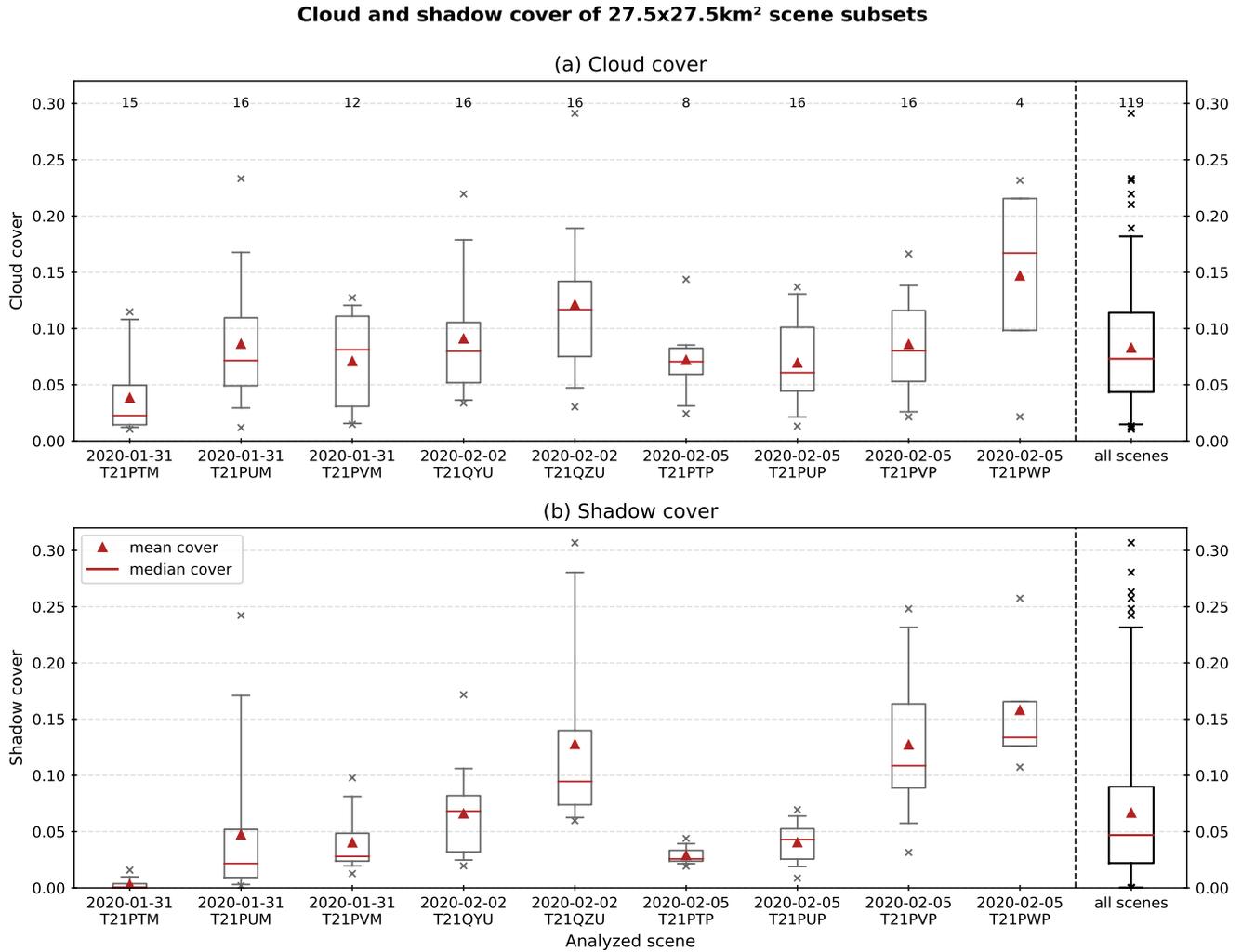


Figure 9: Boxplots of (a) estimated cloud fraction and (b) estimated shadow fraction based on Sentinel-2 measurements for 27.5x27.5 km² subsets of 9 trade cumulus scenes. The mean is indicated by a red line, the boxes shows the 25th and 75th percentile, the whiskers the 5th and 95th percentile.

optically thin cloud parts, especially in the fringes of cumulus clouds, are systematically assigned to the clear category with this method, which leads to considerably lower cloud fractions compared to clear conservative estimates (Zhao and Di Girolamo, 2006; Mieslinger et al., 2021).

The derived cloud fraction differs significantly from climatological mean cloud amounts of low clouds in the region. Ground-based remote sensing measurements using Ceilometer from the Barbados Cloud Observatory (BCO) showed mean cloud fraction of 0.21-0.22 at the height of lifting condensation level with a small seasonality (Nuijens et al., 2014; Stevens et al., 2016). Terra-MODIS datasets showed no 5-day mean cloud fraction below 0.1 (Brueck et al., 2015). Besides the cloud conservative estimation method, one reason for the differences may be the selection of scenes according to the presence of broken cumulus clouds and the absence of stratocumulus and deep convective clouds, which also occurred repeatedly in the Sentinel-2 observations during the campaign period. Thus, the determined cloud fraction can be taken as an approximation for a typical cloud cover of shallow, broken trade cumulus, but not as representative for the total cover of low cloudiness in the observation area. Another reason for the differences is thought to be the overestimation of cloud fraction from broken clouds in long-term satellite observations with

a coarser sensor resolution as in Brueck et al. (2015). The dependence of the determined coverage on detector resolution is presented in more detail in section 5.4. On the other hand, the cloud fraction also slightly overestimates the real cloud base fraction, since

The determined fraction of shadow areas in the Sentinel-2 MSI scenes shows stronger variations between the subsets than cloud fraction. A very small mean value of 0.002 was calculated for the shadow fraction of the subsets in tile T21PTM on 31-Jan-2020, while in some subsets even a larger shadow fraction than cloud fraction was found up to 0.32 (02-Feb-2020, tile T21QZU). The mean value of all determined shadow fractions is 0.067 and 50% of the subsets show fractions within 0.025-0.09, which is slightly below the determined cloud fractions.

The ratio between cloud and shadow fraction is determined by several factors. On the one hand, the overlap of cloud and shadow objects leads to a reduced shadow fraction, especially at low solar zenith angles, while vertically thick cumulus clouds can also cause an increase of the total shadow area. In addition, as shown in section 4.2.3, the accuracy of the determination of the shaded areas depends on the contrast between the illuminated and shaded ocean surfaces. Since the shadow detection has been adapted to reduce the misinterpretation of ocean pixels as shaded areas, a methodological underestimation of the shaded area, such as in tile T21PTM, is possible. Due to the dependence of the shadow fraction on solar and satellite geometry, it has only been determined in the context of method evaluations so far. Nevertheless, knowledge of the shadow fraction of cloud-free pixels is significant for understanding 3D radiative transfer in broken cumulus clouds, and the high-resolution data sets of the Sentinel-2 MSI, in contrast to coarser satellite records, allow for more detailed observations on the topic.

5.2. Cloud base height

Based on the matching algorithm between cloud objects and corresponding objects, a determination of the CBH could be performed for all scenes. In order to avoid misconnections caused by many smaller cloud objects in the edge region of a larger cumulus cloud as well as by the misinterpretation of shaded wave regions as shadow pixels, all objects with an area of less than 100 pixels were excluded in the algorithm. For the calculation of the displacement vector between cloud and shadow object, the sun angles as well as the observation angles of all bands involved in cloud and shadow detection (bands 2, 3, 4, 8, and 12) were averaged and interpolated to the 10 m observation grid. To avoid mismatches, no CBHs were calculated within 200 pixels of the edge of the tile.

Figure 10 displays the statistics of the calculated CBH using the same boxplot measures as in the fraction statistics. The number of successfully paired objects varies between the analyzed scenes, ranging from 187 to 4486 objects. Causes for the different numbers can be found in the success of shadow detection in addition to the smaller data sets in the edge of the observation path and the different cloud fractions and cloud sizes. Especially with the optically thin cloud cover on 31-Jan-2020, numerous cloud objects could not be assigned due to too few detected shadows.

The distribution of the CBH shows a frequency maximum of a few tens to a few hundred meters width for nearly all scenes. Since the number of cloud objects is clearly dominated by the smaller cloud objects (see section 5.3), the use of the frequency maximum can reduce the inclusion of larger and thus deeper cloud objects in the determination of the CBH of a scene. In particular, the mean value is not meaningful due to the positive deviation of the determined CBH for larger objects.

The frequency maximum of CBH of all scenes is at 858 m and the median of the CBH is calculated 933 m. Between observation days, the frequency maxima differed slightly. The lowest CBHs of 550 m-850 m were determined on 31-Jan-2020, while the higher CBHs of 875 m-1025 m were determined on 05-Feb-2020.

