Cloud shadow detection and removal for high spatial resolution optical satellite data

Vorgelegt von

Viktoria Zekoll

Abstract

Optical satellite imagery contain in many cases clouds and cloud shadows, so that an automated classification of objects on the Earth's surface is difficult. Therefore, on the one hand, it is important to create a good cloud and cloud shadow map and, on the other hand, to correct the shadow areas in such a way that a classification via the spectral properties of the objects is uniformly possible. Of course, this also includes the process of atmospheric correction of the optical satellite image data. Especially for land applications the amount of scenes with usable data is of high importance due to the exact timing being significant (crop yield estimation) or because the scene is not free-of-charge and one has to pay for the next acquisition if the current one contains cloud shadow over the location of interest.

The masking of clouds, cloud shadows, water and snow/ice in optical satellite imagery is therefore an essential step in automated processing chains.

Furthermore the exact masking of cloud shadows is a very important task prior to the removal of cloud shadows. Due to the Earth having an annual cloud coverage of approximately 70%, the contamination of multi-spectral satellite imagery is inevitable and scientist will have to work with and around this shortcoming.

For this study, the satellite data from the Sentinel-2 mission is used which provides a five day revisit time at the equator. The swath width of a Sentinel-2 scene is 290 km and the data is acquired in 13 bands with a spatial resolution of 10 m, 20 m and 60 m. For a first comparison of available masking codes, *Function of mask* (Fmask), ATCOR and the scene classification of Sen2Cor are evaluated. All three masking codes use rules that are based on the physical properties such as the *Top of Atmosphere Reflectance* (TOA) in order to differentiate clear pixels from cloud pixels. For the prediction of cloud shadows, the sensor view angle and solar illumination geometry are used.

Furthermore, a special focus is set on the correct and automatic detection of cloud shadows. A new method for cloud shadow detection in multi-spectral satellite images is proposed and compared to current methods. This method is based on the evaluation of *Thresholds*, *Indices* and *Projections*.

Following the detection of cloud shadows, an improved cloud shadow removal algorithm is presented for high spatial resolution optical satellite data over land. It is based on the Matched Filter method which calculates the covariance matrix and a corresponding zero-reflectance matched filter vector. The new cloud shadow map is added to the removal of cloud shadows as well as further evaluations performed on the shadow function to improve the removal algorithm.

Zusammenfassung

Optische Satellitenbilddaten enthalten in vielen Fällen Wolken und Wolkenschatten, sodass eine automatisierte Klassifizierung der Objekte auf der Erdoberfläche erschwert ist. Daher ist es auf der einen Seite wichtig eine gute Wolken- und Wolkenschattenmaske zu erstellen und auf der anderen Seite möglichst die Schattenbereiche so zu korrigieren, dass eine Klassifizierung über die spektralen Eigenschaften der Objekte einheitlich möglich wird. Dazu gehört natürlich auch der Prozess der Atmosphärenkorrektur der optischen Satellitenbilddaten. Insbesondere für Landanwendungen ist die Anzahl der Szenen mit verwertbaren Daten von großer Bedeutung, da der genaue Zeitpunkt von Bedeutung ist (Schätzung von Ernteerträgen) oder weil die Szene nicht kostenlos ist und man für die nächste Erfassung bezahlen muss, wenn die aktuelle Erfassung Wolkenschatten über dem interessierenden Ort enthält.

Die Maskierung von Wolken, Wolkenschatten, Wasser und Schnee/Eis in optischen Satellitenbilddaten ist daher ein wesentlicher Schritt in automatisierten Verarbeitungsketten.

Darüber hinaus ist die genaue Maskierung von Wolkenschatten eine sehr wichtige Aufgabe vor der Entfernung von Wolkenschatten. Da die Erde eine jährliche Wolkenbedeckung von ca. 70% aufweist, ist die Verunreinigung von multispektralen Satellitenbilddaten unvermeidlich, und die Wissenschaftler müssen mit diesem Manko arbeiten und es umgehen.

Für diese Studie werden die Satellitendaten der Sentinel-2-Mission verwendet, die eine fünftägige Wiederholungszeit am Äquator bietet. Die Schwadbreite einer Sentinel-2-Szene beträgt 290 km und die Daten werden in 13 Bändern mit einer räumlichen Auflösung von 10 m, 20 m und 60 m erfasst. Für einen ersten Vergleich der verfügbaren Maskierungscodes, *Function of mask* (Fmask), ATCOR und die Szenenklassifizierung von Sen2Cor ausgewertet. Alle drei Maskierungscodes verwenden Regeln, die auf den physikalischen Eigenschaften wie dem *Top of Atmosphere Reflectance* (TOA) beruhen, um klare Pixel von Wolkenpixeln zu unterscheiden. Für die Vorhersage von Wolkenschatten werden der Blickwinkel des Sensors und die Geometrie der Sonneneinstrahlung verwendet.

Darüber hinaus wird ein besonderer Schwerpunkt auf die korrekte und automatische Erkennung von Wolkenschatten gelegt. Es wird eine neue Methode zur Erkennung von Wolkenschatten in multispektralen Satellitenbilddaten vorgeschlagen und mit aktuellen Methoden verglichen. Diese Methode basiert auf der Auswertung von Schwellenwerten, Indizes und Projektionen.

Nach der Erkennung von Wolkenschatten wird ein verbesserter Algorithmus zur Entfernung von Wolkenschatten für hochauflösende optische Satellitendaten über Land vorgestellt. Er basiert auf der Matched-Filter-Methode, die die Kovarianzmatrix und einen entsprechenden Null-Reflexions-Matched-Filter-Vektor berechnet. Die neue Wolkenschattenkarte wird zur Entfernung von Wolkenschatten hinzugefügt, und es werden weitere Bewertungen der Schattenfunktion durchgeführt, um den Entfernungsalgorithmus zu verbessern.

Acknowledgements

Contents

List of Figures

List of Tables

1 Introduction

The Earth has an annual cloud coverage of approximately 70%. Therefore, satellite imagery is frequently contaminated by low and medium altitude water clouds as well as by high-altitude cirrus clouds in the upper troposphere and in the stratosphere. Such a high cloud coverage in multi-spectral satellite images, as for example given by Sentinel-2, can be seen as a negative effect on many remote sensing tasks. This inevitable contamination limits the ability of a continuous observation of a location on the Earth and degrades the information that can be extracted from a scene. This means, that scientists have to find ways to work around or with the presence of clouds and cloud shadows.

For optical remote sensing of the Earth's surface, clouds and their shadows have always been a major disadvantage, since a lot of remote sensing applications are impacted by their presence and require clear sky pixels as input. These applications involve, for example, image classification, the calculation of surface reflectance or land surface temperature, vegetation indices, etc. Atmospheric correction and land cover classification depend on an accurate cloud map. In addition, maps of water and snow/ice are also indispensable in many applications, i.e. mapping of glaciers and water bodies.

Especially for land applications the amount of scenes with usable data is of high importance. For example because the timing is important (crop yield estimation) or because the scene is not free-of-charge and one has to pay for the next acquisition if the current one contains cloud shadow over the location of interest. The customer of such applications might be interested to determine the ground properties, or use the cloud and cloud shadow free image for geological applications and crop monitoring tasks. Therefore, the correct and exact masking of clouds and cloud shadows is an important preprocessing step for atmospheric correction and also required for the shadow removal of multi-spectral imagery.

Cloud screening is applied to the data in order to retrieve accurate atmospheric and surface parameters as input for further processing steps, either the Atmospheric Correction (AC) itself or higher-level processing such as mosaicking, time-series analysis, or estimation of bio-geophysical parameters. However, a fully automatic detection of these classes is not an easy task, due to the high reflectance variability of earth surfaces. For instance, bright desert surfaces or urban structures can be misclassified as cloud or opposite and shadow surfaces as water. A class assignment for mixed pixels (e.g., semi-transparent cloud over snow) can be problematic because they don't have a spectral signature which clearly belongs to a class. This together will decrease the classification accuracy and shows the need for a performance as-

sessment of classification algorithms.

Several methods of shadow detection have been published over the past years. Most of these methods are applied in the field of computer vision. In remote sensing images, the detection of clouds and their cloud shadows is mostly evaluated with complex multistage processes. But some satellite sensors do not have enough spectral information to properly select appropriate thresholds for cloud shadow detection that will hold for a variety of images.

This proves how important it is nowadays to have a correct masking of clouds and cloud shadows as preprocessing step for atmospheric correction and shadow removal of multi-spectral imagery. Hence, in the past years, more and more cloud and cloud shadow detection and removal approaches have been developed and used to enable various applications.

1.1 Scope of the Dissertation

The objective is the development of an operational method for cloud shadow removal in high spatial resolution (10 – 30 m) optical satellite imagery. The emphasis is on multispectral data (Landsat-8 and Sentinel2), but hyperspectral data should also be tested (DESIS, EnMAP, high altitude 20-km AVIRIS). The first step is an efficient shadow detection algorithm, followed by shadow removal. The current state-of-the-art methods are based on interactive steps not appropriate for automated processing chains. Current alternative methods use multi-temporal datasets and replace shadow areas in a scene with the corresponding parts of a cloud-free scene. This approach has limitations for long time series with large cloud coverage during harvest time or vegetation growth periods, where interpolation in time should not be applied. Therefore, an operational shadow removal method is also needed for mono-temporal scenes.

This thesis is focused on the cloud shadow detection and removal for high spatial resolution optical satellite data. This goal is accomplished by tackling three specific objectives:

1. Objective 1: Evaluation and comparison of Masking Algorithms for Sentinel-Imagery

2. Objective 2: Development of an automated algorithm for Cloud Shadow Detection

3. Objective 3: Development of an operational method for Cloud shadow removal for high spatial resolution optical satellite data

1.2 Guidelines for Reading

The thesis is a cumulative dissertation which is organized as follows:

Chapter 2 gives an introduction to the fundamental knowledge of atmospheric correction, especially cloud and cloud shadow detection, related to this thesis and summarizes the state-of-the-art.

Chapter 3 describes the comparison of three different masking algorithm for atmospheric correction and represents a peer-reviewed journal paper:

1. Zekoll Viktoria, Main-Knorn Magdalena, Louis Jerome, Frantz David, Richter Rudolf, Pflug Bringfried. (2021). Comparison of Masking Algorithms for Sentinel-2 Imagery. Remote Sensing. 13. 137. 10.3390/rs13010137.

Chapter 4 presents the results obtained from the evaluation of a new automated algorithm for cloud shadow detection. This chapter represents the following peer-reviewed journal:

1. Zekoll Viktoria, Reyes Raquel, Richter, Rudolf. (2022). A Newly Developed Algorithm for Cloud Shadow Detection—TIP Method. Remote Sensing. 14. 2922. 10.3390/rs14122922.

Chapter 5 proposes an improved method for cloud shadow removal in optical satellite imagery. It is based on the Matched Filter method with new additions introduced and incoorporates the cloud shadow map presented in Chapter 4. The chapter represents the following submitted paper:

1. Submitted: Viktoria Zekoll. Cloud shadow removal for optical satellite data

Chapter 6 discusses and concludes the cumulative dissertation and gives an outlook on possible future work.

2 Background

> The shadow is the greatest teacher
> for how to come to the light
>
> *Ram Dass*

Very Basic Concept

2.0.1 Preprocessing of Earth Observation data: Radiometry

Preprocessing of the data arriving at the sensor is a step that is fundamental before being able to start with the primary analysis or processing. In digital analysis such preprocessing operations could for example be

1. radiometric preprocessing

2. geometric preprocessing

Radiometric processing helps to adjust the digital values for effects of a hazy atmosphere and geometric preprocessing brings an image into registration with a map or another image. Hence, preprocessing steps improve the image quality for further analysis. It must still be noted that the image data is altered during preprocessing. Hence, even though these changes might seem beneficial, artifacts can be created. [4]

2.0.2 Geometric and Radiometric errors

When image data is recorded by a sensor on satellites and aircraft it can contain errors. These are referred to as **radiometric errors** and **geometric errors**.

Radiometric errors are errors in the measured brightness values of the pixels. They result from the instrumentation used to record data of the wavelength dependence of solar radiation and from the effects of the atmosphere.

Geometric errors arise from the relative motions of the platform, non-idealities in the sensors themselves, the curvature of the earth and uncontrolled variations in the position and from velocity and attitude of the remote sensing platform.[4]

Instrumentation Radiometric Errors

Radiometric distortion can be an effect of mechanisms that affect the measured brightness values of the pixels in an image. The distortion of brightness over an image in a given band can be different from that in the ground scene. Furthermore, the relative brightness of a single pixel from band to band can be distorted compared with the spectral reflectance character of the corresponding region on the ground. Both types of distortion result from the presence of atmosphere as a transmission

medium through which radiation must travel from its source to the sensors. It can also be the result of instrumentation effects.[4]

Sources of Radiometric Distortion

If the radiation detected is ideal, then it should be linear. But real detectors have some degree of non-linearity. There can also be a small signal out, even when no radiation is coming in. This is known as dark current and is the result of residual electronic noise present in the detector at any temperature other than absolute zero. In remote sensing, dark current is known as detector offset. The slope of the detector curve is called its gain. Figure 2.1 shows the transfer characteristics of a detector. The left image (Figure2.1 a) shows an ideal radiation detector, hence a linear radiation detector transfer characteristic. The right image (Figure 2.1 b) demonstrates the hypothetical mismatches in detector characteristics.[4]
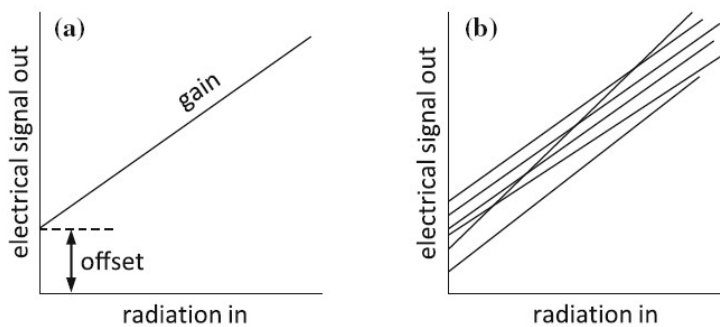
Abbildung 2.1: Transfer characteristics of a detector. a) ideal radiation detector. b) hypothetical mismatches in detector. [4]

Correcting Instrumentation Errors

Errors in relative brightness can be rectified using the following steps: It is first assumed that the detectors used for data acquisition in each band produce statistically similar signals to one another. Then the means and standard deviations are computed for the recorded signals by each of the detectors over the full scene. These should almost be the same and requires the assumption that statistical detail within a band does not change significantly over a distance equivalent to that of one scan covered by the set of the detectors. This is usually a reasonable assumption in terms of the means and standard deviations of pixel brightness. [4]

1. **Striping** is a radiometric error that arises in optical/mechanical scanners. This error appears as a horizontal banding caused by small differences in the sensitivity within the sensor. Within a given band, such differences appear on images as banding where individual scan lines exhibit unusually brighter or darker brightness values that contrast with the background brightness of the normal detectors. Sensor mismatches of this type can be corrected by calculating pixel mean brightness and standard deviation using lines of image data

known to come from a single detector. The correction of radiometric mismatches among the detectors can then be carried out by adopting one sensor as a standard and adjusting the brightness of all pixels recorded by each other detectors. This is called destriping and refers to the application of algorithms to adjust incorrect brightness values to values thought to be near the correct values. The destriping algorithms identify the values generated by the defective detectors by searching for lines that are noticeably brighter or darker than the lines in the remainder of the scene. Then the corrected values for the bad lines are estimated by either local averaging or histogram normalization. In the first strategy, pixels in the defective line are replaced with an average of values of neighboring pixels in adjacent lines. The second strategy of histogram normalization accumulates data from all lines at intervals of six lines (in the case of Landsat MSS). The histogram for defective detectors displays an average different from the others. A correction shifts the values for the defective lines to match the positions for the other lines within the image. Figure 2.2 is a schematic representation of the two destriping strategies.

2. The correction of lost lines of data or lost pixels can also be carried out by averaging over the neighboring pixels. This is called **infilling** and uses the lines on either side for line drop outs or the set of surrounding pixels for pixel drop outs.
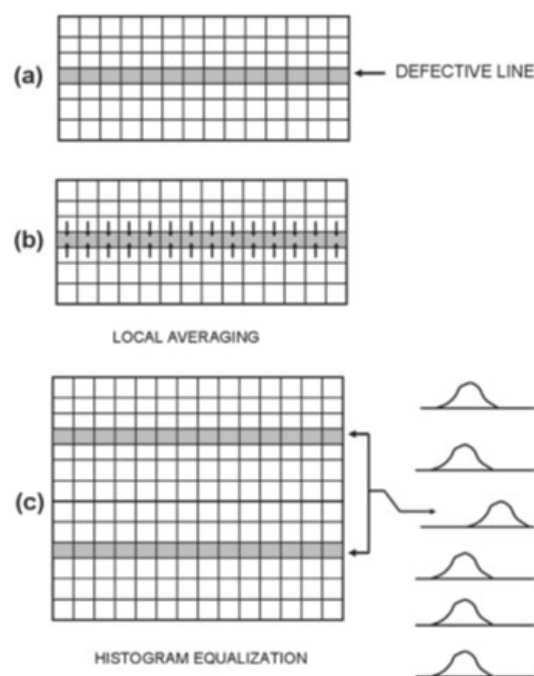
Abbildung 2.2: Two strategies of destriping for the correction of radiometric mismatches among the detectors. a) Defective line is identified with an array of pixels. b) Local averaging. c) Histogram normalization.[4]

2.0.3 Radiometric Preprocessing

Radiometric preprocessing influences the brightness values of an image to correct for sensor malfunctions or to adjust the values to compensate for atmospheric degradation. Any sensor that observes the Earth's surface using visible or near visible radiation will record a mixture of two kinds of brightnesses:

1. the brightness derived from the Earth's surface, which is of interest for remote sensing

2. the brightness of the atmosphere itself, so the effects of atmospheric scattering

This means that an observed digital brightness value can result of part surface reflectance and part atmospheric scattering. The ability to distinguish between the two brightnesses is one objective of atmospheric correction. Hence, to identify and separate these two components.[4]

2.0.4 Solar Curve and the Effect of the Atmosphere

The surface material reflectance spectrum encounters distortion by the spectral dependence of the solar curve and the effect of the atmosphere. This is demonstrated in Figure 2.3. Figure 2.3a shows the detection of the solar curve from a perfectly reflecting surface in the absence of an atmosphere. In Figure 2.3 b we see the effect of the atmosphere on detecting the solar curve and in Figure 2.3 c the detection of the real spectrum distortion by the atmosphere and the solar curve is demonstrated.

Compensation for the Solar Radiation Curve

The wavelength dependence of the radiation falling on the Earth's surface can be compensated by assuming that the Sun is an ideal black body. Most radiometric correction procedures compensate for the solar curve using the actual wavelength dependence above the atmosphere. The common measured solar spectral irradiance of the sun above the earth's atmosphere over the wavelength range is shown in Figure 2.4.[4]

2.0.5 Preprocessing operations

There are three main categories for preprocessing operations to correct for atmospheric degradation:

Preprocessing operations: Radiative Transfer Code

Radiative transfer code (RTC) computer models are procedures that permit the observed brightnesses to be adjusted to approximate true values. This enables to improve the image quality and accuracy's of analyses. The RTC models the physical process of scattering at the level of individual particles and molecules. Even though the model having many advantages with respect to rigor, accuracy and applicability,
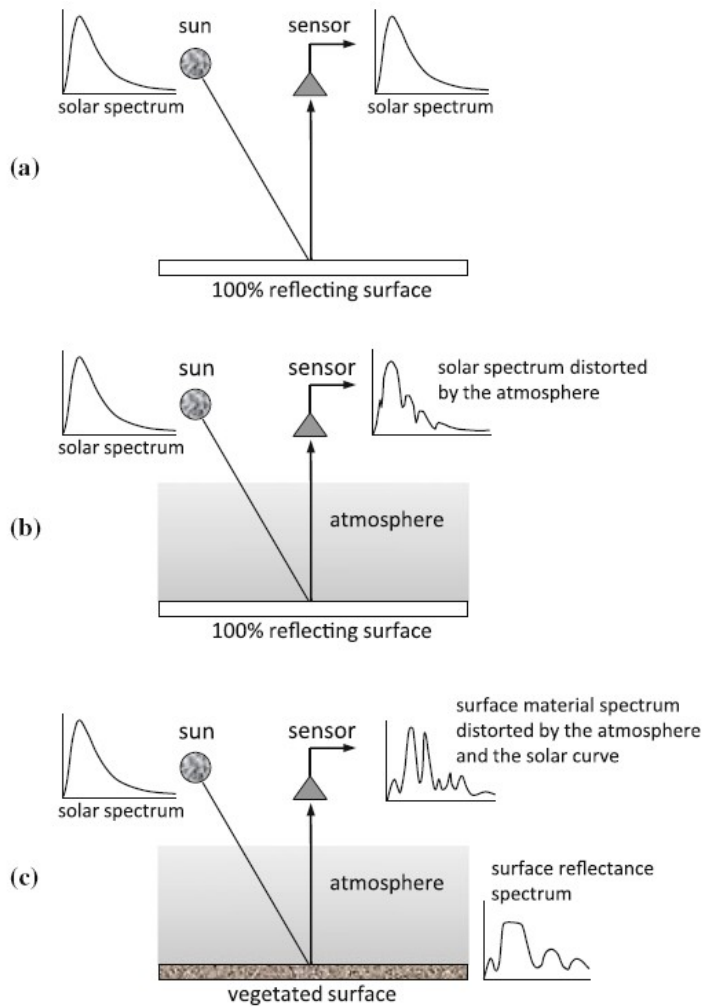
Abbildung 2.3: Solar Curve and effect of the atmosphere. a) solar curve from perfectly reflecting surface without atmosphere. b) solar curve from perfectly reflecting surface with atmosphere. c) real spectrum distorted by atmosphere and solar curve.[4]

it also comes with one significant disadvantage: RTC are often complex and require detailed in situ data to be acquired simultaneously with the image and/or satellite data describing the atmospheric column at the time and place of acquisition of an image.[4]

Preprocessing operations: Image-based Atmospheric Correction

Image-based atmospheric correction is another approach that examines spectra of objects of known or assumed brightness which is recorded by multi-spectral imagery. It relies on the basic principles of atmospheric scattering: scattering is related to wavelength, sizes of atmospheric particles and their abundances. hence, if a known target is observed using a set of multi-spectral measurements, the relationships between values in separate bands can give atmospheric effects. This strategy can be implemented by identifying a very dark object or feature within the scene. This object
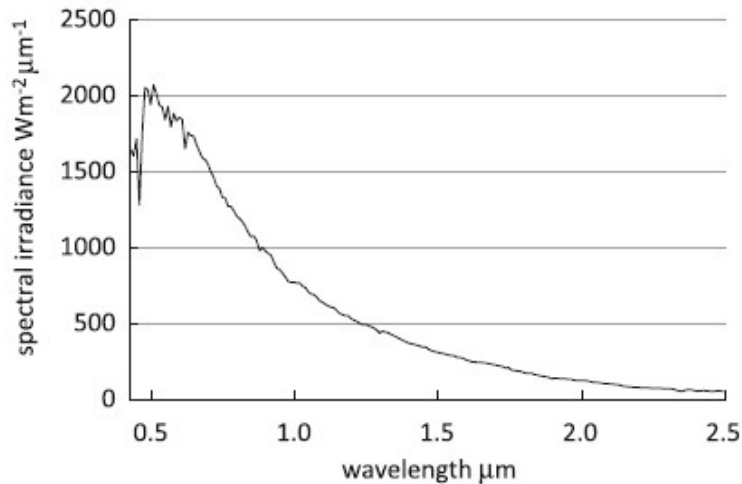
Abbildung 2.4: Solar spectral irradiance of the sun above the Earth's atmosphere vs. wavelength range. [4]

can be a large water body or shadows cast by clouds or large topographic features. In the infrared part of the spectrum, water bodies and shadows have brightness at or very near zero. Hence, analysing the histograms of the digital values for a scene in such areas shows that the lowest values are not zero but some larger value. This value typically differs from one band to the next.[4]

Preprocessing operations: Dark Object Subtraction

It is assumed that the lowest non zero values represent the contribution from atmospheric scattering for each band. They can be subtracted from all digital values for that scene and band. Hence, the lowest value in each band is set to zero. These direct methods for adjusting digital values for atmospheric degradation is known as the histogram minimum method (HMM) or dark object subtraction (DOS) technique. Since this method will remove haze and improves the dynamic range of image intensity it is also referred to as haze removal. Figure 2.5 shows the HMM for correction of atmospheric effects. The lowest brightness value in a given band is taken to indicate the added brightness of the atmosphere to that band and is then subtracted from all pixels in that band. Figure 2.5 a shows the histogram for an image acquired under clear atmospheric conditions. The darkest pixel is near zero brightness. On the other hand, in 2.5 b the darkest pixels are relatively bright, due to the added brightness of the atmosphere. This histogram therefore correspond to an image acquired under hazy atmospheric conditions.
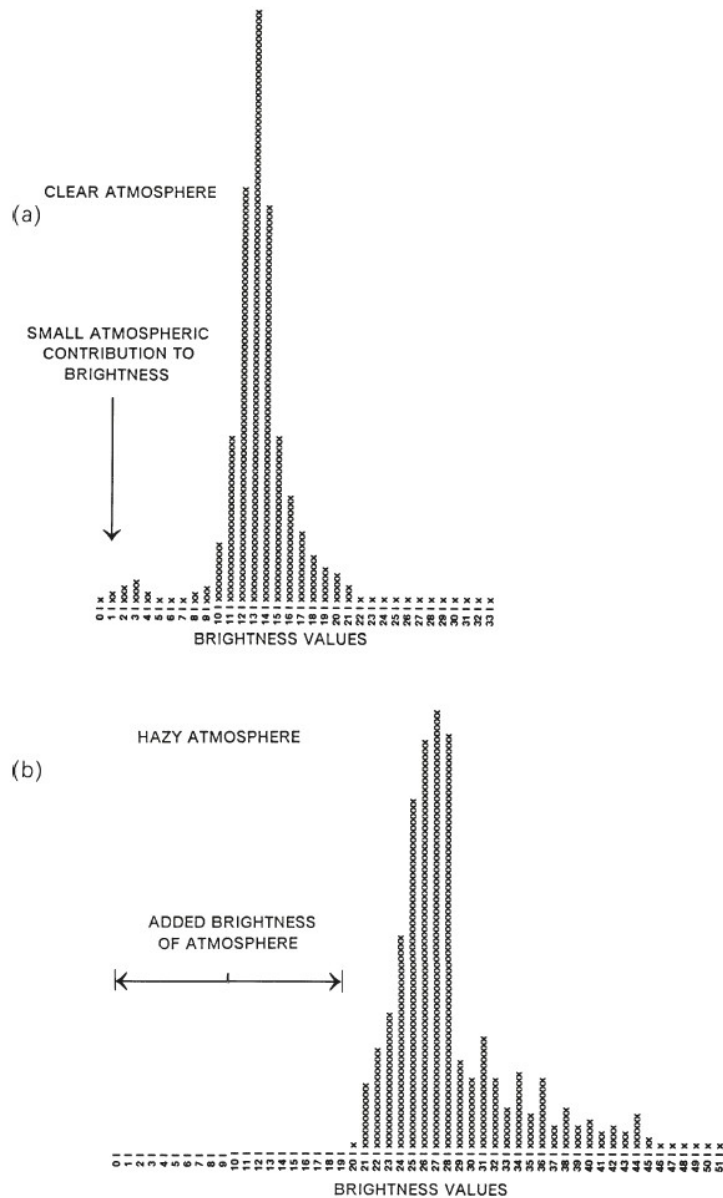
Abbildung 2.5: Histogram minimum method or Dark Object Subtraction for the correction of atmospheric effects. a) Histogram obtained from clear sky atmospheric conditions. b) Histogram obtained from hazy atmospheric conditions.[4]

The DOS procedure has the advantages of simplicity, directness and universal applicability. Nevertheless it must be considered as an approximation and is capable of correction for additive effects of atmospheric scattering but not for multiplicative effects. It must also be noted that the technique of subtraction of a constant from all pixels in a scene will have larger proportional impact on the spectra of dark pixels than on bright pixels. Care must be taken when using this operation for scenes in which spectral characteristics of dark features are present.[4]

Preprocessing operations: Regression Technique

Another approach is known as regression technique. It examines the brightness of objects within each scene but attempts to exploit knowledge of interrelationships between separate spectral bands. [58] came up with a procedure that paired values from each band with values from a NIR spectral channel. This technique can be applied to local areas.[4]

Preprocessing operations: Covariance Matrix Method

The analysis of the variance-covariance matrix, known as covariance matrix method (CMM), is an extension of the regression technique. Both are based on the assumption that within a specified image region, variations in image brightness are due to topographic irregularities and that the reflectivity is constant. Hence, variations in brightness are caused by small-scale topographic shadowing, and the dark regions reveal the contributions of scattering to each band.[4]

2.0.6 Estimation of Top of Atmosphere Reflectance

An accurate measurement of the brightness, whether measured as radiance or as DNs, is not ideal since the values are subject to modification by differences in sun angle, atmospheric effects, and angle of observation. It is much better to observe the reflectance. It gives the proportion of radiation reflected from varied objects relative to the amount of the same wavelengths incident upon the object. The reflectance helps to define the distinctive spectral characteristics of objects.

The reflectance can be described as the relative brightness of a surface as measured for a specific wavelength interval. It is the ratio of the amount of light leaving a target to the amount of light striking the target:

$$Reflectance = \frac{\text{Observed brightness}}{Irradiance} \tag{2.1}$$

Since the reflectance is a ratio it is a dimensionless number between 0 and 1. Often the reflectance is given as percentage. The reflectance is not directly measurable since the observed brightness is normally recorded and not the brightness incident upon the object.

The radiance gives the variable that is directly measured by remote sensing. It is the amount of light the instrument detects from the object being observed. The observed radiance might appear bigger when some light gets scattered by the atmosphere and hence seen by the instrument. On the other hand, the radiance can decrease if light gets absorbed by the atmosphere. The radiance therefore depends on the illumination, the orientation and position of the target and the path of the light through

the atmosphere.[4]

2.1 Atmospheric Correction Background

2.1.1 Basic concept of atmospheric correction

To understand the basic concept of atmospheric correction it is important to realize the importance of the radiation components and the relationship between the at-sensor radiance and the digital number (DN). The at-sensor radiance can also be named top-of-atmosphere (TOA) radiance. In the current state of the art satellite and airborne ATCOR model, the input data are the recorded scene plus a meta file containing the acquisition date, time, solar and sensor view geometry etc. The input data is given as scaled TOA radiance, named Digital Number (DN).

The data is scaled to a 16-bit dynamic range in order to convert the relative pixels DNs coming directly from the sensor into values directly related to absolute TOA radiances. The scaling factor is applied so that the resultant single DN values correspond to 1/100th of a $[W/(m^2 * sr * \mu m)]$. In order to convert the pixel values given by the input data into TOA radiance, the DN has to be multiplied by the radiometric calibration coefficient, c_1, also known as radiometric gain.

