Sentinel-3 OLCI
IdePix
Identification of Pixel features

Algorithm Theoretical Basis Document

Authors:
Jan Wevers, Dagmar Müller, Grit Kirches,
Ralf Quast, Carsten Brockmann

Version: 1.0

Last update:
04.05.2022

Table of Contents

List of figures

List of tables

Version history

Version	Date	Comment
1.0	04.05.2022	Initial version

List of Abbreviations

CRS	Coordinate Reference System
CTH	Cloud Top Height
DEM	Digital Elevation Model
GPT	Graph Processing Tool
HLS	High Latitude Snow/Ice
L1Cb	Level 1b
LLS	Low Latitude Snow/Ice
MERIS	MEdium Resolution Imaging Spectrometer
MSI	Multi-Spectral Imager
NDSI	Normalized Difference Snow Index
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NIR	Near-infrared
NN	Neural Network
OLCI	Ocean and Land Colour Instrument

Proba-V	Project for On-Board Autonomy - Vegetation
S3	Sentinel-3
SAA	Sun Azimuth Angle
SZA	Sun Zenith Angle
SNAP	SeNtinel Application Platform
SRTM	Shuttle Radar Topography Mission
TOA	Top Of Atmosphere
VIIRS	Visible/Infrared Imager Radiometer Suite

1 Introduction

Approximately 66% of the global surface is covered by clouds [Wilson 2016]. Figure *1* illustrates the average cloudiness, expressed as fraction of clouds between 1981 and 2014. While some areas are very rarely covered by clouds some areas have cloud fraction above 80%, giving even satellites with high repetition times only a slight chance to measure a clear surface spectrum.
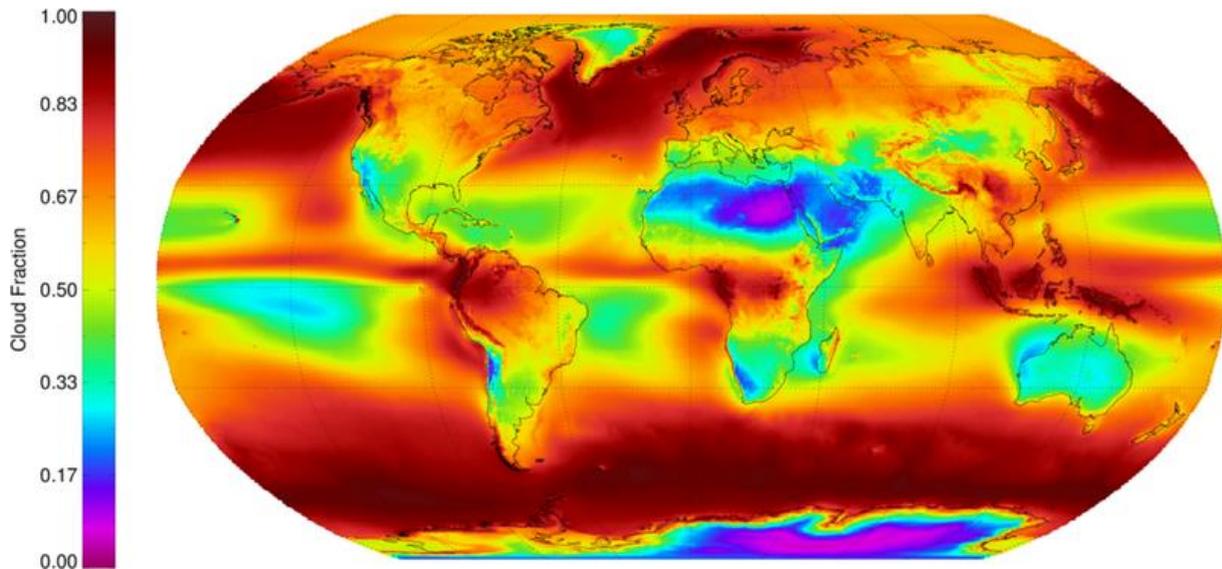
Figure 1: Fraction of clouds over the world, from ESA CCI data (1981-2014 average), CCI Cloud Project[1]

Consequently, most earth observation images in the visible spectral domain include a significant number of cloudy pixels. Such measurements are treated in two opposite ways: either cloud properties are retrieved, e.g., for weather forecast or climate studies (Wylie, D., 1998, Rossow et al, 1999; Liou, 1992), or the focus of the interest is the earth surface – being it land or water – which is then masked by the cloud (Luo,2008). In both cases the presence of the cloud needs to be identified, to ensure correct subsequent processing results.

---

[1] https://climate.esa.int/de/projekte/cmug/case-studies/role-cci-datasets-cloud-modelling/

2   Identification of Pixel properties – the design/overview

The term "pixel identification" refers to a classification of a measurement made by a space borne radiometer. Before processing land or water pixels, or exempting them from further processing, it is necessary to sort them into observation types like clouds, clear sky over land, or clear sky over water. Depending on the downstream processing steps some additional pixel properties like cloud shadow or mountain shadow, snow cover or aerosol contamination are needed.

IdePix (Identification of Pixel properties) is a multi-sensor pixel identification tool available as a SNAP (Sentinel Application Platform) plugin. It provides pixel identification algorithms for a wide variety of sensors like Sentinel-2 MSI, Sentinel-3 OLCI, MERIS, Landsat-8, MODIS, VIIRS, Proba-V or SPOT VGT. It is therefore a one-stop pixel identification tool if you are working with multiple sensors. As a SNAP plugin, it can be used in combination with other algorithms available from S2 & S3 toolboxes, to create powerful processing chains with SNAP or GPT (Graph Processing Tool).