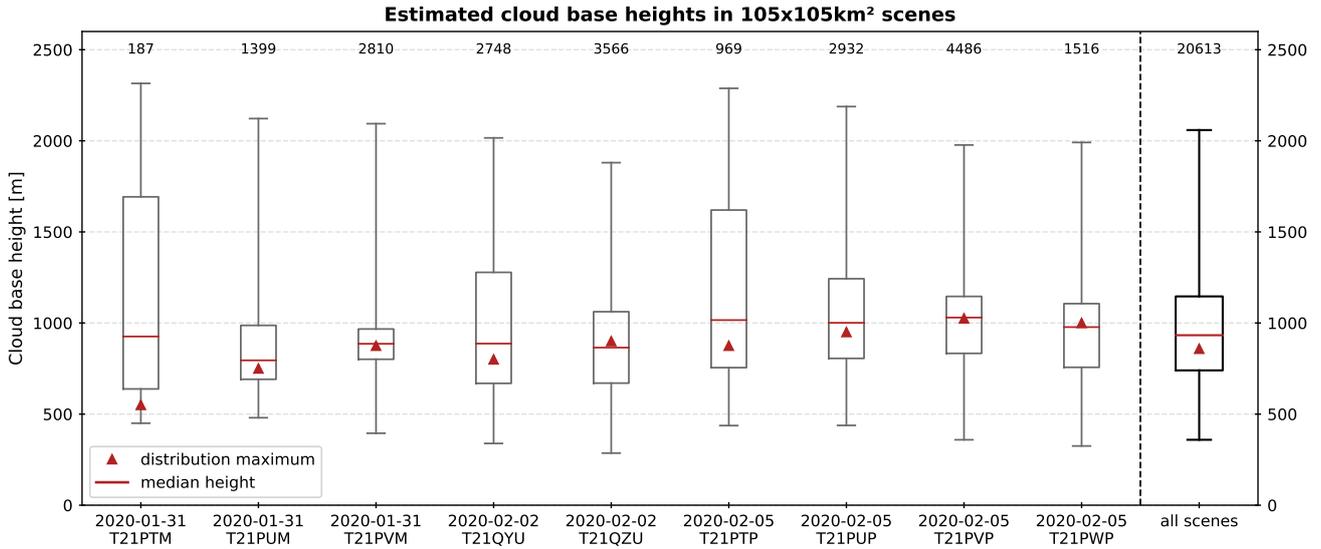


Figure 10: Boxplots of estimated cloud base height of trade cumulus clouds in 9 Sentinel-2 tiles ($105 \times 105 \text{ km}^2$) based on pairing of detected cloud and shadow objects. The mean is indicated by a red line, the boxes shows the 25th and 75th percentile, the whiskers the 5th and 95th percentile of the cloud base height. The numbers indicate the number of paired objects within the tile.

The variability between the observation days agrees with the results from Stephan et al. (2021), who used ascending and descending soundings from all research vessels and the BCO, as well as ceilometer measurements on board the research vessel Meteor. The observed CBH as well as the determined lifting condensation level were 500-600 m on 31-Jan-2020, whereas on 05-Feb-2020, they changed to about 750 m. This comparison highlights the possibility of using Sentinel-2 MSI observations for CBH estimation, although similar to the direct comparison with ceilometer measurements, higher values of CBH were calculated by the satellite algorithm. It must also be noted that the radiosonde and ceilometer measurements were taken up to a few hundred kilometers away from the considered scenes.

However, the estimation of CBH with Sentinel-2 observations also has clear limitations. As highlighted by the large number of observed cloud heights on both sides of the frequency maximum in Figure 10, an increase in cloud fraction or the formation of cloud objects along the displacement vector leads to mismatches between cloud and shadow objects. The change in cloud shadow eccentricity with deeper clouds offers potential for cloud vertical thickness and cloud top height estimations. However, the presence of deeper cumulus clouds in this evaluation results in numerous false CBHs. A further analysis of the dependence between CBH and cloud size could allow for exclusion of deep clouds from base height calculations. Therefore, the best results with this method are expected for very shallow cloud fields with large distances between individual cloud objects. If the shadow detection succeeds despite for the coverage of optically very thin clouds, the Sentinel-2 CBH estimation can provide a lot more information of the distribution of CBH within a region than can be achieved by ground base remote sensing measurements, e.g., using a ceilometer.

5.3. Cloud size statistics

The high spatial resolution of 10 m together with the wide observation path of 290 km allows to observe the abundance of clouds from a wide range of size ranges. For each detected cloud object, the area was converted to an equivalent diameter D using equation 1 and frequency distributions within the scene were established from all calculated diameters. When evaluating the object measures, it must be noted that deep trade cumulus can lead to an increase in the determined cross-sectional area, since the observation was performed with zenith angles up to 14° . For comparability with previous observations of cloud size distributions, two different binning methods, one using constant width bins and one using exponentially increasing bins, were used. Following Mieslinger et al. (2019), these methods are referred to as linear and logarithmic binning, respectively.

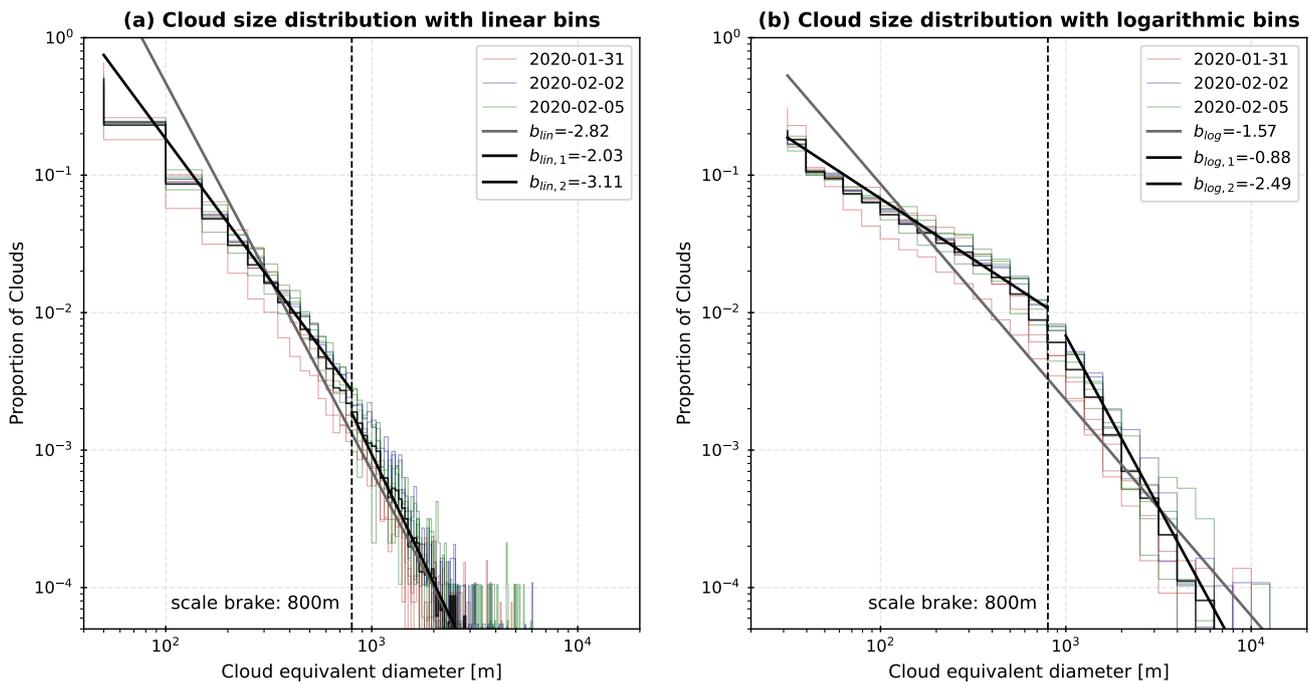


Figure 11: Cloud size distributions of trade cumulus clouds detected in 9 Sentinel-2 MSI scenes. The colored lines show the distribution of cloud size for individual scenes, the black line for all scenes. The distributions were calculated in (a) using a linear binning and a constant width of 50 m and in (b) a logarithmic binning. The thick grey line shows the power-law regression line without using a scale break size. The thick black lines show the splitted power-law regression lines with exponents b_1 and b_2 . The scale break between the two regressions (dashed line) at 800 m diameter was obtained from minimizing the sum of squared residuals.

Figure 11(a) shows the calculated cloud size distribution using linear binning with a constant class width of 50 m. The cloud size distributions of the individual scenes are shown in color, and the distribution of all cloud objects included in the evaluation is shown in black. Since both axes in the diagram are chosen logarithmically and the distributions have one or more straight lines, approximations can be made using power laws. The gray line shows the regression obtained by approximating equation 5 while minimizing the squared residuals for the entire data set. It leads to a scaling parameter $b_{lin} = -2.82$, which means, that the number of cloud objects exponentially decreases with increasing cloud size. In particular, as shown in figure 11(b), the distributions in the middle size range show a change in slope, and can be approximated

much better by 2 independent regressions. Minimizing the sum of all squared residuals of both regressions yielded the best fit for a scale break size D_c of 800 m, which is plotted as a black dashed line in figure 11. The scaling parameters of the double power law distribution are $b_{lin,1} = -2.03$ for $D \leq D_c$ and $b_{lin,2} = -3.11$ for $D \geq D_c$.

Since in the tail of the linear binned distribution of cloud sizes are no longer assigned to all bins, they still have a strong weight in the regression calculation, and can lead to large deviations in the determined regression parameters (Clauset et al., 2009). The cut-off size is therefore essential in linear binning methods and was chosen at 7 km equivalent diameter following Zhao and Di Girolamo (2007) and Mieslinger et al. (2019). The choice of logarithmic binning allows the inclusion of all cloud sizes in the regression calculation. The derived single and double power law distributions with the scaling parameters $b_{log} = -1.57$ and $b_{log,1} = -0.88$, respectively, and $b_{log,2} = -2.49$ around the scale break of 800 m are shown in Figure 11(b). According to equation 6: $b(lin(D)) \propto b(log(D)) - 1$, the scaling parameters differ slightly from those determined with linear binning. In contrast to linear binning, the area under the function in logarithmic binning gives the contribution to the total equivalent diameter.

The equidistant size classes in the logarithmic plot also allow comparison of individual scenes. The most scenes vary only slightly in number density up to the middle size range by a maximum factor of 1.5. Only one scene at 31-Jan-2020 January (T21PTM) shows a very high number of smallest cloud objects, which causes the number density of all other size classes to be significantly lower. All scenes of this day also show a slightly larger slope in the upper part, and a smaller scale break size of about 600 m. Minimizing the quadratic residuals to determine the scale break size therefore resulted in similarly successful regressions for all scale breaks in a size range between 600 m and 900 m as well.

Table 4 compares the scaling parameters and scale breaks obtained in this evaluation with those of previous publications that used high-resolution satellite remote sensing. In particular, comparing two evaluations based on large Terra-ASTER datasets (Zhao and Di Girolamo, 2007; Mieslinger et al., 2019), similar scaling parameters b , b_1 , and b_2 were obtained using both binning methods, with slightly steeper distributions in this evaluation than in Mieslinger et al. (2019).