For each spectral band of a sensor, a linear equation exists which describes the relationship between the recorded DN (or brightness) and the TOA radiance, L. It is obtained with the radiometric offset, c_0, and gain, c_1 as follows and has units of $[W/(m^2 * sr * \mu m)]$:

$$L = c_0 + c_1 DN \tag{2.2}$$

For some instruments, the data can also be given as top-of-atmosphere (TOA) reflectance ρ_{TOA} also called apparent or at-sensor reflectance, ρ^*:

$$\rho_{TOA} = \rho^* = \frac{\pi L d^2}{E_s \cos \theta_s} \tag{2.3}$$

where d is the Earth-Sun distance in astronomical units, E_s is the extraterrestrial solar irradiance, and θ_s is the solar zenith angle. This TOA reflectance is also called apparent reflectance, since it gives the combined Earth-atmosphere behavior with respect to the solar radiation for the specified view geometry. Hence, the ratio between the upwelling and downwelling flux.[57] Solar irradiance, E (W /m2), is defined as the power/area in the incident direction, whereas the flux, F (W/m2), is on a horizontal surface, i.e $F = E * \cos \theta_s$. For TOA there is no incoming diffuse flux and the incoming direct flux on the horizontal surface is $E_s \cos \theta_s$. The mea-

sured radiation in equation 2.36 is radiance, L, from a certain direction and not a hemispherical flux.

In ATCOR, if the input data is TOA reflectance, it will be converted into TOA radiance before starting the process of atmospheric correction. This conversion is given by equation 2.4, solving equation 2.36 for L, and will further be analysed in section 5.2.1.

$$L = \frac{\rho^* \, E_s \, cos\theta_s}{\pi \, d^2} \tag{2.4}$$

For Sentinel-2, the input data is TOA reflectance (ρ^*), but for most sensors the input data is radiance. So equation 2.4 is only applied to Sentinel-2 in the loading function. The radiometric coefficients c_0 and c_1 include some sensor-specific scaling factors.

2.1.2 Visibility and optical thickness

The visibility and optical thickness are two parameters that are often used to describe the atmosphere. As visibility we understand the maximum horizontal distance a human eye can distinguish an object or light from its background. The visibility, or meteorological range, is defined through a relation between two parameters:

- the extinction coefficient, β, in units of $[km^{-1}]$ and at 550 nm

- a contrast threshold of 0.02 giving the minimum observable contrast between a dark object against the bright horizon sky

This relation defining the visibility is also called the Koschmieder equation:

$$VIS = \frac{1}{\beta} \ln \frac{1}{0.02} = \frac{3.912}{\beta} \tag{2.5}$$

The optical thickness is given by a pure number and is usually evaluated at a wavelength of 550 nm (which corresponds to a typical yellow light). The vertical optical thickness of the atmosphere, δ, is given by the product of the extinction coefficient and the path length, x:

$$\delta = \beta x \tag{2.6}$$
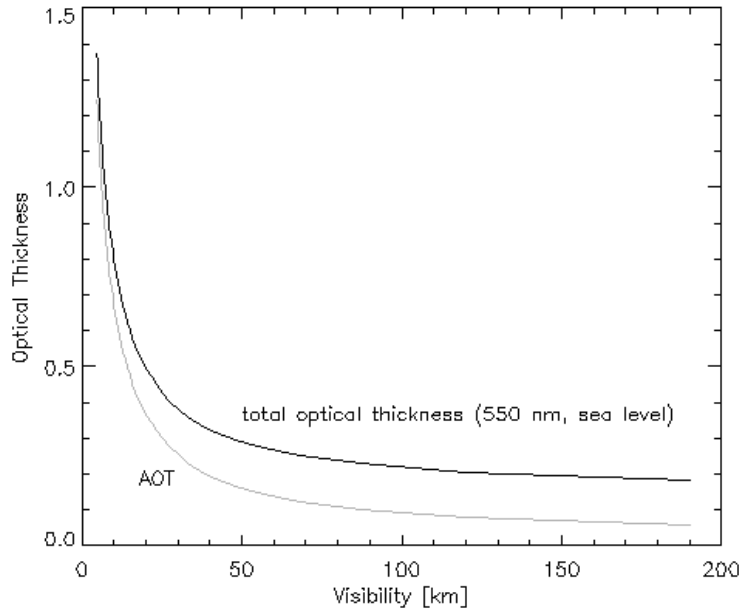
Abbildung 2.6: Relationship between the horizontal visibility and the vertical optical thickness of the atmosphere for a path from sea level to space.[72]

Figure 2.6 shows the relationship between the horizontal visibility and the vertical optical thickness of the atmosphere for a path from sea level to space and for the wavelength of 550 nm.

Since the atmosphere has different constituents, the total optical thickness is given by the sum of each individual contributor. These contributors are molecular (Rayleigh) scattering, aerosol scattering and molecular absorption. Hence:

$$\delta = \delta(\text{molecular scattering}) + \delta(\text{aerosol scattering}) +$$
$$\delta(\text{molecular absorption})$$

(2.7)

The aerosol content is scaled by the MODTRAN$^{©}$ visibility parameter into the boundary layer of 0 to 2 km altitude. For visibilities greater than 100 km the total optical thickness asymptotically approaches a value of about 0.17 which (at 550 nm) is the sum of the molecular thickness ($\delta = 0.0973$) plus ozone thickness ($\delta = 0.03$) plus a very small amount due to trace gases, plus the contribution of residual aerosols in the higher atmosphere (2 - 100 km) with $\delta = 0.04$.

The total optical thickness (sea level to space) at 550 nm asymptotically approaches a value of ~ 0.17. This value results from the sum of the molecular thickness, ozone thickness, a small amount from trace gases and the residual aerosols present in the higher atmosphere (2-100 km).

A minimum optical thickness and hence the highest visibility is obtained at the Rayleigh limit, when the air contains no aerosol particles. The optical thickness is then due to molecular thickness and ozone. δ(molecular scattering) coming from nitrogen and oxygen is dependent on the pressure level and can therefore be calculated if the ground elevation is known. The contribution from ozone is small at 550 nm and hence a climatological and geographic average can be taken.[72]

So the most common parameter used to describe the atmosphere is the aerosol optical thickness (AOT) at 550 nm. The aerosol contribution varies strongly in space and time and is therefore the most important component apart from the atmospheric water vapor.

If a beam goes through the atmosphere it will encounter transmittance and absorption, depending on the spectral regions. The atmospheric transmittance for a vertical path through the atmosphere is given as:

$$\tau = e^{-\delta} \tag{2.8}$$

A high transmittance corresponds to a spectral region with small or no absorption and is called a "spectral window".

2.1.3 Radiation components and surface reflectance: Flat terrain

The radiation signal in the solar region (0.35-2.5 μm) arriving at the sensor is due to three different components in a flat terrain condition (see Figure 2.7:

- Path radiance (L_p): from photons that did not have contact with the ground and are scattered into the field-of-view of the sensor

- Reflected radiation from a pixel (L_g): the fraction of the diffuse and direct solar radiation incident on the pixel that is getting reflected from the surface

- Reflected radiation from the surrounding (L_a): the fraction of the solar radiation reflected from the neighborhood and scattered by the air volume into the field-of-view of the sensor. This radiation is also called adjacency radiation.
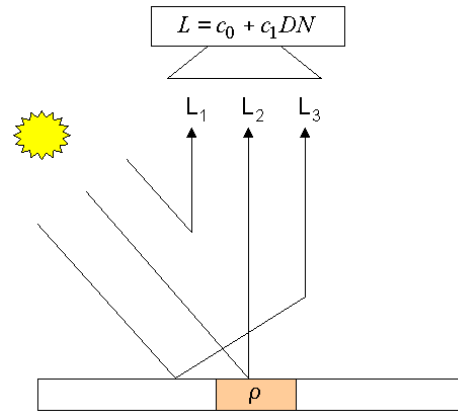
Abbildung 2.7: Solar radiation components arriving at the sensor. L_1: path radiance, L_2: reflected radiance, L_3: adjacency radiation.[72]

The three components show that only the reflected radiation from a pixel contains the necessary information about the viewed pixel. Hence, in atmospheric correction, it is important to remove the other components and to retrieve the correct ground reflectance from the pixel of interest.

If we now combine all three components of the radiation to get the total radiation arriving at the sensor we can write:

$$L = L_p + L_g + L_a \tag{2.9}$$

All components have different dependencies with respect to the spectral range. The path radiance decreases with increasing wavelength and is very small for $\lambda > 800$ nm. The adjacency radiation depends on the brightness difference between the pixel of interest and the surrounding. Its value also decreases with wavelength and is very small for $\lambda > 1.5\ \mu$m.

An iterative approach for the atmospheric correction in a flat terrain is necessary, since the surface reflectance and large-scale (1-2 km neighborhood) background reflectance are unknown. Therefore, ATCOR computes the ground reflectance in three steps. These are explained in 2.1.3, 2.1.3 and 2.1.3. In section 2.1.5 the surface reflectance retrieval is explained in more detail.

Ground reflectance calculation: step 1
In the first iteration step of the ground reflectance calculation, the adjacency effect from the neighborhood is neglected. The total radiation can then be simplified as follows:

$$L = L_p + L_g = L_p + \frac{\tau \rho E_g(\rho_r = 0.15)}{\pi}$$
$$= c_0 + c_1 DN \tag{2.10}$$

where τ is the atmospheric transmittance from the ground to the sensor with values in the range from [0 - 1], ρ is the surface reflectance (Bottom-of-atmosphere BOA) and E_g is the global flux on the ground in units of $[mW\,cm^{-2}\mu m^{-1}]$ calculated for a reference background reflectance. The global flux consists of the direct solar and diffuse solar fluxes on the ground, and the surface is assumed to be a Lambertian i.e., perfectly isotropic, reflector. The path radiance is given in units of $[mW\,cm^{-2}sr^{-1}\mu m^{-1}]$.

If we rearrange the radiation equation for the surface (BOA) reflectance we get:

$$\rho^{(1)} = \frac{\pi(d^2(c_0 + c_1 DN) - L_p)}{\tau E_g(\rho_r = 0.15)} \tag{2.11}$$

The factor of d^2 has to be added in order to account for the Sun-to-Earth distance, where d is given in astronomical units. In ATCOR the Lookup Tables (LUT's) for the path radiance and the global flux are calculated using d=1.

Ground reflectance calculation: step 2

The next steps calculates an average surface reflectance and, as opposed to the previous step, in a large neighborhood where each pixel is set to a default range of 1 km. The surface reflectance can be written as follows:

$$\bar{\rho} = \frac{1}{N^2} \sum_{i,j=1}^{N} \rho_{i,j}^{(1)} \tag{2.12}$$

where N is the number of pixels in the selected range of the adjacency effect. Equation 2.12 has a range independent weighting. This can of course be changed into a range dependent function as is demonstrated in [72]. Here I will only state the range independent function. Since the adjacency effect is only a small second-order effect, the differences to the range dependent approach is very small. The surface reflectance after step 2 can now be written as:

$$\rho^{(2)}(x,y) = \rho^{(1)}(x,y) + q(\rho^{(1)} - \bar{\rho}(x,y)) \tag{2.13}$$

where q is a measure for the strength of the adjacency effect.

Ground reflectance calculation: step 3

The last step of the ground reflectance calculation includes the spherical albedo effect on the global flux. Instead of using the reference background reflectance of $\rho_r = 0.15$ as in section 2.1.3, it is now adapted to the scene dependent value of section 2.1.3, $\bar{\rho}$. This is done by correcting for the difference between the two values as follows:

$$\rho^{(3)}(x,y) = \rho^{(2)}(x,y)(1 - (\bar{\rho}(x,y) - \rho_r)s) \qquad (2.14)$$

2.1.4 Radiation components and surface reflectance: Rugged terrain

For the retrieval of the surface reflectance in mountainous terrain, a fourth radiation component has to be taken into account. In rugged terrain, the pixel-dependent radiation is specified through a *digital elevation model* (DEM) since the incident radiation is strongly dependent on the topographic variation. Hence, the radiation signal arriving at the sensor is not only due to the path radiance,(L_p), the ground reflected radiation,(L_g), and the radiation reflected from the surrounding,(L_a), but also due to the reflected radiance from opposite mountains, (L_t). So the total measured at-sensor radiance can be written as follows:

$$L(x,y,z,\lambda) = L_p(x,y,z,\lambda) + L_g(x,y,z,\lambda) + L_a(x,y,z,\lambda) + L_t(x,y,z,\lambda) \qquad (2.15)$$

where the planar and height coordinates are denoted with x,y and z and λ denotes the wavelength dependency. For brevity, the spatial coordinates and wavelength will be omitted.

The additional terrain radiance component, L_t, is evaluated with the terrain view factor, V_t, which depends on the sky view factor, V_{sky}, as follows:

$$V_t = 1 - V_{sky} \qquad (2.16)$$

The sky view factor can be calculated through a simple approximation based on the local DEM slope angle, θ_n (see equation 2.17).

$$V_{sky}(x,y) = cos^2(\theta_n(x,y)/2) = \frac{1 + cos\theta_n(x,y)}{2} \qquad (2.17)$$

The radiation component due to the reflected terrain radiance can then be written

as:

$$L_t = TE_t^{(1)}V_t\rho_t/\pi \tag{2.18}$$

where T represents the total ground-to-sensor transmittance, $E_t^{(1)}$ the global terrain flux and V_t the terrain-view factor.

Again, as performed in section 2.1.3, the adjacency effect from the neighborhood is neglected. Hence, the ground reflectance, ρ_g, which depends on L_g is calculates as follows:

$$\rho_g = \frac{\pi((d^2(c_0 + c_1 DN) - L_p)}{T(bE_s cos\beta + E_d^*)} \tag{2.19}$$

where b=0 for shadow pixels, else b=1, E_s represents the direct solar beam irradiance at the ground and E_d^* represents the total diffuse flux due to the slope surface. E_d^* is calculated as shown in equation 2.20.

$$E_d^* = E_d[b\tau_s cos\beta/cos\theta_s + 1 - b\tau_s V_{sky}] \tag{2.20}$$

Here, E_d represents the diffuse solar flux on an horizontal plane and τ_s denotes the sun-to-ground direct beam transmittance. The first part of equation 2.20 describes the anisotropic circumsolar radiation and the second part the isotropic sky radiation.

The first order correction with the terrain influence is included in the first iteration of equation 2.19 as follows:

$$\rho^{(1)} = \frac{\pi((d^2(c_0 + c_1 DN) - L_p)}{T(bE_s cos\beta + E_d^* + E_t^{(1)})} \tag{2.21}$$

This iteration can be done a second time in order to account and correct for many terrain reflection effects with additional approximations included in E_t. This will update the ground reflectance accordingly and the steps performed in section 2.1.3 and section 2.1.3 can be performed with this new ground reflectance.

2.1.5 Surface Reflectance Retrieval

Reflectance and reflectance factor

As mentioned in section 2.1.1, the spectral radiance is defined as the radiant flux in a beam per unit wavelength and per unit area and solid angle of that beam [75]. It has SI units of $[Wm^{-2}sr^{-1}nm^{-1}]$. The reflectance is then given by the ratio of

the observed brightness or upwelling band-integrated irradiance ($[Wm^{-2}]$) with the downwelling irradiance ($[Wm^{-2}]$) and has values that lie within the interval of [0-1]. In case of spectral irradiance or spectral flux we have the SI unit ($W\,m^{-2}\,\mu m^{-1}$). The reflectance factor is defined as the ratio of

- the radiation flux reflected by a surface and

- the radiation flux reflected into the same reflected-beam geometry and wavelength range by an ideal (lossless) and diffuse (Lambertian) standard surface, irradiated under the same conditions

The values of the reflectance factor can be greater than 1. This is especially the case for surfaces of high reflectance such as snow [62].

If an isotropic behavior (ideal or lossless) is assumed, then a spherical surface will radiate with same intensity in ($[Wm^{-2}]$) in all directions. If on the other hand a diffuse or Lambertian behavior is assumed, then this will imply a flat reflective surface. In such a case, the intensity will decrease as the cosine of the observation angle decreases with respect to the surface normal. Hence, the intensity will be greatest if the normal vector and the observation angle point in the same direction ($\cos 0 = 1$). This is also known as the Lambert's law. Furthermore, this diffuse behavior also implies that the radiance, L, is independent of direction. But is must be noted that the ratio between the upwelling irradiance ($[Wm^{-2}]$) and the radiance, L in $[Wm^{-2}sr^{-1}]$, of this Lambertian surface gives a factor of π. This is due to the cosine factor in the definition of the radiance [63].
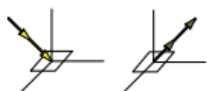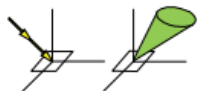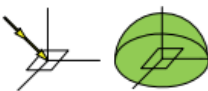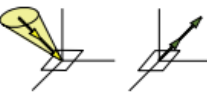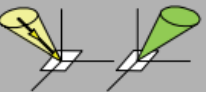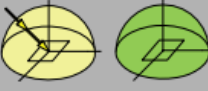
Incoming/Reflected	Directional	Conical	Hemispherical
Directional	Bidirectional CASE 1	Directional–conical CASE 2	Directional–hemispherical CASE 3
Conical	Conical–directional CASE 4	Biconical CASE 5	Conical–hemispherical CASE 6
Hemispherical	Hemispherical–directional CASE 7	Hemispherical–conical CASE 8	Bihemispherical CASE 9

Abbildung 2.8: Terminology of the reflectance quantities, which are separated into nine different cases according to the relation of the incoming and reflected radiation. Grey fields represent the quantities that can be measured and white fields are conceptual quantities. [75]

The angular behavior of the incoming radiance is named and classified into nine cases. This naming is according to [57] and is illustrated in Figure 2.8. The naming of the term follows a simple order of first naming the angular characteristic of the incoming radiance which is then followed by the angular characteristic of the reflected radiance.

The bidirectional reflectance distribution function (BRDF)

When the surface reflectance is getting measured remotely it is dependent on the illumination and observation geometry. Its angular behavior is described by the bidirectional reflectance distribution function (BRDF). This corresponds to the first case illustrated in Figure 2.8. If we now integrate over the small finite incidence and observation angles we get the fifth case of Figure 2.8, called biconical reflectance factor (BCRF).

The Hemispherical directional reflectance factor (HDRF)

Hemispherical illumination conditions are the conditions under which remote measurements of the earth's surface are performed. This corresponds to a diffuse radiation combined with a directional solar irradiance. Hence, to best define the retrieved reflectance quantity, the hemispherical directional reflectance factor (HDRF) is used. This corresponds to case 7 of the terminology in Figure 2.8. The HDRF can only be used in an isotropic hemispherical radiation field and is defined for an infinitesimal solid angle. For measurements with a finite angle, the hemispherical conical reflectance factor (HCRF) is used. This is illustrated in Figure 2.8 as case 8.

Lambertian reflectance surface

For a perfectly isotropic, Lambertian, 100% reflecting surface, the HDRF, HCRF and BCRF are equivalent and equal to the BRDF times a factor of π.

2.2 Methods

2.2.1 Matched Filter Method

Main steps of MF

The main steps of the matched filter (MF) approach are shadow area detection and then restoration. The detection is done through model-based and shadow feature-based methods. In multi-temporal data, non-shadow pixels from one image are used to replace shadow pixels in the other image.

So the two main steps of the MF Method are:

- Atmospheric correction of scene

- MF algorithm uses sequence of 8 processing steps: surface reflectance; exclude cloud and water; matched filter vector; unscaled shadow function Φ;

scaled shadow function Φ^*; threshold Φ_T (core shadow function); expand shadow mask; de-shadowing with Φ^*

Introduction of MF

De-shadowing is the compensation process which uses an estimate of the fraction of direct solar irradiance for a fully or partially shadowed pixel.

The MF method needs at least one channel in visible and at least one spectral band in NIR. The bands used in the MF are: blue, green, red, NIR, SWIR1 and SWIR2, if existing.

The method starts with the calculation of the surface reflectance image cube, ρ. The surface reflectance, ρ, is computed with the assumption of fully solar illumination, excluding water and clouds. Then the covariance matrix, $C(\rho)$, is calculated where ρ represents the surface reflectance vector of the three selected bands (see equation 2.22).

$$V_{MF} = \frac{C^{-1}(\rho_t - \overline{\rho})}{(\rho_t - \overline{\rho})^T C^{-1}(\rho_t - \overline{\rho})} \qquad (2.22)$$

The MF vector, V_{MF}, is tuned to a certain target reflectance spectrum, ρ_t, to be detected. $\overline{\rho}$ is the scene-average spectrum without water and cloud pixels. For the shadow target, a target reflectance spectrum of $\rho_t = 0$ is selected which will give the simplified version of the shadow MF vector, V_{sh}, as follows:

$$V_{sh} = -\frac{C^{-1}\overline{\rho}}{\overline{\rho}^T C^{-1}\overline{\rho}} \qquad (2.23)$$

Method: surface reflectance and covariance matrix

The MF methods starts, as already mentioned, with the calculation of the surface reflectance image cube, ρ. ρ is computed with the assumption of fully solar illumination and with the exclusion of water and cloud pixels. Then the covariance matrix, $C(\rho)$, is calculated where ρ represents the surface reflectance vector of the three selected bands. The MF vector (see equatiion 2.22) is tuned to a certain target reflectance spectrum, ρ_t, to be detected. $\overline{\rho}$ is the scene-average spectrum without water and cloud pixels. For a shadow target, $\rho_t = 0$ is selected which will give the simplified version of the shadow MF vector, V_{sh} (see equation 2.23).

The covariance is a measure of how much two random variables vary together such as for example the height of a person and the weight of a person in a population. The covariance, $\sigma(x, y)$, of two random variables x and y with n samples is given by:

$$COV(X,Y) = \frac{\sum_{i=1}^{n}(X_i - \overline{X})((Y_i - \overline{Y})}{n-1} \tag{2.24}$$

With the covariance it is possible to calculate the entry of the covariance matrix which is a square matrix. The covariance matrix is given as $C_{i,j} = \sigma(x_i, y_i)$, where $C \in \mathbb{R}^{dxd}$ and d is the dimension or number of random variables of the data. $C_{i,j}$ is a symmetric matrix and the diagonal entries are called variance and all other entries are called covariance.

The covariance matrix can represent the relation between all variables while inverse covariance shows the relation of the element with their neighbours.

Method: unscaled shadow function

The MF shadow vector, V_{sh}, can be applied to the non-water and non-cloud part of the scene to give the un-normalized values, Φ, also called unscaled shadow function. Φ gives a relation measure of the fractional direct illumination (see equation 2.25).

$$\Phi(x,y) = V_{sh}^{T}(\rho(x,y) - \overline{\rho}) \tag{2.25}$$

The MF calculates a minimum RMS shadow target abundance for the entire scene. The values of Φ can be both, positive and negative. Therefore, Φ is rescaled into the physical range from 0 (full shadow) to 1 (full direct illumination).

Method: rescaling and scaled shadow function

The MF calculates a minimum RMS shadow target abundance for the entire scene. The values of Φ can be both, positive and negative. Therefore, Φ is rescaled into the physical range from 0 (full shadow) to 1 (full direct illumination). The histogram of the unscaled shadow function is used for rescaling (see Figure 2.9).

The first peak of the histogram of Φ represented in Figure 2.9, Φ_2, represents the shadow pixels. On the other hand, the highest peak of the histogram in Figure 2.9, ϕ_{max}, represents the fully illuminated areas.

The rescaling of Φ is done by linear mapping of the Φ values from the unscaled interval (Φ_{min}, Φ_{max}) onto the physically scaled interval (0,1). Hence, the scaled shadow function, Φ^{*}, is calculated as follows:

$$\Phi^{*} = \frac{\Phi - \Phi_{min}}{\Phi_{max} - \Phi_{min}} \quad \text{if} \quad \Phi \leq \Phi_{max}$$
$$\Phi = 1 \quad \text{if} \quad \Phi > \Phi_{max} \tag{2.26}$$
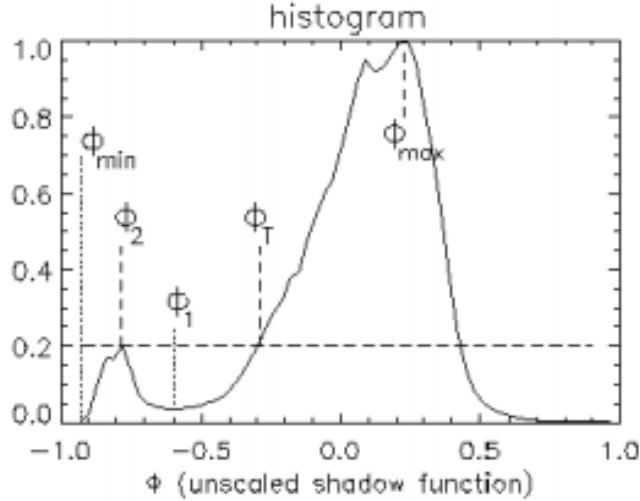
Abbildung 2.9: Histogram of the unscaled shadow function, Φ. [72]

The scaling and normalizing of the MF vector into the (0,1) interval is based on the assignment of:

- min and max direct sun fraction in the shadow regions (a_{min} and a_{max}; the defaults are $a_{min} = 0.20$ and $a_{max} = 0.95$)

- corresponding shadow thresholds Φ_{min} and Φ_{max} obtained from the normalized histogram of Φ

The normalized and scaled shadow function is given as:

$$\Phi_n = \frac{(\Phi_{max} - \Phi_{min})(\Phi - \Phi_{min})}{a_{max} - a_{min}} \quad (2.27)$$

Method: iteration

The potential shadow pixels are those which satisfy $\phi_n < 1$, but as can be seen from equation 2.27, the value strongly depends on Φ_{max}. Therefore an iterative strategy is applied.

De-shadowing reflectance equation

The scaled shadow function, Φ^*, represents the fraction of the direct illumination for each pixel in the surface reflectance vector, ρ. The MF method tries to find the core shadows and them subsequently expands the core regions. This enables a smooth shadow to clear transition. The scaled shadow function is only applied to the pixels in the final mask. The core shadow mask is defined by the pixels with $\Phi(x,y) < \Phi_T$.

The de-shadowing is then performed by multiplying the direct illumination, E_{dir}, with the pixel-dependent Φ^*. This reduces the direct solar term and increases the

brightness of a shadow pixel, since it is located in the denominator of the de-shadowing equation (see equation 2.28).

$$\rho_i(x,y) = \frac{\pi(d^2(c_0(i) + c_1(i)DN_i(x,y)) - L_{p,i})}{\tau_i(E_{dir,i}\Phi^*(x,y) + E_{dif,i})} \qquad (2.28)$$

Equation 2.28 gives the surface reflectance which can be converted back to DN as follows:

$$DN_{ds} = \frac{L_p + (\frac{\tau\rho_{ds}}{\pi})(E_{dir} + E_{dif})/d^2 - c_0}{c_1} \qquad (2.29)$$

2.2.2 Adler-Golden MF and Deshadowing Method

The de-shadowing method by [12] consists of five steps:

- atmospheric correction of the original radiance data to reflectance

- application of a MF to the reflectance data for a black surface

- re-balancing of data to correct skylight skewing effect

- re-applying the MF to refine the shadow fraction

- de-shadow the image by using the shadow fraction

Endmember unmixing Method

In the de-shadowing method given by [12], the MF for a zero reflectance target is applied to the reflectance image. The pixel reflectance spectrum is then given as the mixture of linearly independent and fully illuminated endmember spectra $\mathbf{e_i}$ as follows:

$$y = \sum_i w_{iy}\mathbf{e_i} \qquad (2.30)$$

where w_{iy} is the ith endmember weight of pixel.

In the endmember unmixing method, shadows are zero reflectance endmembers. Hence, the shadow fraction, f, is given as:

$$f_y = \sum_i w_{ia} - \sum_i w_{iy} = -\sum_i (w_{iy} - w_{ia}) \qquad (2.31)$$

In equation 2.31, a represents the fully illuminated pixel.

In unconstrained least-squares unmixing, the ith endmember weight is extracted

from each spectrum by taking its dot product with a filter vector, $\mathbf{v_i}$. The filter vector is orthogonal to all endmembers except $\mathbf{e_i}$. Hence, equation 2.31 can be rewritten as follows:

$$f_y = -\sum_i \mathbf{v_i^T}(\mathbf{y} - \mathbf{a}) = \mathbf{g^T}(\mathbf{y} - \mathbf{a}) \tag{2.32}$$

where $\mathbf{g}$ represents the filter vector. If it is assumed that shadows are rare, the scene mean spectrum is used for $\mathbf{a}$.

MF Method

Equation 2.32 can be unmixed by applying a MF. Hence, instead of endmember construction, the covariance matrix, $\mathbf{C}$, is used to calculate the MF $\mathbf{q}$ (See equation 2.33).

$$\mathbf{q} = \frac{\mathbf{C^{-1}}(\mathbf{t} - \mathbf{a})}{(\mathbf{t} - \mathbf{a})^\mathbf{T} \mathbf{C^{-1}}(\mathbf{t} - \mathbf{a})} = \frac{\mathbf{C^{-1}} \mathbf{a}}{\mathbf{a^T} \mathbf{C^{-1}} \mathbf{a}} \tag{2.33}$$

where t represents the target spectrum and is equal to 0 for shadows and a is the scene mean spectrum. In the limit of $t = 0$, the output of the endmember-based filter, $\mathbf{g}$, and the MF, $\mathbf{q}$, should be zeor and equivalent.

Spectral Rebalancing

The MF yields a first estimate of the shadow fraction for each pixel. In a next step, the spectra are rebalanced in order to simulate the illumination by a spectrally uniform source. This is done by multiplying (1-f) with the direct sun plus sky spectrum. This will give the rebalanced spectra, $F(\lambda)$, which is wavelength dependent.

Re-application of the MF and de-shadowing

As a final step, the MF is applied to the rebalanced spectra which results in the refined shadow fraction. The steps of rebalancing and MF application are iterated a few times in order to insure convergence.

Once the refined shadow fraction is computed, the de-shadowed spectrum for each pixel can be calculated as follows:

$$Deshadowed_s pectrum = F(\lambda) x \frac{1}{g} \tag{2.34}$$

2.3 State-of-the-Art

2.3.1 detection algorithms: mono- and multi-temporal approaches

Generally, detection algorithms can be grouped into mono- and multi-temporal approaches. Mono-temporal observations are acquired at one specific point in time. This can be a disadvantage for the classification of classes which change in spectral and textural appearance over the course of a cycle, such as vegetation and crops [73]. Multi-temporal approaches usually achieve better results, especially for cloud masking because transient changes are easy to detect [31, 97, 34, 37]. However, this is only the case when a high temporal dataset frequency is available. Furthermore, the data processing complexity increases significantly, and most Level-2 processors for atmospheric correction rely on mono-temporal processing.

2.3.2 System Configuration description Language

Validation of *System Configuration description Language* (SCL) comprises verification of the cloud screening and classification accuracy to clarify uncertainties of L2A products for their applications.

CMIX

Cloud Masking Inter-comparison Exercise (CMIX) [1] was a recent state-of-the art inter-comparison of a set of cloud detection algorithms for Sentinel-2 and Landsat-8 representative for sensors in the 10-30 m range. However, CMIX was limited to differentiate only cloudy and cloudless pixels. Reference [14] is limited to valid and invalid pixels too. Valid pixels are cloudless pixels like land, water and snow and invalid are clouds and cloud shadows. Cloud masks from the *MACCS ATCOR Joint Atmospheric Correction* (MAJA) algorithm using multi-temporal information are compared with mono-temporal classification by Sen2Cor and Fmask [95]. The comparison is done twice: Once for cloud masks of all three processors dilated around clouds and second for all processors with non-dilated cloud masks. This means, that there is no comparison on original processor outputs. Overall accuracies for all three algorithms are nearby at 90-93% in case of non-dilated cloud mask. Mono-temporal Fmask gave equivalent classification performance as multi-temporal MAJA for dilated masks and Sen2Cor was on average 6% worse on these. However, dilation of Sen2Cor cloud mask is not recommended with the present processor version because it is a known issue, that misclassification of bright objects as clouds in urban area leads to commission of clouds and even more if dilation is applied. It is still to mention that the data set used in reference [14] excluded sites which are frequently cloudy, where mono-temporal algorithms have advantages over multi-temporal.