IdePix classifies pixels into a series of categories (flags) for further processing using a mono-temporal approach and background information. Its uniqueness consists of a certain set of flags, which are calculated for all instruments (common flags), complemented by instrument specific flags (instrument flags). The technical design of all IdePix is instrument specific and can include decision trees, probabilistic combination of calculated features or neural networks.

Sentinel-3 OLCI IdePix derives cloud flags on two confidence levels, as well as cloud shadow, mountain shadow, snow/ice, and water flags. Cloud boundary pixels are flagged using a dilation filter. In principle, cloud boundaries are regarded as neighbour pixels of a cloud as identified before by the processor; thus, a buffer is set around the cloud. The width of this boundary (in number of pixels) can be set by the user.

In contrast to many other pixel identification tools, the final IdePix classification is non-exclusive and therefore allows multiple classes to be set for a single pixel. This means a single pixel can have multiple properties like land and cloud (semi-transparent cloud over land), land and snow (land covered with snow), or land, snow, and cloud (semi-transparent cloud over snow covered land). This type of implementation allows the most versatile usage of the flagging and combinations according to users' needs compared to a standard integer flag allowing a single status per pixel.

3 IdePix algorithm

The Sentinel-3 OLCI IdePix algorithm consists of multiple modules. While most of the classification is based on a neural network (NN) and is done in the main classification module, some classification like mountain and cloud shadow are done in post processing, since they rely on prior process flags or auxiliary data.

3.1 Static land/water and coastline mask

In a first processing step land and coastline pixels are flagged. Within the Sentinel-3 OLCI IdePix only land pixels are provided as a Boolean mask to the user, called IDEPIX_LAND. Since this is a Boolean mask, the inversion of this mask can provide a water mask. In addition, coastline pixels are provided through the IDEPIX_COASTLINE mask. The IDEPIX_LAND and IDEPIX_COASTLINE masks are based on auxiliary data and therefore are static maps.

There are two options implemented that can be selected by the user. In the default setting the Sentinel-3 OLCI L1b masks for land and coastline pixel are forwarded to the respective IdePix masks. Defining the masks as follows:

$$IDEPIX\_LAND = quality\_flags\_land[2]$$

$$IDEPIX\_COASTLINE = quality\_flags\_coastline[3]$$

In addition to the default setting, the user can also choose to use the SRTM water body dataset[4] (SWBD) as input for land and coastline definition. The SWBD is implemented in SNAP[5] within the Fractional Land/Water Mask Processor. The processor provides percentile water fraction $f_{SWBD} \in$ [0,100] for each pixel.

If the SRTM option is selected, the IdePix land and coastline masks are defined as follows:

$$IDEPIX\_LAND = (f_{SWBD} = 0)$$

$$IDEPIX\_COASTLINE = (f_{SWBD} > 0) \wedge (f_{SWBD} < 100)$$

Besides the differences in the input data between the default setting and using the SWBD dataset, selecting the SRTM option (SWBD dataset) will also lead to a difference in the mask definition. While the OLCI L1b land mask includes all inland water as land, the SWBD mask excludes inland water from the land mask. Therefore, the IDEPIX_LAND mask using default settings includes all inland water, whereas the SRTM option excludes inland water from the IDEPIX_LAND mask.

3.2 O$_2$ harmonization

As described in Aktaruzzaman, M. D. (2008), "*the spectral smile, also as 'smile' or 'frown' curve, is a spectral distortion that is typically found in push-broom sensors. It is a shift in wavelength in the spectral domain, which is a function of across-tack pixel (column) in the swath.*"

---

[2] OLCI Level 1b mask
[3] OLCI Level 1b mask
[4] https://lpdaac.usgs.gov/products/srtmswbdv003/
[5] https://step.esa.int/main/toolboxes/snap/

Preusker et al. (2021), describe the effect for OLCI in detail as follows:

"OLCI is known to exhibit small systematic shifts of the spectral detector rows along the spatial CCD dimension, sometimes referred as smile. This causes a variation of the central wavelength and even the spectral width of each of the 3700 pixels across OLCI's field of view. Within one camera, the shifts are approximately 0.8 nm between the left and the right camera side. Further on, the cameras differ slightly from each other, thus discontinuities of up to 1.8 nm occur at the camera interfaces."

And continue:

"On top of the smile, a significant temporal drift of the OLCI wavelengths have been observed during dedicated spectral campaign measurements. For these campaigns, OLCI's band setting is changed in order to observe several spectral features, like solar Fraunhofer lines and the Oxygen A absorption line with the highest possible spectral resolution and thus allow a spectral characterisation of the imager. The campaigns are executed regularly (around every 4-6 months) by EUMETSAT and, eventually their results […] are used to form a spectral temporal evolution model for all OLCI's spectral band characteristics. The model uses a logarithmic fit to parametrize central wavelength and bandwidth for all bands at all pixels as a function of the orbit number. It is available on the European Space Agency (ESA) S3 website."