There are differences with the evaluation of Koren et al. (2008), who derived a scaling parameter of $b(log(A)) = -1.3$ size distributions of the cloud area based on 5 Landsat-EMT+ scenes, which corresponds to a $b(lin(D)) = -3.6$ according to equation 9. They concluded from the strongly negative scaling parameter that the frequency distribution of the total area of clouds with a certain diameter decreases strictly. This follows from equation 10 and $\frac{dF}{dD} < 0$ for $b < -2$. Thus the smallest clouds would contribute most to the total area and cloud fraction, respectively. However, the analysis of the Sentinel-2 datasets leads to a different sign in the expected increase of the cloud fraction distribution for all cloud diameters below the scale break, since the scaling parameter is around $b = 2$ depending on the binning method. The strongly negative slope parameters b_2 in all publications therefore indicate, that larger clouds contribute less to the total cloud area.

The calculated scale break size is slightly larger than that of the other evaluations. The physical interpretation of the scale break in the cloud size distribution remains a subject of research. Neggers et al. (2003) used Large Eddy Simulation (LES) simulations to show that the vertical thickness of the boundary layer below the CBH as well as the vertical wind shear have an influence on the size of the scale break. They postulated that the cloud size distribution below the scale break is dominated by the coherent turbulence structure below the cloud, whereas the cloud size distribution above the scale break is dominated by the conversion of larger clouds into smaller ones. The largest individual convective cells also lie in this size range (Cahalan and Joseph, 1989). Since the determined CBHs of the scenes - taking

Study	N_{scenes}	Methods		Scaling parameters			
		Binning	Size limit	b	b_1	b_2	D_s
Cahalan and Joseph (1989)	19	$\log(\sqrt{A})$	-	-2.5	-1.6	-3.3	0.5 km
Zhao and Di Girolamo (2007)	152	lin(D)	7 km	-2.85	-1.88	-3.18	0.6 km
Koren et al. (2008)	5	log(A)	-	-3.6	-	-	-
Mieslinger et al. (2019)	1158	lin(D)	7 km	-	-1.6	-3.23	0.59 km
		log(D)	-	-2.55	-1.68	-3.12	
Sentinel-2 MSI	9	lin(D)	7 km	-2.82	-2.03	-3.11	0.8 km
		log(D)	-	-2.57	-1.88	-3.49	

Table 4: Cloud-size statistics from previous observation studies using high-resolution satellite sensors. The table shows the Number of analyzed satellite scenes (N_{scenes}), the binning scheme (linear or logarithmic) and the cloud measure (cloud equivalent diameter D or cloud area A), the maximum cloud equivalent diameter used in the linear regressions, the scaling parameters of the single power law regression b and double power-law regression b_1 and b_2 as well as the scale break size D_s . All values are converted to natural scaling parameters using equations 8 and 9

into account the systematic overestimation - are in the range of the scale break, and the scale break in the scenes of 31-Jan-2020 already occurs at smaller equivalent diameters of 600 m, this evaluation also strengthens the arguments for the influence of the vertical thickness of the lower cloud layer. However, Zhao and Di Girolamo (2007) point out that the scale break may also be a consequence of the regression method. This follows from the inclusion of rare but large clouds, which can influence the determination of the regression. Mieslinger et al. (2019) didn't find significant variations in cloud field properties with the vertical wind shear and conclude a minor importance. Similarly, both model and satellite sensor resolution can influence the strength and location of the scale break (Neggers et al., 2003). Therefore, a closer look at the properties of clouds and cloud size distributions as a function of sensor resolution is given below.

5.4. Resolution effects on cloud optical properties

Trade cumulus observations are possible for the first time at a spatial resolution of 10 m using Sentinel-2 MSI observations. Although many publications on the cloud fraction and cloud size distribution of trade cumulus are based on high-resolution satellite data with spatial resolutions of 15 m-30 m, the higher resolution allows theoretical considerations on the cloud properties below the sensor resolution to be supported by observations. Previous analyses by Koren et al. (2008) and Zhao and Di Girolamo (2007) based on satellite observations and by Neggers et al. (2003) based on LES model simulations differ with respect to the influence of small cumulus clouds on total cloud fraction. On the one hand, based on cloud size distributions with a scaling parameter of $b = -3.6$ and assuming a continuous cloud size, Koren et al. (2008) postulated that small cumulus clouds below the detector resolution constitutes the largest contribution to the total cloud fraction. On the other hand, Zhao and Di Girolamo (2007) and Neggers et al. (2003) showed the presence of an intermediate dominating cloud size equivalent to a cloud diameter in the scale break region.

Figure 12(a) shows the distribution of the total area of clouds of an equivalent diameter calculated from Sentinel-2 observations as a fraction of the total cloud cover for the highest sensor resolution. Contrary to the assumption of a dominant cloud size in the scale break region, all size classes up to an equivalent diameter of 500 m show a similar contribution. No single scene shows a significant increase in cloud cover

fraction with increasing cloud size in the size classes below the scale break, but a single scene (T21PTM) shows a decrease in this size range. Consequently, the Sentinel-2 observations support the hypothesis of Koren et al. (2008) of a large contribution to the total cloud fraction from small cloud sizes. However, not that only the smallest detectable cloud objects make this contribution. Figure 12(b) displays the same distributions of the contribution to the total cloud fraction as blue lines, but as a cumulative function. The importance of small cumulus clouds to the determined total cloud fraction is also evident here, as 50% of the contribution comes from clouds less than 1 km in diameter and 70% of the contribution comes from clouds less than 2 km in diameter. The contribution by smaller clouds is thus on average even stronger in comparison to Zhao and Di Girolamo (2007).

To investigate the influence of detector resolution on cloud size distribution, the resolution of the reflectance data sets was converted to coarser resolutions by reducing several pixels into a mean reflectance value, and then scene classification was applied to these data sets. The observation of the scenes by the Terra-MODIS or Geostationary Operational Environmental Satellite (GOES)-Advanced Baseline Imager (ABI) sensor was simulated according to Koren et al. (2008) by computing data sets with the spatial resolution of 240 m (corresponding to the highest resolution of MODIS VIS bands), 960 m (MODIS cloud mask resolution), and 1920 m (GOES-ABI cloud mask resolution) using the local reflectance mean. Figures 12(c-e) display the mean reflectance and the scene classification of a trade cumulus scene in tile T21PUP on 05-Feb-2020 for three different resolutions. On the one hand, it can be noticed in the comparison of the 240 m resolution with the 10 m resolution, that a worse detector resolution leads to the reduction of the mean reflectance of smallest cloud objects by averaging in surrounding ocean reflectances below the threshold for cloud detection. On the other hand, minima of the reflectance neighboring area of large reflectances are raised and included in the cloud mask, thereby grouping neighboring clouds into larger objects. As it is visible in the 1920 m scene, it is clear that a coarser resolution reduces the number of cloud objects to a few very large cloud objects.

As a result, the cumulative function of the fractions of clouds of a given size in the total fraction in figure 12(b) also shows a shift toward larger equivalent diameters with coarser detector resolution. This means that when the detector resolution is changed to 20 m, only 45% of the total cloud fraction is contributed by clouds with $D \leq 1 \text{ km}$, and when the scenes are detected using the highest MODIS channel resolution, only 10% of the total fraction comes from this size range. In contrast, clouds above a diameter of 15 km are not detected using 10 m resolution, but contribute with 10% at 240 m resolution, and 90% at a resolution of 1920 m to the total cloud cover. The consequence is a shift in the overall number distribution of cloud sizes towards a strong decrease in the number of small clouds and a slight increase in the number of very large clouds with increasing resolution. Accordingly, the scaling parameter of the cloud size distribution is smaller, which can lead misrepresentations e.g. in parameterizations which uses cloud size distributions from coarser detectors.

In addition to the number and area distribution of cloud sizes, the total area of all cloud objects and thus the cloud fraction in the scene shown in Figure 12(c-e) also increases from 0.085 to 0.212 when the resolution is changed to 1920 m. On average across all scenes, cloud fraction changes from 0.08 to 0.135 at a resolution of 960 m and to 0.16 at a resolution of 1920 m, as shown in Figure 12(f). This determined change is below the more than twofold overestimation of the cloud fraction found in the comparison of Terra-ASTER (15 m resolution) and MODIS data (1 km resolution) (Zhao and Di Girolamo, 2006). A direct comparison of Sentinel-2 observations with observational datasets from other satellite sensors is therefore an important further step of consideration, but is beyond the scope of this work.

However, the change in the cloud fraction of the individual scenes differs significantly. Up to a detector