In contrary to radiometric validation, the SCL validation is limited due to the lack of suitable reference datasets. Existing references like CORINE land cover classification ([6]) consider only land cover classes and build on either local ground surveys (performed rarely within existing projects on very detailed information level but on limited geographical range) or satellite-based mapping (provided at regional to global scale) with low spatial resolution level > 100m. Cloud screening validation is even more challenging due to high variability of clouds over time. CMIX is based on four classification reference data bases for Sentinel-2 data. Hollstein data set [38] was created by photointerpretation and manually labelling selected polygons. It labels the classes clear sky, cloud, cirrus, shadow, water and snow. Unfortunately there is no separation between terrain shadow and cloud shadow [14]. The data set can be downloaded from the internet. Another available data base used in CMIX is called CESBIO data set. It contains reference images classified at 60 m spatial resolution which is too coarse for our investigation. Additionally it doesn't provide reference data for regions with frequent cloud cover because it is using the same active learning method as was used in [14]. Two more reference data sets used in CMIX are not publically available: PixBox data set contains manually classified single pixels at 10 m spatial resolution and GSFC data contains polygons of cloud covered area determined from sky images and photo-interpretation. There is neither a quality assessment and comparison of these classification references nor an investigation what is the best way to create them.

2.3.3 Cloud Shadow Detection

Several methods of shadow detection have been published over the past years. Most of these methods are applied to the field of computer vision. In the work proposed by [86], the shadows appearing in an image are detected by using derivatives of the input image and color invariant images. Reference [94] detects shadows from monochromatic natural images using shadow-variant and shadow-invariant cues from illumination, textural and odd order derivative characteristics.

In remote sensing images, the detection of clouds and their cloud shadows is mostly evaluated with complex multistage processes ([10], [44], [16], [27], [96]). But some satellite sensors do not have enough spectral information to properly select appropriate thresholds for cloud shadow detection that will hold for a variety of images. For example the GaoFen-1 and Proba-V images only have three bands in the visible and one band in the near-infrared available ([44], [52]). Reference [88] tries to overcome this lack of spectral information by introducing the cloud and cloud shadow detection into deep convolutional neural networks (DCNNs). Furthermore, [93] proposes a method for cloud shadow detection relying on land cover data support using a shadow probability generation method. In [92], the detection of cloud and cloud

Tabelle 2.1: Sentinel-2 spectral bands and spatial resolution

Sentinel-2 bands	resolution (m)
band 1 (0.433-0.453)	60
band 2 (0.458-0.523)	10
band 3 (0.543-0.578)	10
band 4 (0.650-0.680)	10
band 5 (0.698-0.713)	20
band 6 (0.733-0.748)	20
band 7 (0.765-0.785)	20
band 8 (0.785-0.900)	10
band 8a (0.855-0.875)	20
band 9 (0.930-0.950)	60
band 10 (1.365-1.385)	60
band 11 (1.565-1.655)	20
band 12 (2.100-2.280)	20

shadows is done solely by evaluating spectral indices (CSD-SI algorithm).

2.3.4 Sentinel-2 Mission

The Sentinel-2 mission consists of two polar-orbiting satellites, Sentinel-2A and Sentinel-2B, providing a five day revisit time at the equator. The swath width of a Sentinel-2 scene is 290 km and data is acquired in 13 bands with spatial resolutions of 10 m, 20 m, and 60 m [23] (see Table 2.1). Sentinel-2 images are open access data, offer high quality radiometric measurements and include a dedicated cirrus detection band. The free data access, frequent coverage of territories, wide swath, and many spectral bands are reasons for the wide-spread use of this kind of data in many applications.

For each spectral band of a sensor, a linear equation exists which describes the relationship between the DN (or brightness) and the TOA radiance, L. It is obtained with the radiometric offset, c_0, and gain, c_1 as follows and has units of $[W/(m^2 * sr * \mu m)]$:

$$L = c_0 + c_1 DN \qquad (2.35)$$

For some instruments, e.g. Sentinel-2, the data is given as top-of-atmosphere (TOA) reflectance ρ_{TOA} also called apparent or at-sensor reflectance, ρ^*:

$$\rho_{TOA} = \rho^* = \frac{\pi L d^2}{E_s \cos \theta_s} \qquad (2.36)$$

where d is the Earth-Sun distance in astronomical units, E_s is the extraterrestrial solar irradiance, and θ_s is the solar zenith angle. [57]

In the current state of the art PACO/ATCOR model, the input data are the recorded scene plus a meta file containing the acquisition date, time, solar and sensor view geometry etc. The input data is given as scaled TOA radiance, named Digital Number (DN). [72]

3 Comparison of Masking Algorithms for Sentinel-2 Imagery

This chapter describes the peer reviewed journal paper [91]. It evaluates the performance of three different masking algorithms which provide the user with a classification map for further atmospheric correction steps.

3.1 Problem Statement

Satellite imagery is frequently contaminated by low and medium altitude water clouds as well as by high-altitude cirrus clouds in the upper troposphere and in the stratosphere. Many operations require clear sky pixels as input, such as agriculture related products [41, 18], the retrieval of surface reflectance within atmospheric correction [37, 54] and the co-registration with other images [87, 67].

Atmospheric correction and land cover classification depend on an accurate cloud map [15, 21, 38]. In addition, maps of water and snow/ice are also indispensable in many applications, i.e. mapping of glaciers [66] and water bodies [24].

Cloud screening is applied to the data in order to retrieve accurate atmospheric and surface parameters as input for further processing steps, either the *Atmospheric Correction* (AC) itself or higher-level processing such as mosaicking, time-series analysis, or estimation of bio-geophysical parameters.

However, a fully automatic detection of these classes is not an easy task, due to the high reflectance variability of earth surfaces. For instance, bright desert surfaces or urban structures can be misclassified as cloud or opposite and shadow surfaces as water. A class assignment for mixed pixels (e.g., semi-transparent cloud over snow) can be problematic because they don't have a spectral signature which clearly belongs to a class. This together will decrease the classification accuracy and shows the need for a performance assessment of classification algorithms.

Our contribution evaluates the performance of the Fmask FORCE version [30], AT-COR, and Sen2Cor original classification outputs and not only for valid and invalid pixels, but in more detail for six consolidated classes given below. This gives more insight into their strengths and weaknesses.

In contrary to radiometric validation, the SCL validation is limited due to the lack of suitable reference datasets. The primary objective of the chapter is the classifications at the actual point of time with focus on clouds resp. cloudless surface area and cloud shadows. Our contribution evaluates the performance of three widely used mono-temporal masking codes on Sentinel-2 imagery.

Our first masking code is *Function of mask* (Fmask) [96]. It was originally designed for Landsat imagery, but later extended for Sentinel-2 data [95]. Here we use the improved version of reference [30] (Fmask FORCE version 3.0), which is able to separate clouds from bright surfaces exploiting parallax effects. The individual bands of MSI-sensor have slightly different viewing directions alternating from forward view to backward view between adjacent bands. The second code is the latest version of ATCOR (v 9.3.0) which contains an improved masking algorithm [68] as a necessary preprocessing part before starting the atmospheric correction. The third code is the scene classification of Sen2Cor (version 2.8.0). Sen2Cor is an atmospheric correction processor for Sentinel-2 (S2) data provided by the *European Space Agency* (ESA), which contains a preprocessing scene classification step preceding atmospheric correction [45]. Whereas the atmospheric correction module of Sen2Cor was developed in heritage of ATCOR, the scene classification is completely independent. Scene classification of Sen2Cor makes use of some external auxiliary data from Climate Change Initiative. It is still to mention that Fmask uses a 300 m dilation buffer for cloud, and 60 m for cloud shadow, while ATCOR uses 100 m and 220 m, respectively, and Sen2Cor (version 2.8.0) uses no dilation buffers. Fmask applies also a 1 pixel buffer for snow.

This chapter is organized as follows: section 3.2 presents an overview over the S2 scenes used for the exercise. Section 3.3 describes the approach to define the reference ("truth") mask (validation procedure). Section 3.4 presents the classification results in terms of user's, producer's and overall accuracy [19]. The conclusion and possible further improvements are given at the end of the chapter.

3.2 Methods (processors) and Data

Twenty S2 scenes are processed with three codes. A list of the investigated Sentinel-2 scenes is given in Table 5.1. The scenes were selected to cover all continents, different climates, seasons, weather conditions, and land cover classes (Figure 3.1). They represent flat and mountainous sites with cloud cover from 1% to 62% and include the presence of cumulus, thin and thick cirrus clouds and snow cover. Additionally, the scenes are representative for different land cover types such as desert, urban, cropland, grass, forest, wetlands, sand, coastal areas, and glaciers. The range of solar zenith angles is from $18°$ to $62°$. For the scene classification processing and validation, all S2 bands with 10 m and 60 m are resampled to a common 20 m pixel size. All processors used *Digital Elevation Models* (DEMs), usually from SRTM (90 m) (downloaded from the USGS website (https://earthexplorer.usgs.gov/)), except for the scenes number 1, 6, and 16, which used Planet DEM's [3]. Also all DEM's are resampled to a common 20 m resolution.
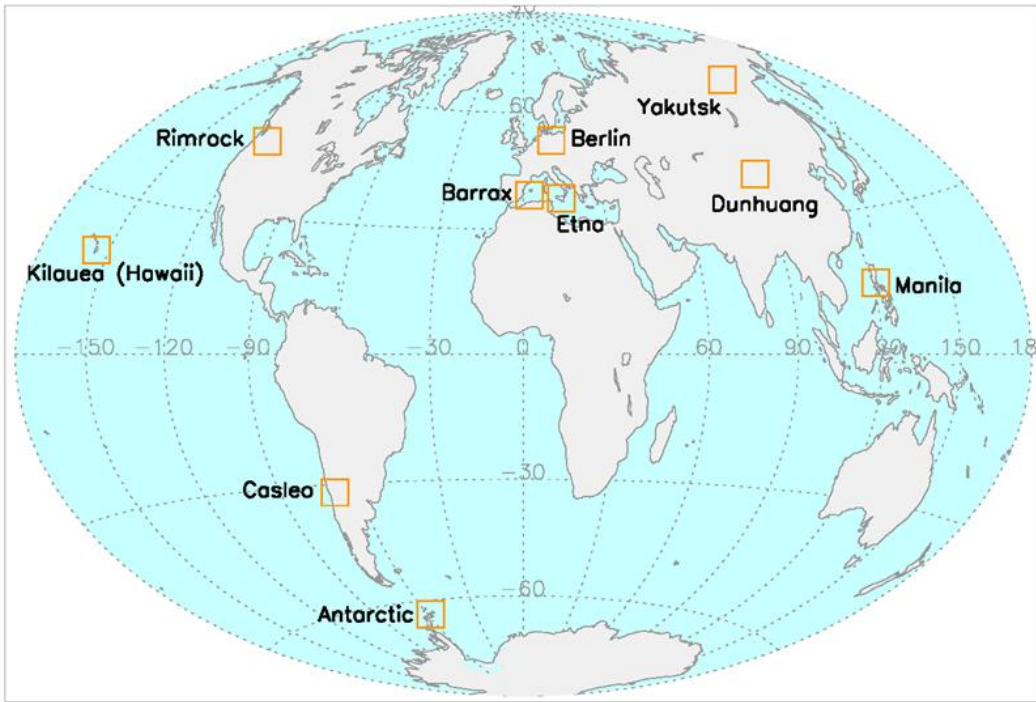
Abbildung 3.1: Geographical distribution of 20 test sites selected for validation (orange squares)

The SCL algorithm of Sen2Cor aims to detect clouds with their shadows and to generate a scene classification map. The latter raster map consists of 12 classes, including 2 classes for cloud probabilities (medium, and high), thin cirrus, cloud shadows, vegetated pixels, non-vegetated pixels, water, snow, dark feature pixels, unclassified, saturated or defective pixels and no data. This map is used internally in Sen2Cor in the atmospheric correction module to distinguish between cloudy pixels, clear land pixels and water pixels, and it does not constitute a land cover classification map in a strict sense [33]. The scene classification map is delivered at 60 m and 20 m spatial resolution, with associated *Quality Indicators* (QI) for cloud

Tabelle 3.1: Sentinel-2 level L1C test scenes (SZA = Solar Zenith Angle)

Scene	Location	Date	Tile	SZA	Cloud cover	Desert	Ice/Snow	Non-veg	Veg	Water	Mountains	Rural	Urban
1	Antarctic	2019/02/03	T21EVK	54.9°	28%		X	X		X			
2	Argentina, Buenos Aires	2018/08/27	T21HUC	51.5°	0%					X		X	X
3	Australia, Lake Lefroy	2018/08/19	T51JUF	51.5°	0%			X					
4	Bolivia, Puerto Siles	2018/09/06	T19LHF	30.6°	0%				X	X		X	
5	China, Dunhuang	2018/01/22	T46TFK	62.3°	24%	X	X					X	
6	Estonia, Tallin	2018/07/14	T35VLG	39.0°	2%					X		X	X
7	Germany, Berlin	2018/05/04	T33UUU	38.0°	1%			X	X	X		X	X
8	Italy, Etna	2017/03/09	T33UUU	45.1°	7%		X			X		X	X
9	Kazakhstan, Balkhash	2018/07/30	T43TFM	30.7°	7%	X				X		X	
10	Mexico, Cancun	2018/05/27	T16QDJ	18.4°	7%				X	X		X	
11	Morocco, Quarzazate	2018/08/30	T29RPQ	27.2°	2%	X		X	X	X			
12	Mosambique, Maputo	2018/07/13	T36JVS	54.4°	0%					X		X	X
13	Netherlands, Amsterdam	2018/09/13	T31UFU	49.7°	5%				X	X		X	X
14	Philipines, Manila	2018/03/19	T51PTS	27.4°	1%					X		X	X
15	Russia, Sachalin	2018/05/09	T54UVC	35.5°	0%		X	X		X			
16	Russia, Yakutsk	2017/08/08	T52VEP	45.9°	6%		X	X	X	X			
17	Spain, Barrax-1	2017/05/09	T30SWH	24.1°	18%				X		X	X	
18	Spain, Barrax-2	2017/05/19	T30SWH	22.0°	2%				X		X	X	
19	Switzerland, Davos	2019/04/17	T32TNS	37.7°	25%		X	X	X	X	X		
20	USA, Rimrock	2018/05/12	T11TMM	30.4°	1%		X			X		X	X

Tabelle 3.2: Comparison of individual masking outputs and common mask classification

Label	Classes	Definition for reference	Fmask	ATCOR	Sen2Cor
1	Clear land	Clear pixels over land	Clear	Clear	Vegetation; not vegetated; unclassified
2	Semi-transparent cloud	$0.01 <$ TOA rho $(1.38\mu\text{m}) < 0.04$; also haze, smoke or any kind of cloud which transparency enables to recognize the background features	Cloud	Semi-transparent cloud	Thin cirrus
3	Cloud	Cumulus cloud; thick clouds (also thin cirrus)	Cloud shadow	Cloud	Cloud medium and high probability
4	Cloud shadow	Shadow thrown by the clouds over land	Snow/ice	Shadow	Cloud shadow
5	Clear water	Clear pixels over water	Water	Water	Water
6	Clear snow/ice	Clear pixels over snow and ice	-	Snow/ice	Snow and ice
7	Topographic shadow	Self-shadow and/or cast-shadows	-	Topographic shadow	Dark feature
0	Background		-	Geocoded background	No data; saturated or defective

and snow probabilities. The QIs provide the probability measure (0-100%) that the Earth surface is obstructed either by clouds or by snow. Class dark area pixels can contain dark features like burned area, topographic shadows or cast shadows, but also very dark water bodies and vegetation. Thin cirrus may also be other transparent cloud and the transition from medium to high probability cloud is impossible to validate. Pixels assigned to unclassified are mostly pixels with low probability of clouds or mixed pixels which don't fit into any of the other classes.

3.3 Validation procedure

Validation of SCL comprises verification of the cloud screening and classification accuracy to clarify uncertainties of masking products for their applications. Comparison of different masking algorithms requires first to map all the individual masking outputs to a common set of classes. Table 3.2 shows the seven classes used as a common set for Fmask, ATCOR and Sen2Cor.

Semi-transparent cloud is defined as optically thin cirrus cloud and thin lower altitude clouds or haze, smoke, etc. To detect thin cirrus clouds, the processors use the TOA reflectance in the cirrus band 10, lying below 0.04 but above 0.01. The lower threshold is used to avoid classifying all pixels as semi-transparent. The label cloud comprises optically thick (opaque) water cloud and also cirrus cloud with ρ (TOA, band 10) > 0.04.

The focus of the present paper is not only validation of the scene classification provided by the three processors, but its comparison. The best way to do that comparison is to generate two each other complementing reference maps – one for the 'difference area' and another for the 'same area'. The difference area is the part of the classification images where the classification maps provided by the three processors disagree. Validation statistics over the difference area is perfect to achieve a relative comparison between processors pointing on strengths and weaknesses

much sharper than interpreting statistics over an entire image. The same area is the remaining part of the images where all three classifications give the same result. Validation statistics over the same area enables to compute the absolute classification performance of the processors if the number of pixels in the reference map for the difference and the same area is balanced with regard to the ratio of corresponding pixels in the entire image. The challenge for validation of SCL is generation of high quality reference maps which gives the 'truth'. Generation of the reference maps for the performed comparison of Fmask, ATCOR and Sen2Cor outputs relies on visual inspection, supplemented by meteorological data, if available. The following procedure is repeated for each image of the validation data set.

First, stratified random sampling [19] is applied to the difference mask between three processors to get the sample points for visual labelling. Stratification serves to get the sample size balanced between all classes present in the image, thus to guarantee statistical consistency and to avoid exclusion of spatially limited classes from the validation. Our aim is an amount of 1000 randomly selected samples per image with the minimum number of 50 samples for the smallest class (for reference please see following authors: [28, 59, 77, 83]). Visual inspection by human interpreter results in labelling of either one pixel only or labelling a polygon drawn around an adjacent area of pixels of the same class to assign the correct class and create the reference ("truth") map. All labelled pixels are used to create the reference classification image typically resulting in an average number of 5000 pixels per scene. (Figure 3.2)

Visual inspection by the expert human interpreter is supported by:

- Visual checks of the true color image (bands 4, 3, 2),near infrared false color (bands 8A, 6, 5), and short-wave infrared false color (bands 12, 11, 8A).

- Check of L1C cirrus (band 10) concerning semi-transparent cirrus regions.

- Check of BOA reflectance spectral profiles from Level-2 Sen2Cor products..

- Comparison with imagery archive from GoogleEarth^{TM}.

The created reference classification map is finally compared to the consolidated classification maps from Fmask, ATCOR, and Sen2Cor and a confusion matrix is obtained for each classification. Finally classification accuracy statistics is computed from confusion matrixes. The same procedure is repeated for the same area to allow computation of absolute classification accuracy statistics of the 3 classifications. Generation of the reference mask for Cloud Screening and Classification Validation was performed. Figure 3.3 shows an example of a true color (RGB = bands 4, 3, 1) composite of scene 19 (Davos) of Table 5.1, a false color composite
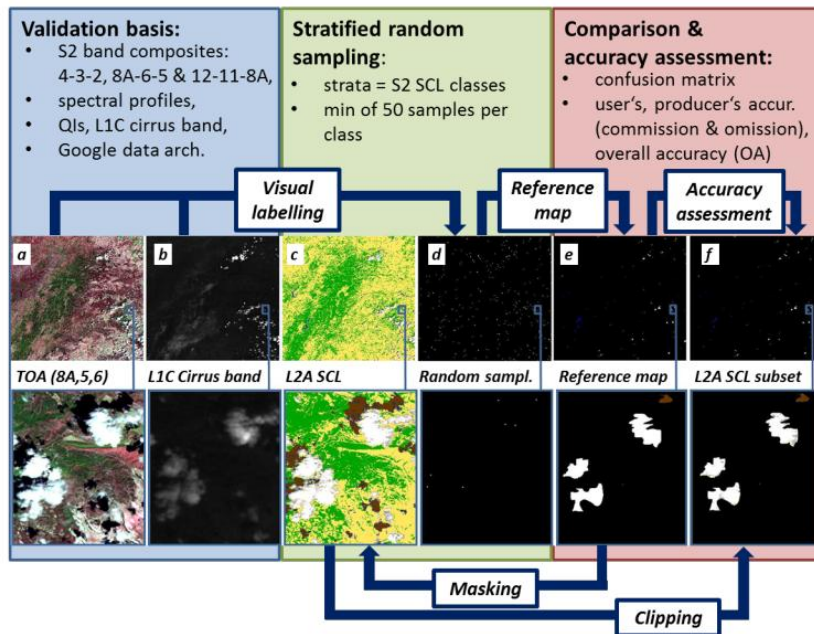
Abbildung 3.2: Schema for Cloud Screening and Classification Validation of Sen2Cor scene classification (SCL) product over Barrax test-site (Spain), acquired on May 19, 2017. This example represents various topography (flat and rough) and land-cover (vegetated, non-vegetated, water), as well as cloud cover dominated by the cumulus clouds. Top image cubes: single granule; bottom image cubes: zoomed area.

using RGB (SWIR1, NIR, red) and some typical BOA reflectance spectra of snow, clouds, clear (vegetation), and water. Obviously, snow/ice and clouds cannot easily be discerned in the true color image. Therefore, the human interpreter also uses other band combinations, in this case with band 11 (SWIR1), where snow/ice (colored blue) is clearly recognized. In addition, BOA reflectance spectra are evaluated for a polygon if a class assignment is not obvious.

The CORINE land cover classification of 2018 is not used instead (or additionally) to Google Earth for several reasons. Both don't help for the primary focus of the classifications investigated which is cloud screening and cloud shadow detection at the time of satellite acquisition. The comparison of classifications is performed on a global scale because Sentinel-2 operates globally. CORINE land cover classification is restricted to Europe and Google Earth is available globally. CORINE land cover classification contains 44 classes which is too detailed for this purpose and the spatial resolution is too low. The S2 true color composite (10 m resolution) coupled with high spatial resolution GoogleEarth^{TM} imagery ensures that the reference
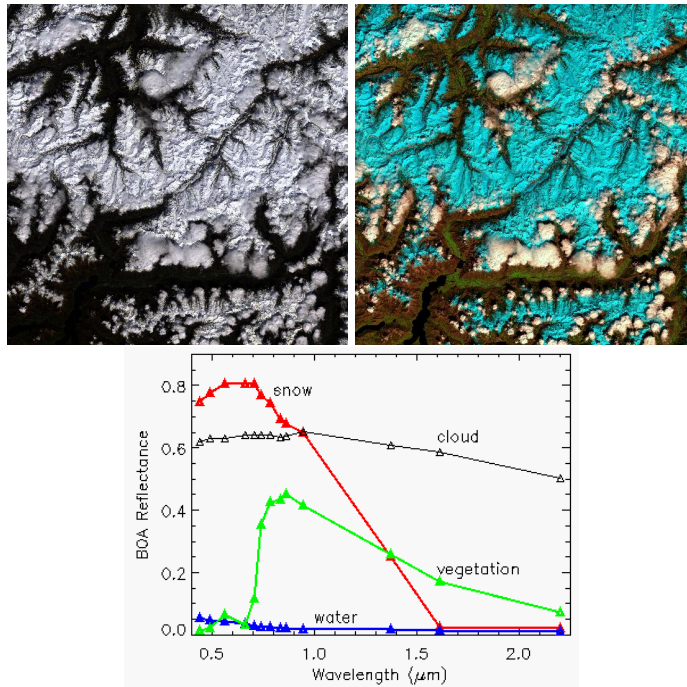
Abbildung 3.3: Left to right: true color (RGB = 665,560,443nm) composite of scene ID 19, SWIR1 (RGB = 1600,860,660nm) composite, and example spectra.

classification is of higher spatial quality than the 20 m classification maps with the three methods. Since the reference and image classification are based on the same single-date images there is no geolocation error and no temporal change between reference map and the classified maps. The reference polygons are selected as internally (spectrally) homogeneous areas in terms of the map classification and the minimum mapping unit is a 20 m pixel.

The procedure applied for generation of our reference classification map is similar to the way used to create the references for the Hollstein and PixBox data sets. The new point is, that we split the validation into creating a reference for the difference area and another for the same area for comparison of classification tools. Obtained reference maps are not perfect. The manual labelling includes some amount of subjectivity. Most of all visual interpretation and labelling of transparent clouds is challenging. Subjectivity of the method was tested with 4 persons creating a reference map for the same 2 products. The test revealed quite stable results with 5-6% differences in *overall accuracy* (OA) using the reference maps for computation of classification accuracy statistics. Another limitation of our classification reference maps comes from the stratified random sampling. The stratification between classes has to be oriented itself on one classification which was Sen2Cor in our case. If Sen2Cor classification fails, then the reference map becomes imbalanced. Even if this is not the case, then the reference maps are not perfectly balanced for the other classifications and the validation results are little biased.

Classification accuracy statistics it represented by three parameters calculated from the error confusion matrix [19, 60, 78]. If the number of classes is n, then the confusion matrix C is a $n \times n$ matrix, and the *user's accuracy* of class i (percentage of the area mapped as class i that has reference class i) is defined as

$$UA(i) = 100\, C_{ii} / \sum_{j=1}^{n} C(i,j) \qquad j = column\ number \qquad (3.1)$$

The second parameter is the *producer's accuracy* of class i (percentage of the area of reference class i that is mapped as class i)

$$PA(i) = 100\, C_{ii} / \sum_{j=1}^{n} C(j,i) \qquad j = row\ number \qquad (3.2)$$

The last is the *overall accuracy* :

$$OA = 100\, \frac{\sum_{j=1}^{n} C(j,j)}{\sum_{i=1}^{n} \sum_{j=1}^{n} C(i,j)} \qquad (3.3)$$

Another commonly used measure is the kappa coefficient of agreement [19]. Consistent with the recommendations in [60, 80] the use of kappa is strongly discouraged despite its widespread use, due to its high correlation with the overall accuracy.

Besides the OA, UA and PA measures a detailed visual inspection supported the analysis of the confusion within and between classes per processor. Comparison was performed per processor and class over difference area, including recognition rates, misclassification rates of particular class as well as its confusion potential with other classes (the proportion of one mistaken by other class).

3.4   Results

Validation results consist of confusion matrix with the number of correctly classified pixels in the validation set. Confusion matrix is the basis for computation of *user's accuracy* (UA) and *producer's accuracy* (PA) and *overall accuracy* (OA) of classification. Table 3.3 and Table 3.4 provide results for absolute validation of classifications comparable to results present in the literature.Table 3.3 provides a summary of the UA and PA per class,i.e. the average over all 20 scenes, and Table 3.4 contains the OA per scene. Boldface numbers indicate the method with the best performance, but if the values differ less than about 1% then two methods are marked correspondingly.

Tabelle 3.3: Summary of classification accuracy (percent) for difference area (F=Fmask, A=ATCOR, S=Sen2Cor; bold face numbers indicate the best performances).

Class	UA (F)	UA (A)	UA (S)	PA (F)	PA (A)	PA (S)
clear	**82.3**	67.4	77.5	61.5	68.3	**72.2**
semi-transp. cloud	**74.9**	59.3	54.8	2.8	**33.5**	33.4
cloud	19.2	**59.9**	48.4	**89.2**	52.9	64.9
cloud shadow	55.6	70.1	**85.9**	46.4	**53.6**	26.3
water	**93.5**	36.1	70.3	64.3	58.8	**75.5**
snow/ice	48.8	52.7	**57.2**	48.6	64.7	**79.0**
topographic shadows	**72.7**	11.9	3.6	1.8	2.6	**44.9**

Tabelle 3.4: Summary of overall accuracy (percent) (F=Fmask, A=ATCOR, S=Sen2Cor).

Scene	Location	OA Difference Area			OA Same Area	OA Total Area		
		F	A	S		F	A	S
ID	Average (all scenes)	54	59	**62**	97	79	79	79
1	Antarctic	36	51	**56**	98.1	81	87	**89**
2	Argentina, Buenos Aires	**89**	59	59	98.1	**97**	93	93
3	Australia, Lake Lefroy	59	49	**67**	99.6	61	50	**68**
4	Bolivia, Puerto Siles	**90**	22	58	99.6	**99**	95	97
5	China, Dunhuang	56	40	**74**	97.4	91	89	**96**
6	Estonia, Tallin	40	71	**78**	97.9	92	95	**97**
7	Germany, Berlin	57	**76**	67	99.9	96	**98**	97
8	Italy, Etna	32	**70**	**71**	99.8	97	**98**	**98**
9	Kazakhstan, Balkhash	47	**75**	45	91.3	88	**90**	88
10	Mexico, Cancun	44	59	**66**	98.5	89	92	**93**
11	Morocco, Quarzazate	73	**86**	42	99.9	80	**90**	63
12	Mosambique, Maputo	**75**	35	45	84.8	**84**	82	83
13	Netherlands, Amsterdam	38	**63**	**64**	95.6	92	93	**94**
14	Phillipines, Manila	46	**69**	67	97.5	92	**95**	94
15	Russia, Sachalin	**85**	51	63	98.4	**94**	80	85
16	Russia, Yakutsk	**64**	49	53	96.5	**88**	83	85
17	Spain, Barrax-1	25	**74**	62	98.6	94	**97**	96
18	Spain, Barrax-2	64	**92**	90	99.9	66	**92**	90
19	Switzerland, Davos	16	42	**46**	97.1	65	81	**82**
20	USA, Rimrock	50	52	**63**	99.2	96	96	**97**

For space reasons, we cannot present detailed results for each scene. The example of scene 4 (Bolivia, Puerto Siles) in Figure 3.4 serves as an example to demonstrate the difference mask validation. The image contains no clouds, but water with different color and sediment, bright soil and burned area. This image is the example with the smallest difference area resp. opposite the one with the largest agreement between Fmask, ATCOR and Sen2Cor classifications. This is also underlined with high absolute OA over complete image of 99%, 95% and 97%. There is only a small difference between classifications in PA for class clear land with 100%, 94% and 97% representing what is visible in Figure 3.4 – a different amount of burned area is classified as water. User accuracy of class water for the total image is 99% for both Fmask and Sen2Cor classifications hiding differences clearly to see in the figure. Statistics over difference area gives a much more detailed insight into classification performance. OA over difference area is 90%, 22% and 58% for Fmask, ATCOR and Sen2Cor. Differences in PA for class clear land are now more highlighted with values 97%, 18% and 57%. User accuracy of class water for the difference image now is different between Fmask and Sen2Cor with 74% resp. 80%. Whereas Fmask identifies 84% of clear pixels in the difference area as clear, ATCOR and Sen2Cor do it for less than 50% of pixels. ATCOR largely misclassifies burned area as water. However, this is less critical than misclassification of burned area as topographic shadows by Sen2Cor. Both clear land and water are valid pixels, free of clouds and shadows. The problem shown for Sen2Cor with misclassification of clear land as topographic shadow has its origin in the transformation of Sen2Cor classification outputs to the consolidated mask. Consolidated class topographic shadow corresponds to Sen2Cor class dark area, which can contain dark features like burned area, topographic shadows or cast shadows, but also very dark water bodies and vegetation. Therefore the classification output of Sen2Cor for this example can be interpreted as much more correct as given by the statistics. A planned update of class definition of Sen2Cor class dark area to only topographic or cast shadow will solve this confusion.

To furthermore compare the classification performance of Fmask, ATCOR and Sen2Cor, details are given for three selected cases: the best and worst case scenarios, and an average case.