For algorithms using the Sentinel-3 OLCI O_2-A bands, this smile effect poses a problem. A smile correction as implemented in the OLCI ground segment processor it is based on a first order Taylor expansion, approximating the spectral sensitivity with a simple difference quotient of neighbouring bands and modifying the bands to be consistent with their nominal wavelength. For the OLCI bands in the O2A band, namely bands 13, 14 and 15, such a standard smile correction is not possible, since the spectral sensitivity cannot be approximated using neighbourhood bands. This is undesirable, since concurrently the impact of OLCI's smile is largest at these bands. Eventually, the smile effect impedes a simple usage of the oxygen bands for cloud detection. To mitigate this problem, an algorithm was developed by Preusker et al. (2021). The *"approach is to carry out an upstream modification of the O2-A band measured radiance to a radiance at their nominal position (i.e., theoretical wavelength as defined in the OLCI Mission requirements), referred herein as spectral harmonization."* Preusker et al. (2021) further explain, *"the background of the harmonization is the sensitivity of the apparent transmission to deviations of the band central wavelength and bandwidth with respect to their nominal values. The core is a look up table that contains transmissions for the nominal band characteristics as well as for all sensible modifications of band characteristics."*

For a detailed description of the Sentinel-3 OLCI O_2-A band harmonization please read Preusker et al. (2021) and Preusker and Fischer (2021).

3.3   Single pixel-based processing

The main part of the IdePix classification algorithm for Sentinel-3 OLCI is based on a neural network (NN). Compared to processing done in the subsequent chapters, single pixel processing means that no information about surrounding pixels is required. Making this processing step quite inexpensive.

3.3.1   Invalid test

In a first step all invalid pixels are identified. Let ρ_i denote the TOA reflectance measured for OLCI channel $i \in [1, \dots 21]$. Then the following Boolean test is executed to identify invalid pixels:

$$\text{INVALID} = \text{quality\_flags\_invalid}^6 \lor \lnot \left(\bigwedge_{i \in [1,\dots,21]} \rho_i \geq 0 \right)$$

3.3.2   Internal tests

Two tests are used later in the cloud classification process to aid the neural network (NN) classification. This includes a brightness test on Oa17_radiance as well as the use of the L1b included glint flag.

Four different brightness tests are used to aid the classification of opaque and semi-transparent clouds over land or water.

Land:

$$\text{isBrightLand1} = \text{Oa17\_radiance} > \text{BrightThreshLand1}$$

$$\text{isBrightLand2} = \text{Oa17\_radiance} > \text{BrightThreshLand2}$$

Water:

$$\text{isBrightWater1} = \text{Oa17\_radiance} > \text{BrightThreshWater1}$$

$$\text{isBrightWater2} = \text{Oa17\_radiance} > \text{BrightThreshWater2}$$

The following thresholds defined in Table 1 are used.

Table 1: Brightness test thresholds

Threshold name	Threshold value
BrightThreshLand1	0.3
BrightThreshLand2	0.25
BrightThreshWater1	0.2
BrightThreshWater2	0.08

Glint areas are identified by using the Boolean L1b glint risk mask and therefore the internal IdePix glint mask is defined as:

$$\text{isGlint} = \text{quality\_flags\_sun\_glint\_risk}^7$$

---

[6] Level 1b flag
[7] Level 1b flag quality_flags_sun_glint_risk

3.3.3 Cloud/snow separation using harmonized O₂ bands

3.3.3.1 OLCI excess of scattering (EOS) algorithm

While in general clouds can be best detected by their white and bright colour, these features totally fail for the separation over snow/ice. In contrast, the height of scattering surface is independent and thus sensitive to clouds regardless of the whiteness and brightness of the underlying surface (as long as the underlying surface is reflecting a minimum amount of light, see below). OLCI allows to estimate the height of the scattering surface over the height of the surface, expressed as "excess of transmission".

OLCI includes three bands in the O_2-A absorption band. The downwelling solar irradiance is largely absorbed in these bands due to oxygen absorption. The absorption, or its inverse, the transmission, depends on the amount of oxygen along the light path. Inversely, the measurement of reflectance in an O2-A absorption band can be used to estimate the amount of oxygen in the light path, and from this amount, together with the sun and viewing conditions, the height of the surface which reflected the sun light can be estimated.

Since the mixing ratio of air is constant, the amount of oxygen depends (a) on the height of surface which reflects the sun light and (b) on the amount of oxygen within the medium which forms the surface.

- In the case of clear atmosphere over land, the amount of oxygen depends only on altitude of the land surface and the air pressure. Thus, the measurement can be converted into a surface height if the air pressure is known, or, if the altitude is known, it can be converted into surface pressure.
- In the case of an opaque cloud, the amount of oxygen depends on the airmass above the cloud, i.e., the cloud top height, and the absorption of light within the cloud. The latter can be well modelled for opaque clouds, and thus the measurement in the O2-A band is a reliable measure for the cloud top height.
- In case of a semi-transparent cloud, the amount of oxygen is a mixture of airmass above the cloud, absorption within the cloud, and airmass below the cloud. In any case, there will be less oxygen along the path compared to clear sky conditions but more compared to opaque cloud conditions. If the measurement is converted into a height value, this value will be something between the cloud top height and the land surface altitude.

The above discussion leads directly to the principle how the O2A measurements of OLCI are used for cloud detection: the measurement in the O2A band can be converted into transmission value, which can be converted into a height value. This height is either equal to the altitude of the pixel (within the uncertainty of the method) in which case there are clear sky conditions, or it is larger, taking the uncertainty into account, a cloud (semi-transparent or opaque) is present.