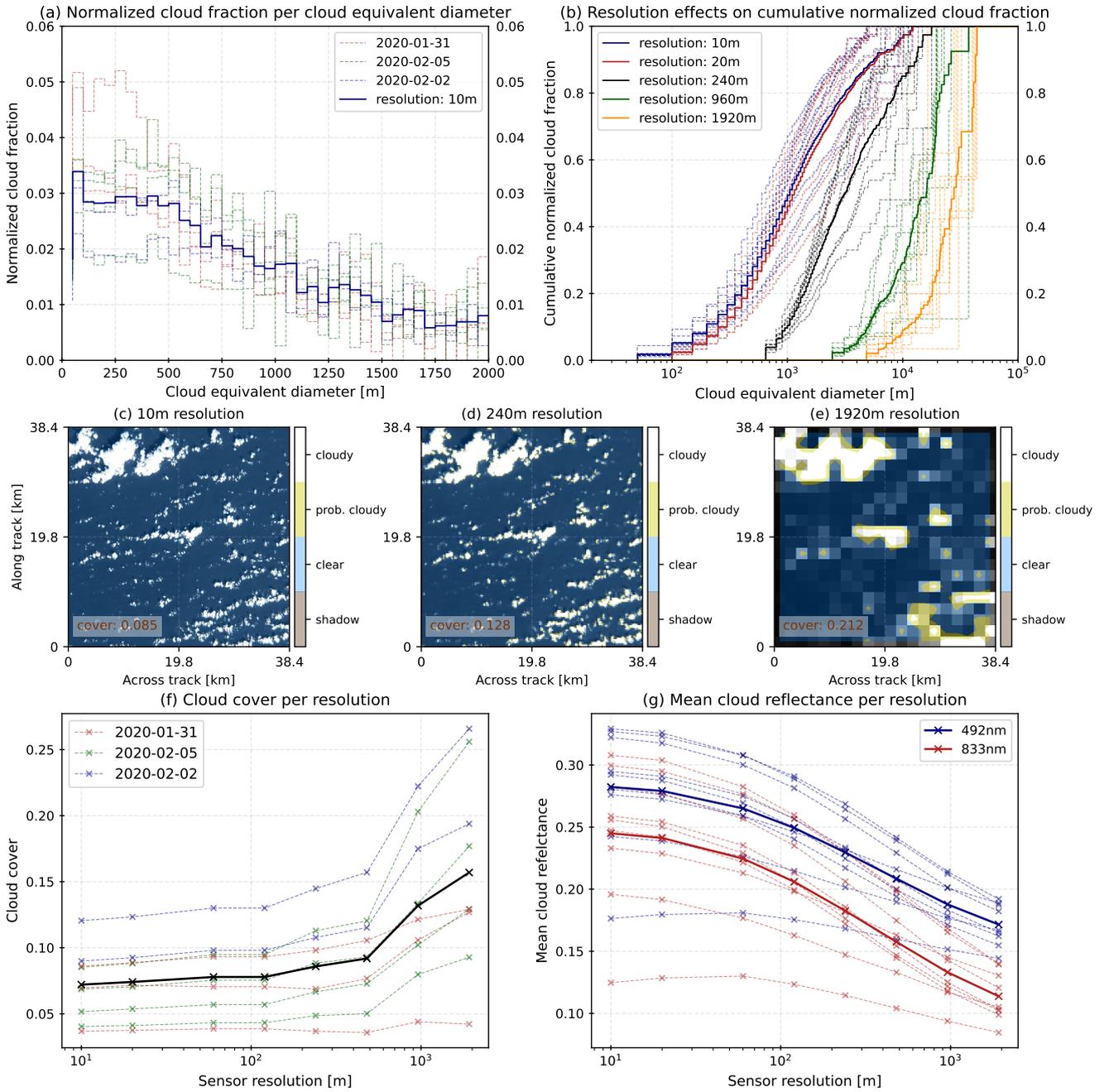


Figure 12: (a) Cumulative distribution of normalized cloud fraction detected in 9 Sentinel-2 MSI scenes as a function of cloud equivalent diameter for different sensor resolutions, (b-d) Mean reflectance, scene classification and estimated cloud cover of an example cumulus scene (05-Feb-2020, T21PUP) for 3 different sensor resolutions, (e) Estimated cloud cover as function of sensor resolution, (f) Estimated mean reflectance of clouds in band 2 (492 nm) and band 8 (833 nm) as function of sensor resolution

resolution of 240 m, the determined cloud fractions of all scenes increase only slightly, and the cloud fractions of some scenes on 31-Jan-2020 also remain constant. With further coarse graining, the change in cloud fraction differs between observation days: it is weaker in the scenes of 31-Jan-2020 than in the scenes of 02-Feb-2020 and 05-Feb-2020.

One reason for this is assumed to be the presence of optically very thin cumulus clouds in the scenes of 31-Jan-2020. For interpretation, the mean reflectance of all cloud objects is displayed as a function of sensor resolution for 2 channels in the blue and red visible wavelength ranges in figure 12(g). Almost all scenes show a strictly decreasing curve of cloud reflectance down to 50%-60% of the 10 m mean reflectance when resolution changes to 1920 m, since at coarser resolution more ocean pixels with a low reflectance are mixed into the cloud objects. However, some scenes initially show a slight increase in mean reflectance. It is believed that in the case of optically thin cumulus clouds, a reduction in resolution initially leads to an exclusion of pixels at the cloud edge that slip below the cloud mask threshold. This increases the average reflectance of the remaining pixels. With a further coarse graining, however, more and more optically thin clouds are thus also classified as clear and the cloud fraction does not increase further due to the lack of cloud regions with a high reflectance. Thus, the change in the cloud amount with resolution probably depends on the ratio of the mean cloud reflectance to the selected cloud detection threshold, which should be taken into account when interpreting different cloud detection algorithms. In addition, a relationship to the spatial organization of the clouds is expected, since, for example, the sensitivity differs between tiles T21QZU (02-Feb-2020) and T21PVP (05-Feb-2020) with similar initial cloud fraction and mean reflectance. In tile T21PVP the deeper and thus brighter clouds are distributed more randomly in the scene, whereas in tile T21QYU they begin to cluster in the "Flower" organizational form. Therefore, the large reflectance of the large clouds in tile T21PVP dominates all areas of the scene at coarse resolution, whereas in tile T21QYU more cloud objects outside the clusters fall below the detection thresholds.

Thus, the change in the cloud fraction with detector resolution is not a simple function that can be applied to correct for overestimation by coarse-resolution satellites, but is determined by the observed cloud scene and the cloud detection algorithm itself. The decrease in mean cloud reflectance with increasing detector resolution also indicates that the derivation of optical cloud properties with coarser resolution satellite sensors such as the GOES-ABI using bispectral reflectance functions (Nakajima and King, 1990) introduces reflectances that are too low, caused by the lack of representation of subpixel variability. It should also be noted that these evaluations were performed with a constant cloud detection algorithm, constant sensor geometry, and under simple spatial averaging of reflectances. Data sets from other satellite sensors may differ significantly due to different bands, as well as different sensor observation angles for both cloud and ocean reflectances.

6. Conclusion and Outlook

This thesis has used Sentinel-2 MSI satellite imagery for the first time to characterize the macrophysical properties of tropical marine trade cumulus clouds in terms of their cloud fraction, shadow fraction, cloud size distribution, and cloud base height. The extension of the satellite observations to the maritime area east of Barbados (9.5°-17°N, 53.5°-59.5°W) within the EUREC⁴A campaign in January and February 2020 forms the basis for the investigation of 9 trade cumulus scenes with a size of 110x110 km². The high spatial resolution of the MSI channels of up to 10 m provides the basis for the study of very small to medium sized clouds, which were assigned to the mesoscale organization type "Sugar" according to Stevens et al. (2020).

One objective of this work was to develop an algorithm for the satellite-based estimation of cloud base height of trade cumulus using the parallax of clouds and cloud shadows. Since this work is the first effort to investigate cloud properties over ocean surfaces using Sentinel-2 MSI observations, a cloud-conservative cloud detection algorithm developed for Terra-ASTER by Werner et al. (2016) was adapted. The decision tree for classifying pixels as confidently cloudy, probably cloudy, probably clear, or confidently clear includes two limit tests and two ratio tests. The thresholds were adjusted for the sensor characteristics in this evaluation. Ocean areas were separated into shaded and sunlit areas using the cloud shadow detection index (CSDI) developed by Amin et al. (2012). The high variability of reflectances of sunlit ocean surfaces as a consequence of the staged sensor configuration necessitated the introduction of a variable CSDI threshold calculated for each individual scene from the frequency distribution of the CSDI.

A visual evaluation of the scene classification using 3 scenes with different cloud optical thickness and ocean reflectance showed a high detection accuracy even in the presence of optically very thin cloud cover, but some misclassifications in shaded cloud parts and in the cloud edge region as a consequence of horizontal photon transport. Compared to the Sentinel-2 Fmask by Zhu et al. (2015), the scene classification showed similar detection accuracy. Compared to the operational Sen2Cor scene classification, it showed improved detection accuracy. In particular, the detection of cloud shadows over the dark ocean surface was improved. However, it showed difficulties where the contrast between shaded and sunlit surface is too low. The accuracy of the scene classification meets the requirements for cloud height estimation, but requires a preselection of the satellite images by the presence of low clouds and the absence of high cloud cover. Application of the classification to more extensive data sets will therefore require an accuracy analysis on a broader range of marine low clouds.

The estimation of cloud base height applies the horizontal distance between a cloud object and its associated shadow object. The high variability of object shapes required the adjustment of the matching method from a shape-based approach (Berendes et al., 1992) to an area-based approach according to Zhu and Woodcock (2012). The cloud base height was calculated at the location of highest match similarity between moved cloud object and associated shadow object, and the estimation of cloud base height succeeded for a large number of cloud objects. However, this matching method requires the assumption of a vertically unextended cloud, because the vertical structure of the cumulus cloud results in an extension of the cloud shadow along the displacement vector. Therefore, a direct comparison with ship-based ceilometer measurements revealed too high values of cloud base height calculated from the matching method, especially for deeper trade cumulus. It is argued here, correction of the overestimation is possible by simplifying the object geometry and knowing the eccentricity of the paired objects. The matching methods also show a limit in scenes with higher cloud fraction, as this causes more shadow areas to be obscured by neighboring clouds, leading to an increase in misattribution. Cloud base height estimation at high cloud

fractions require new estimation methods, e.g., based on the parallax of an elevated object with knowledge of the wind speed (Frantz et al., 2018). The cloud base height of more than 20,000 analyzed cloud objects in this work was determined to be 858 m on average. The differences in cloud base height between the observation days follow the data from radiosondes and surface based remote sensing instruments.

Analysis of all 9 scenes yielded a mean cloud fraction of 0.083 and a mean shadow fraction of 0.067 and showed large variability (5th to 95th percentile) of 0.02-0.19 and 0.01-0.24, respectively, within 119 subsets with a size of 27.5x27.5 km². The determined cloud fraction and shadow fraction differs, because parts of the cloud shadows are superimposed by neighboring clouds. The mean values of cloud fraction for shallow trade cumulus scenes are consistent with previous analyses of high-resolution satellite data (Zhao and Di Girolamo, 2007; Mieslinger et al., 2019), but cannot be compared with mean values of the total low cloud fraction in the region because of the selection of scenes according to the presence of small- to medium-sized cumulus clouds. Large differences are also presented to clear-conservative estimates of cloud fraction as a consequence of a small shift of the separation threshold towards cloud reflectances (e.g. Zhao and Di Girolamo, 2006; Mieslinger et al., 2021). The relatively large width of the observation path of both Sentinel-2 satellites offers the possibility to investigate on the total marine low cloud fraction in this region in further studies.