Figure 4.3 shows the best case (highest accuracy) scenario of all analyzed scenes from Table 3.4. It is scene number 18 from Spain (Barrax) taken on the 19th of March 2017 with a zenith angle of 22.0° and a azimuth angle of 143.2°. In Figure 4.4 a subset of scene ID 18 can be found. It nicely illustrates the differences between Fmask, ATCOR and Sen2Cor. The cloud percentage is strongly overestimated in Fmask, while ATCOR and Sen2Cor classifications are very similar and close to the
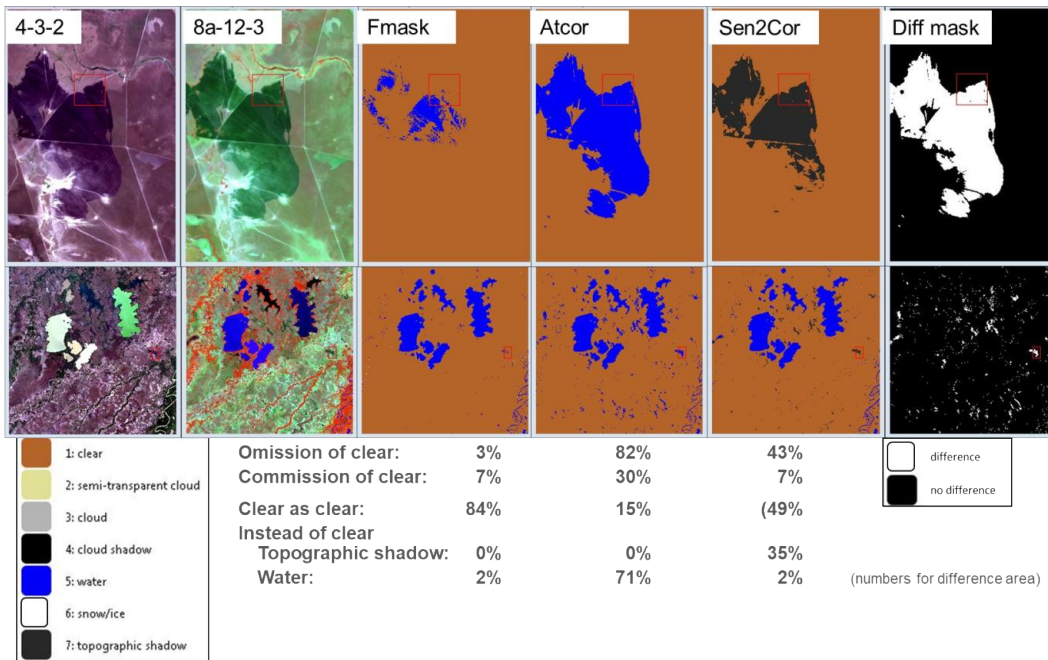
Abbildung 3.4: Difference area validation on example of scene 4 (Bolivia, Puerto Siles). Bottom row: Sentinel-2 Scene; top row: zoom of image showing a region with burned area; From left to right: Natural color composite of bands 2,3,4; false color composites of bands 8a, 12, 3 helpful for discrimination between dark classes, vegetation types and clouds; Classification map from Fmask; Classification output of ATCOR; Classification map from Sen2Cor; Difference area map

reference.

The overall worst case scenario of the 20 scenes analyzed is illustrated in Figure 3.7. This scene from Switzerland (Davos) was acquired on the 4th of April 2019 at a zenith and azimuth angle of $37.7°$ and $158.5°$, respectively. This scene is difficult to classify correctly for all the processors due to the high reflectivity of the snow and complex topography. The snow is often misclassified as cloud, especially in Fmask hence also the low overall accuracy of 16.4%.

Firgure 3.7 shows a subset of scene ID 19. As in the previous case (scene ID 18 from Spain) Fmask overestimates the percentage of cloud coverage whereas ATCOR and Sen2Cor show a more accurate cloud mask. An inspection of a zoom area (see Fig.3.8) reveals that Sen2Cor sometimes falsely classifies cloud shadows as water.

A scene showing an average case scenario (i.e. no complex topography, small percentage of cloud cover and bright objects) for all classification methods is the one from the USA (Rimrock). It was taken on the 12th of Mai 2018 at a zenith angle of $30.4°$ and an azimuth angle of $153.5°$. Figure 3.9 shows the entire area of the scene with the three different classification maps, whereas Figure 3.10 only illustrates a subset of scene ID 20. Most of the scene is clear with some clouds and snow/ice in
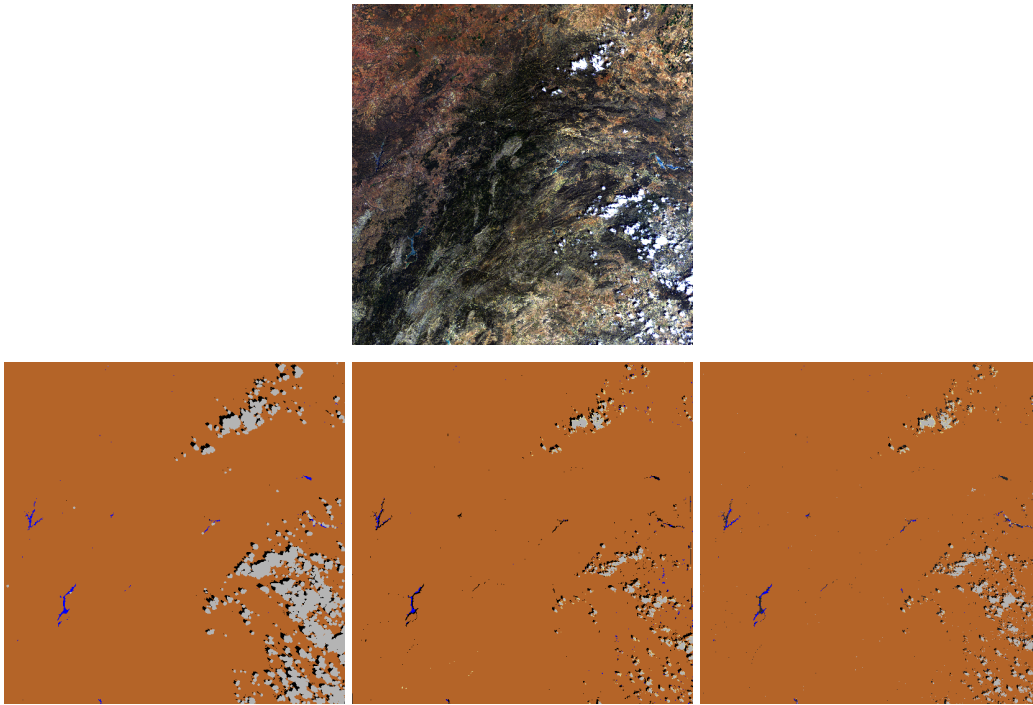
Abbildung 3.5: Top row: true color (RGB = 665,560,443nm) composite of scene ID 18 (Barrax-2). Bottom row (left to right):Fmask, ATCOR and Sen2Cor classification maps.
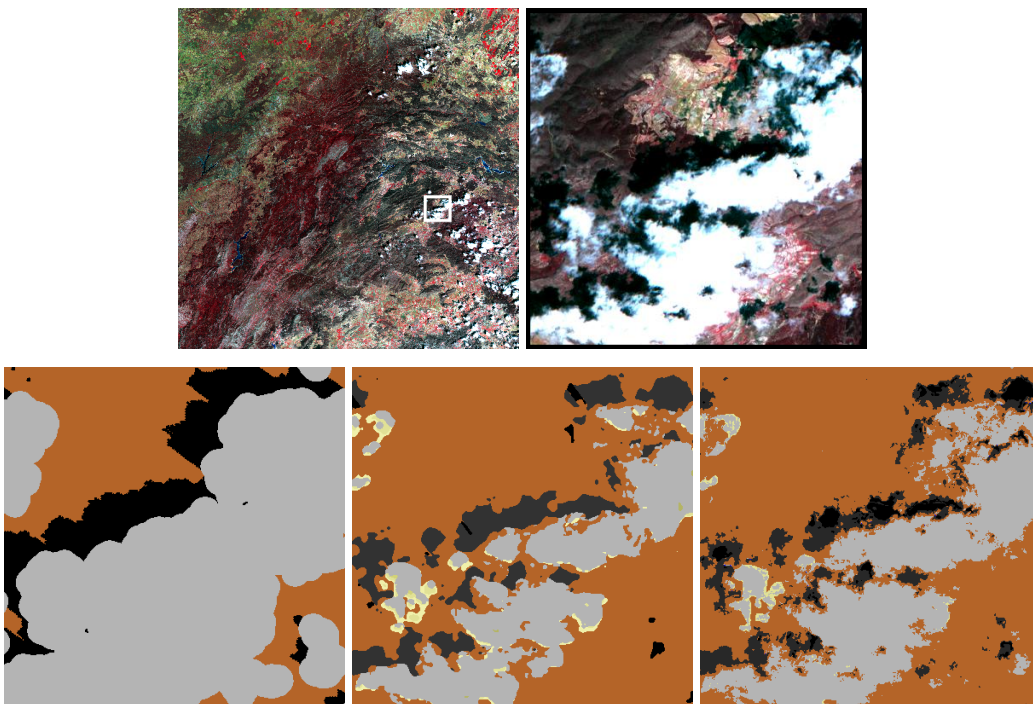
Abbildung 3.6: Top row (left to right): CIR (RGB = 865,665,560nm) composite and CIR subset of scene ID 18 (Barrax-2). Bottom row (left to right): Fmask, ATCOR and Sen2Cor classification maps of the subset.

the southern part. Additionally the river is accurately mapped by all processors.

The subset (Figure 3.10) demonstrates the difficulties Sen2Cor faces when distin-
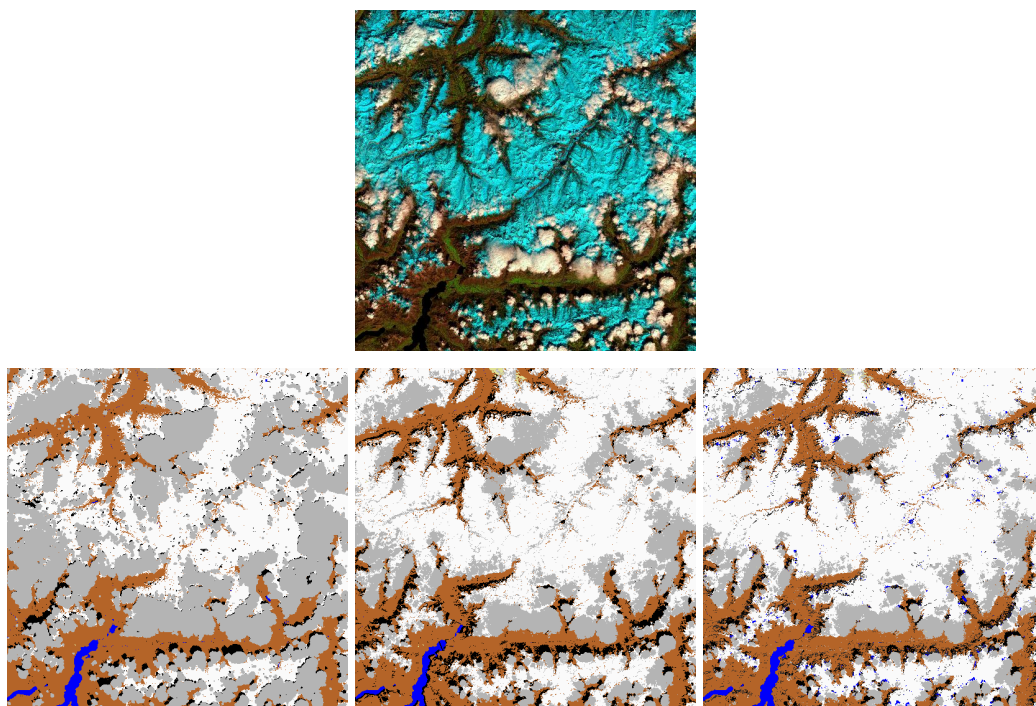
Abbildung 3.7: Top row: SWIR1/NIR/red composite of scene ID 19 (Davos). Bottom row (left to right): Fmask, ATCOR and Sen2Cor classification maps.
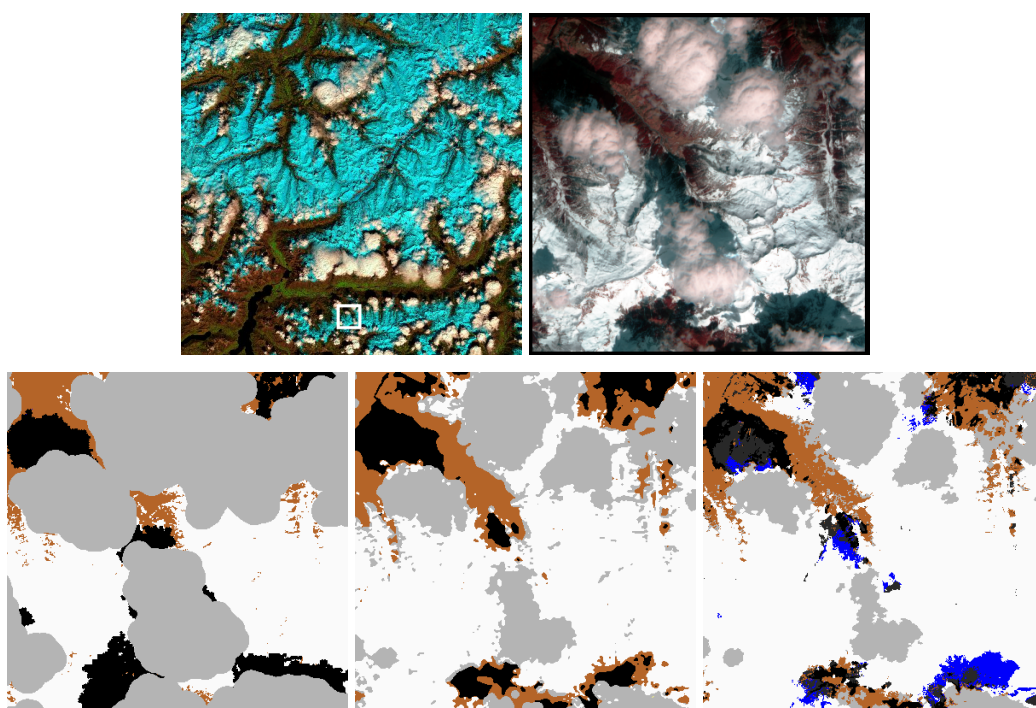
Abbildung 3.8: Top row (left to right): SWIR1/NIR/red composite and CIR (RGB = 865,665,560nm) subset of scene ID 19 (Davos). Bottom row (left to right): Fmask, ATCOR and Sen2Cor classification maps.

guishing between urban areas or bright ground objects and clouds. ATCOR on the otherhand misinterprets dark water for shadow. But if both classes have about the same probability, then ATCOR's preference is shadow.
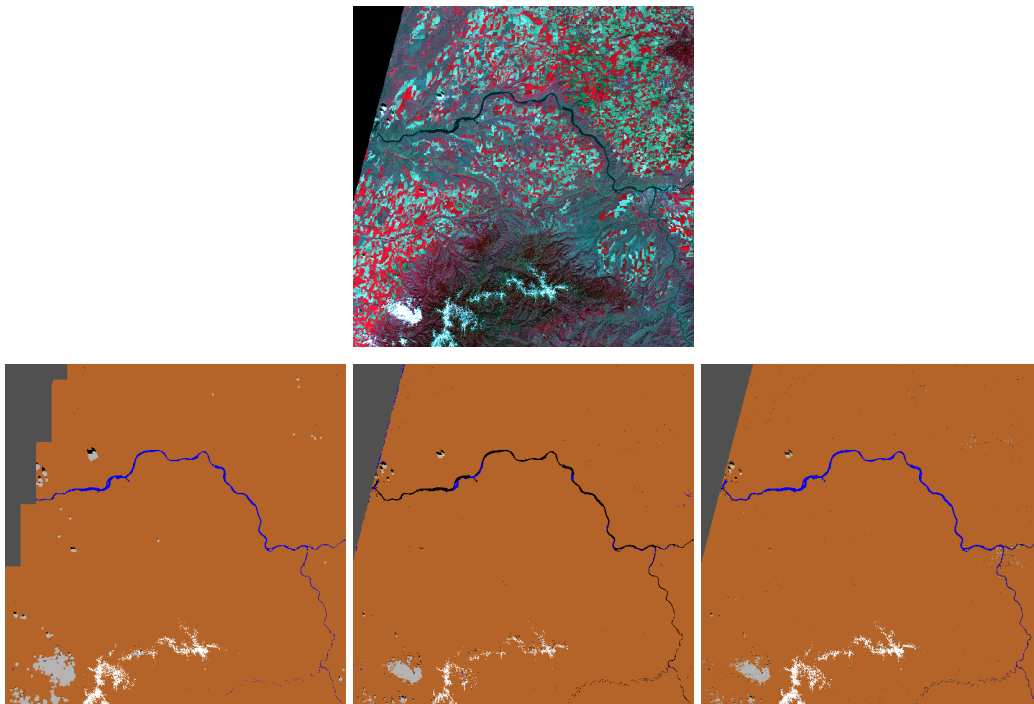
Abbildung 3.9: Top row: CIR (RGB = 865,665,560nm) composite of scene ID 20 (USA Rimrock). Bottom row (left to right): Fmask, ATCOR and Sen2Cor classification maps.
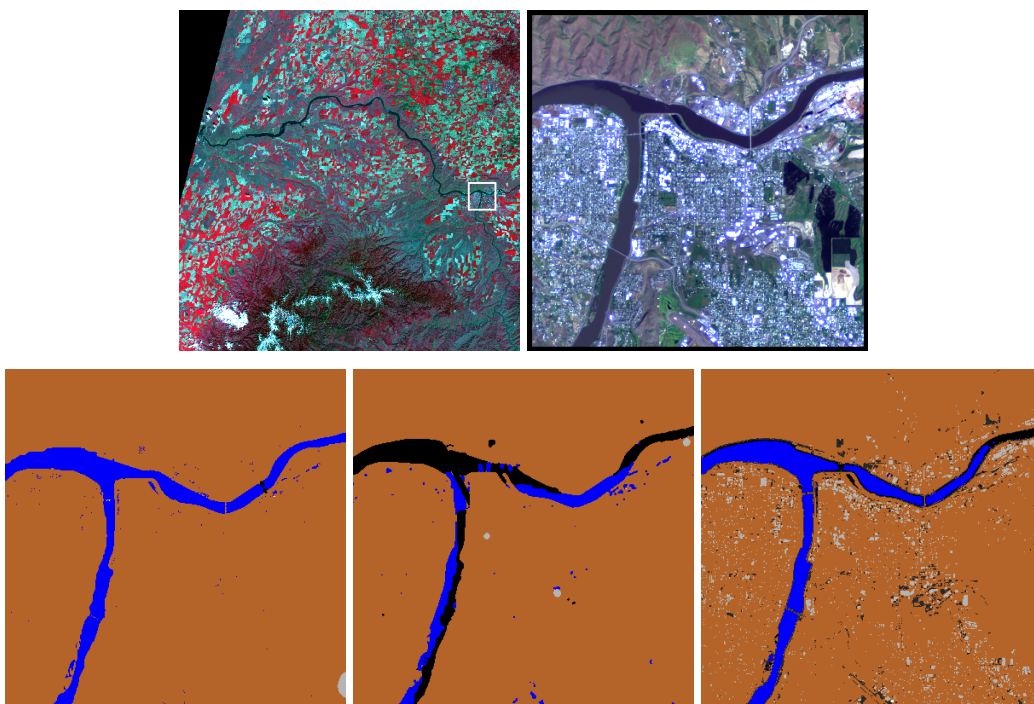
Abbildung 3.10: Top row (left to right): CIR (RGB = 865,665,560nm) composite and true color (RGB = 665,560,443nm) subset of scene ID 20 (USA Rimrock). Bottom row (left to right): Fmask, ATCOR, and Sen2Cor classification maps.

As can be taken from Table 3.3, up to 70% of clear pixels were correctly classified by Sen2Cor, whereas Fmask and ATCOR recognize 61.5% and 68.3% correctly.

The highest share of clear pixel misclassification was found by clouds for Fmask, water for ATCOR and topographic shadows for Sen2Cor. Semi-transparent clouds were recognized up to 33.5% and 33.4% for ATCOR and Sen2Cor, respectively, while the omitted pixels were mainly distributed between classes clear and clouds by ATCOR and clear and snow by Sen2Cor. Fmask only classifies 2.8% of semi-transparent cloud pixels correctly and mostly missclassifies the omitted pixels as clouds. Fmask performs best for the classification of cloud pixels (89.2%), while ATCOR and Sen2Cor have a recognition rate of 52.9% and 64.9%, respectively. The highest share of the cloud omission was found by class clear for Fmask and Sen2Cor and by class cloud shadows for ATCOR. Cloud shadows have low recognition rate (26.3%) and high confusion with class clear in the case of Sen2Cor. Fmask and ATCOR have lowest recognition for the class topographic shadows with a rate of 1.8% and 2.6%, respectively. Sen2Cor performs slightly better with 44.9%. Their omission is distributed mainly between classes clear and cloud shadows. The highest recognition rates (and lowest confusion to other classes) were found for clouds (89.2%) and water (64.3%), clear (68.3%) and snow (64.7%), and water (75.5%) and snow pixels (79%), for Fmask, ATCOR and Sen2Cor, respectively. Surprisingly,the proportion of snow pixels being mistaken toward clouds was low for ATCOR and Sen2Cor (9% and 8%, respectively), whereas Fmask misclassifies 61.6%.

The confusion within and between classes can be additionally illustrated using the proportion of the individual class omissions (Figure 3.11, 3.12 and 3.13).
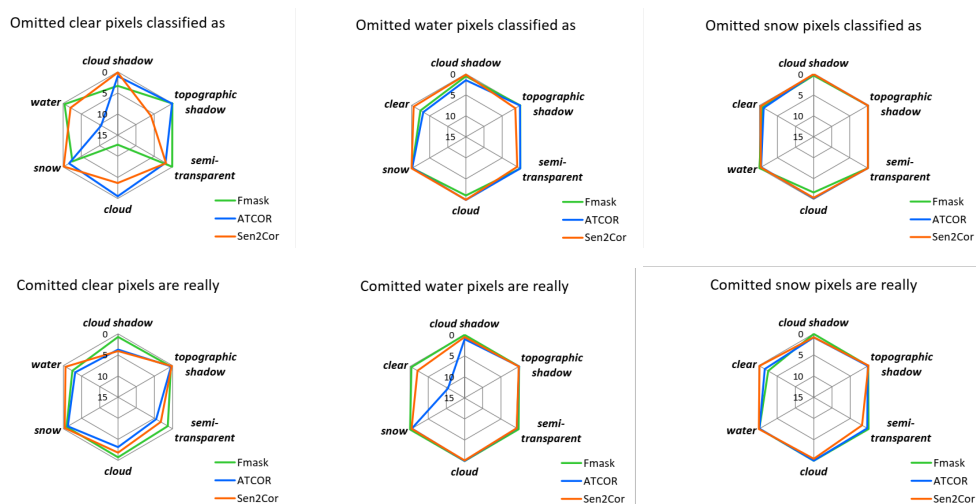
Abbildung 3.11: Omission and Commission per Class for clear classes clear land, water and snow.

Figure 3.11 illustrates spider diagrams for the omission and comission for the clear classes clear land, water and snow. Fmask, ATCOR and Sen2Cor are represented by the colors green, blue and orange. Looking at the left upper plot of Figure 3.11,

it can be noted that Fmask has a large omission of clear land towards clouds. This is due to dilation. ATCOR has a omission of clear land to water, which is uncritical for pure cloud masking. Sen2Cor confuses most clear land pixels with topographic shadows due to the difficulty to correctly classify dark features. In the left lower plot of Figure 3.11 we see that Sen2Cor and ATCOR have more difficulties with cloud shadow than Fmask, which on the otherhand, classifies more clear land as water. All three masking codes show comission of water pixels towards the same direction of clear land but with different amounts. A similar shape can be seen for the omission of snow pixels. The comission of snow, on the other hand, is in different directions. Fmask classifies some cloud shadows as snow, which is uncritical for clear/cloud mask. ATCOR classifies some cloud shadows as snow and Sen2Cor classifies some semi-transparent clouds as snow.
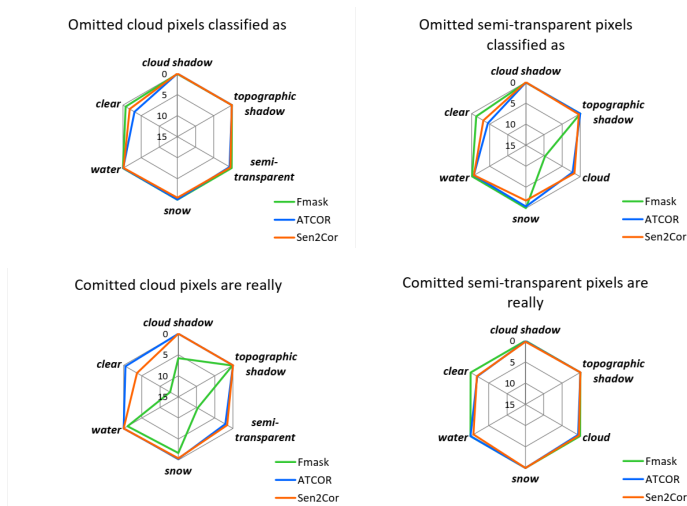
Abbildung 3.12: Omission and Commission per Class for cloud classes cloud and semi-transparent cloud.

Figure 3.12 illustrates spider diagrams for the omission and comission of the classes cloud and semi-transparent cloud. The upper left image shows that all three masking codes have a omission of cloud pixels towards the class clear. The comission of cloud is on the other hand different. Fmask shows the largest comission of cloud pixels towards clear, semi-transparent cloud and cloud shadow. Sen2Cor classifies some clear and semi-transparent pixels as cloud and ATCOR shows a slight comission of semi-transparent pixels towards cloud. For the class semi-transparent cloud, the largest omission comes from Fmask, which confuses most semi-transparent clouds as cloud. Fmask and Sen2Cor show a comission of semi-transparent pixels towards the class clear.

Figure 3.13 illustrates spider diagrams for the omission and comission of the shadow classes cloud shadows and topographic shadows. From the left upper image of figure 3.13 it can be noted that Fmask has largest omission of cloud shadow towards
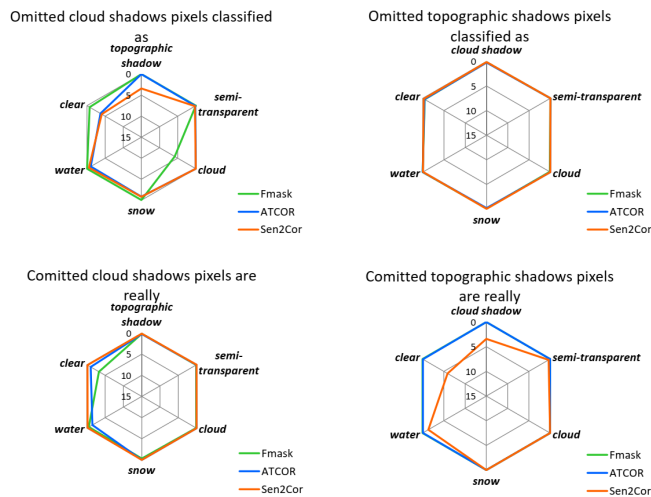
Abbildung 3.13: Omission and Commission per Class for shadow classes cloud shadows and topographic shadows.

the class cloud. ATCOR and Sen2Cor confuse cloud shadows mostly with clear pixels. All three masking codes show a similar comission of cloud shadows towards clear pixels. Except for the class definition problem of Sen2Cor for topographic shadows, the upper and lower right images show good agreement between the processors. Sen2Cor shows a large comission of topographic shadow pixels towards the class clear due to its definition of dark pixels. Fmask classifies few topographic shadow pixels as clear land and ATCOR and Fmask few topographic shadows as snow.

3.5 Summary

The performance of three classification methods (Fmask, parallax version), AT-COR, and Sen2Cor was evaluated on a set of 20 Sentinel-2 scenes covering all continents, different climates, seasons and environments. The reference maps with six classes (clear, semi-transparent cloud, cloud, cloud shadow, water, snow/ice) were created by an experienced human interpreter. The average overall accuracy is 54%, 59%, and 62% for Fmask, ATCOR, and Sen2Cor, respectively. High values of producer's accuracy ($> 80\%$) were achieved for cloud and snow/ice, and lower values for the other classes typically range between 30% and 70%. Future improvements could involve texture measures and convolutional neural networks.

4 A newly developed algorithm for Cloud Shadow Detection - TIP Method

This chapter describes the peer reviewed journal paper [90]. It follows the evaluation performed in Chapter 3 and presents an improved cloud shadow detection algorithm based on thresholds, indeces and projections.

4.1 Problem Statement

The Earth has an annual cloud coverage of approximately 70% [25]. Such a high cloud coverage in multi-spectral satellite images, as for example given by Sentinel-2, can be seen as a negative effect on many remote sensing tasks [76]. This inevitable contamination limits the ability of a continuous observation of a location on the Earth and degrades the information that can be extracted from a scene [44]. Especially for land applications the amount of scenes with usable data is of high importance. For example because the timing is important (crop yield estimation) or because the scene is not free-of-charge and one has to pay for the next acquisition if the current one contains cloud shadow over the location of interest. The customer of such applications might be interested to determine the ground properties ([64], [39],[85], [20]), or use the cloud and cloud shadow free image for geological applications ([82]) and crop monitoring tasks ([76]). Therefore, the correct and exact masking of clouds and cloud shadows is an important preprocessing step for atmospheric correction and also required for the shadow removal of multi-spectral imagery.

In this chapter the cloud shadow detection is performed on multispectral satellite images covering the spectrum from the blue to the short-wave infrared 2.2 μm. The newly presented method can overcome the disadvantages of only being applicable to one sensor, as is the case for [29], since only specific spectral channels have to be present (VNIR and SWIR) and does not have to undergo complex multistage processes as done by [10], [44], [16], [27] and [96].

In this chapter, the PACO atmospheric processor is used. This is the python-based version of ATCOR. For PACO the input data is in L1C radiances in units of $[mW/(cm^2 * sr * \mu m)]$. If the input scene is given in terms of TOA reflectance, it has to be converted into TOA radiance solving eq. 2.36 for L.

PACO performs the atmospheric correction using the spectral information in all bands by resampling them to a 20m cube, yielding an image cube of 13 bands with a size of 5490 x 5490 pixels. This so called "merged cube"will be a Sentinel-2 TOA cube considered in the rest of this study.

4.2 Methods

4.2.1 Current PACO shadow masking

PACO performs a pre-classification prior to atmospheric correction, which classifies the pixels of the input data into different classes. This is an important step since the following atmospheric correction relies on a correct pre-classification. PACO uses the spectral channels in the visible, near infrared (NIR) and short-wave infrared (SWIR). Additionally, empirical thresholds on the TOA reflectance are used to determine the selection of the pre-classification classes. In PACO, the basic classes are land, water, snow/ice, cloud (non-cirrus), cirrus cloud, haze, shadow and topographic shadow.[72]

The class of shadow is currently calculated using spectral and geometric criteria, whereas the topographic shadows are classified with the digital elevation model (DEM) and the solar geometry.

So far the cloud shadow masking incorporated in PACO performs very well under good conditions but fails for Sentinel-2 images that are located in desert regions and have features such as dunes and bright sand and for scenes mainly covered by thick ice. In this paper a new cloud shadow masking method for PACO is presented which will overcome these disadvantages. In a first part, the new TIP method is introduced (see section 4.2.2). In a second part, the results obtained from the new method are shown in Section 5.4. For the final part of the paper, the new method implemented in PACO will be compared with the cloud shadow detection of ATCOR through the validation done by [91] (see section 4.4).