The uncertainty of the transmission or height retrieval is an important quantity for the sensitivity of the method. Only if the difference between the surface altitude (i.e., DEM altitude of the pixel) and the retrieved one exceeds the uncertainty, a cloud can be identified.

One major contributor to the uncertainty is the exact spectral position of the O2A band. This was already discussed in section 3.2 (O_2 harmonisation).

The O_2 harmonisation results in a measurement of the transmission of the atmosphere within OLCI O_2 absorption wavebands $i \in [13,14,15]$. Since it does not matter, in principle, which waveband is considered, we consider $i = 13$ only. This value is determined along the path from sun to the scattering surface, and then towards the sensor. In order to be used for a cloud detection, the measured transmission T_i^m needs to be normalised to standard viewing and illumination conditions, $\bar{T}_i^m$. The normalisation puts the sun in zenith and the sensor into nadir view conditions by considering the air mass

$$f_{\mathrm{AMF}} = \frac{1}{\cos\theta_s} + \frac{1}{\cos\theta_v}$$

$$\bar{T}_i^m = \kappa_l\, f_{\mathrm{AMF}}\, T_i^m$$

where κ_l is a correction factor to account for residual smile found in camera l (see Figure 2).
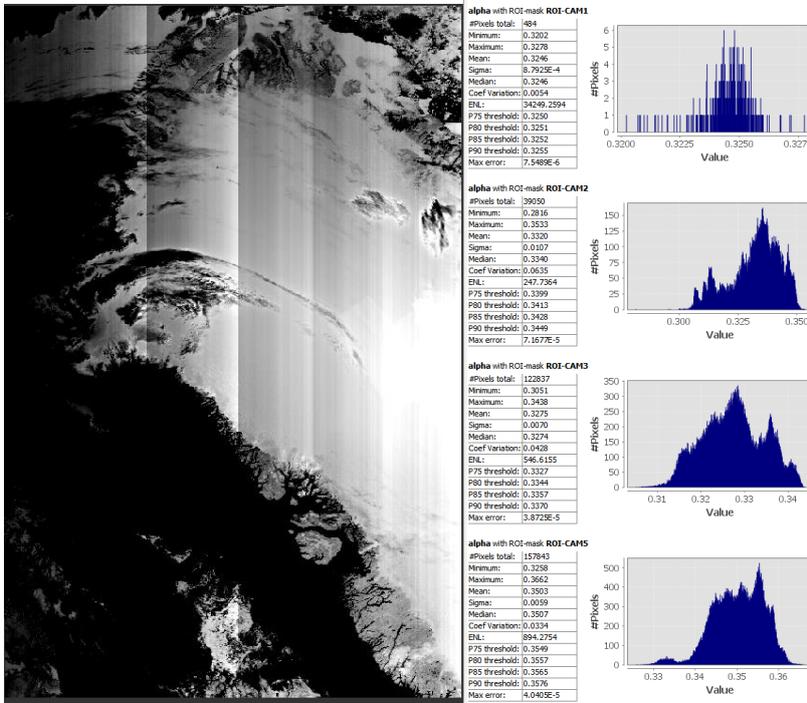
Figure 2: Correction factor κ for an OLCI scene (28.07.17, Greenland), and its distribution per camera. For each camera l a mean value κ_l can be used.

In a next step the expected transmission, $\bar{T}_i^{cs}$, is calculated. $\bar{T}_i^{cs}$ is the transmission which would be measured over clear sky in waveband i for the nominal wavelength and nadir looking conditions and an altitude z of the pixel. According to Bouguer's law, T_i^{cs} is an exponential function of the optical depth, which scales linearly with the height.

$$\bar{T}_i^{cs} = a\, e^{bz} + c$$

The coefficients a, b and c can be approximated from RT calculations or found experimentally using OLCI images.

Over clear sky, $\bar{T}_i^{cs}$ and $\bar{T}_i^m$ shall agree within a tolerance ε given by the combined uncertainty of both quantities, i.e., the following inequality holds:

$$|\bar{T}_i^{cs} - \bar{T}_i^m| < \varepsilon$$

If the inequality is not fulfilled, $\bar{T}_i^{\mathrm{m}}$ is significantly greater than $\bar{T}_i^{\mathrm{cs}}$, which indicates an excess of transmission and thus the presence of a cloud or other type of elevated atmospheric scattering layer.