The number-size distribution of all cloud objects were mathematically described with a double power law. The scaling parameters were estimated using linear fits to log-log histograms. The scaling parameters $b_1 = -2.03$ for smaller cloud sizes and $b_2 = -3.11$ for larger cloud sizes were calculated using a linear binning scheme and a cut-off size of 7 km. The application of a logarithmic binning scheme resulted in scaling parameters $b_1 = -1.88$ and $b_2 = 3.49$. The estimated scaling parameters are consistent to previous publications. A scale break was found at a typical cloud size of 0.8 km, which is slightly higher than in previous evaluations of high-resolution observations. Differences in the length of the scale break between scenes are consistent with differences in cloud base height, providing support for subcloud layer height as a physical control parameter of the scale break. For comparison with model observations, however, it must be noted that the observation angles deviates from the zenith by up to 14°. Deep cumulus clouds thus lead to an increase in the projected cloud area. The estimated parameters of the cloud size distribution are also sensitive to the the fitting strategy.

Another goal of this work was to demonstrate the dependence of cloud property retrievals on detector resolution. To achieve this, a coarse graining was used to reduce the satellite scenes to 1920 m resolution by local averaging. The scene classification was applied to the scenes and distributions of total cloud area, cloud fraction, and mean cloud reflectance were calculated. It was shown that half of the total cloud fraction is provided by clouds with size up to 1 km. However, the results contradict both previous analyses of Zhao and Di Girolamo (2007) and Neggers et al. (2003) on the presence of an area-dominant cloud size in the scale break region. Furthermore the results contradict Koren et al. (2008) on the area dominance of smallest clouds below detector resolution, showing a largely homogeneous distribution of the area contribution of individual cloud sizes below 500 m. In contrast, a coarser detector resolution shifts the area contributions towards a few very large cloud objects. A coarser detector resolution leads to an increase of the cloud fraction of the whole scene by reducing the mean cloud reflectance. However, the sensitivity of cloud fraction and mean cloud reflectance to the detector resolution showed differences between individual scenes. It is argued here, the organisation of the clouds cause the differences, because it determines how many pixels still remain above the reflectance threshold of the cloud mask after averaging. Since the use of different cloud masks and sensor geometries can also lead to deviations of retrieved cloud properties, this analysis shows the need to present systematic errors of the detection of cloud properties

with lower detector resolutions in a direct comparison with GOES-16 ABI or MODIS data sets during the campaign period in a further study. Similarly, the influence of the mesoscale organization of trade cumulus on the systematic differences in cloud properties between high- and medium-resolution satellites as well as geostationary satellites should be investigated in more detail, since the distribution of cumulus clouds in the boundary layer and their mesoscale organization represent key issues in understanding shallow cumulus cloud feedbacks (Vial et al., 2017).

In conclusion, this thesis successfully demonstrated the macrophysical properties of marine trade cumulus clouds from satellite observations of the Sentinel-2 MSI. However, validation and improvement of the analyses require comparisons with other measurement methods over longer time scales than the EUREC⁴A campaign period, for which satellite data is currently only available from coastal ocean areas. Nevertheless, the 290 km wide observation paths of both Sentinel-2 satellites allow the application of the data not only to high-resolution observations of the spatial distribution on the micro- β to meso- β scale, but also of the temporal evolution of trade cumulus on larger time scales. The combination of high spatial resolution and sensor geometry also offers the potential of deriving other cloud properties such as the motion vector of trade cumulus even at very small meteorological size ranges. The use of the Sentinel-2 satellite for atmospheric observations could thus provide an additional data set for improving the understanding of trade cumulus clouds in a warming climate.

A. Appendix

A.1. Single cloud scenes

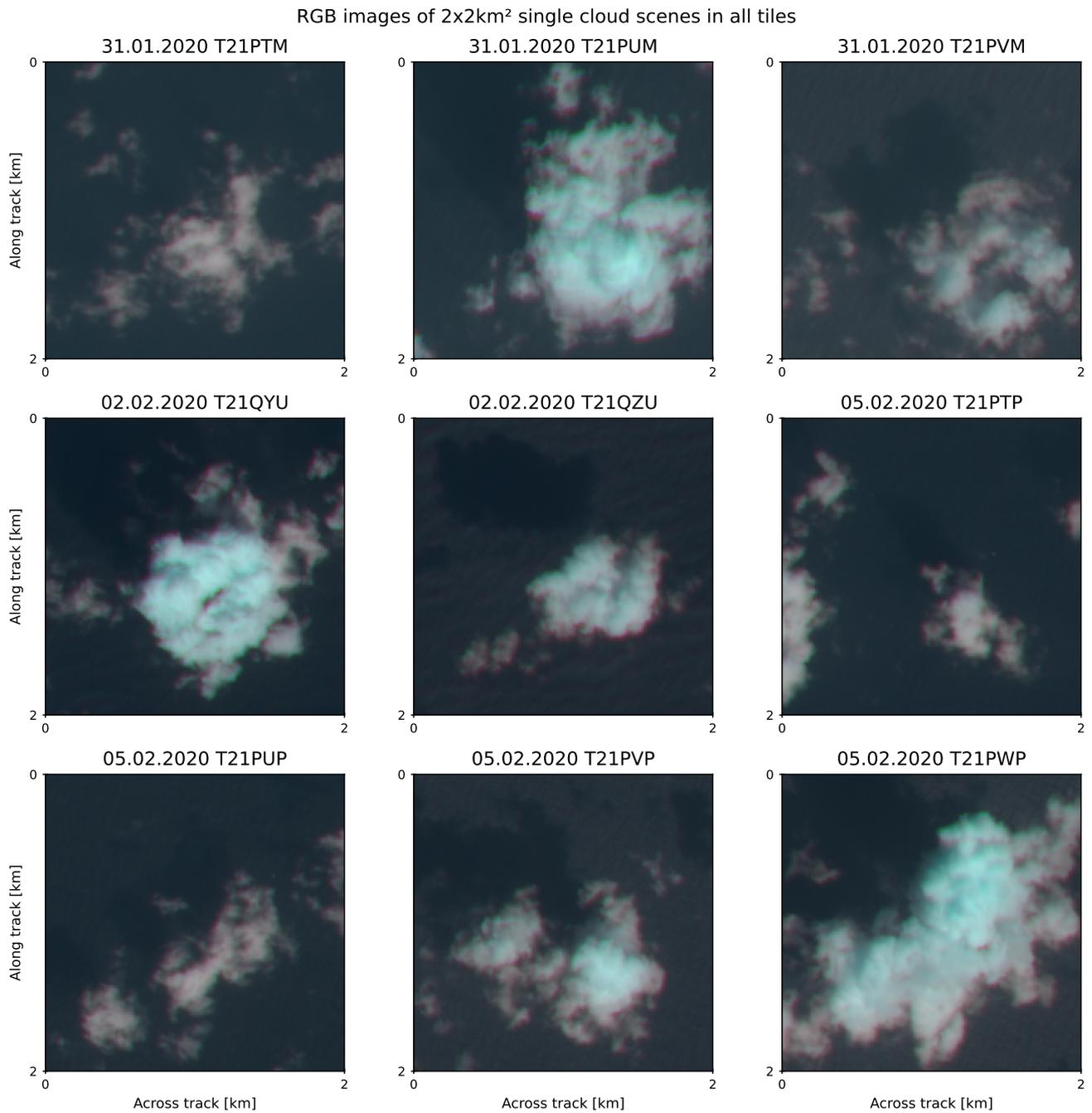


Figure 13: RGB (band 12, band 8, band 4) images of the 2x2 km² single cloud scenes which are used for threshold adaption in section 4.1.3 and section 4.2.2