4.2.2 TIP Method

This study describes the newly developed masking criteria that overcome the drawbacks of the current cloud shadow masking. The TIP method is named after its step by step computation of the cloud shadow map. These steps include the evaluation of *Thresholds*, *Indices* and *Projections*. The exact order of the cloud shadow detection within the TIP method is listed below and will be explained in detail in the sections 4.2.2 to 4.2.2.

- **T**hreshold selection - (T)

- Difference **I**ndex SWIR-NIR (DISN) - (I)

- Normalized Difference Water **I**ndex (NDWI-green) for water correction - (I)

- Elimination of small isolated pixels or patches and smoothing of borders

- Elimination of shadows not corresponding to the present cloud **P**rojection (P)

TIP input image and preparation for masking

In this new approach, the RGB bands (RGB = 665,560,443nm) are extracted from the full spectral TOA reflectance cube (ρ^*). Then the RGB image is re-ordered into a BGR image (with B = 490 nm being the first band) and the channels are mapped in the range from 0 up to 255 through linear stretching between the minimum and maximum range.

TIP masking: Threshold selection

The first step of the shadow masking is the calculation of the mean and the standard deviation of the BGR channels of the converted input image. These mean and standard deviations are calculated only including the classes of cloud shadow pixels and clear pixels. Hence, an accurate threshold selection for the cloud shadow pixels can be determined from the signal pixel distribution histogram which only includes the distribution from clear and cloud shadow pixels. In order to perform this step as accurately as possible, by only evaluating the signal arriving from cloud shadow pixels and clear pixels, the background, cloud and topographic shadow pixels are included in a mask and excluded from the calculation.

The mean and standard deviation are calculated for each band (B,G and R) across all remaining pixels and the threshold is calculated as done in the LAB algorithm described by [55] but slightly adapted to the Sentinel-2 Multi Spectral Instrument (MSI), an instrument covering the spectral range from 0.44 μm to 2.2 μm. The threshold is calculated for a confidence interval of ~ 1 σ, hence for 66.67% of the values lying within this interval. This corresponds to a fraction of $\frac{1}{3}$. The threshold per channel, i, can be written as follows:

$$T_i = mean_i - \text{standard deviation}(\frac{i}{3}) \tag{4.1}$$

where i stands for the blue (B), green(G) and red(R) bands.

Now the shadow pixel selection is done for two conditions:

- If the sum of the mean values in the green and red channel is less or equal to 45, then all the pixels that lie between 0 and 100, 100 and T_{red} (for the B, G and R channel, respectively) are selected as shadow pixels and others as non-shadow pixels.

- Else, all pixels that lie between 0 and T_{blue}, 255, T_{red} (for the B, G and R channel, respectively) are selected as shadow pixels and others as non-shadow pixels.

The threshold of 100 for the blue and green channel is selected since the values in

the shadow region are lower, as mentioned by [55]. To not loose any cloud shadows in scenes with a lot of water bodies, a threshold of 45 for the sum of the mean values in the green and red channel was determined.

After profound evaluations Sentinel-2 scenes it was clear that most cloud shadows lie below the threshold of 100, and bright pixels, such as cloud pixels, lie above the threshold of 100. The threshold selection alone does not give satisfying results. Dark vegetation and water pixels are typical examples of miss-classified pixels if no additional processing steps are included. Hence in the next sections, further fine tuning steps have been implemented to improve the shadow map.

TIP masking: DISN for dark vegetation correction

In order to avoid the miss-classification of dark vegetation pixels, a new spectral index is presented for the cloud shadow detection. This index is defined as the *difference deshadowing index SWIR-NIR* (DISN). The two bands were chosen since healthy vegetation shows a high reflectance in the NIR spectrum and dark areas, that might be due to cloud shadows, show a high reflectance in the SWIR spectrum [8].

The spectral index calculation for the shadow map is evaluated and created through trials with multiple scenes that will be discussed in section 5.4. The best result was obtained when calculating the difference of the TOA reflectance, ρ^*, between a SWIR2 band around 2200 nm and a NIR band around 860 nm. (See equation 4.2)

$$DISN = \rho^*_{SWIR2} - \rho^*_{NIR} \tag{4.2}$$

Once the DISN has been computed for the original image, the peak of the histogram of the DISN is evaluated. With this peak, a threshold is defined as follows:

$$threshold_{DISN} = peak_{DISN} + |peak_{DISN} * 0.30| \tag{4.3}$$

Now all shadow pixels defined in 4.2.2 that satisfy DISN $<$ $threshold_{DISN}$, are omitted from the shadow mask and classified as non-shadow pixels. It can happen that there is more than one peak in the histogram of the DISN. If this case applies, then the lower peak is selected for computation.

TIP masking: NDWI-green for water correction

For the water pixel detection and elimination, the histogram of the green normalized difference water index ($NDWI_{green}$) is used. The NDWI allows to differentiate water from dry land. Water bodies are known to have a low reflectance and strong absorption in the NIR wavelength range. The green band has a relatively high reflectance

for water bodies compared to the reflectance from cloud shadows [48]. Hence the NDWI uses the infra-red and green bands of remote sensing images [7]. The NDWI green is given as [17]:

$$NDWI_{green} = \frac{\rho^*_{green} - \rho^*_{NIR}}{\rho^*_{green} + \rho^*_{NIR}} \tag{4.4}$$

For each scene, a pixel distribution of the $NDWI_{green}$ is made. All pixels of the scene are included in the $NDWI_{green}$ computation. The distribution of this index is expected to be divided into two populations with two peaks. The pixels with the values around the first peak (low NDWI) include the shadow pixels and the pixels with a $NDWI_{green}$ value around the second peak (higher NDWI) are water pixels. Hence, this makes it possible to differentiate between dark water pixels and shadow pixels. The value of the valley to the right of the main peak is detected and all pixels with a $NDWI_{green}$ value above this valley are omitted from the cloud shadow pixels and assumed to be part of the water classification.

If more then two peaks are present, the differentiation between water and shadows is done using the valley to the right of the main peak.

Examples of the water pixel elimination can be found in section 4.3.2.

TIP masking: small isolated pixels or patches and smoothing of borders
This fine tuning step of the shadow masking removes all dark fields, vegetation or dark pixels that are too small to be part of a cloud shadow.

This is done by using the current shadow map and splitting it into separate patches ("labels") [42]. The size of each patch is calculated, so that the number of shadow pixels contained in each patch is known. Only patches that contain an area larger than 100x100 m, corresponding to 25 pixels at 20 m of Sentinel-2, are selected as shadows and the rest is excluded from the shadow map. Finally the shadow map is smoothed for the final cloud shadow border pixel correction. The smoothing is performed by selecting 5 pixels into each direction around the cloud shadow pixels to be smoothed.

TIP masking: Cloud Projection
In order to remove the rest of the shadows that are wrongly classified as cloud shadows, the cloud projection is evaluated for each classified cloud shadow pixel. If the projected cloud belongs to a cloud pixel or to a pixel outside of the scene, then the shadow is considered as cloud shadow. The rest of the shadow pixels are excluded from the cloud shadow mask.

As shown by [84], the orthographic position of the cloud and the cloud shadow can

be expressed through geometric equations which include the cloud height, H, the satellite view zenith and azimuth angles, θ_v and ϕ_v, respectively, and the sun zenith and azimuth angles, θ_s and ϕ_s, respectively.

Since the cloud projection step in the TIP method calculates the cloud projection pixel for each detected cloud shadow pixel the equations given by [84] are rearranged for the corresponding cloud projection pixel on the scene (see equation 4.5).

$$
\begin{aligned}
x_{projection} &= X_{shadow} - H * (tan(\theta_v)sin(\phi_v) - tan(\theta_s)sin(\phi_s)) \\
y_{projection} &= Y_{shadow} - H * (tan(\theta_v)cos(\phi_v) - tan(\theta_s)cos(\phi_s))
\end{aligned}
\tag{4.5}
$$

As recently shown by [40], the direction of the cloud shadow can be given by the apparent solar azimuth, ϕ_a.

To fully estimate the correct location of the cloud with respect to the cloud shadow, the distance, d, between the pixel of the cloud shadow and the cloud projection on the image plane has to be calculated. For the distance estimation, the sun and viewing angles and the cloud height are needed [40].

The apparent solar azimuth and the distance, d, are calculated for all currently estimated cloud shadow pixels and the mean is evaluated. In order to get the distance for the x and y coordinates of the cloud shadow pixel, d_x and d_y are calculated as follows (equation 4.6):

$$
\begin{aligned}
d_x &= sin(\phi_a) * d_{mean} \\
d_y &= cos(\phi_a) * d_{mean}
\end{aligned}
\tag{4.6}
$$

For the final location of the cloud pixel with respect to the estimated cloud shadow the distance is added to each cloud shadow pixel (see equation 4.7).

$$
\begin{aligned}
X_{final} &= x_{projection} + \frac{d_x}{H} \\
Y_{final} &= y_{projection} + \frac{d_y}{H}
\end{aligned}
\tag{4.7}
$$

The estimation of the correct cloud height, H, is done with an iterative approach, since the cloud altitude is unknown. The cloud projection is calculated over a range of cloud altitudes (from 0.5 to 10 km) and stops, when the projected cloud shadow

Tabelle 4.1: Sentinel-2 level L1C test scenes. Information on scene climate, main surface cover, rural/urban. (SZA = Solar Zenith Angle)

Scene	Location	Date	Tile	SZA	Desert	Ice/Snow	Nonveg	Veg	Water	Mountains	Rural	Urban
1	Africa, Gobabeb	2019/03/06	T33KWP	32.5°	X					X	X	
2	Antarctic	2018/01/26	T34DFH	58.8°		X	X		X			
3	France, Arcachon	2017/11/15	T30TXQ	63.9°			X	X	X		X	X
4	France	2016/01/16	T31TFJ	66.8°			X	X	X	X	X	X
5	Morocco, Quarzazate	2018/08/30	T29RPQ	27.2°	X		X	X	X			
6	Netherlands, Amsterdam	2018/09/13	T31UFU	49.7°				X	X		X	X

pixels have its maximum match with the cloud and/or cirrus mask given by PACO.

4.3 Results

4.3.1 Data and Material for training set

To test the new TIP cloud shadow masking on a set of data, six Sentinel-2 (S2) scenes are chosen. A list of the investigated Sentinel-2 A and B scenes is given in Table 5.1. The scenes were selected to cover different regions of Europe, Africa and Antarctica. Hence, the very different desert, ice and continental climates are investigated. They represent flat and mountainous sites with cloud cover from 3% to 80% and include the presence of cumulus, thin and thick cirrus clouds. Additionally, the scenes represent different land cover types such as desert, urban, cropland, grass, forest, wetlands, sand and coastal areas. The range of solar zenith angles is from 27° to 67°. Only the results of 3 scenes (Gobabeb, Arcachon, Barrax-2) are shown in section 4.3.2 as examples.

4.3.2 Masking Sequence of TIP Method

This section will present the results of the TIP method explained in section 4.2.2. Figure 4.1 shows the results of the cloud shadow masking of the TIP Method on the Gobabeb scene. The left image shows the cloud shadow mask provided by the current ATCOR masking algorithm. The middle image shows the cloud shadow mask obtained through the TIP method. The image on the right of Figure 4.1 shows the original Gobabeb scene which was stretched into the RGB=665/560/490 nm bands for better optical comparison. As can be seen in Figure 4.1, the TIP method does improve the cloud shadow masking in the desert scene. The cloud shadows present in the right lower corner are not detected by ATCOR but detected with the TIP method. Furthermore, ATCOR faces some difficulties in the top part of the cloud shadow mask. This can be seen in form of stripes and will lead to multiple misclassified pixels.

To demonstrate the importance of the TIP method masking step that corrects for misclassified water pixels using the NDWI-green threshold as mentioned in section 4.2.2, a Sentinel-2 scene from Arcachon is used. This scene acquired on the 15
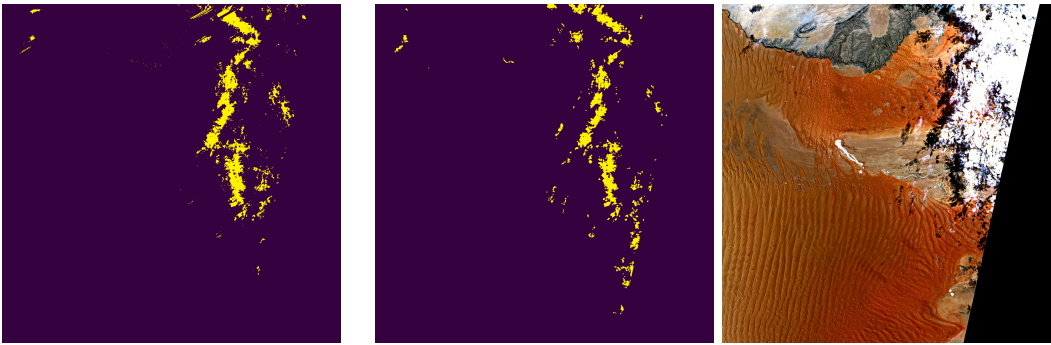
Abbildung 4.1: Gobabeb shadow map. Left: cloud shadow map of ATCOR; Middle: TIP cloud shadow map; Right: Gobabeb RGB=665/560/490 nm with linear histogram stretching.
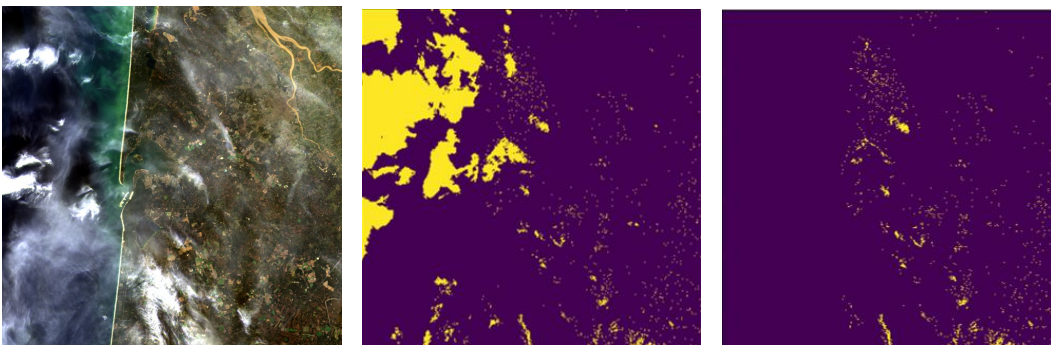
Abbildung 4.2: Arcachon Scene analysis of the NDWI. The left image represents the original TOA radiance image. The figure in the middle represents the shadow mask before the water correction. The figure on the right represents the shadow mask after the water correction.

November 2017 is located on the seaside and has a lot of water bodies (see Figure 4.2).

The left image of Figure 4.2 shows the true color image of the Scene. The middle image represents the shadow mask before the water correction and the right image the shadow mask after the water correction. All major water bodies seen in yellow in the middle image of Figure 4.2, not excluded previously, are excluded from the shadow map.

4.4 Validation of Results

4.4.1 Validation Statistic

In order to quantitatively validate the results obtained from the TIP method, the validation done by [91] was used, where twenty Sentinel-2 scenes are processed with three different mono-temporal masking codes. These masking codes are *Function of mask* (Fmask FORCE version 3.0), ATCOR (version 9.3.0) and Sen2Cor (version 2.8.0). The scenes were were selected to cover all continents in order to

represent different climates, seasons and weather conditions. This enables to cover a wide range of land cover types (desert, urban, cropland, grass, forest, wetlands, sand, coastal areas, glaciers). The range of solar zenith angles is from 18° to 62°. No solar zenith angle above 70° was chosen since ATCOR and Sen2Cor have this angle as the upper limit. This is due to extreme BRDF effects for angles above this limit and the Scenes are mostly covered by shadows from buildings, trees, etc. Furthermore, the validation aims for an amount of 1000 randomly selected samples per image with minimum number of 50 samples for the smallest class ([28], [60], [78], [83]).

The validation in this section compares the performance obtained by the ATCOR algorithm in [91] with the newest masking code of PACO (revision 1432) which includes the TIP method for the cloud shadow class.

The reference data obtained has been slightly changed since more pixels were validated than in the comparison performed by [91]. Therefore, the exact same validation steps were performed and the statistic was repeated for the ATCOR masking algorithm. The results consist of a confusion matrix with the number of correctly classified pixels in the validation set. This confusion matrix enables the computation of the *user accuracy* (UA), *producer accuracy* (PA) and *overall accuracy* (OA) of classification.

User, Producer and Overall accuracy

The UA for a specific class gives the percentage of the area mapped as this class that has reference of the class. The PA of a specific class gives the percentage of the area of the reference class that is mapped as this class. The OA of a specific class is the total amount of correctly classified pixel of this class divided by the total amount of reference pixel of this same class ([19], [79]). Hence, the UA and PA are the important accuracies to fully understand the statistic. The OA will not add any new statistical information if UA and PA are known. Hence, one can claim that the OA will partially hide the lack of accuracies present in the UA or PA.

4.4.2 Sentinel-2 Validation Results

Table 4.2 and Table 4.3 provide the results for absolute validation of classifications of the ATCOR and PACO masking algorithms, comparable to the results presented in [91].

A masking algorithm which will satisfy not only the producer but also the user will aim to obtain high values in both, PA and UA. Such an algorithm produces a classification map that detects as many pixels of each class as possible but misclassifies as few pixels as possible per class.

Tabelle 4.2: Summary of classification accuracy (percent) for difference area   (
A=ATCOR, P=PACO; bold face numbers indicate the best performances).

Class	UA (A)	UA (P)	PA (A)	PA (P)
clear	74.6	**75.7**	67.1	**80.4**
semi-transp. cloud	**61.1**	53.8	35.8	**38.8**
cloud	62.2	**80.5**	**67.4**	47.6
cloud shadow	**69.0**	57.3	64.9	**75.3**
water	48.2	**91.5**	**82.3**	81.3
snow/ice	**60.5**	51.4	**67.2**	66.9
topographic shadows	**32.0**	20.6	1.2	**2.4**

Table 4.2 provides a summary of the UA and PA per class, i.e. the average over all 20 scenes. If we concentrate our error matrix in Table 4.2 on the classification of cloud shadows we can make following conclusion for the classification done by PACO: The cloud shadow classification is 75.3% accurate. This high value could lead a producer to conclude that this classified Sentinel-2 image is sufficiently accurate for his needs. However, a user will be disappointed to find that only 57.3% of the image identified as cloud shadows on the classification map are actually cloud shadows. Hence, 75.3% of the cloud shadows has been correctly identified as such, but only 57.3% of those areas identified as cloud shadows are actually cloud shadows while 42.7% of those areas identified as cloud shadows belong to another class.

It can be noted that Table 4.3 is divided into three sections. The top sections contains the OA of the class cloud shadow of all the scenes for each masking algorithm. The middle sections contains the OA of the class cloud shadow per scene for each masking algorithm and for the scenes where cloud shadows are present. The bottom section of Table 4.3 shows the false positive percentage for the scenes that do not contain any cloud shadow. For this type of scenes the statistics references have been changed and, since no cloud pixels are expected, the accuracy of the false positives is evaluated: 100.0% means that no pixels are wrongly classified as cloud shadows. Depending on the misclassified pixels, the OA will deviate from 100.0%. Boldface numbers indicate the method with the best performance. Overall, PACO gives the highest OA over all scenes with 76.6%. Following PACO is ATCOR with an OA of 70.4%. Hence, PACO has improved the ATCOR results in cloud shadows.

Only the scenes of Mexico and Antarctica show worse results of PACO compared to ATCOR. This is due to the lack of correct cloud and water maps.

Table 4.4 shows the UA and PA of the cloud shadows for each scene and for ATCOR (A) and PACO (P) respectively. The Table shows, that PACO performs best for the PA of the class cloud shadow and has a high OA since the values for the UA perform better than ATCOR in many cases. ATCOR has high UA for the class cloud shadow

Tabelle 4.3: Summary of cloud shadow class overall accuracy (percent) (A=ATCOR, P=PACO; bold face numbers indicate the best performances).

		OA Difference Area	
Scene	Location	A	P
ID	Average (all scenes)	70.4	**76.6**
1	Antarctic	0	0
6	Estonia, Tallin	**85.0**	84.5
8	Italy, Etna	63.5	**83.5**
9	Kazakhstan, Balkhash	57.5	**78.5**
10	Mexico, Cancun	**64.5**	42.5
11	Morocco, Quarzazate	91.5	**94.5**
12	Mosambique, Maputo	7.5	**46.5**
13	Netherlands, Amsterdam	65.0	**76.0**
14	Phillipines, Manila	66.0	**71.0**
16	Russia, Yakutsk	73.0	**88.0**
17	Spain, Barrax-1	67.0	**78.5**
18	Spain, Barrax-2	76.0	**87.5**
19	Switzerland, Davos	52.5	**55.5**
20	USA, Rimrock	48.5	**67.5**
2	Argentina, Buenos Aires	**99.8**	93.7
3	Australia, Lake Lefroy	**100.0**	**100.0**
4	Bolivia, Puerto Siles	**99.8**	98.4
5	China, Dunhuang	92.3	**92.4**
7	Germany, Berlin	**99.7**	94.0
15	Russia, Sachalin	**99.2**	**99.2**

Tabelle 4.4: Summary of cloud shadow class user and producer accuracy (percent)
(A=ATCOR, P=PACO; bold face numbers indicate the best performances).

Scene	Location	UA(A)	UA(P)	PA(A)	PA(P)	annotated cloud shadow pixels
1	Antarctic	0	0	0	0	35
6	Estonia, Tallin	**93.0**	87.0	77.0	**82.0**	1415
8	Italy, Etna	**88.0**	80.0	39.0	**87.0**	1446
9	Kazakhstan, Balkhash	76.0	**93.0**	39.0	**64.0**	809
10	Mexico, Cancun	**91.0**	60.0	**38.0**	25.0	755
11	Morocco, Quarzazate	88.0	**93.0**	95.0	**96.0**	32318
12	Mosambique, Maputo	**4.0**	3.0	11.0	**90.0**	82
13	Netherlands, Amsterdam	**77.0**	68.0	53.0	**84.0**	857
14	Phillipines, Manila	**74.0**	58.0	58.0	**84.0**	1076
16	Russia, Yakutsk	80.0	**96.0**	66.0	**80.0**	958
17	Spain, Barrax-1	87.0	**93.0**	47.0	**64.0**	1592
18	Spain, Barrax-2	73.0	**86.0**	79.0	**89.0**	25115
19	Switzerland, Davos	61.0	**62.0**	44.0	**49.0**	2561
20	USA, Rimrock	45.0	**69.0**	52.0	**66.0**	847
2	Argentina, Buenos Aires	**99.6**	87.4	**100.0**	**100.0**	0
3	Australia, Lake Lefroy	**100.0**	**100.0**	**100.0**	**100.0**	0
4	Bolivia, Puerto Siles	**99.5**	96.8	**100.0**	**100.0**	0
5	China, Dunhuang	84.6	**84.8**	**100.0**	**100.0**	0
7	Germany, Berlin	**99.3**	88.1	**100.0**	**100.0**	0
15	Russia, Sachalin	98.3	**98.5**	**100.0**	**100.0**	0
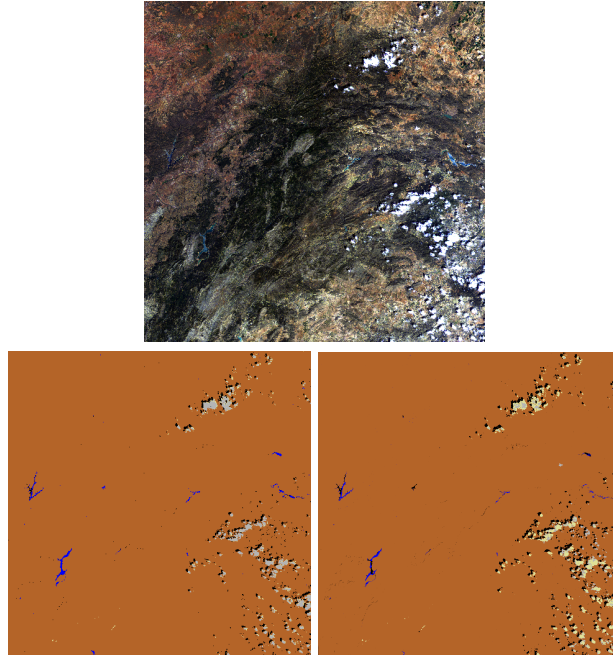
Abbildung 4.3: Top row: true color (RGB = 665,560,443nm) composite of scene ID 18 (Barrax-2). Bottom row (left to right):ATCOR and PACO classification maps (with clear (brown), cloud (grey), water (blue), shadow (black) and cirrus cloud (yellow) pixels).

but lower PA than PACO in all scenes apart from Mexico. The right column of Table 4.4 furthermore indicates the amount of cloud shadow pixels that have been annotated in the reference data. The scenes that do not contain any cloud shadows are listed in a separate section within Table 4.4. In this section, the UA and PA for the false positive pixels of the class cloud shadows is shown. It can be seen that only for the UA of all masking algorithm, pixels are wrongly classified as cloud shadows and hence the accuracy deviates from 100.0%.

To furthermore compare the classification performance of ATCOR and PACO the classification maps for a specific scene and a subset taken from that scene are shown in Figure 4.3 and 4.4. Figure 4.3 shows scene number 18 from Spain (Barrax) taken on the 19th of March 2017 with a zenith angle of $22.0°$ and a azimuth angle of $143.2°$. In Figure 4.4 a subset of scene ID 18 can be found. It nicely illustrates the improvements that the TIP method adds to PACO with respect to the current ATCOR version. It also demonstrates, where each of the different algorithms lies its focus on. While ATCOR shows a very high PA in the cloud mask, PACO has the highest PA for the cloud shadows. This is due to high importance of the TIP method to reach high PA in the class cloud shadow for further cloud shadow removal. AT-COR, on the other hand, focuses more on a high UA for cloud shadows and hence does not classify cloud shadow very closely to the cloud shadow borders. PACO is designed to classify all the border pixels as cloud shadows.
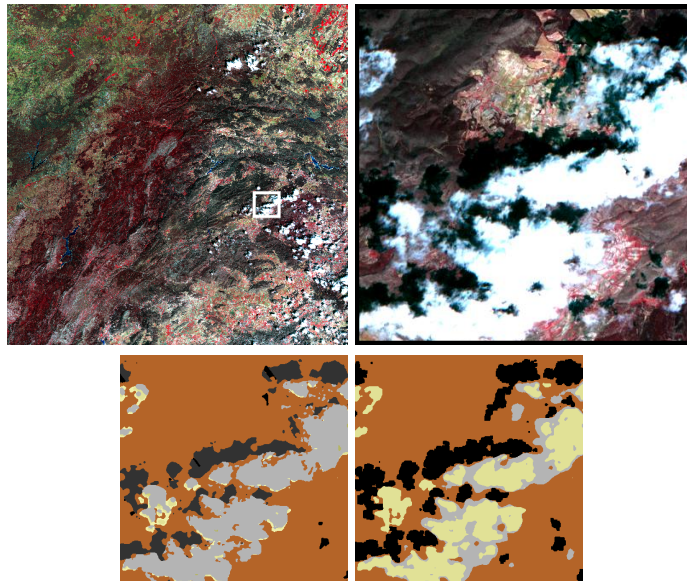
Abbildung 4.4: Top row (left to right): CIR (RGB = 865,665,560nm) composite and CIR subset of scene ID 18 (Barrax-2). Bottom row (left to right): ATCOR and PACO classification maps of the subset (with clear (brown), cloud (grey), shadow (black) and cirrus cloud (yellow)).

The confusion within and between classes can be illustrated in spider diagrams using the proportion of the individual class omissions for the difference area (Figure 4.5). The difference area is the part of the classification images where the classification maps provided by ATCOR and PACO disagree. Validation statistics over the difference area gives a good comparison between the strengths and weaknesses of each processor [91].

Figure 4.5 illustrates spider diagrams for the omission and commission for difference area of the classes clear and cloud shadows.

From the left upper image of figure 4.5 it can be noted that ATCOR has the highest omission of clear pixels towards water pixels. PACO has omission of clear pixels toward cloud shadows. This is due to the higher PA of PACO of cloud shadows. Due to PACO wanting to classify as many cloud shadow pixels also from the shadow border the PA is higher than its UA. This can also be seen in the lower left image of 4.5 where PACO has commission of clear pixels towards cloud shadows.

Furthermore, the difference of the UA and PA between ATCOR and PACO can be seen in the two images on the right of 4.5. The top right image shows that ATCOR has a higher omission of cloud shadow pixel towards clear pixels. This corresponds to its higher UA. The bottom right image shows that PACO has higher commission of cloud shadow pixels toward clear pixels which on the other hand corresponds to its higher PA. PACO has therefore a less conservative cloud shadow mask than ATCOR. This is due to PACO wanting to reach high values in PA in order to pro-
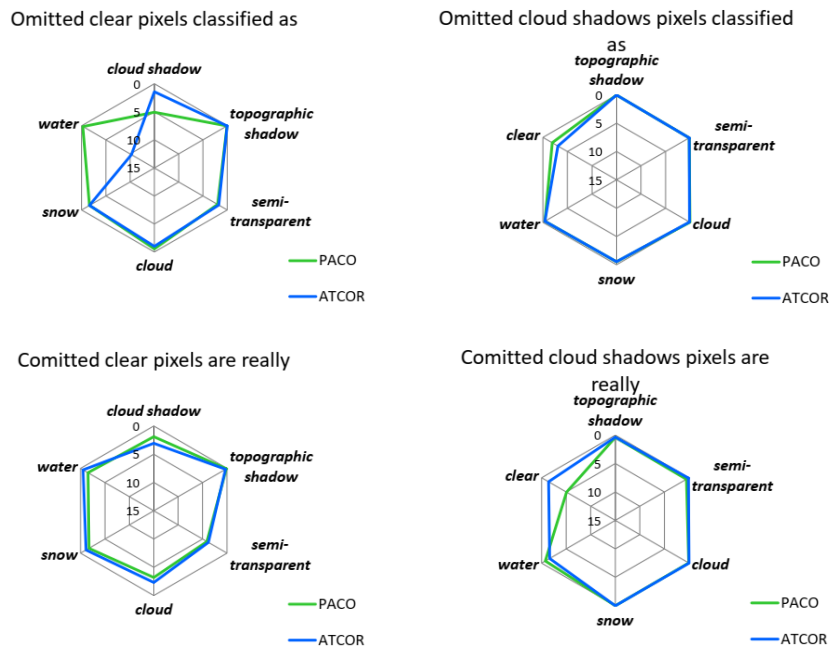
Abbildung 4.5: Omission and commission per Class for difference area for the classes clear and cloud shadows. PACO is represented in the color green and ATCOR in the color blue.

ceed with a shadow removal process. ATCOR, as proven by the omission of cloud shadow pixels, has a more conservative cloud shadow map. Hence, it has a lower PA value and will not be able to detect all cloud shadow pixels around the border of the shadow.

4.5 Summary

A newly developed cloud shadow masking method, called TIP, which has been tested in the PACO masking algorithm was proposed. Cloud shadows of multispectral satellite images covering the spectrum 0.44 μm to 2.2 μm are found through the steps including the evaluation of thresholds, indices and projections. This method is intended to improve the PA of the class cloud shadow of the current ATCOR and PACO masking algorithms, because it is oriented to be the previous step to a deshadowing method. Furthermore the performance of the class cloud shadows of the improved PACO version was statistically compared with ATCOR for S2 scenes. The overall accuracy of the class cloud shadow is 70.4% and 76.6% for ATCOR and PACO, respectively. Hence, the newly developed TIP method for cloud shadows has improved the current cloud shadow mask. The TIP method encounters limitations for scenes where the cloud and water maps are wrongly classified prior to the cloud shadow detection. Future improvements could involve the fine tuning of the cloud and water masks which will improve the results of the TIP method for cloud shadows. Additionally, the TIP method can not be fully implemented for

VNIR sensors like DESIS [13] but is applicable for VNIR-SWIR sensors such as EnMAP [35], PRISMA [9] and Landsat-8 [49], -9 [51].