The tolerance ε was determined experimentally (Figure 3): in an OLCI scene the difference $\bar{T}_i^{\mathrm{cs}} - \bar{T}_i^{\mathrm{m}}$ was calculated. Within the scene, a large cloud free area was detected manually. The histogram of the differences for all pixels in this area showed a gaussian distribution centred around 0 and a maximum value of 0.23. This maximum value can be used as a conservative value for the uncertainty ε. The histogram for all pixels shows that value separates the "clear" peak of the distribution from the rest. It may be slightly too conservative.
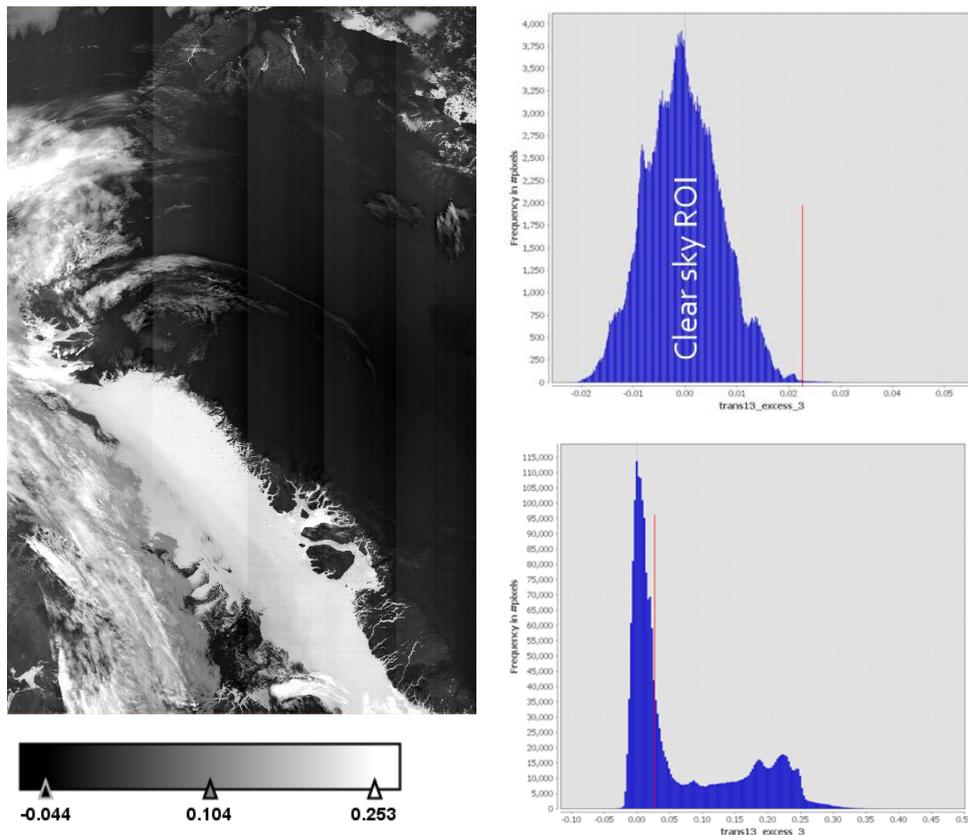

Figure 3: Excess of transmission in band 13 (left), and the histogram for the clear sky area (top right) and the whole image (bottom right). The uncertainty is indicated as red line.

In an update of the algorithm, the normalization as well as the calculation of the clear sky transmittance were also calculated with look-up tables based on radiative transfer simulations. In this context, look-up tables were also generated to calculate the surface pressure at the pixel, based on DEM altitude and ECMWF pressure available in the OLCI L1b products, and a so-called effective pressure, which is the pressure level at which an opaque cloud would be placed in order to correspond to the measured transmittance. These latter two quantities allow to perform the cloud detection in pressure units, but also to get an estimate of a cloud top pressure if an opaque cloud is confirmed (e.g., from other tests in the VIS-NIR spectral domain).

3.3.3.2   OLCI 1020 nm reflectance algorithm

The OLCI excess of scattering (OES) algorithm works fine within its uncertainty limits but is missing low level clouds (i.e., within the uncertainty) and does not work if the underlying surface is dark, i.e., it has to be disabled over sea. Thus, it is advantageous to combine the OES algorithm with a spectral test

which is simple and exploits the lower reflectance of snow compared to clouds towards the SWIR (see Figure 9). The longest wavelength band of OLCI is band 21 centred at 1020 nm. It was found that a simple global threshold applied to the TOA radiance converted into reflectance identified a good portion of the clouds over snow.

3.3.3.3   Combined algorithm

It was found that the OLCI 1020 nm reflectance algorithm has a large overlap with the cloud pixels identified already by OES, however, it also identifies additional pixels of low level or thin cloud not detectable by OES. Thus, a logical "or" combination (i.e., a logical disjunction) of the two tests improves the overall performance.
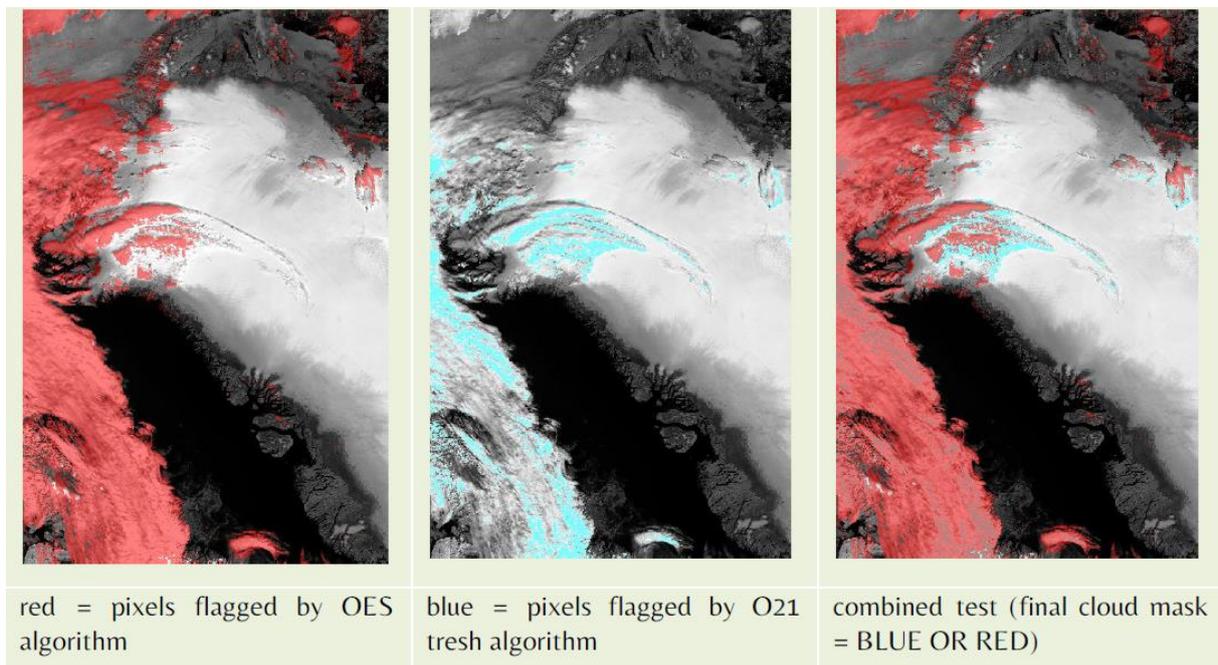