A.2. Overview images of analyzed scenes

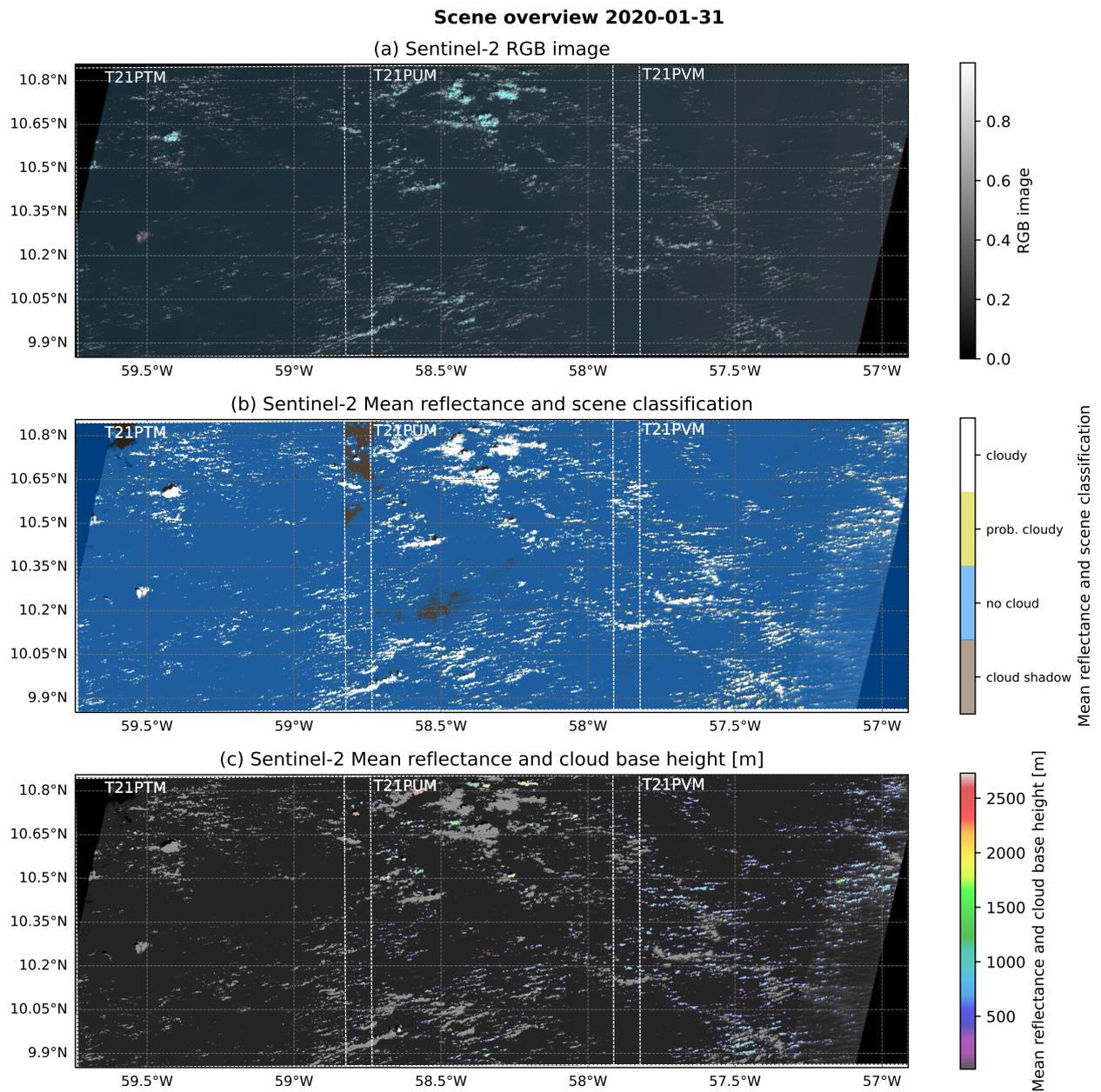
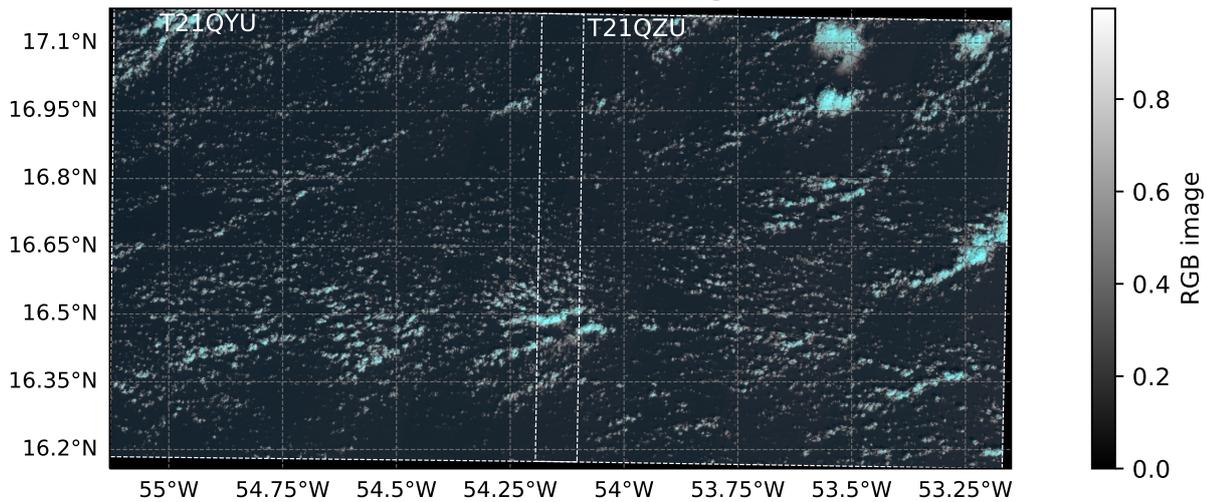


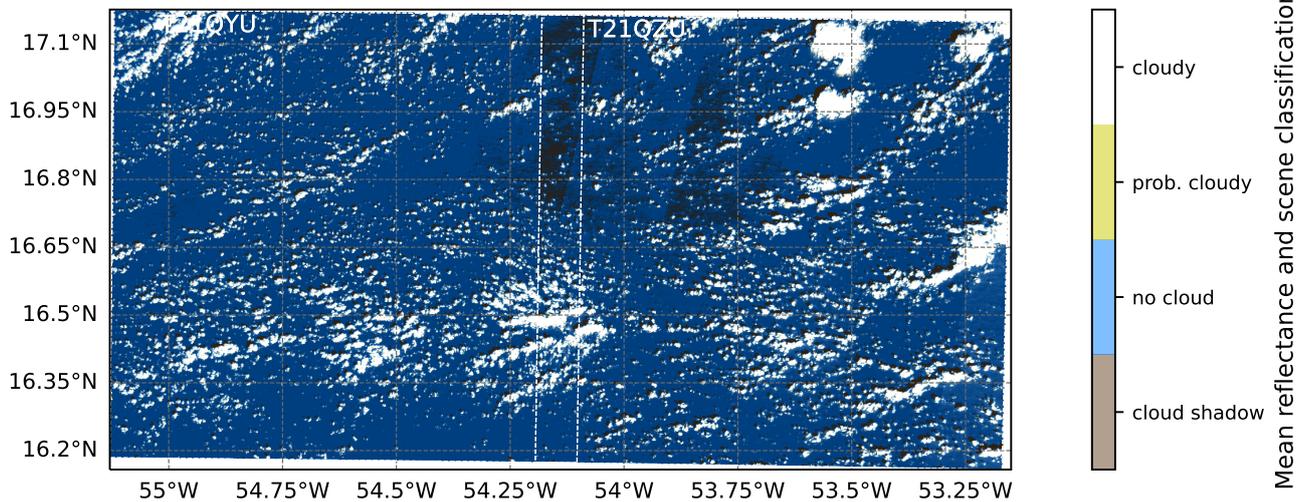
Figure 14: Cross-section of satellite path and analyzed tiles on 31-Jan-2020. (a) shows the RGB composit (band 12, band 8 and band 4) of the scenes, (b) shows the mean reflectance as well as the scene classification, and (c) shows the mean reflectance and the estimated cloud base height.

Scene overview 2020-02-02

(a) Sentinel-2 RGB image



(b) Sentinel-2 Mean reflectance and scene classification



(c) Sentinel-2 Mean reflectance and cloud base height [m]

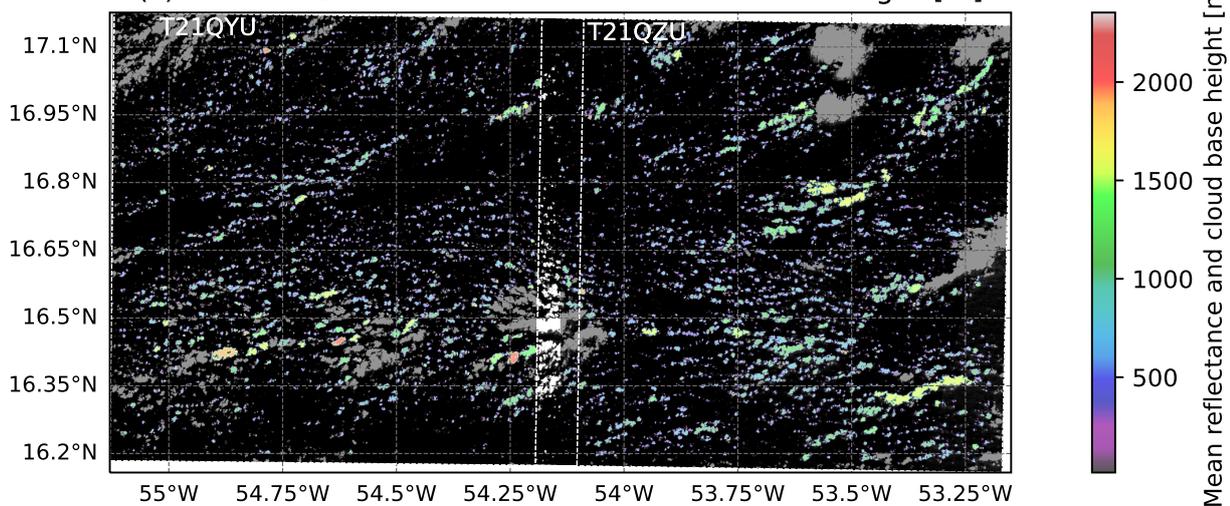


Figure 15: Cross-section of satellite path and analyzed tiles on 02-Feb-2020. (a) shows the RGB composit (band 12, band 8 and band 4) of the scenes, (b) shows the mean reflectance as well as the scene classification, and (c) shows the mean reflectance and the estimated cloud base height.