5 Cloud shadow removal for high spatial resolution optical satellite data

This chapter describes the final part of the dissertation. After the comparison of masking algorithms (Chapter 3) and the evaluation of a new cloud shadow detection algorithm (Chapter 4), an improved cloud shadow removal algorithm for high spatial resolution optical satellite data is presented. This chapter describes the submitted peer reviewed journal paper [89].

5.1 Problem Statement

For optical remote sensing of the Earth's surface, clouds and their shadows have always been a major disadvantage, since a lot of remote sensing applications are impacted by their presence. These applications involve, for example, radiation, image classification, the calculation of surface reflectance or land surface temperature, vegetation indices, etc. [84] [76]

The annual cloud coverage of the Earth lies around approximately 70% [25]. Therefore it is inevitable to get a continuous cloud and cloud shadow free observation of a specific location on Earth and the information that can be extracted from a scene will have a high percentage of degradation [44]. This means, that scientists will have to find ways to work around or with the presence of clouds and cloud shadows. This is especially crucial for land applications for which the amount of usable data per scene and specific timing is of high importance, for example for crop yield estimation [50]. Furthermore, the use of cloud and cloud shadow free images enable to determine ground properties of the Earth's surface ([64], [39],[85], [20]) and facilitate crop monitoring tasks ([76]). Even geological applications ([82]) are disturbed if clouds and their shadows cover parts of high spatial resolution optical satellite data. This proves how important it is nowadays to have a correct and exact masking of clouds and cloud shadows as preprocessing step for atmospheric correction and shadow removal of multi-spectral imagery. Hence, in the past years, more and more cloud and cloud shadow detection and removal approaches have been developed ([47],[53],[26],[65],[43],[32],[81],[22]) and used to enable various applications.

In order to undergo a cloud shadow removal algorithm, the cloud and their shadows have to be detected and mapped. The detection of clouds can be done by studying each satellite scene separately by using a mono-temporal approach ([95],[96],[11], [46], [38], [15], [92], [56]) or through a multi-temporal methodology ([74], [36]) and hence studying a time series of images. For the detection of the correct location and geometry of a cloud shadow, the direction of observation is crucial since they

represent projections of clouds in an image [84]. In this paper, a mono-temporal cloud shadow detection approach is used as preprocessing step for cloud shadow removal, called TIP method [90].

Due to remotely sensed optical imagery of the Earth's surface being contaminated by cloud and cloud shadows, the surface information underneath a cloud covered region cannot be retrieved with optical sensors. The surface information underneath a cloud shadow, on the other hand, can be retrieved since the ground reflected solar radiance is a small non-zero signal. Now the total radiation signal that is measured at the sensor is composed out of a direct beam and a diffuse, reflected skylight component. This means, that even if there is no direct solar beam arriving at the sensor from the shadow region, there will still be some information arriving from the reflected diffuse flux. [72]

The proposed shadow removal method works with this knowledge and uses the estimate of the fraction of direct solar irradiance for a fully or partially shadowed pixel as basis for the removal algorithm. The aim is to provide an improved cloud shadow removal algorithm based on the current version of the Matched Filter proposed by ATCOR [72]. As opposed to the IDL ATCOR algorithm, the new cloud shadow removal algorithm is implemented into the Python-based Atmospheric Correction (PACO) software.

In this chapter, the Multispectral Instruments (MSIs) of the Copernicus Sentinel-2 satellites are used [2]. The MSIs are sensors on-board of the satellite, which allow free access to the data and a high revisit time [40]. Furthermore, the PACO atmospheric processor is used. This is the python-based version of ATCOR. For PACO the input data is in L1C radiances in units of $[mW/(cm^2 * sr * \mu m)]$. If the input scene is given in terms of TOA reflectance, it has to be converted into TOA radiance. PACO performs the atmospheric correction using the spectral information in all bands by resampling them to a 20m cube, yielding an image cube of 13 bands with a size of 5490 x 5490 pixels. This so called "merged cube"will be a Sentinel-2 TOA cube considered in the rest of this study.

5.2   Basic Concept of atmospheric correction

5.2.1   Radiation components and surface reflectance

The radiation signal in the solar region (0.35-2.5 μm) arriving at the sensor is due to four different components ([69], [70]):

- Path radiance: from photons that did not have contact with the ground and are scattered into the field-of-view of the sensor
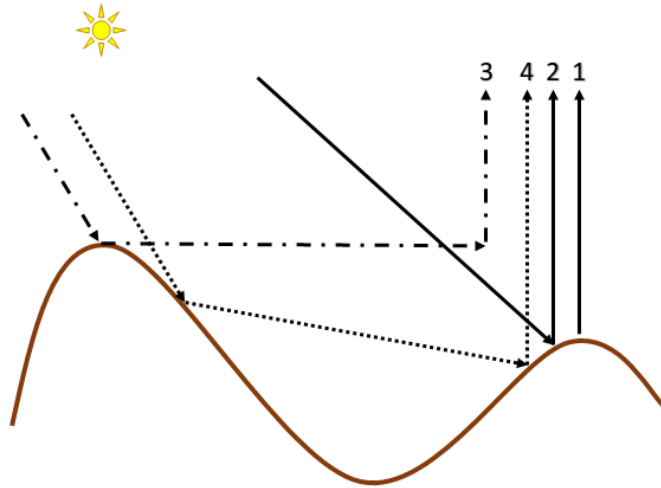
Abbildung 5.1: Radiation components arriving at the sensor. The path radiance is represented by component number 1. The ground reflected radiation is represented by component number 2. The reflected radiation from the surrounding is represented by component number 3. The reflected terrain radiance is represented by component number 4.

- Ground reflected radiation from a pixel: the fraction of the diffuse and direct solar radiation incident on the pixel that is getting reflected from the surface

- Reflected radiation from the surrounding: the fraction of the solar radiation reflected from the neighborhood and scattered by the air volume into the field-of-view of the sensor. This radiation is also called adjacency radiation.

- Reflected terrain radiance from opposite mountains

Figure 5.1 shows the four different components arriving at the sensor. For a full evaluation of the radiation component in rugged terrain, please refer to [70] and section 6.2 in [72].

From the four components, only the reflected radiation from a pixel contains the necessary information about the viewed pixel. Hence, in atmospheric correction, it is important to remove the other components and to retrieve the correct ground reflectance from the pixel of interest.

If we now combine all four components of the radiation to get the total radiation arriving at the sensor we can write:

$$L = L_{path} + L_{ground} + L_{adj} + L_{terrain} \tag{5.1}$$

5.3 Methods

5.3.1 Previous Method: Matched Filter

MF Method: surface reflectance and covariance matrix

Deshadowing is the compensation process which uses an estimate of the fraction of direct solar irradiance for a fully or partially shadowed pixel.

The MF method needs at least one channel in visible and at least one spectral band in the NIR. The bands used in the MF are: blue, green, red, NIR, SWIR1 and SWIR2, if existing.

The method starts with the calculation of the surface reflectance image cube, ρ. The surface reflectance, ρ, is computed with the assumption of fully solar illumination, excluding water and clouds. Then the covariance matrix, $C(\rho)$, is calculated where ρ represents the surface reflectance vector of the three selected bands (see equation 5.2).

$$V_{MF} = \frac{C^{-1}(\rho_t - \overline{\rho})}{(\rho_t - \overline{\rho})^T C^{-1}(\rho_t - \overline{\rho})} \tag{5.2}$$

The MF vector, V_{MF}, is tuned to a certain target reflectance spectrum, ρ_t, to be detected. $\overline{\rho}$ is the scene-average spectrum without water and cloud pixels. For the shadow target, a target reflectance spectrum of $\rho_t = 0$ is selected which will give the simplified version of the shadow MF vector, V_{sh}, as follows ([12]):

$$V_{sh} = -\frac{C^{-1}\overline{\rho}}{\overline{\rho}^T C^{-1}\overline{\rho}} \tag{5.3}$$

MF Method: unscaled shadow function

The MF shadow vector, V_{sh}, can be applied to the non-water and non-cloud part of the scene to give the un-normalized values, Φ, also called unscaled shadow function. Φ gives a relation measure of the fractional direct illumination (see equation 5.4).

$$\Phi(x,y) = V_{sh}^T(\rho(x,y) - \overline{\rho}) \tag{5.4}$$

MF Method: rescaling and scaled shadow function

The MF calculates a minimum RMS shadow target abundance for the entire scene. The values of Φ can be both, positive and negative. Therefore, Φ is rescaled into the physical range from 0 (full shadow) to 1 (full direct illumination). The histogram of the unscaled shadow function is used for rescaling and an illustration can be found in Figure 3 of [71].

The first peak of the histogram of Φ, Φ_2, represents the shadow pixels. On the other hand, the highest peak of the histogram, ϕ_{max}, represents the fully illuminated areas.

The rescaling of Φ is done by linear mapping of the Φ values from the unscaled interval (Φ_{min}, Φ_{max}) onto the physically scaled interval $(0,1)$. Hence, the scaled shadow function, Φ^*, is calculated as follows:

$$\Phi^* = \frac{\Phi - \Phi_{min}}{\Phi_{max} - \Phi_{min}} \quad \text{if} \quad \Phi \leq \Phi_{max}$$
$$\Phi = 1 \quad \text{if} \quad \Phi > \Phi_{max} \tag{5.5}$$

The scaling and normalizing of the MF vector into the $(0,1)$ interval is based on the assignment of:

- min and max direct sun fraction in the shadow regions (a_{min} and a_{max}; the defaults are $a_{min} = 0.20$ and $a_{max} = 0.95$)

- corresponding shadow thresholds Φ_{min} and Φ_{max} obtained from the normalized histogram of Φ

The normalized and scaled shadow function is given as:

$$\Phi_n = a_{min} + \frac{(\Phi - \Phi_{min})(a_{max} - a_{min})}{\Phi_{max} - \Phi_{min}} \tag{5.6}$$

MF Method: iteration

The potential shadow pixels are those which satisfy $\phi_n < 1$, but as can be seen from equation 5.6, the value strongly depends on Φ_{max}. Therefore, an iterative strategy is applied and the exact steps of the ATCOR MF iteration method can be found in [71].

Deshadowing reflectance equation

The scaled shadow function, Φ_n, represents the fraction of the direct illumination for each pixel in the surface reflectance vector, ρ. The MF method tries to find the core shadows and then subsequently expands these core regions. This enables a smooth shadow to clear transition. The scaled shadow function is only applied to the pixels in the final mask. The core shadow mask is defined by the pixels with $\Phi(x,y) < \Phi_T$.

The final deshadowing is performed by multiplying the direct illumination, E_{dir}, with the pixel-dependent Φ_n. This reduces the direct solar term and increases the brightness of a shadow pixel, since it is located in the denominator of the deshado-

wing equation (see equation 5.7).

$$\rho_i(x,y) = \frac{\pi(d^2(c_0(i) + c_1(i)DN_i(x,y)) - L_{p,i})}{\tau_i(E_{dir,i}\Phi_n(x,y) + E_{dif,i})} \qquad (5.7)$$

5.3.2 Proposed Method: Cloud Shadow removal MF Method with new additions

The proposed cloud shadow removal algorithm was created for data acquired by satellite/airborne sensors in multispectral and hyperspectral imagery. It is computed in order to create a fully automated shadow removal algorithm and based on the main concept of the MF method ([72], [71]).

The MF method was first implemented from the IDL ATCOR version into the python-based PACO software. The new cloud shadow removal algorithm incorporates the main equations from the MF method used by the ATCOR deshadowing algorithm after [72]. To improve the results, additions have been added and are proposed in this paper.

Figure 5.2 shows a flow chart of the eight main steps performed during the new cloud shadow removal algorithm.

The first step of the new method is the calculation of the surface reflectance for the bands required to perform the MF evaluation (blue, green, red, NIR, SWIR1 and SWIR2 if available). Constant atmospheric conditions, a standard atmosphere and a fixed visibility are assumed. Hence, the default of the visibility is set to 30 km which corresponds to an aerosol optical thickness at 550 nm of 0.32 for sea level.

The new additions are the cloud shadow map from the TIP method developed in [90] and a newly defined iteration over the scaled minimum shadow function, a_{min}.

The TIP cloud shadow detection is based on thresholds, indices and projections and has been able to improve the results of the previous ATCOR cloud shadow map calculation. The detailed evaluation of the TIP cloud shadow map calculation can be found in [90]. Having included a better cloud shadow map will improve the results of the shadow removal algorithm.

The second newly implemented addition is performed within the computation of the matched filter vector Φ. The normalization of Φ into the physical range between 0 and 1 is evaluated with the minimum and maximum direct sun fraction in the cloud shadow area, a_{min} and a_{max}, and their corresponding shadow thresholds Φ_{min} and Φ_{max}. Φ_{min} and Φ_{max} are obtained from the normalized histogram of Φ. The default values of a_{max} is set to 0.95 and, as opposed to the previous method, the starting default value of a_{min} is set to 0.01. Then the normalized scaled shadow function is
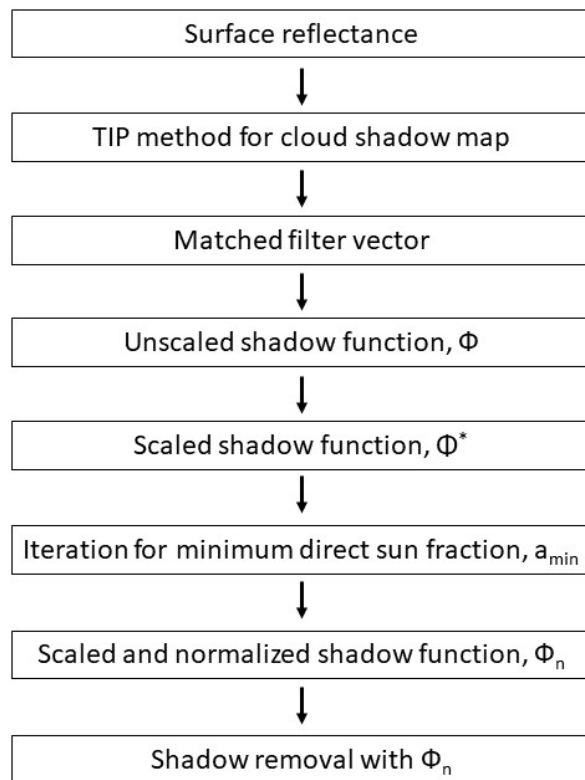
Abbildung 5.2: Flow Chart of the 8 main steps of the proposed cloud shadow removal method.

calculated using equation 5.6.

The proposed iteration calculates the mean reflectance of all shadow pixels (Φ_{sh}) and all sun-lit pixels (Φ_{sun}) for all bands, j, and terminates when the overall absolute difference of these two values, $\sum D(j)$, is less than the difference of the previous iteration, $\sum D(j)_{previous}$ (see equation 5.8).

$$\sum D(j) = abs(\Phi_{sh}(j) - \Phi_{sun}(j)) < \sum D(j)_{previous} \qquad (5.8)$$

When this limit is reached, then the two signals are as closely as possible. If the condition is not reached for a_{min} in the range 0.01-0.3, then the iteration stops and takes the upper limit of a_{min}=0.3 as default value. The iteration condition is performed for the total reflectance, hence taking into account all bands.

The final calculation of the reflectance is done using equation 5.7 with the corresponding scaled shadow function.

5.4 Results

The results of the deshadowing algorithm presented in this paper are first shown through three selected scenes from the data set where a visual and spectral comparison is performed (see section 5.4.2). Additionally, a metric quantitative comparison for all evaluated scenes is shown and discussed in section 5.5.

5.4.1 Data and Material for training set

To test the new cloud shadow removal method on a set of data, twenty-five Sentinel-2 (S2) scenes are chosen. A list of the investigated Sentinel-2 A and B scenes is given in Table 5.1. The scenes were selected to cover a wide variety of regions over the entire globe (see Figure 5.3). This enables to validate the shadow removal algorithm for different continents, climates, seasons, weather conditions and land cover classes. Furthermore, they have been selected to represent flat and mountainous sites with a cloud cover from 3% to 80% and they include the presence of cumulus, thin and thick cirrus clouds. The land cover types represented are: desert, urban, cropland, grass, forest, wetlands, sand and coastal areas. The range of solar zenith angles is from 27° to 67°.

5.4.2 Cloud shadow removal Results

In the following section, three scenes (scene ID 18,16 and 10 from Table 5.1) are chosen to show the results of the cloud shadow removal algorithm. For each scene, the deshadowing results are given for a subset in order to better evaluate the results spectrally and visually. In each subset a clear pixel and a shadow pixel are selected

Tabelle 5.1: Sentinel-2 level L1C test scenes. Information on scene climate, main surface cover, rural/urban. (SZA = Solar Zenith Angle)

Scene	Location	Date	Tile	SZA	Desert	Ice/Snow	Nonveg	Veg	Water	Mountains	Rural	Urban
1	Africa, Gobabeb	2019/03/06	T33KWP	32.5°	X					X	X	
2	Africa, Namibia			°	X					X	X	
3	Antarctic-1	2019/02/03	T21EVK	54.9°		X	X		X			
4	Argentina, Buenos Aires	2018/08/27	T21HUC	51.5°					X		X	X
5	Australia, Lake Lefroy	2018/08/19	T51JUF	51.5°			X					
6	Bolivia, Puerto Siles	2018/09/06	T19LHF	30.6°				X	X		X	
7	China, Dunhuang	2018/01/22	T46TFK	62.3°	X	X					X	
8	France, Arcachon-1	2017/11/15	T30TXQ	63.9°			X	X	X		X	X
9	France, Arcachon-2	2017/02/18		°								
10	France-3	2016/01/16	T31TFJ	66.8°			X	X	X	X	X	X
11	Estonia, Tallin	2018/07/14	T35VLG	39.0°					X		X	X
12	Germany, Berlin	2018/05/04	T33UUU	38.0°			X	X	X		X	X
13	Italy, Etna	2017/03/09	T33UUU	45.1°		X			X		X	X
14	Kazakhstan, Balkhash	2018/07/30	T43TFM	30.7°	X				X		X	
15	Mexico, Cancun	2018/05/27	T16QDJ	18.4°				X	X		X	
16	Morocco, Quarzazate	2018/08/30	T29RPQ	27.2°	X		X	X	X			
17	Mosambique, Maputo	2018/07/13	T36JVS	54.4°					X		X	X
18	Netherlands, Amsterdam	2018/09/13	T31UFU	49.7°				X	X		X	X
19	Philipines, Manila	2018/03/19	T51PTS	27.4°					X		X	X
20	Russia, Sachalin	2018/05/09	T54UVC	35.5°		X	X		X			
21	Russia, Yakutsk	2017/08/08	T52VEP	45.9°		X		X	X			
22	Spain, Barrax-1	2017/05/09	T30SWH	24.1°				X		X	X	
23	Spain, Barrax-2	2017/05/19	T30SWH	22.0°				X		X	X	
24	Switzerland, Davos	2019/04/17	T32TNS	37.7°		X	X	X	X	X		
25	USA, Rimrock	2018/05/12	T11TMM	30.4°		X			X		X	X
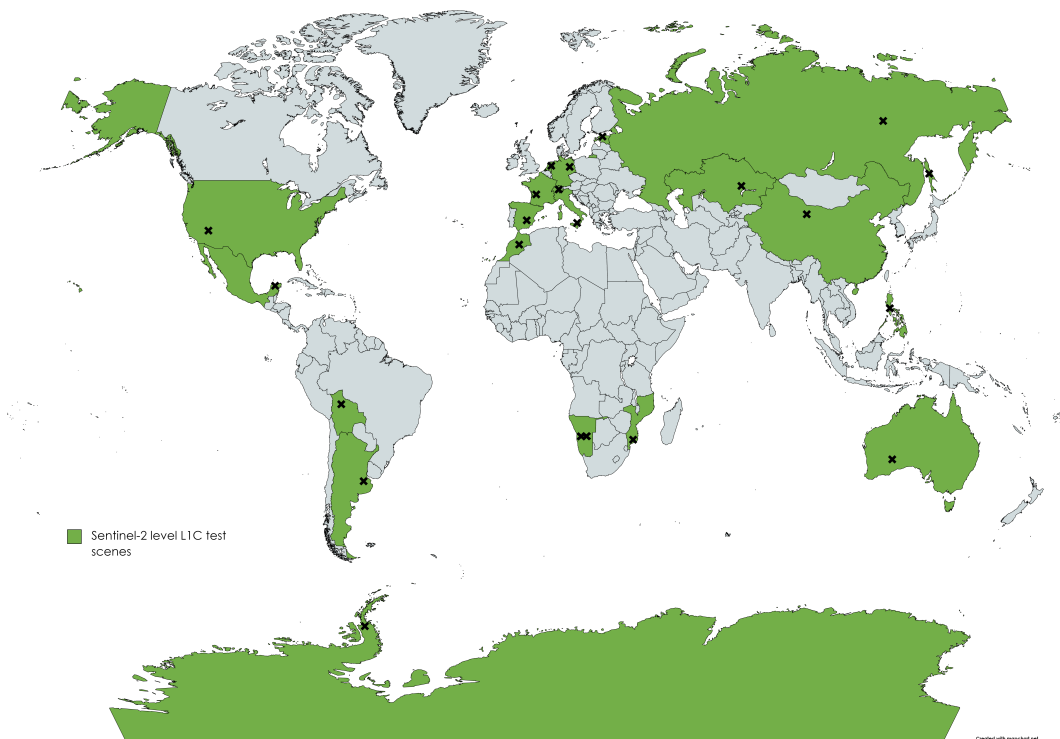
Abbildung 5.3: Geographical distribution of the Sentinel-2 selected test sites. Black crosses represent the exact location of the test sites.

which are located close by and represent the same ground properties. The visual and spectral comparison is done between the original scene, the new presented cloud shadow removal algorithm from PACO and the cloud shadow removal algorithm as given by ATCOR.

Netherlands, Amsterdam (Scene ID 18)

Figure 5.4 shows the first scene results to be analyzed in this paper. It is scene number 18 of Table 5.1, located in Amsterdam, Netherlands, and taken on the 13th of September 2018 with a zenith angle of $49.7°$. In order to show the visual results of the deshadowing algorithm, a subset from scene ID 18 was chosen. The exact location of the subset is shown in the top row of Figure 5.4. The left image of Figure 5.4 represents the original Amsterdam scene stretched into the RGB=665/560/490 nm bands for better optical comparison. The right image shows the chosen subset.

To compare the new method visually with the previous version, Figure 5.4 provides the subset from the original scene and the two deshadowed subsets from PACO (new version) and ATCOR (old version) in the middle row, respectively. For each subset the same zoom is provided (see Figure 5.4 bottom row). In this image zoom, a clear pixel and a cloud shadow pixel are chosen. In order to not only provide visual results, the spectra of the clear pixel, the cloud shadow pixel, the PACO deshadowed cloud shadow pixel and the ATCOR deshadowed cloud shadow pixel are presented in Figure 5.5.

The black curve of Figure 5.5 represents the reflectance spectrum obtained from the cloud shadow pixel without correction. The pink curve represents the chosen clear pixel close by. The orange curve represents the reflectance spectrum of the cloud shadow pixel after deshadowing with the new PACO version. Finally the blue curve represents the cloud shadow pixel after being deshadowed with ATCOR. As can be seen from Figure 5.5 both methods nicely deshadow the cloud shadow pixel and get very close to the true clear pixel reflectance spectrum. ATCOR slightly overshoots the compensation whereas the new PACO method closely follows the pink curve but slightly lower.

Morocco, Quarzazate (Scene ID 16)

In order to prove the promising results of the cloud shadow removal algorithm presented in this paper, a second scene is illustrated in Figure 5.6. Figure 5.6 shows scene number 16 located in Quarzazate, Morocco. The scene was taken on the 30th of August and has a zenith angle of $27.2°$. The top left image of Figure 5.6 shows the RGB=665/560/490 nm true color composite of Quarzazate and the top right image the chosen subset from this scene.

The same evaluation as done in Section 5.4.2 for scene ID 18 is performed on the
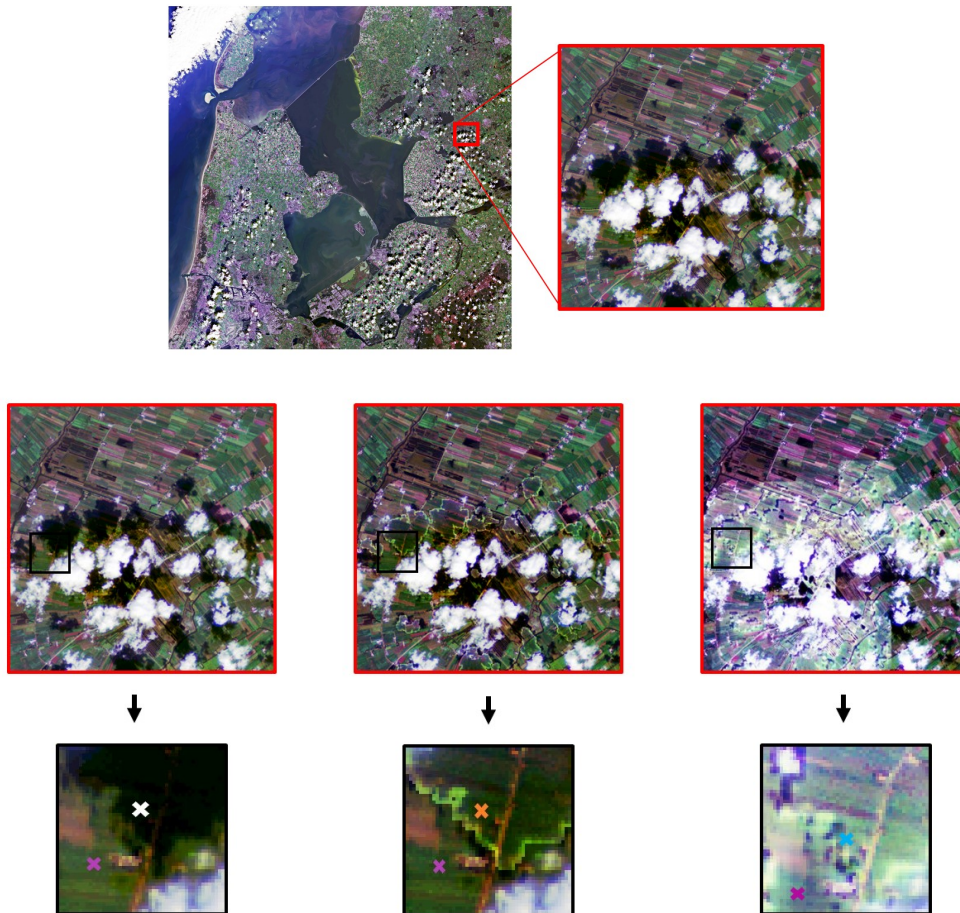
Abbildung 5.4: Top row (left): Amsterdam (Netherlands) RGB=665/560/490 nm with equalization stretching. Top row (right): subset of scene ID 18 with equalization stretching. Middle row (left to right): Subset of Amsterdam (Netherlands) RGB=665/560/490 nm with equalization stretching; PACO deshadowed subset of Amsterdam with equalization stretching; ATCOR deshadowed subset of Amsterdam with equalization stretching. Bottom row (left to right): zoom of subset of Amsterdam (Netherlands) RGB=665/560/490 nm with equalization stretching (white cross = cloud shadow pixel; pink cross = clear pixel); zoom of PACO deshadowed subset of Amsterdam with equalization stretching (orange cross = PACO deshadowed cloud shadow pixel; pink cross = clear pixel); zoom of ATCOR deshadowed subset of Amsterdam with equalization stretching (blue cross = ATCOR deshadowed cloud shadow pixel; pink cross = clear pixel).
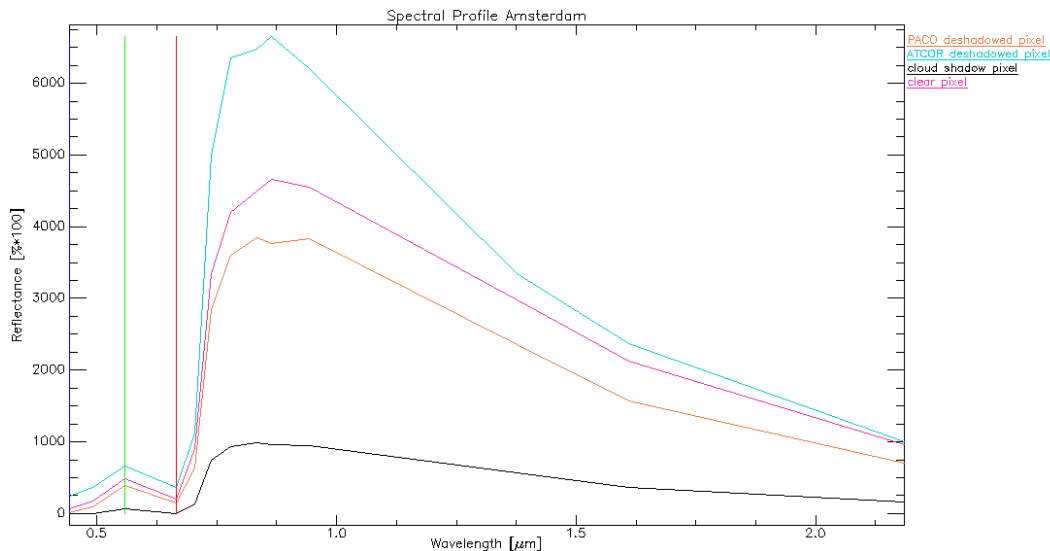
Abbildung 5.5: Spectral profile of scene ID 18 (Amsterdam). Orange = PACO deshadowed cloud shadow pixel; Blue = ATCOR deshadowed cloud shadow pixel; Black = original cloud shadow pixel spectrum; Pink = clear pixel spectrum.

Morocco example. Figure 5.6 shows the scene subsets of the original, PACO deshadowed and ATCOR deshadowed scene in the middle row and a zoom of the subset in the bottom row.

In Figure 5.7 the reflectance spectra for the clear and cloud shadow pixel of the original scene, the PACO deshadowed scene and the ATCOR deshadowed scene are given. The reflectance spectra again show how the ATCOR deshadowing algorithm rather overcompensates the cloud shadow pixel whereas the PACO deshadowing algorithm follows the reflectance spectrum of the clear pixel very nicely for low wavelengths and then slowly becomes smaller.

France (Scene ID 10)

As a final example, scene ID 10 is chosen. It represents a scene from France taken on the 16th January 2016 and has a zenith angle of 66.8°. The top row of Figure 5.8 shows on the left the original scene stretched into the RGB=665/560/490 nm bands for better optical comparison and the chosen subset on the right. For this example, two image zooms of the subset were taken.