red = pixels flagged by OES algorithm

blue = pixels flagged by O21 tresh algorithm

combined test (final cloud mask = BLUE OR RED)

Figure 4: O21 Thresh cloud mask, OES cloud mask and combined cloud mask applied to OLCI, 28.07.2017, Greenland

3.3.4   Sea ice limitation tests

The NN has a small tendency to detect snow/ice in some very cloudy tropical regions. Therefore, a climatology is introduced to limit the later NN based snow/ice test to polar regions. The polar regions (Arctic and Antarctic) are stored as polygons. The pixel latitude and longitude for each processed pixel are checked to be inside or outside of these regions:

isSeaIcePossible = "pixel lat/lon is within Arctic or Antarctic polygon"

3.3.5   NN classification for cloud and snow

The cloud and snow value features are predicted by a neural network (NN) model or to put it more succinctly, a feed-forward neuronal network with one input and one output layer. The NN has been trained with a back-propagation learning algorithm by using a PixBox[8] dataset. The cloud and snow value features are calculated through the NN by using the SENTINEL 3 OLCI TOA reflectance spectrum as input. The output value feature has a range from 0 through 6 (see Table 2) and corresponds to the trained classes clear land, clear water, clear snow/ice, opaque cloud, semi-transparent cloud, and spatially mixed cloud (spatially mixed cloud/water or cloud/land). The output of the NN is represented

---

[8] PixBox is a proprietary software developed by Brockmann Consult to conduct expert pixels collections.

by floating point values between these feature values. User defined thresholds (or cut values) are used to define the best possible separation of the classes.

> **Excursion**
>
> *PixBox: The overarching goal of the so called "PixBox" is to enable a quantitative assessment of the quality of a pixel classification produced by an automated algorithm/procedure. Pixel classification is defined as assigning a certain number of attributes to an image pixel, such as cloud, clear sky, water, land, inland water, flooded, snow etc. These pixel classification attributes are typically used to further guide higher level processing.*

Table 2: NN output features

Feature value	Feature property
1	clear snow/ice
2	opaque cloud
3	semi-transparent cloud
4	spatially mixed cloud
5	clear Land
6	clear water

The following threshold are used to define intermediate classes:

Intermediate class names	Surface	Lower bound threshold	Upper bound threshold
CLEAR_SNOW_ICE	All	0.0	1.1
OPAQUE_CLOUD	All	1.1	2.75
SEMI_TRANS_CLOUD	All	2.75	3.5
SPATIAL_MIXED_LAND	Land	3.5	3.85
SPATIAL_MIXED_WATER_GLINT	Water with glint	3.5	3.5
SPATIAL_MIXED_WATER_NOGLINT	Water no glint	3.5	3.75

The intermediate classes are then used to define the internal IdePix flags called "isCloudSure", "isCloudAmbiguous" and "isSnowIce". Prior to using the NN values for cloud classification additional spectral tests defined in section 3.4.2 are used to refine the NN behaviour. The refinement is done separately over water and land. Furthermore, the commissioning of sea ice is limited using the test defined in section 3.3.4.

To separate the different approaches over land and water surfaces the Level 1b flag quality_flags_land is used to define an internal Boolean "isLand" flag.

$$isLand \ = \ quality\_flags\_land[9]$$

If isLand is false, the equations for water are used else the land equations are used.

---

[9] Level 1b flag

Water:

isCloudSure = OPAQUE_CLOUD ∧ isBrightWater1

isCloudAmbiguous =

$$(\text{isGlint}^{10} \wedge (\text{SEMI\_TRANS\_CLOUD} \vee \text{SPATIAL\_MIXED\_WATER\_GLINT})$$
$$\wedge \text{ isBrightWater2}) \vee (\neg\text{isGlint}^{11} \wedge (\text{SEMI\_TRANS\_CLOUD}$$
$$\vee \text{ SPATIAL\_MIXED\_WATER\_NOGLINT}) \wedge \text{ isBrightWater2})$$

isSnowIce = isSeaIcePossible ∧ CLEAR_SNOW_ICE

Land:

isCloudSure = OPAQUE_CLOUD ∧ isBrightLand1

isCloudAmbiguous =

(SEMI_TRANS_CLOUD ∨ SPATIAL_MIXED_LAND) ∧ isBrightWater2

isSnowIce = CLEAR_SNOW_ICE

3.3.6   Final cloud and snow definition

In a final step the internal cloud and snow flags are used to define the IdePix user flags (masks).