Scene overview 2020-02-05

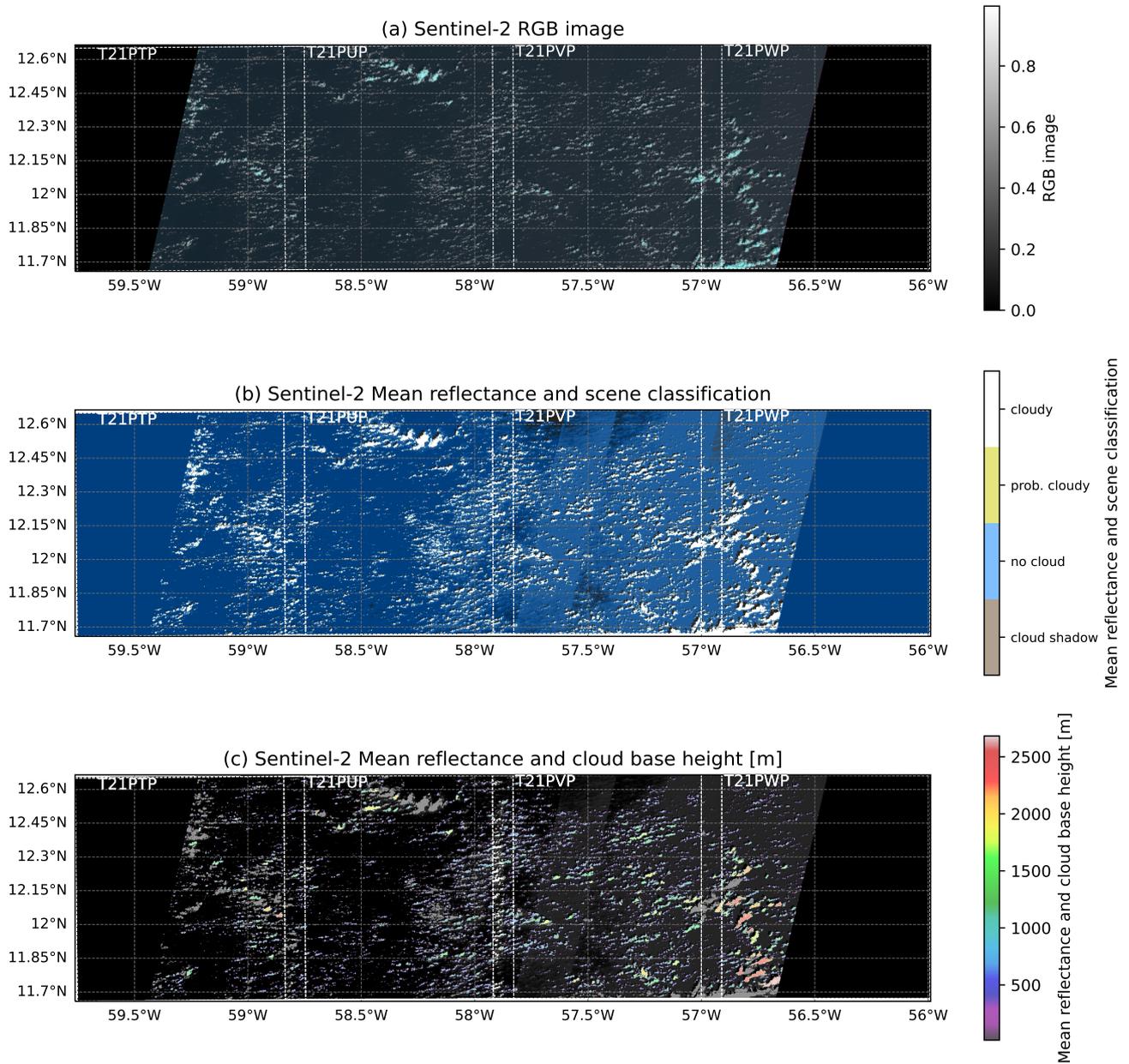


Figure 16: Cross-section of satellite path and analyzed tiles on 05-Feb-2020. (a) shows the RGB composit (band 12, band 8 and band 4) of the scenes, (b) shows the mean reflectance as well as the scene classification, and (c) shows the mean reflectance and the estimated cloud base height.

B. References

- Ackermann, S., Strabala, K., Menzel, P., Frey, R., Moeller, C., Gumley, L., Baum, B., Seemann, S., and Zhang, H. (2006). *Discriminating clear-sky from cloud with MODIS algorithm theoretical basis document*. Madison, WI: Coop. Inst. Meteorol. Satell. Stud.
- Amin, R., Gould, R., Hou, W., Arnone, R., and Lee, Z. (2012). Optical algorithm for cloud shadow detection over water. *IEEE transactions on geoscience and remote sensing*, 51(2):732–741.
- Baum, B. A., Menzel, W. P., Frey, R. A., Tobin, D. C., Holz, R. E., Ackerman, S. A., Heidinger, A. K., and Yang, P. (2012). Modis cloud-top property refinements for collection 6. *Journal of applied meteorology and climatology*, 51(6):1145–1163.
- Berendes, T., Sengupta, S. K., Welch, R. M., Wielicki, B. A., and Navar, M. (1992). Cumulus cloud base height estimation from high spatial resolution landsat data: A hough transform approach. *IEEE transactions on geoscience and remote sensing*, 30(3):430–443.
- Böhm, C., Sourdeval, O., Mülmenstädt, J., Quaas, J., and Crewell, S. (2019). Cloud base height retrieval from multi-angle satellite data. *Atmospheric Measurement Techniques*, 12(3):1841–1860.
- Bony, S. and Dufresne, J.-L. (2005). Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. *Geophysical Research Letters*, 32(20).
- Bony, S., Stevens, B., Ament, F., Bigorre, S., Chazette, P., Crewell, S., Delanoë, J., Emanuel, K., Farrell, D., Flamant, C., et al. (2017). Eurec 4 a: a field campaign to elucidate the couplings between clouds, convection and circulation. *Surveys in Geophysics*, 38(6):1529–1568.
- Bony, S., Stevens, B., Frierson, D. M., Jakob, C., Kageyama, M., Pincus, R., Shepherd, T. G., Sherwood, S. C., Siebesma, A. P., Sobel, A. H., et al. (2015). Clouds, circulation and climate sensitivity. *Nature Geoscience*, 8(4):261–268.
- Brueck, M., Nuijens, L., and Stevens, B. (2015). On the seasonal and synoptic time-scale variability of the north atlantic trade wind region and its low-level clouds. *Journal of the Atmospheric Sciences*, 72(4):1428–1446.
- Cahalan, R. F. and Joseph, J. H. (1989). Fractal statistics of cloud fields. *Monthly weather review*, 117(2):261–272.
- Clauset, A., Shalizi, C. R., and Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM review*, 51(4):661–703.
- Dey, S., Di Girolamo, L., and Zhao, G. (2008). Scale effect on statistics of the macrophysical properties of trade wind cumuli over the tropical western atlantic during rico. *Journal of Geophysical Research: Atmospheres*, 113(D24).
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., et al. (2012). Sentinel-2: Esa’s optical high-resolution mission for gmes operational services. *Remote sensing of Environment*, 120:25–36.

- Emeis, S., Münkler, C., Vogt, S., Müller, W. J., and Schäfer, K. (2004). Atmospheric boundary-layer structure from simultaneous sodar, rass, and ceilometer measurements. *Atmospheric environment*, 38(2):273–286.
- Fletcher, K. (2012). *SENTINEL 2: ESA’s Optical High-Resolution Mission for GMES Operational Services*. European Space Agency.
- Frantz, D., Haß, E., Uhl, A., Stoffels, J., and Hill, J. (2018). Improvement of the fmask algorithm for sentinel-2 images: Separating clouds from bright surfaces based on parallax effects. *Remote sensing of environment*, 215:471–481.
- Gascon, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J., Lonjou, V., Lafrance, B., Massera, S., Gaudel-Vacaresse, A., et al. (2017). Copernicus sentinel-2a calibration and products validation status. *Remote Sensing*, 9(6):584.
- Goodman, A. and Henderson-Sellers, A. (1988). Cloud detection and analysis: A review of recent progress. *Atmospheric Research*, 21(3-4):203–228.
- Hagolle, O., Huc, M., Desjardins, C., Auer, S., and Richter, R. (2017). Maja algorithm theoretical basis document. *Development*, pages 1–39.
- Hartmann, D. L., Ockert-Bell, M. E., and Michelsen, M. L. (1992). The effect of cloud type on earth’s energy balance: Global analysis. *Journal of Climate*, 5(11):1281–1304.
- Hasler, A. (1981). Stereographic observations from geosynchronous satellites: An important new tool for the atmospheric sciences. *Bulletin of the American Meteorological Society*, 62(2):194–212.
- Heinemann, G. (2015). Jenoptick chm15k nimbus ceilometer data sheet.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049.
- Klein, S. A., Hall, A., Norris, J. R., and Pincus, R. (2017). Low-cloud feedbacks from cloud-controlling factors: A review. *Shallow Clouds, Water Vapor, Circulation, and Climate Sensitivity*, pages 135–157.
- Koren, I. and Joseph, J. H. (2000). The histogram of the brightness distribution of clouds in high-resolution remotely sensed images. *Journal of Geophysical Research: Atmospheres*, 105(D24):29369–29377.
- Koren, I., Oreopoulos, L., Feingold, G., Remer, L., and Altaratz, O. (2008). How small is a small cloud? *Atmospheric Chemistry and Physics*, 8(14):3855–3864.
- Louis, J., Debaecker, V., Pflug, B., Main-Knorn, M., Bieniarz, J., Mueller-Wilm, U., Cadau, E., and Gascon, F. (2016). Sentinel-2 sen2cor: L2a processor for users. In *Proceedings Living Planet Symposium 2016*, pages 1–8. Spacebooks Online.
- Mieslinger, T., Horváth, Á., Buehler, S. A., and Sakradzija, M. (2019). The dependence of shallow cumulus macrophysical properties on large-scale meteorology as observed in aster imagery. *Journal of Geophysical Research: Atmospheres*, 124(21):11477–11505.
- Mieslinger, T., Stevens, B., Kölling, T., Brath, M., Wirth, M., and Buehler, S. A. (2021). Optically thin clouds in the trades. *Atmospheric Chemistry and Physics Discussions*, pages 1–33.