As performed in sections 5.4.2 and 5.4.2, Figure 5.8 and 5.10 show the scene subsets of the original, PACO deshadowed and ATCOR deshadowed scene in the middle row and the first zoom of the subset in the bottom row. In Figure 5.9 the reflectance spectra for the clear and cloud shadow pixel of the original scene, the PACO deshadowed scene and the ATCOR deshadowed scene are given for zoom number 1. As can be seen from the visual zoom in Figure 5.8 and the reflectance spectra given in Figure 5.9, ATCOR is not able to deshadow this area of the cloud shadow. The
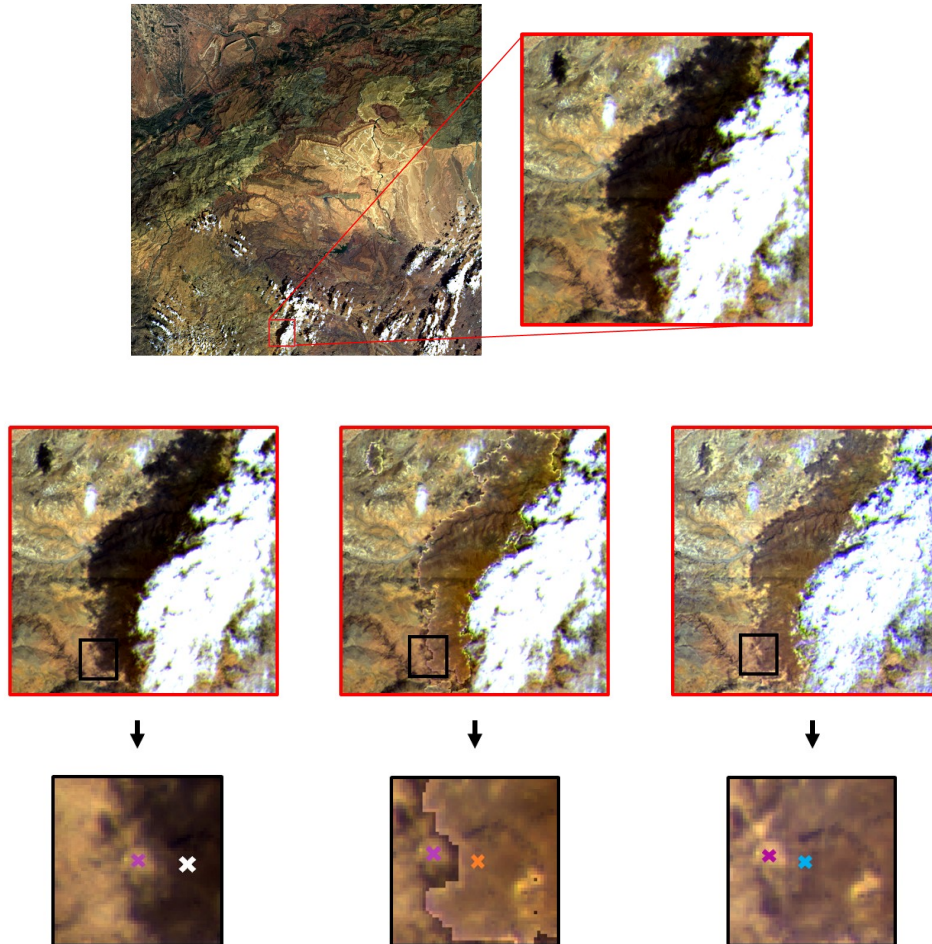
Abbildung 5.6: Top row (left): Quarzazate (Morocco) RGB=665/560/490 nm true color composite. Top row (right): RGB=665/560/490 nm true color composite subset of scene ID 16. Middle row (left to right): Subset of Quarzazate (Morocco) RGB=665/560/490 nm true color composite; PACO deshadowed subset of Quarzazate; ATCOR deshadowed subset of Quarzazate. Bottom row (left to right): zoom of subset of Quarzazate (Morocco) RGB=665/560/490 nm true color composite (white cross = cloud shadow pixel; pink cross = clear pixel); zoom of PACO deshadowed subset of Quarzazate (orange cross = PACO deshadowed cloud shadow pixel; pink cross = clear pixel); zoom of ATCOR deshadowed subset of Quarzazate (blue cross = ATCOR deshadowed cloud shadow pixel; pink cross = clear pixel).
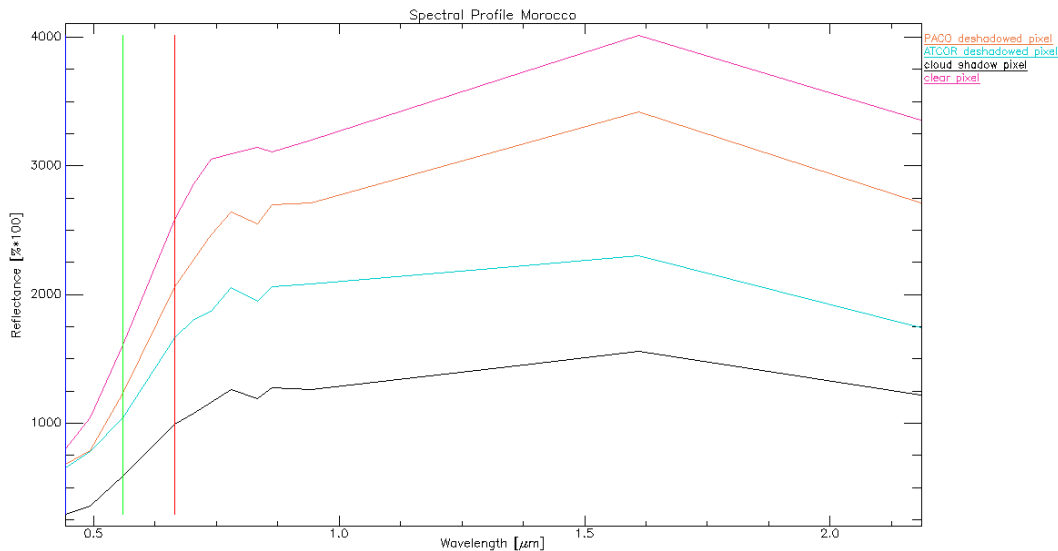
Abbildung 5.7: Spectral profile of scene ID 16 (Quarzazate). Orange = PACO deshadowed cloud shadow pixel; Blue = ATCOR deshadowed cloud shadow pixel; Black = original cloud shadow pixel spectrum; Pink = clear pixel spectrum.

new version on the other hand performs similar to the previous examples and nicely recovers most of the reflectance spectrum.

Figure 5.10 shows the scene subsets of the original, PACO deshadowed and ATCOR deshadowed scene in the middle row and the second zoom of the subset in the bottom row. In Figure 5.11 the reflectance spectra for the clear and cloud shadow pixel of the original scene, the PACO deshadowed scene and the ATCOR deshadowed scene are given for zoom number 2. In this example of France, ATCOR performs better visually as seen in Figure 5.11 but looking at the reflectance spectras, the new deshadowing algorithm gives better results.

5.5 Validation of Data Set

In order to evaluate the data set, a metric for the comparison of the surface reflectance retrieval without deshadowing and the deshadowed reflectance is calculated. This metric is represented by the relative ratio of the mean reflectance vectors over all spectral bands with deshadowing and without deshadowing. Hence, the mean reflectance of all the clear pixels is divided by the mean reflectance of all cloud shadow pixels. For perfect deshadowing, the value of the relative ratio should lie as close as possible to the perfect value of +1. Depending on the disagreement between clear pixels and cloud shadow pixels, the relative ratio will deviate from +1.

The computation is done with the following three steps and is performed for each scene and for PACO and ATCOR:

- calculation of the relative ratio R1: ratio of the mean reflectance vector of the
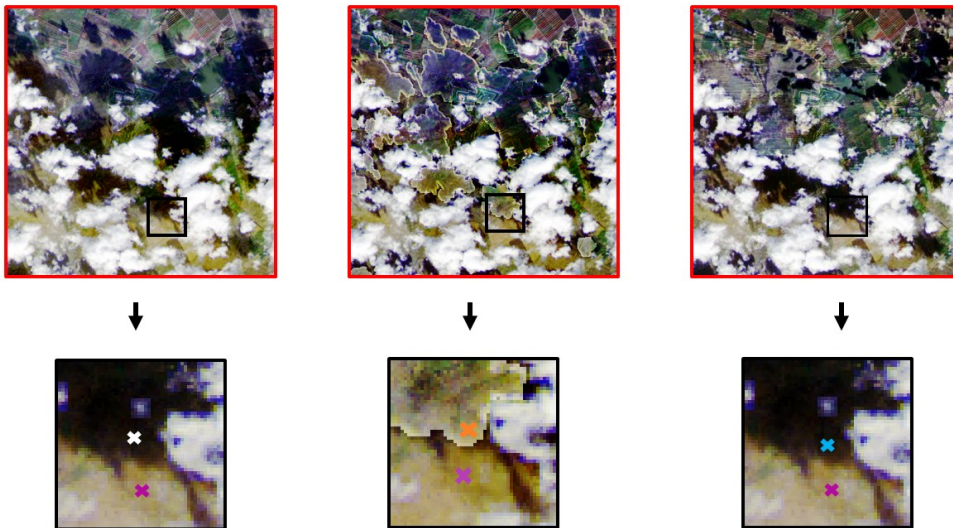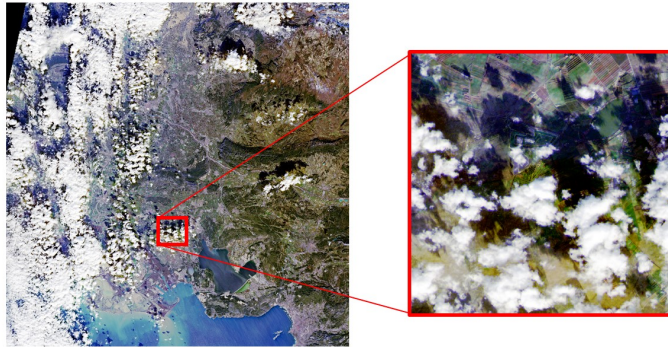
Abbildung 5.8: Top row (left): France RGB=665/560/490 nm with equalization stretching. Top row (right): subset of scene ID 10 with equalization stretching. Middle row (left to right): Subset 1 of France RGB=665/560/490 nm with equalization stretching; PACO deshadowed subset 1 of France with equalization stretching; ATCOR deshadowed subset 1 of France with equalization stretching. Bottom row (left to right): zoom of subset 1 of France RGB=665/560/490 nm with equalization stretching (white cross = cloud shadow pixel; pink cross = clear pixel); zoom of PACO deshadowed subset 1 of France with equalization stretching (orange cross = PACO deshadowed cloud shadow pixel; pink cross = clear pixel); zoom of ATCOR deshadowed subset 1 of France with equalization stretching (blue cross = ATCOR deshadowed cloud shadow pixel; pink cross = clear pixel).
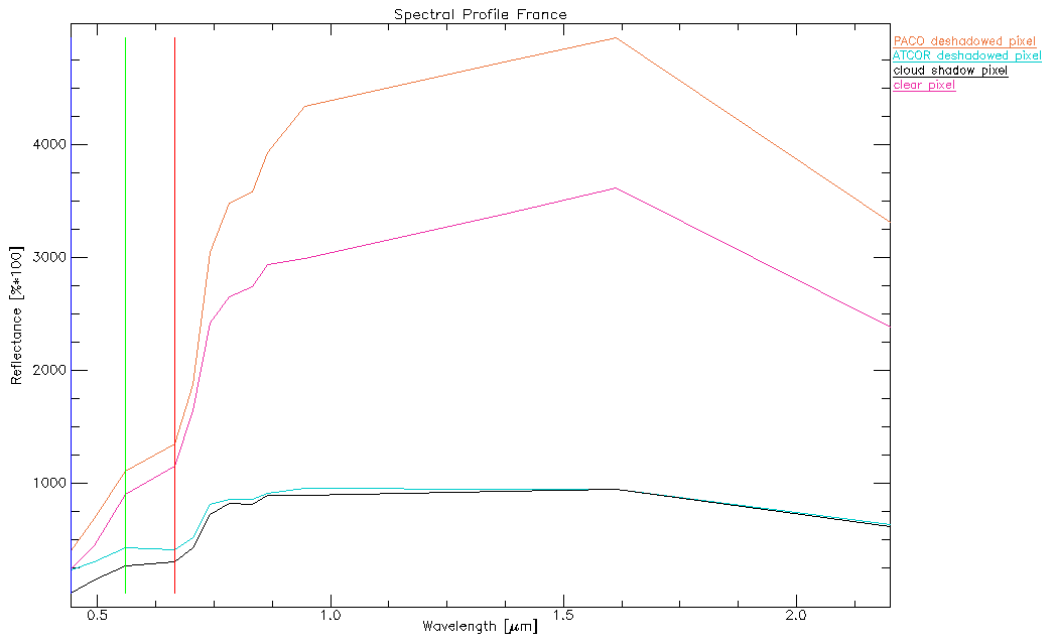
Abbildung 5.9: Spectral profile of scene ID 10 (France). Orange = PACO deshadowed cloud shadow pixel; Blue = ATCOR deshadowed cloud shadow pixel; Black = original cloud shadow pixel spectrum; Pink = clear pixel spectrum.

clear scene pixels with atmospheric correction versus the mean reflectance of the shadow pixels but without deshadowing

- calculation of the relative ratio R2 for PACO and ATCOR: ratio of the mean reflectance vector of the clear scene pixels versus the mean reflectance of the shadow pixels after deshadowing

- comparison of R1 and R2 (see Table 5.2)

For an improvement in the reflectance vector after deshadowing, the relative ratio R2 should be closer to the value +1 than the relative ratio R1. Table 5.2 gives the values of R1 and R2 for each scene. R2 was evaluated for the new method, PACO, and for the previous ATCOR method. The bold face numbers of Table 5.2 indicate the correlation coefficient with a value closest to +1 and hence with the best correlation. No cloud shadows are present in scene ID 5 and 15, hence no values for R1 and R2 are obtained.

As can be deduced from Table 5.2, all of the relative ratios are improved by the deshadowing algorithm for PACO apart from the case of scene ID 8. The best performance of PACO is obtained with scene ID 1 located in Gobabeb (Africa) with a value of R2=0.999. The worst performance of PACO is obtained with scene ID 23 located in Barrax (Spain) with a value of R2=0.596. This is due to the value of R1 for this scene to be the worst outlier and hence represents a hard scene to be deshadowed.
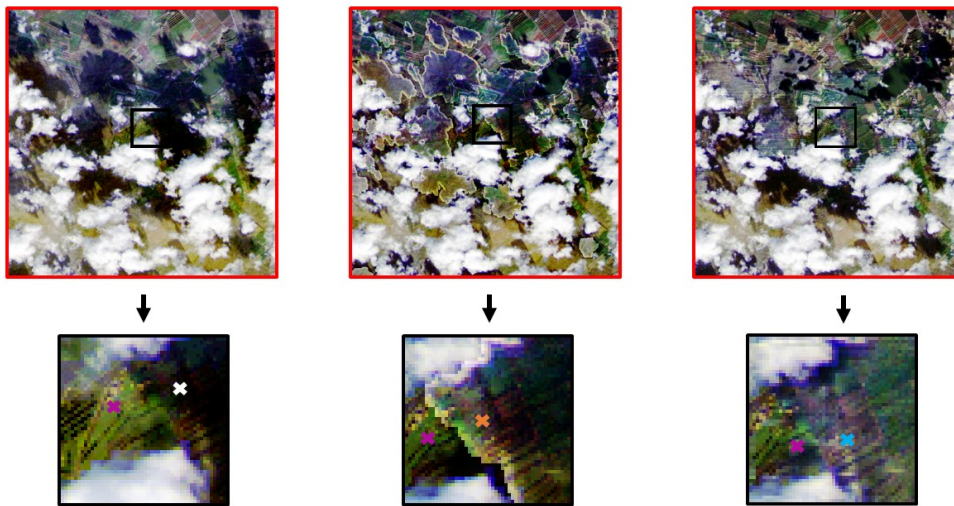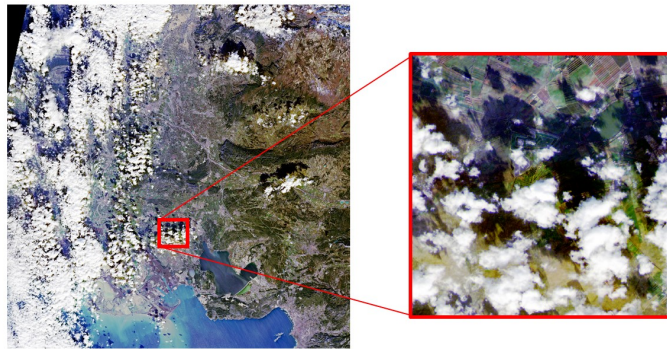
Abbildung 5.10: Top row (left): France RGB=665/560/490 nm with equalization stretching. Top row (right): subset of scene ID 10 with equalization stretching. Middle row (left to right): Subset 2 of France RGB=665/560/490 nm with equalization stretching; PACO deshadowed subset 2 of France with equalization stretching; ATCOR deshadowed subset 2 of France with equalization stretching. Bottom row (left to right): zoom of subset 2 of France RGB=665/560/490 nm with equalization stretching (white cross = cloud shadow pixel; pink cross = clear pixel); zoom of PACO deshadowed subset 2 of France with equalization stretching (orange cross = PACO deshadowed cloud shadow pixel; pink cross = clear pixel); zoom of ATCOR deshadowed subset 2 of France with equalization stretching (blue cross = ATCOR deshadowed cloud shadow pixel; pink cross = clear pixel).

Tabelle 5.2: Sentinel-2 level L1C test scenes. Information on the relative ratio between the clear and shadow pixels with (R2) and without (R1) deshadowing. Bold face numbers indicate the best performance.

| Scene	Location	R1	R2 (PACO)	R2 (ATCOR)
1	Africa, Gobabeb	2.310	**0.999**	0.705
2	Africa, Namibia	1.136	**0.995**	1.175
3	Antarctic	0.469	**1.245**	1.582
4	Argentina, Buenos Aires	1.991	**0.938**	1.219
5	Australia, Lake Lefroy	-	-	-
6	Bolivia, Puerto Siles	0.926	**1.016**	0.979
7	China, Dunhuang	1.105	**1.040**	1.064
8	France, Arcachon-1	**1.161**	0.694	-
9	France, Arcachon-2	0.829	**0.911**	-
10	France-3	3.826	**1.124**	2.495
11	Estonia, Tallin	1.663	**0.614**	-
12	Germany, Berlin	1.350	**0.689**	-
13	Italy, Etna	3.895	**0.586**	2.945
14	Kazakhstan, Balkhash	1.312	**0.875**	1.435
15	Mexico, Cancun	-	-	-
16	Morocco, Quarzazate	1.560	**0.924**	1.530
17	Mosambique, Maputo	1.072	**0.967**	-
18	Netherlands, Amsterdam	2.939	**0.825**	2.131
19	Philipines, Manila	1.699	**0.635**	1.756
20	Russia, Sachalin	1.278	**1.083**	1.229
21	Russia, Yakutsk	3.074	**1.59**	1.608
22	Spain, Barrax-1	2.886	**0.983**	1.621
23	Spain, Barrax-2	12.434	**0.596**	1.651
24	Switzerland, Davos	3.446	**1.175**	1.523
25	USA, Rimrock	2.968	**0.976**	1.632
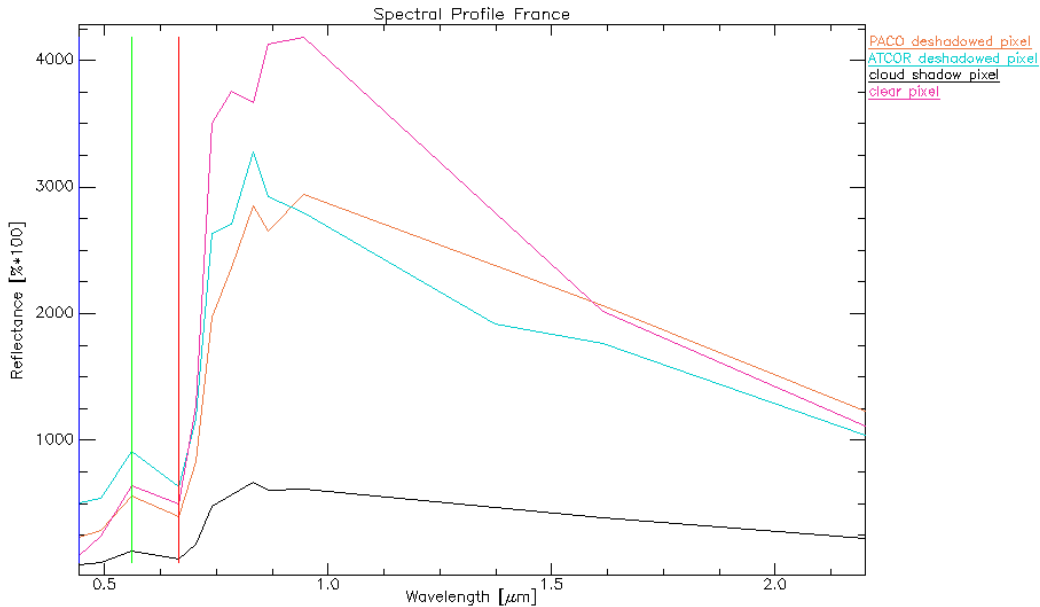
Abbildung 5.11: Spectral profile of scene ID 10 (France). Orange = PACO deshadowed cloud shadow pixel; Blue = ATCOR deshadowed cloud shadow pixel; Black = original cloud shadow pixel spectrum; Pink = clear pixel spectrum.

For the case of scene ID 8 located in Arcachon (France) a relative ratio close to +1 for R1 is obtained, since the scene is covered by a film of haze. Hence, the overall scene appears brighter in the reflectance spectrum, even the shadows. This results in a ratio of R1 close to +1. When the deshadowing is done for a cloud shadow with a bit of haze on top, the corrected deshadowed image does not have a film of haze covering the deshadowed area. Hence the reflectance spectrum of this cloud shadow will appear lower. This has as a consequence, that the value of R2 is a bit less than the value of R1.

In the case of the deshadowing algorithm in ATCOR, no values were obtained for the scenes 8, 9, 11, 12 and 17. For these scenes, the MF deshadowing was turned off during the atmospheric correction due to not enough pixels present or due to a problem in the matrix inversion. To summarize this section, it can be seen, that the PACO deshadowing algorithm with the TIP cloud shadow masking and its additions to the MF method highly improves the deshadowing of the Sentinel-2 data.

5.6 Summary

An improved cloud shadow removal algorithm for high spatial resolution optical satellite data was presented. It is based on the Matched Filter method with the addition of an improved cloud shadow masking and an iterative process for the final reflectance value calculation. Through visual and spectral inspection it was shown, that the new method improves the previous de-shadowing algorithm. This was further-

more presented through an evaluation of the relative ratio between the reflectance of clear and cloud shadow pixel and showed promising values for the new method. Future work will have to include the new cloud shadow detection improvements of the TIP method and the evaluation of the cloud shadow border correction.

6 Discussion and Conclusion

6.1 Masking Algorithms for Sentinel-2 Imagery

The correct classification of satellite imagery is a crucial preprocessing step for atmospheric correction. Three masking algorithms (Fmask, ATCOR and Sen2Cor) were evaluated and compared with one another on a set of 20 Sentinel-2 scenes. These scenes were selected to cover all continents, different climates, seasons and environments in order to make the comparison as variable as possible. The masking outputs of all three algorithms was set to a common set of 8 classes. The common features of the three codes on a spectral point of view are the spectral threshold of the top of atmosphere reflectance, the band ratios, the normalized difference vegetation index and the normalized difference snow index. Furthermore, all have cloud masks and geometric criteria to calculate cloud shadows. The differences of the three masking algorithms are the differently set thresholds and the different criteria to define potential shadows. It is to note as well that the buffering of cloud is different. This was especially noticeable for the Fmask processor version, which has a very present cloud buffer.

High overall accuracies for all 3 masking codes were obtained but each code provides the user with advantages and disadvantages.

ATCOR has a robust classification. An accurate cloud mask over urban, arid and desert area. An accurate water mask where it has a shadow preference when classifying the shadows over water. And finally, it has an adequate shadow mask. The disadvantages are that it has a rather conservative mask of water cloud. So it underestimates the clouds over water. And it has a conservative cloud shadow mask, so underestimates the cloud shadows. Sen2Cor has a robust classification for all scenes. The cloud mask is accurate in many cases where the scene has a medium brightness. On the other hand, the cloud mask fails in urban areas and in arid or desert areas, where it cannot distinguish between bright ground object and cloud. The cloud shadow mask also consistently underestimates the shadow. These criteria will be corrected in the next update of Sen2Cor. Fmask dilates the cloud mask which increases the accuracy of clear pixels. It is adequate in desert and arid regions. Furthermore, Fmask has a high user accuracy for semitransparent cloud, water and topographic shadow. On the other hand, urban structures are usually classified as cloud. Strange circular or elliptical blobs of the cloud mask appear due to the expanded cloud buffer. And lastly it classifies bright water as snow or ice and often misclassifies snow/ ice as cloud.

Since the reference and classified maps are based on the same dataset, i.e., a perfect

match of geometry and acquisition time, the main uncertainty of the reference map classification is the use of a human interpreter [60]. Experiences with similar experiments using several human analysts report an average interpretation variability of ~5-7% [96, 61] for cloud masks. In order to reduce the influence of the interpreter a reference polygon should have homogeneous BOA reflectance properties per class, i.e., heterogeneous areas with mixed pixels are excluded [78]. The area homogeneity can be checked visually per band and it also shows if pixel spectra of a polygon have a large dispersion, e.g. for cloud border regions or snow areas below semi-transparent cloud. Although the variability within a polygon should be small, large differences can exist between different polygons of the same class, e.g. in case of different cloud types or fresh and polluted snow.

Table 3.3 presents the class-specific *user's accuracy* (UA) and *producer's accuracy* (PA) for the three methods averaged over the 20 scenes. High PA values are only achieved for the classes cloud (Fmask) and snow/ice (Sen2Cor) indicating how difficult a classification is for all other classes. The low values for semi-transparent cloud are most likely caused by the interpreter and his visual assessment which does not agree with the selected physical criterion ($0.01 < \rho(\text{TOA}, 1.38 \ \mu m) < 0.04$) of the three methods. Another known classification problem concerns the distinction of water and cloud shadow if no external maps are included. Both classes can be spectrally very similar. Additionally, there can be cloud shadow over water, but since a pixel can only belong to one class in our evaluation, the setting of the preference rule adds another uncertainty.

Nevertheless, a comparison with S2 classification results obtained by the standard Fmask [95] (applied to seven scenes) demonstrates that all three methods yield better overall accuracies than presented in reference [95] (Fig. 6). This is even more remarkable because our approach uses six classes instead of four, and an increase of the number of classes tends to decrease the overall accuracy. Table 3.3 allows a selection of the best method depending on location and cloud cover:

- Fmask can best be applied for scenes in moderate climate, excluding arid and desert regions as well as areas with a large snow/ice coverage.

- ATCOR can best be applied for urban (bright surfaces), arid and desert scenes.

- Sen2Cor can best be applied for rural scenes in moderate climate, and also in scenes with snow and cloud.

Again, a reminder is needed: the Fmask results shown here pertain to the Fmask parallax version [30], not the available standard version [95]. Furthermore, Sen2Cor uses an additional external ESACCI-LC data package, which improves the classi-

fication accuracy over water, urban and bare areas and enables a better handling of false detection of snow pixels [5]. Therefore, Sen2Cor profits of a certain advantage compared to Fmask and ATCOR. During this investigation we also found out that the performance of Fmask (parallax version) can be improved if the current cloud buffer size of 300 m is reduced to 100 m. The performance of Sen2Cor (version 2.8.0) can be slightly improved with an additional cloud buffer of 100 m (instead of no buffer), whereas an additional 100 m cloud buffer is almost of no influence on the ATCOR performance.

To sum up, we can say that the overall accuracy is very high for all 3 masking codes and near together and the balanced OA is equal. ATCOR finds most valid pixels, has the highest PA and lowest UA for valid pixels. Sen2Cor finds less valid pixels due to its class definition of dark area. Fmask finds least valid pixels due to dilation of cloud mask. Opposite has Fmask lowest cloud omission and clear comission.

Therefore it can be concluded that all of the methods have their advantages and disadvantages and it is up to the user to see what the important criteria are that they want to classify.

6.2 Cloud Shadow Detection

Due to the high contamination of satellite data by clouds and their shadows, a new developed algorithm for cloud shadow detection was presented and tested in the PACO masking algorithm. The TIP method only relies on the evaluation of thresholds, indices and projections and can be applicable to more than one sensor due to its advantage of only depending on the spectral channels. Hence, cloud shadows of multispectral satellite images covering the spectrum of 0.44 μm to 2.2 μm are found. The presented TIP method in Chapter 4 was able to improve current results by 6%.

Remotely sensed optical images have a lot more details to be accounted for and need a lot more fine tuning than in computer vision applications. The LAB method alone as done by [55] might be a good method for simple close-range images taken by a RGB camera. However, it is not possible to adapt the LAB color scale to detect shadows in a Sentinel-2 satellite image. Nevertheless, the threshold selections implemented in the TIP method have opened new doors for the shadow detection. The presented TIP method adapts thresholds based on histogram analysis and introduces new spectral criteria to improve the shadow mask. Furthermore it incorporates projection methods for the cloud shadows which has so far only been used for the detection of clouds.

To compare how the cloud shadow detection of the TIP method incorporated in

PACO differentiates from other masking algorithms, the analysis done by [91] has been performed once again to include the PACO masking algorithm next ATCOR. This validation done by human eye has to be seen as a first guidance for a solid comparison between different algorithms. Furthermore, each algorithm has a different focus for the user and producer accuracy. The cloud shadow mask from the TIP method included in PACO is designed to detect as many cloud shadows as realistic in order to perform a subsequent cloud shadow correction. This is also the reason why the PA (75.3%) is much higher than the UA (57.3%) (see Table 4.2). A lower UA has as a consequence, that clear pixels close to the border of the cloud shadows will be deshadowed in shadow removal algorithms following the shadow masking. On the other hand, Sen2Cor, as evaluated by [91], shows the highest UA for the class cloud shadow of 81.0%, since it is set to obtain as few false positive pixels as possible and has therefore a more conservative cloud shadow masking. But this happens at the expenses of the PA percentage which is lower than the PA of PACO.

Another criterion which differs between masking algorithms is when multiple classes lie on top of each other. For example if cloud shadows are over water, then the pixels are identified as water for PACO. This is because the priority in PACO is that water bodies are correctly identified and not excluded if parts of it are covered by shadows. Hence, the TIP method excludes all water pixels before starting the cloud shadow calculation. This has also the consequence, that if a water body is wrongly classified in the steps prior to the TIP masking, then cloud shadows might be wrongly excluded from the final cloud shadow mask. This was the case of scene number 10 from Mexico taken on the 27th of May 2018. Not all cloud shadows have been detected by the TIP method due to wrong cloud and water classification.

Another critical scenario for all cloud shadow masking algorithms is when the sensor incidence angle is very low. The scene of Antarctica is such a case. Here the incidence angle of the sensor is 4.9°, close to nadir. Due to the low incidence angle, the scene appears very dark in regions where no snow and clouds are present. But since the scene has a lot of ice and clouds, these areas appear very bright with respect to areas where no ice or clouds are present. PACO has difficulties to determine the clouds that lie above the ice. Hence, due to the cloud projection step present in the TIP method, the shadows from the clouds above the ice gets removed from the cloud shadow map. A correct classification of clouds is important for the cloud shadow detection within the TIP method. A wrong cloud mask will lead to errors in the cloud shadow detection.

Furthermore, one has to note that the TIP method can be used for VNIR-SWIR sensors such as EnMAP [35], PRISMA [9], Landsat-8 [49] and -9 [51], but for VNIR sensors, such as for example the DESIS sensor [13],the method can only

be implemented partially. In the case of DESIS, the DISN correction explained in section 2.2.3 will have to be excluded.

The newly presented TIP method overcomes the disadvantage of only being applicable to one sensor, as is the case of Fmask4P which is only applicable for Sentinel-2 [29]. Fmask5, on the other hand, can be applied to Landsats 4-8 and Sentinel-2 [95]. The newly presented method only needs the presence of specific spectral channels (VNIR and SWIR) and can be applied to hyperspectral sensors such as PRISMA [9] and EnMAP [35]. Furthermore, the TIP method does not have to undergo complex multistage processes as is currently done in most state-of-the-art methods ([10], [44], [16], [27], [96]). Additionally, the method combines the advantages of multiple spectral indices that can be used for the detection of cloud shadows.