Water:

IDEPIX_CLOUD = isCloudSure ∨ isCloudAmbiguous

IDEPIX_CLOUD_SURE = isCloudSure

IDEPIX_CLOUD_AMBIGUOUS = isCloudAmbiguous

IDEPIX_SNOW_ICE = isSnowIce

Land:

The NN based cloud and snow flags of land surfaces over Greenland and Antarctica are refined using the S3-SNOW approach, described in section 3.3.3, to counteract the behaviour of overestimation of clouds over snow covered land areas.

The following test is defined to identify clouds over snow covered land surfaces more precisely compared to the NN results:

$$\text{isCloudOverSnow} = \left((\rho_{21} > 0.5) \wedge (\bar{T}_{13}^{\text{cs}} - \bar{T}_{13}^{\text{m}} < 0.01)\right) \vee (\rho_{21} > 0.76)$$

---

[10] L1b quality_flags_sun_glint_risk
[11] L1b quality_flags_sun_glint_risk

With $\bar{T}_{13}^{cs}$ being the theoretical transmittance from a clear surface for OLCI band 13, and $\bar{T}_{13}^{m}$ being the normalised measured transmittance from the O_2 harmonized OLCI band 13, and ρ_{21} denoting the TOA reflectance measured for OLCI channel 21.

Since this test is only executed over Greenland and Antarctica, these areas are defined by:

$$isInsideGreenland = \text{"pixel lat/lon is within Greenland polygon"}$$

$$isInsideAntarctica = \text{"pixel lat/lon is within Antarctica polygon"}$$

If the following condition, for location of a pixel being in Greenland or Antarctica, is fulfilled

$$isInsideGreenland \lor isInsideAntarctica$$

the following refinement is executed to change the Boolean status of the internal flags:

If isCloudOverSnow

       isCloudSure $=$ true

       isCloud $=$ true

       isSnowIce $=$ false

else

       if isCloud

              isSnowIce $=$ true

              isCloud $=$ false

              isCloudSure $=$ false

              isCloudAmbiguous $=$ false

If the previous condition for Greenland and Antarctica is not fulfilled, the status of previously defined internal flags ("isCloudSure", "isCloudAmbiguous" and "isSnowIce") remains unchanged.

Now the IdePix user flags are defined as:

IDEPIX_CLOUD $=$ isCloudSure $\lor$ isCloudAmbiguous

IDEPIX_CLOUD_SURE $=$ isCloudSure

IDEPIX_CLOUD_AMBIGUOUS $=$ isCloudAmbiguous

IDEPIX_SNOW_ICE $=$ isSnowIce

3.3.7 L1 flag forwarding

The L1b quality_flags_bright is forwarded for internal use and provided as an IdePix flag to identify bright pixels

3.4   Cloud buffer

For pixels identified as cloudy (IDEPIX_CLOUD), a cloud buffer of specified width can be set by the user to finally obtain a more conservative cloud mask. This is done in a post-processing step after the cloud classification has been applied on the full image. The cloud buffer algorithm works on pixel windows of size $(2N + 1)^2$ where N is the width of the cloud buffer. Note that the cloud buffer is only applied to cloud-free pixels, i.e., cloudy pixels are not flagged as cloud buffer of cloudy neighbouring pixels. The cloud buffer is optional and can be omitted.

3.5   Mountain shadow

The algorithm for computing Slope, Aspect and Orientation for a Sentinel-3 OLCI product is based on the hillshade theory (e.g., Schuckmann 2020, ESRI 2016).

Orientation gives the angle between the North direction and the y-direction in a 3x3 macro pixel (same x-pixel-value in satellite coordinates), which is defined by two points: Point 1 with satellite coordinates (x_0, y_1) and geographic latitude and longitude coordinates (ϑ_1, φ_1) and Point 2 (x_2, y_2) with (ϑ_2, φ_2). From these coordinates orientation (or bearing, see [1]) can be derived as follows:

$$X = \cos(\vartheta_2) \sin(\varphi_2 - \varphi_1)$$

$$Y = \cos(\vartheta_1) \sin(\vartheta_2) - X \sin(\vartheta_1)$$

$$\text{orientation} = \text{atan2}(X, Y)$$

Spatial resolution is either calculated from the great circle distance of two points or it relies on the CRS geo-coding.

The illumination direction has to be corrected. Following the same logic, which has been applied in the cloud shadow algorithm (section 3.6.1), elevated objects are only positioned correctly on the satellite image, if they were observed from nadir view. If the view is tilted, the object appears to be further away in that viewing direction, whereas shadows (of clouds) seem to indicate a change in illumination direction across the swath. The apparent sun azimuth angle is calculated to correct for the tilted observation and the expected direction of a shadow in the satellite image is met by this adjustment.

The same logic has to be applied to the slope of elevation as well: looking at a single mountain, its peak appears to be further away from its actual position (if seen from nadir) on the satellite grid in a tilted observation. Therefore, if uncorrected, all hill and mountain slopes for view zenith angles $\theta_v > 0$ are too flat and a lot of shadows cannot be detected.

To derive a meaningful slope, the spatial resolution of the satellite pixel grid has to be corrected for the tilted view, in x and in y-direction of the grid according to the zenith and azimuthal viewing geometry θ_v, ϕ_v and the orientation (angle between y direction and North).