- Nakajima, T. and King, M. D. (1990). Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. part i: Theory. *Journal of Atmospheric Sciences*, 47(15):1878–1893.
- Neggers, R., Jonker, H., and Siebesma, A. (2003). Size statistics of cumulus cloud populations in large-eddy simulations. *Journal of the atmospheric sciences*, 60(8):1060–1074.
- Neggers, R. A., Neelin, J. D., and Stevens, B. (2007). Impact mechanisms of shallow cumulus convection on tropical climate dynamics. *Journal of Climate*, 20(11):2623–2642.
- Nuijens, L., Medeiros, B., Sandu, I., and Ahlgrim, M. (2015). Observed and modeled patterns of covariability between low-level cloudiness and the structure of the trade-wind layer. *Journal of Advances in Modeling Earth Systems*, 7(4):1741–1764.
- Nuijens, L., Serikov, I., Hirsch, L., Lonitz, K., and Stevens, B. (2014). The distribution and variability of low-level cloud in the north atlantic trades. *Quarterly Journal of the Royal Meteorological Society*, 140(684):2364–2374.
- Orlanski, I. (1975). A rational subdivision of scales for atmospheric processes. *Bulletin of the American Meteorological Society*, pages 527–530.
- Pahlevan, N., Sarkar, S., Franz, B., Balasubramanian, S., and He, J. (2017). Sentinel-2 multispectral instrument (msi) data processing for aquatic science applications: Demonstrations and validations. *Remote sensing of environment*, 201:47–56.
- Plank, V. G. (1969). The size distribution of cumulus clouds in representative florida populations. *Journal of Applied Meteorology and Climatology*, 8(1):46–67.
- Platnick, S., King, M. D., Ackerman, S. A., Menzel, W. P., Baum, B. A., Riédi, J. C., and Frey, R. A. (2003). The modis cloud products: Algorithms and examples from terra. *IEEE Transactions on geoscience and Remote Sensing*, 41(2):459–473.
- Richter, R., Louis, J., and Berthelot, B. (2011). Sentinel-2 msi – level 2a products algorithm theoretical basis document.
- Rossow, W. B. (1989). Measuring cloud properties from space: A review. *Journal of Climate*, 2(3):201–213.
- Seiz, G., Davies, R., and Grün, A. (2006). Stereo cloud-top height retrieval with aster and misr. *International Journal of Remote Sensing*, 27(9):1839–1853.
- Stephan, C. C., Schnitt, S., Schulz, H., Bellenger, H., De Szoeki, S. P., Acquistapace, C., Baier, K., Dauhut, T., Laxenaire, R., Morfa-Avalos, Y., et al. (2021). Ship-and island-based atmospheric soundings from the 2020 eurec 4 a field campaign. *Earth System Science Data*, 13(2):491–514.
- Stephens, G. L. and Kummerow, C. D. (2007). The remote sensing of clouds and precipitation from space: A review. *Journal of the Atmospheric Sciences*, 64(11):3742–3765.
- Stevens, B., Bony, S., Brogniez, H., Hentgen, L., Hohenegger, C., Kiemle, C., L’Ecuyer, T. S., Naumann, A. K., Schulz, H., Siebesma, P. A., et al. (2020). Sugar, gravel, fish and flowers: Mesoscale cloud patterns in the trade winds. *Quarterly Journal of the Royal Meteorological Society*, 146(726):141–152.

- Stevens, B., Bony, S., Farrell, D., Ament, F., Blyth, A., Fairall, C., Karstensen, J., Quinn, P. K., Speich, S., Acquistapace, C., et al. (2021). Eurec 4 a. *Earth System Science Data*, 13(8):4067–4119.
- Stevens, B., Farrell, D., Hirsch, L., Jansen, F., Nuijens, L., Serikov, I., Brüggmann, B., Forde, M., Linne, H., Lonitz, K., et al. (2016). The barbados cloud observatory: Anchoring investigations of clouds and circulation on the edge of the itcz. *Bulletin of the American Meteorological Society*, 97(5):787–801.
- Szantoi, Z. and Strobl, P. (2019). Copernicus sentinel-2 calibration and validation.
- Tarrio, K., Tang, X., Masek, J. G., Claverie, M., Ju, J., Qiu, S., Zhu, Z., and Woodcock, C. E. (2020). Comparison of cloud detection algorithms for sentinel-2 imagery. *Science of Remote Sensing*, 2:100010.
- Vial, J., Bony, S., Stevens, B., and Vogel, R. (2017). Mechanisms and model diversity of trade-wind shallow cumulus cloud feedbacks: a review. *Shallow Clouds, Water Vapor, Circulation, and Climate Sensitivity*, pages 159–181.
- Werner, F., Wind, G., Zhang, Z., Platnick, S., Girolamo, L. D., Zhao, G., Amarasinghe, N., and Meyer, K. (2016). Marine boundary layer cloud property retrievals from high-resolution aster observations: case studies and comparison with terra modis. *Atmospheric Measurement Techniques*, 9(12):5869–5894.
- Wielicki, B. A. and Welch, R. M. (1986). Cumulus cloud properties derived using landsat satellite data. *Journal of Applied Meteorology and Climatology*, 25(3):261–276.
- Wood, R. and Bretherton, C. S. (2006). On the relationship between stratiform low cloud cover and lower-tropospheric stability. *Journal of climate*, 19(24):6425–6432.
- Yang, Y. and Di Girolamo, L. (2008). Impacts of 3-d radiative effects on satellite cloud detection and their consequences on cloud fraction and aerosol optical depth retrievals. *Journal of Geophysical Research: Atmospheres*, 113(D4).
- Zekoll, V., Main-Knorn, M., Alonso, K., Louis, J., Frantz, D., Richter, R., and Pflug, B. (2021). Comparison of masking algorithms for sentinel-2 imagery. *Remote Sensing*, 13(1):137.
- Zhao, G. and Di Girolamo, L. (2006). Cloud fraction errors for trade wind cumuli from eos-terra instruments. *Geophysical research letters*, 33(20).
- Zhao, G. and Di Girolamo, L. (2007). Statistics on the macrophysical properties of trade wind cumuli over the tropical western atlantic. *Journal of Geophysical Research: Atmospheres*, 112(D10).
- Zhong, B., Chen, W., Wu, S., Hu, L., Luo, X., and Liu, Q. (2017). A cloud detection method based on relationship between objects of cloud and cloud-shadow for chinese moderate to high resolution satellite imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(11):4898–4908.
- Zhu, Z., Wang, S., and Woodcock, C. E. (2015). Improvement and expansion of the fmask algorithm: Cloud, cloud shadow, and snow detection for landsats 4–7, 8, and sentinel 2 images. *Remote Sensing of Environment*, 159:269–277.
- Zhu, Z. and Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in landsat imagery. *Remote sensing of environment*, 118:83–94.

Zhu, Z. and Woodcock, C. E. (2014). Automated cloud, cloud shadow, and snow detection in multitemporal landsat data: An algorithm designed specifically for monitoring land cover change. *Remote Sensing of Environment*, 152:217–234.

List of Figures

1.	Satellite observation zenith and azimuth angles	8
2.	Decision tree of the Werner et al. (2016) cloud detection algorithm	13
3.	Probability density functions of Reflectance and adjustment of cloud detection thresholds	14
4.	Cloud shadow detection index and adjustment of shadow masking thresholds	17
5.	Comparison of scene classification with Fmask and operational Sen2Cor classification	19
6.	Sketch of the cloud base height estimation	21
7.	Cloud shadow matching example with 3 shallow cumulus clouds	23
8.	Comparison of cloud base height estimation with shipbased Ceilometer measurements	25
9.	Scene analysis of cloud fraction and shadow fraction	27
10.	Scene analysis of estimated cloud base height	29
11.	Cloud size distribution of the analyzed scenes with two different binning schemes	30
12.	Sensitivity of cloud properties to detector resolution	34
13.	RGB images of the single cloud scenes	39
14.	Scene overview and characteristics at 31-Jan-2020	40
15.	Scene overview and characteristics at 02-Feb-2020	41
16.	Scene overview and characteristics at 05-Feb-2020	42

List of Tables

1.	Sentinel-2 MSI spectral band characteristics	7
2.	Sentinel-2 MSI dataset overview table	10
3.	Cloud detection thresholds	15
4.	Cloud size statistics	32

C. List of acronyms

ABI	Advanced Baseline Imager
AOD	Aerosol optical depth
ASB	Spatial adaptive sliding box
ASTER	Advanced Spaceborn Thermal Emission and Reflection Radiometer
BOA	Bottom of atmosphere
CBH	Cloud base height
CSDI	Cloud shadow detection index
ESA	European Space Agency
EUREC⁴A	Elucidating the role of clouds-circulation coupling in climate
Fmask	Function of mask
GMES	Global Monitoring for Enviroment and Security
GOES	Geostationary Operational Enviromental Satellite
HALO	High-flying aircraft
HICO	Hyperspectral Imager for the coastal ocean
ISS	International Space Station
LES	Large Eddy Simulation
LIDAR	Light Detection and Ranging
MODIS	Moderate Resolution Imaging Spectrometer
MSI	Multi Spectral Imager
NIR	Near-infrared
PDF	Probability density function
RGB	Red-green-blue image
SWIR	Shortwave infrared
TOA	Top of atmosphere
VIS	Visible
VNIR	Visible and near-infrared

D. Acknowledgement

First of all, I would like to thank Dr. Hartwig Deneke for the intensive supervision and support during the preparation of this thesis. I thank you especially for the time you took in the final phase for the detailed review and comments on the thesis. I would also like to thank Prof. Andreas Macke and Prof. Johannes Quaas for the final review of my master thesis and its defense.

This thesis contains Sentinel-2 data of Year 2020. For providing the Sentinel-2 data via Open Access Hub: <https://scihub.copernicus.eu/dhus/home> I would like to thank the Copernicus Service. The Ceilometer data used in this publication was gathered in the EUREC⁴A field campaign and is made available through AERIS data server: <https://observations.ipsl.fr/aeris/eurec4a/>

Furthermore, I would like to thank Dr. Anja Huehnerbein and Dr. Sebastian Bley for the many informative discussions and comments as well as Dr. Rico Hengst for the technical support especially at the beginning of the master thesis. Thanks go of course also to all other members of the Satellite Remote Sensing group for the nice conversations in presence as well as in the digital coffee breaks.

I thank my partner Florian Neuhaeuser for the patience, the emotional support and the encouraging words, without which this thesis would not have been finished. I also thank my fellow students and friends for the fantastic years of study in Leipzig and the successful distraction during the work breaks of the last months!

Finally, I would like to thank my parents Helga Ritter and Gerd Ritter and all other members of my family. You made it possible for me to study my desired subject of meteorology in Leipzig and you supported me fully at all times.

E. Declaration of Authorship

Hereby, I declare that I have composed the presented thesis independently on my own and without any other resources than the ones indicated. All thoughts taken directly or indirectly from external sources are properly denoted as such.

Oscar Ritter
Leipzig, 07th March 2022