Even though the TIP method has only been tested for Sentinel-2 scenes with a geometric resolution of 20m, it can be assumed to work for Landsat-8 which has a resolution of 30m, since Fmask5 [95] works without any issues for Landsat-8 and Sentinel-2. One has to note, that for Sensors with ground sampling distance of 1-5m additional problems will arise even when all VNIR and SWIR channels are present. Hence, more research will have to be done in the future in order to be able to include such sensors.

6.3 Claud Shadow Removal

This new method to evaluate cloud shadows was furthermore included into the shadow removal algorithm presented in Chapter 5. It was shown that including the TIP method and an iterative process for the final reflectance value calculation to the Matched Filter method improved the results of the cloud shadow removal algorithm.

The improved deshadowing algorithm implemented in the python atmospheric correction (PACO) has shown very promising results and was able to improve the previous Matched Filter version implemented in ATCOR through its additions. Relative ratios close to the value of +1 are reached as shown in Table 5.2 with a range between 0.596 and 1.245 with deshadowing, R2 (PACO), and 0.469 and 12.434 without deshadowing, R1. The high values that are reached without going through the cloud shadow removal algorithm can be explained by the presence of haze covering the scene (scene ID 8) or when the scene has a lot of water bodies which are part of the clear pixels (scene ID 17). The cloud shadow removal algorithm performed within PACO is done without any atmospheric correction on the haze particles. This is one additional step that can be taken into account to further improve the deshadowing algorithm. Hence, to fully correct the scene for the visibility and then perform the cloud shadow removal.

The visual and spectral comparison with the deshadowing results obtained through PACO and ATCOR are able to show the achieved improvements, but also the weaknesses that will have to be changed in the future work. One of the weaknesses of the presented cloud shadow removal algorithm is the correction at the borders of the shadows. The visual results show, that ATCOR is able to better correct for the transition region between shadow pixels and clear pixels. Nevertheless, better results are obtained overall for PACO and this is also proven by the reflectance spectra shown.

The advantages of the deshadowing method presented in this paper are that it is performed through a fully automatic algorithm which is based on the Matched Filter Method, the TIP cloud shadow masking and a iterative process of the scaled shadow function. It works for multispectral and hyperspectral imagery over land acquired by satellite/airborne sensors. Since the deshadowing algorithm relies on the TIP method after [90] it must be noted that it is applicable for VNIR-SWIR sensors such as EnMAP [35], PRISMA [9], Landsat-8 [49] and Landsat-9 [51], but for VNIR sensors, such as for example the DESIS sensor [13],the method can only be implemented partially.

So far the new deshadowing algorithm has been tested for Sentinel-2 scenes having a geometric resolution of 20m. Nevertheless, it can be assumed, that the method will perform similar for Landsat-8, having a resolution of 30m. For Sensors with a ground sampling distance less than 5m additional problems will arise, due to the TIP method still having to perform test for these cases.

6.4 Future Work

Even though the proposed and described methods presented show promising results, further research could add additional improvements, for instance:

- Texture measures and convolutional neural networks

- Fine tuning of the cloud and water masks in order to improve the results of the TIP method

- Further evaluation and smoothing of the pixels around cloud shadow borders to get even more accurate results of the cloud shadow map

- Testing all results with more sensors such as for example Landsat-8 [49] and EnMAP [35]

A   Appendix

Zekoll Viktoria, Main-Knorn Magdalena, Louis Jerome, Frantz David, Richter Rudolf, Pflug Bringfried. (2021). Comparison of Masking Algorithms for Sentinel-2 Imagery. Remote Sensing. 13. 137. 10.3390/rs13010137.

B   Appendix

Zekoll Viktoria, Reyes Raquel, Richter, Rudolf. (2022). A Newly Developed Algorithm for Cloud Shadow Detection—TIP Method. Remote Sensing. 14. 2922. 10.3390/rs14122922.

C   Appendix

Submitted: Viktoria Zekoll. Cloud shadow removal for high spatial resolution optical satellite data

Bibliography

[1] Earth.esa.int. 2020. acix ii - cmix 2nd ws. URL: `https://earth.esa.int/web/sppa/meetings-workshops/hosted-and-co-sponsored-meetings/acix-ii-cmix-2nd-ws`.

[2] Level-1c. URL: `https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/level-1c/algorithm`.

[3] Planet dem. URL: `https://www.planetobserver.com/products/planetdem/planetdem-30/`.

[4] Preprocessing of eo data radiometry. URL: `https://eo4geocourses.github.io/GEOF_Preprocessing-of-EO-data/#/1`.

[5] S2 mpc - sen2cor configuration and user manual. URL: `http://step.esa.int/thirdparties/sen2cor/2.8.0/docs/S2-PDGS-MPC-L2A-SUM-V2.8.pdf2`, `doi:Ref.S2-PDGS-MPC-L2A-SUM-V2.8.Issue2.2019-02-05`.

[6] Copernicus land monitoring service. 2020. URL: `https://land.copernicus.eu/`.

[7] Ndwi: Normalized difference water index in agriculture, Oct 2021. URL: `https://eos.com/make-an-analysis/ndwi/`.

[8] Vegetation indices as a satellite-based add-on for agri solutions, Mar 2022. URL: `https://eos.com/blog/vegetation-indices/`.

[9] Nicola Acito, Marco Diani, Gregorio Procissi, and Giovanni Corsini. Atmospheric compensation of prisma data by means of a learning based approach. *Remote Sensing*, 13(15), 2021. URL: `https://www.mdpi.com/2072-4292/13/15/2967`, `doi:10.3390/rs13152967`.

[10] Steven A. Ackerman, Robert E. Holz, Richard A. Frey, Edwin W. Eloranta, Brent Maddux, and Matthew J. McGill. Cloud detection with modis. part ii: Validation. *Journal of Atmospheric and Oceanic Technology*, 25:1073–1086, 2008.

[11] Steven A Ackerman, Kathleen I Strabala, W Paul Menzel, Richard A Frey, Christopher C Moeller, and Liam E Gumley. Discriminating clear sky from clouds with modis. *Journal of Geophysical Research: Atmospheres*, 103(D24):32141–32157, 1998.

[12] Steven Adler-Golden, Michael Matthew, Gail Anderson, Gerald Felde, and James Gardner. An algorithm for de-shadowing spectral imagery. *Procee-*

dings of SPIE - The International Society for Optical Engineering, 4816:10, 07 2002. `doi:10.1117/12.451691`.

[13] Kevin Alonso, Martin Bachmann, Kara Burch, Emiliano Carmona, Daniele Cerra, Raquel de los Reyes, Daniele Dietrich, Uta Heiden, Andreas Hölderlin, Jack Ickes, Uwe Knodt, David Krutz, Heath Lester, Rupert Müller, Mary Pagnutti, Peter Reinartz, Rudolf Richter, Robert Ryan, Ilse Sebastian, and Mirco Tegler. Data products, quality and validation of the dlr earth sensing imaging spectrometer (desis). *Sensors*, 19(20), 2019. URL: `https://www.mdpi.com/1424-8220/19/20/4471`, `doi:10.3390/s19204471`.

[14] Louis Baetens, Camille Desjardins, and Olivier Hagolle. Validation of copernicus sentinel-2 cloud masks obtained from maja, sen2cor, and fmask processors using reference cloud masks generated with a supervised active learning procedure. *Remote. Sens.*, 11:433, 2019.

[15] Andrea Baraldi and Dirk Tiede. Autocloud+, a "universal" physical and statistical model-based 2d spatial topology-preserving software for cloud/cloud–shadow detection in multi-sensor single-date earth observation multi-spectral imagery—part 1: Systematic esa eo level 2 product generation at the ground segment as broad context. *ISPRS International Journal of Geo-Information*, 7(12):457, 2018.

[16] Justin D. Braaten, Warren B. Cohen, and Zhiqiang Yang. Automated cloud and cloud shadow identification in landsat mss imagery for temperate ecosystems. *Remote Sensing of Environment*, 169:128–138, 2015. URL: `https://www.sciencedirect.com/science/article/pii/S0034425715300948`, `doi:https://doi.org/10.1016/j.rse.2015.08.006`.

[17] Bo cai Gao. Ndwi—a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3):257–266, 1996. URL: `https://www.sciencedirect.com/science/article/pii/S0034425796000673`, `doi:https://doi.org/10.1016/S0034-4257(96)00067-3`.

[18] Jan G.P.W. Clevers and Anatoly A. Gitelson. Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on sentinel-2 and -3. *Int. J. Appl. Earth Obs. Geoinformation*, 23:344–351, 2013.

[19] Russell G. Congalton. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37(1):35–46, 1991. URL: `https://www.sciencedirect.com/science/article/`

pii/003442579190048B, doi:https://doi.org/10.1016/0034-4257(91)90048-B.

[20] Michele Dalponte, Sebastian Marzini, Yady Solano Correa, Giustino Tonon, Loris Vescovo, and Damiano Gianelle. Mapping forest windthrows using high spatial resolution multispectral satellite images. *International Journal of Applied Earth Observation and Geoinformation*, 93:102206, 12 2020. doi:10.1016/j.jag.2020.102206.

[21] Pierre Defourny, Sophie Bontemps, Nicolas Bellemans, Cosmin Cara, Gérard Dedieu, Eric Guzzonato, Olivier Hagolle, Jordi Inglada, Laurentiu Nicola, Thierry Rabaute, Mickael Savinaud, Cosmin Udroiu, Silvia Valero, Agnès Bégué, Jean-Francois Dejoux, Abderrazak El Harti, Jamal Ezzahar, Nataliia Kussul, Kamal Labbassi, Valentine Lebourgeois, Zhang Miao, Terrence Newby, Adolph Nyamugama, Norakhan Salh, Andrii Shelestov, Vincent Simonneaux, Pierre Sibiry Traoré, Souleymane Sidi Traore, and Benjamin Koetz. Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the sen2-agri automated system in various cropping systems around the world. *Remote Sensing of Environment*, 2019.

[22] Felipe Díaz, H Montero, Daniel Santana, Gustavo Montero, Eduardo Rodríguez, L Mazorra Aguiar, and Albert Oliver. Improving shadows detection for solar radiation numerical models. *Applied Mathematics and Computation*, 319:71–85, 2018.

[23] Matthias Drusch, Umberto Del Bello, S. Carlier, Oliver Colin, Valérie Fernandez, Ferran Gascon, Bianca Hoersch, Claudia Isola, Paolo Laberinti, Philippe Martimort, Aimé Meygret, François Spoto, Omar Sy, Franco Marchese, and Pier L. Bargellini. Sentinel-2: Esa's optical high-resolution mission for gmes operational services. *Remote Sensing of Environment*, 120:25–36, 2012.

[24] Yun Du, Yihang Zhang, Feng Ling, Qunming Wang, Wenbo Li, and Xiao Li. Water bodies' mapping from sentinel-2 imagery with modified normalized difference water index at 10-m spatial resolution produced by sharpening the swir band. *Remote Sensing*, 354:1–19, 04 2016. doi:10.3390/rs8040354.

[25] Brian Dunbar. Nasa's icesat satellite sees changing world affecting many. URL: https://www.nasa.gov/vision/earth/lookingatearth/icesat_light.html.

[26] Robert Eckardt, Christian Berger, Christian Thiel, and Christiane Schmullius. Removal of optically thick clouds from multi-spectral satellite images

using multi-frequency sar data. *Remote Sensing*, 5(6):2973–3006, 2013. URL: https://www.mdpi.com/2072-4292/5/6/2973, doi:10.3390/rs5062973.

[27] Adrian Fisher. Cloud and cloud-shadow detection in spot5 hrg imagery with automated morphological feature extraction. *Remote Sensing*, 6(1):776–800, 2014. URL: https://www.mdpi.com/2072-4292/6/1/776, doi:10.3390/rs6010776.

[28] Giles M. Foody. Sample size determination for image classification accuracy assessment and comparison. *International Journal of Remote Sensing*, 30(20):5273–5291, 2009. arXiv:https://doi.org/10.1080/01431160903130937, doi:10.1080/01431160903130937.

[29] David Frantz. Force—landsat + sentinel-2 analysis ready data and beyond. *Remote Sensing*, 11(9), 2019. URL: https://www.mdpi.com/2072-4292/11/9/1124, doi:10.3390/rs11091124.

[30] David Frantz, Erik Haß, Andreas Uhl, Johannes Stoffels, and Joachim Hill. Improvement of the fmask algorithm for sentinel-2 images: Separating clouds from bright surfaces based on parallax effects. *Remote Sensing of Environment*, 2018.

[31] David Frantz, Achim Roder, Thomas Udelhoven, and Michael Schmidt. Enhancing the detectability of clouds and their shadows in multitemporal dryland landsat imagery: Extending fmask. *IEEE Geoscience and Remote Sensing Letters*, 12:1242–1246, 2015.

[32] A. Gafurov and A. Bárdossy. Cloud removal methodology from modis snow cover product. *Hydrology and Earth System Sciences*, 13(7):1361–1373, 2009. URL: https://hess.copernicus.org/articles/13/1361/2009/, doi:10.5194/hess-13-1361-2009.

[33] Mireia Gascon, Wilma Zijlema, Cristina Vert, Mathew P. White, and Mark J. Nieuwenhuijsen. Outdoor blue spaces, human health and well-being: A systematic review of quantitative studies. *International Journal of Hygiene and Environmental Health*, 220(8):1207–1221, nov 2017. URL: https://doi.org/10.1016{%}2Fj.ijheh.2017.08.004, doi:10.1016/j.ijheh.2017.08.004.

[34] Nicholas Goodwin, Lisa Collett, Robert Denham, Neil Flood, and Dan Tindall. Cloud and cloud shadow screening across queensland, australia: An automated method for landsat tm/etm+ time series. *Remote Sensing of Environment*, 134:50–65, 2013.

[35] Luis Guanter, Hermann Kaufmann, Karl Segl, Saskia Foerster, Christian Rogass, Sabine Chabrillat, Theres Kuester, André Hollstein, Godela Rossner, Christian Chlebek, et al. The enmap spaceborne imaging spectroscopy mission for earth observation. *Remote Sensing*, 7(7):8830–8857, 2015.

[36] Olivier Hagolle, Mireille Huc, Camille Desjardins, Stefan Auer, and Rudolf Richter. Maja algorithm theoretical basis document, 2017. `doi:10.5281/zenodo.1209632`.

[37] Olivier Hagolle, Mireille Huc, David Villa Pascual, and Gérard Dedieu. A multi-temporal method for cloud detection, applied to formosat-2, venμs, landsat and sentinel-2 images. *Remote Sensing of Environment*, 114:1747–1755, 2010.

[38] André Hollstein, Karl Segl, Luis Guanter, Maximilian Brell, and Marta Enesco. Ready-to-use methods for the detection of clouds, cirrus, snow, shadow, water and clear sky pixels in sentinel-2 msi images. *Remote Sensing*, 8(8):666, 2016.

[39] Jie Hu, Jie Peng, Yin Zhou, Dongyun Xu, Zhao Ruiying, Qingsong Jiang, Fu Tingting, Fei Wang, and Zhou Shi. Quantitative estimation of soil salinity using uav-borne hyperspectral and satellite multispectral images. *Remote Sensing*, 11, 04 2019. `doi:10.3390/rs11070736`.

[40] Elsy Ibrahim, Jingyi Jiang, Luisa Lema, Pierre Barnabé, Gregory Giuliani, Pierre Lacroix, and Eric Pirard. Cloud and cloud-shadow detection for applications in mapping small-scale mining in colombia using sentinel-2 imagery. *Remote Sensing*, 13(4), 2021. URL: `https://www.mdpi.com/2072-4292/13/4/736`, `doi:10.3390/rs13040736`.

[41] Markus Immitzer, Francesco Vuolo, and Clement Atzberger. First experience with sentinel-2 data for crop and tree species classifications in central europe. *Remote. Sens.*, 8:166, 2016.

[42] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–. URL: `http://www.scipy.org/`.

[43] Jun Li, Chian-Yi Liu, Hung-Lung Huang, T.J. Schmit, Xuebao Wu, W.P. Menzel, and J.J. Gurka. Optimal cloud-clearing for airs radiances using modis. *IEEE Transactions on Geoscience and Remote Sensing*, 43(6):1266–1278, 2005. `doi:10.1109/TGRS.2005.847795`.

[44] Zhiwei Li, Huanfeng Shen, Huifang Li, Guisong Xia, Paolo Gamba, and Liangpei Zhang. Multi-feature combined cloud and cloud

shadow detection in gaofen-1 wide field of view imagery. *Remote Sensing of Environment*, 191:342–358, 2017. URL: https://www.sciencedirect.com/science/article/pii/S003442571730038X, doi:https://doi.org/10.1016/j.rse.2017.01.026.

[45] J. Louis. Sentinel 2 msi - level 2a product definition. URL: https://sentinel.esa.int/documents/247904/1848117/Sentinel-2-Level-2A-Product-Definition-Document.pdf, doi:Issue4.4.2016-08-12.

[46] Jérôme Louis, Vincent Debaecker, Bringfried Pflug, Magdalena Main-Knorn, Jakub Bieniarz, Uwe Mueller-Wilm, Enrico Cadau, and Ferran Gascon. Sentinel-2 sen2cor: L2a processor for users. In *Proceedings living planet symposium 2016*, pages 1–8. Spacebooks Online, 2016.

[47] Haitao Lv, Yong Wang, and Yang Shen. An empirical and radiative transfer model based algorithm to remove thin clouds in visible bands. *Remote Sensing of Environment*, 179:183–195, 2016. URL: https://www.sciencedirect.com/science/article/pii/S0034425716301286, doi:https://doi.org/10.1016/j.rse.2016.03.034.

[48] Shengfang Ma, Yuting Zhou, Prasanna H. Gowda, Jinwei Dong, Geli Zhang, Vijaya G. Kakani, Pradeep Wagle, Liangfu Chen, K. Colton Flynn, and Weiguo Jiang. Application of the water-related spectral reflectance indices: A review. *Ecological Indicators*, 98:68–79, 2019. URL: https://www.sciencedirect.com/science/article/pii/S1470160X18308215, doi:https://doi.org/10.1016/j.ecolind.2018.10.049.

[49] Brian Markham, Julia Barsi, Geir Kvaran, Lawrence Ong, Edward Kaita, Stuart Biggar, Jeffrey Czapla-Myers, Nischal Mishra, and Dennis Helder. Landsat-8 operational land imager radiometric calibration and stability. *Remote Sensing*, 6(12):12275–12308, 2014. URL: https://www.mdpi.com/2072-4292/6/12/12275, doi:10.3390/rs61212275.

[50] Michael Marshall, Mariana Belgiu, Mirco Boschetti, Monica Pepe, Alfred Stein, and Andy Nelson. Field-level crop yield estimation with prisma and sentinel-2. *ISPRS Journal of Photogrammetry and Remote Sensing*, 187:191–210, 2022. URL: https://www.sciencedirect.com/science/article/pii/S0924271622000739, doi:https://doi.org/10.1016/j.isprsjprs.2022.03.008.

[51] Jeffrey Masek, Michael Wulder, Brian Markham, Joel McCorkel, Christopher Crawford, James Storey, and Del Jenstrom. Landsat 9: Empowering open science and applications through continuity. *Remote Sensing of Environment*, 248:111968, 10 2020. `doi:10.1016/j.rse.2020.111968`.

[52] Gonzalo Mateo-García, Luis Gómez-Chova, and Gustau Camps-Valls. Convolutional neural networks for multispectral image cloud masking. In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 2255–2258, 2017. `doi:10.1109/IGARSS.2017.8127438`.

[53] E Menaka, S Suresh Kumar, and M Bharathi. Cloud removal using efficient cloud detection and removal algorithm for high–resolution satellite imagery. *International Journal of Computer Applications in Technology*, 51(1):54–61, 2015.

[54] Uwe Müller-Wilm, Jérôme M. B. Louis, Rudolf Richter, Ferran Gascon, and Marc Niézette. Sentinel-2 level 2a prototype processor: Architecture, algorithms and first results. 2013.

[55] Saritha Murali and V. K. Govindan. Shadow detection and removal from a single image using lab color space. *Cybernetics and Information Technologies*, 13(1):95–103, 2013. URL: `https://doi.org/10.2478/cait-2013-0009`, `doi:doi:10.2478/cait-2013-0009`.

[56] Tatiana Nazarova, Pascal Martin, and Gregory Giuliani. Monitoring vegetation change in the presence of high cloud cover with sentinel-2 in a lowland tropical forest region in brazil. *Remote Sensing*, 12(11):1829, 2020.

[57] F. E. Nicodemus. Geometrical considerations and ... - stanford university. URL: `http://www.graphics.stanford.edu/courses/cs448-05-winter/papers/nicodemus-brdf-nist.pdf`.

[58] American Society of Photogrammetry. Fall Technical Meeting and American Society of Photogrammetry. *Proceedings of the American Society of Photogrammetry Fall Technical Meeting*. The Society, 1975. URL: `https://books.google.de/books?id=KBY5AAAAIAAJ`.

[59] Pontus Olofsson, Giles M. Foody, Martin Herold, Stephen V. Stehman, Curtis E. Woodcock, and Michael A. Wulder. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, 148:42–57, 2014.

[60] Pontus Olofsson, Giles M. Foody, Martin Herold, Stephen V. Stehman, Curtis E. Woodcock, and Michael A. Wulder. Good practices for estimating

area and assessing accuracy of land change. *Remote Sensing of Environment*, 148:42–57, 2014.

[61] Lazaros Oreopoulos, Michael J. Wilson, and Tamás Várnai. Implementation on landsat data of a simple cloud-mask algorithm developed for modis land bands. *IEEE Geoscience and Remote Sensing Letters*, 8:597–601, 2011.

[62] Thomas Painter and Jeff Dozier. Measurements of the hemispherical-directional reflectance of snow at fine spectral and angular resolution. *Journal of Geophysical Research*, 109, 01 2004.

[63] James M. Palmer. Radiometry and photometry faq - wp.optics.arizona.edu. URL: https://wp.optics.arizona.edu/jpalmer/wp-content/uploads/sites/65/2017/03/rpfaq.pdf.

[64] Chaitanya B. Pande, Sunil A. Kadam, Rajesh Jayaraman, Sunil Gorantiwar, and Mukund Shinde. Prediction of soil chemical properties using multispectral satellite images and wavelet transforms methods. *Journal of the Saudi Society of Agricultural Sciences*, 21(1):21–28, 2022. URL: https://www.sciencedirect.com/science/article/pii/S1658077X21000874, doi:https://doi.org/10.1016/j.jssas.2021.06.016.

[65] Keshav Prasad Paudel and Peter Andersen. Monitoring snow cover variability in an agropastoral area in the trans himalayan region of nepal using modis data with improved cloud removal methodology. *Remote Sensing of Environment*, 115(5):1234–1246, 2011. URL: https://www.sciencedirect.com/science/article/pii/S0034425711000198, doi:https://doi.org/10.1016/j.rse.2011.01.006.

[66] Frank Paul, Solveig H. Winsvold, Andreas Kääb, Thomas Nagler, and Gabriele Schwaizer. Glacier remote sensing using sentinel-2. part ii: Mapping glacier extents and surface facies, and comparison to landsat 8. *Remote. Sens.*, 8:575, 2016.

[67] B. Srinivasa Reddy and Biswanath N. Chatterji. An fft-based technique for translation, rotation, and scale-invariant image registration. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, 5 8:1266–71, 1996.

[68] D. Richter, R.; Schläpfer. Atcor theoretical background document. *DLR Report DLR-IB 564-03/2019*, 2019. URL: https://www.rese-apps.com/software/atcor/manual-papers.html.

[69] R. Richter. A fast atmospheric correction algorithm applied to landsat tm images. *Int. J. Remote Sensing*, 11:159–166, 1990. LIDO-Berichtsjahr=1990,. URL: https://elib.dlr.de/34458/.

[70] R. Richter. Correction of satellite imagery over mountainous terrain. *Applied Optics*, 37:4004–4015, 1998. LIDO-Berichtsjahr=2000,. URL: https://elib.dlr.de/8870/.

[71] R. Richter and A. Müller. De-shadowing of satellite/airborne imagery. *International Journal of Remote Sensing*, 26(15):3137–3148, 2005. arXiv:https://doi.org/10.1080/01431160500114664, doi:10.1080/01431160500114664.

[72] R Richter and D Schläpfer. Atmospheric and topographic correction (atcor theoretical background document). URL: https://www.rese-apps.com/pdf/atcor_ATBD.pdf.

[73] Marc Rußwurm and Marco Körner. Multi-temporal land cover classification with long short-term memory neural networks. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 551–558, 2017.

[74] Alber Hamersson Sanchez, Michelle Cristina A Picoli, Gilberto Camara, Pedro R Andrade, Michel Eustaquio D Chaves, Sarah Lechler, Anderson R Soares, Rennan FB Marujo, Rolf Ezequiel O Simões, Karine R Ferreira, et al. Comparison of cloud cover detection algorithms on sentinel–2 images of the amazon tropical forest. *Remote Sensing*, 12(8):1284, 2020.

[75] G. Schaepman-Strub, M.E. Schaepman, T.H. Painter, S. Dangel, and J.V. Martonchik. Reflectance quantities in optical remote sensing—definitions and case studies. *Remote Sensing of Environment*, 103(1):27–42, 2006. URL: https://www.sciencedirect.com/science/article/pii/S0034425706001167, doi:https://doi.org/10.1016/j.rse.2006.03.002.

[76] Michal Segal-Rozenhaimer, Alan Li, Kamalika Das, and Ved Chirayath. Cloud detection algorithm for multi-modal satellite imagery using convolutional neural-networks (cnn). *Remote Sensing of Environment*, 237:111446, 2020. URL: https://www.sciencedirect.com/science/article/pii/S0034425719304651, doi:https://doi.org/10.1016/j.rse.2019.111446.

[77] Stephen V. Stehman. Sampling designs for accuracy assessment of land cover. *International Journal of Remote Sensing*, 30(20):5243–5272, 2009.

arXiv:https://doi.org/10.1080/01431160903131000, doi:10.1080/01431160903131000.

[78] Stephen V. Stehman. Sampling designs for accuracy assessment of land cover. *International Journal of Remote Sensing*, 30(20):5243–5272, 2009. arXiv:https://doi.org/10.1080/01431160903131000, doi:10.1080/01431160903131000.

[79] Matt Story and Russell G. Congalton. Accuracy assessment: a user's perspective. *Photogrammetric Engineering and Remote Sensing*, 52:397–399, 1986.

[80] A H Strahler, L Boschetti, G M Foody, M A Friedl, M C Hansen, and M Herold. *Global land cover validation: Recommendations for evaluation and accuracy assessment of global land cover maps*. 2006.

[81] Din-Chang Tseng, Hsiao-Ting Tseng, and Chun-Liang Chien. Automatic cloud removal from multi-temporal spot images. *Applied Mathematics and Computation*, 205(2):584–600, 2008. Special Issue on Advanced Intelligent Computing Theory and Methodology in Applied Mathematics and Computation. URL: https://www.sciencedirect.com/science/article/pii/S0096300308003378, doi:https://doi.org/10.1016/j.amc.2008.05.050.

[82] F.D. van der Meer, H.M.A. van der Werff, and F.J.A. van Ruitenbeek. Potential of esa's sentinel-2 for geological applications. *Remote Sensing of Environment*, 148:124–133, 2014. URL: https://www.sciencedirect.com/science/article/pii/S0034425714001084, doi:https://doi.org/10.1016/j.rse.2014.03.022.

[83] John E. Wagner and Stephen V. Stehman. Optimizing sample size allocation to strata for estimating area and map accuracy. *Remote Sensing of Environment*, 168:126–133, 2015. URL: https://www.sciencedirect.com/science/article/pii/S0034425715300572, doi:https://doi.org/10.1016/j.rse.2015.06.027.

[84] Tianxing Wang, Jiancheng Shi, Letu Husi, Tianjie Zhao, Dabin Ji, Chuan Xiong, and Bo Gao. Effect of solar-cloud-satellite geometry on land surface shortwave radiation derived from remotely sensed data. *Remote Sensing*, 9(7), 2017. URL: https://www.mdpi.com/2072-4292/9/7/690, doi:10.3390/rs9070690.

[85] Por Alexandre Wiefels and Carlos Baroja. Red edge detects vegetative stress earlier in plant growth cycle,

Fey 2022. URL: https://mundogeo.com/en/2022/02/09/red-edge-detects-vegetative-stress-earlier-in-plant-growth-cycle/.

[86] Li Xu, Feihu Qi, and Renjie Jiang. Shadow removal from a single image. In *Sixth International Conference on Intelligent Systems Design and Applications*, volume 2, pages 1049–1054, 2006. doi:10.1109/ISDA.2006.253756.

[87] Lin Yan, David P. Roy, Hankui K. Zhang, Jian Li, and Haiyan Huang. An automated approach for sub-pixel registration of landsat-8 operational land imager (oli) and sentinel-2 multi spectral instrument (msi) imagery. *Remote. Sens.*, 8:520, 2016.

[88] Zhiyuan Yan, Menglong Yan, Hao Sun, Kun Fu, Jun Hong, Jun Sun, Yi Zhang, and Xian Sun. Cloud and cloud shadow detection using multilevel feature fused segmentation network. *IEEE Geoscience and Remote Sensing Letters*, 15(10):1600–1604, 2018. doi:10.1109/LGRS.2018.2846802.

[89] Viktoria Zekoll. Cloud shadow removal for high spatial resolution optical satellite data.

[90] Viktoria Zekoll, Raquel de los Reyes, and Rudolf Richter. A newly developed algorithm for cloud shadow detectionmdash;tip method. *Remote Sensing*, 14(12), 2022. URL: https://www.mdpi.com/2072-4292/14/12/2922, doi:10.3390/rs14122922.

[91] Viktoria Zekoll, Magdalena Main-Knorn, Kevin Alonso, Jerome Louis, David Frantz, Rudolf Richter, and Bringfried Pflug. Comparison of masking algorithms for sentinel-2 imagery. *Remote Sensing*, 13(1), 2021. URL: https://www.mdpi.com/2072-4292/13/1/137, doi:10.3390/rs13010137.

[92] Han Zhai, Hongyan Zhang, Liangpei Zhang, and Pingxiang Li. Cloud/shadow detection based on spectral indices for multi/hyperspectral optical remote sensing imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 144:235–253, 2018. URL: https://www.sciencedirect.com/science/article/pii/S0924271618301989, doi:https://doi.org/10.1016/j.isprsjprs.2018.07.006.

[93] Xueying Zhou and L. Sun. A new cloud shadow detection algorithm based on prior land type database support. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B3-2020:849–851, 08 2020. doi:10.5194/isprs-archives-XLIII-B3-2020-849-2020.

[94] Jiejie Zhu, Kegan G. G. Samuel, Syed Z. Masood, and Marshall F. Tappen. Learning to recognize shadows in monochromatic natural images. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 223–230, 2010. doi:10.1109/CVPR.2010.5540209.

[95] Zhe Zhu, Shixiong Wang, and Curtis E. Woodcock. 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, 2015. URL: https://www.sciencedirect.com/science/article/pii/S0034425714005069, doi:https://doi.org/10.1016/j.rse.2014.12.014.

[96] Zhe Zhu and Curtis E. Woodcock. Object-based cloud and cloud shadow detection in landsat imagery. *Remote Sensing of Environment*, 118:83–94, 2012. URL: https://www.sciencedirect.com/science/article/pii/S0034425711003853, doi:https://doi.org/10.1016/j.rse.2011.10.028.

[97] Zhe Zhu and Curtis E. Woodcock. 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, 2014.