From a 3x3 macro pixel, the slope is calculated from elevation H with corrected spatial resolution:

$$b = \frac{H[2] + 2H[5] + H[8] - H[0] - 2H[3] - H[6]}{8}$$

$$c = \frac{H[0] + 2H[1] + H[2] - H[6] - 2H[7] - H[8]}{8}$$

$$\text{spatialResolutionXadjust} = \frac{\text{spatialResolution}}{\cos \theta_v} \times \sin\big(360° - (\phi_v + \text{orientation})\big)$$

$$\text{spatialResolutionYadjust} = \frac{\text{spatialResolution}}{\cos\theta_v} \times \cos\big(360° - (\phi_v + \text{orientation})\big)$$

$$\text{slope} = \text{atan}\left(\sqrt{\left(\frac{b}{\text{spatialResolutionXadjust}}\right)^2 + \left(\frac{c}{\text{spatialResolutionYadjust}}\right)^2}\right)$$

$$\text{aspect} = \text{atan2}(-b, -c)$$

If the sun stands in the North (mostly for southern hemisphere, depending on season), the aspect has to be turned by 180°. This happens, if the sun azimuth angle $\phi_s > 270°$ or $\phi_s < 90°$; in these cases, aspect is subtracted by 180°. Afterwards, the aspect is mapped from [-180°, 180°] onto [0°, 360°].

A pixel is flagged as mountain shadow if the following condition is met:

$$\cos\beta = \cos(\theta_s)\cos(\text{slope}) + \sin(\theta_s)\sin(\text{slope})\cos\big(\phi_s - (\text{aspect} + \text{orientation})\big)$$

Mountain shadow condition: $\cos\beta < \text{mntShadowExtent} - 1$

Mountain shadow extent is set in default to 0.9, the parameter excepts values between 0 and 1, corresponding to no shadow extension at 0 and maximum shadow extension at 1.

3.5.1 Known issues
- On the southern hemisphere: either Western part of swath from nadir has too little mountain shadow pixels, or Eastern part too many.
- Some shadows can never be detected by this algorithm via slopes. If a steep, high mountain casts a shadow into a flat valley, the true origin of the shadow is not recognized, and the small slopes will not allow the shadow to be identified. (This problem intensifies for higher spatial resolutions.)

3.6 Cloud shadow

In a postprocessing step, after the pixel identification procedure has already found out cloud pixels in the scene, the cloud shadow for these clouds can be calculated. The general idea consists of the following steps:

- Start at a cloud free pixel
- Check along the illumination path in direction towards the sun, up to a maximum distance, whether a cloud pixel lies in this path.
  - Calculate the geometry of this triangle: distance between start point and current position with an identified cloud. The sun zenith angle at the start point defines the height of the illumination path at the current position.
  - If the cloud top height is larger than the current height, the illumination path is obstructed by the cloud and the start point is covered by cloud shadow.
  - If cloud shadow is found, stop the search process, and move to the next cloud free pixel.
  - Otherwise iterate along the illumination path, until a cloud pixel with the correct height or the maximum distance is reached.

3.6.1 Calculating the direction of illumination – apparent sun azimuth angle

The direction of illumination needs very careful consideration. The observation and sun geometry is specified at each pixel in the OLCI product, giving zenith angles counting from local zenith position downwards and azimuths counting counter clock-wise from the direction towards North. Though the sun azimuth angle defines the direction of illumination, the cloud as the object which casts a shadow is imaged in the plain of observation depending to the view zenith angle: the larger the OZA, the further away from the clouds actual position above ground it appears in the observation. As we use the cloud pixel position in the observation plain to calculate the cloud shadow, the sun azimuth angle has to be corrected, so that it becomes the projection of the illumination direction on the surface.

If the cloud pixel positions are used for the calculation of the cloud shadow, their projection on the surface by the viewing geometry must be corrected. An elevated object appears in the scene only at its true position above ground if it is looked at in nadir view. Otherwise, its apparent position moves further away from the true position in the direction of the observation geometry. Instead of using the CTH estimation and calculating the true position of a cloud pixel, the sun azimuth angle is geometrically corrected.

There are two cases to be considered:

- East of nadir view (view azimuth negative) $\phi_v < 0$
- West of nadir view (view azimuth positive) $\phi_v > 0$

For $\phi_v < 0 : \Delta\phi = 360° - |\phi_v| - \phi_s$; and the apparent sun azimuth is smaller than TP_SAA.

For $\phi_v > 0: \Delta\phi = \phi_s - \phi_v$; and the apparent sun azimuth is larger than TP_SAA.

Derived from the geometry on a planar surface, the correction for the sun azimuth angle Δ can be calculated per pixel:

$$\cos\Delta = \frac{\tan\theta_s - \tan\theta_v \cos\Delta\phi}{\sqrt{\tan^2\theta_v + \tan^2\theta_s - 2\tan\theta_s \tan\theta_v \cos\Delta\phi}}$$

If the illumination is directed towards the South, the sign of the correction Δ has to switch:

For $\phi_s > 270°$ or $\phi_s < 90°$: $\Delta = -\Delta$

And depending on the observation azimuth angle (looking towards east or west), the sign of the correction has to be adjusted accordingly:

For $\phi_v < 0$: $\phi_{s,app} = \phi_s - \Delta$

For $\phi_v > 0$: $\phi_{s,app} = \phi_s + \Delta$

(θ_s = SZA; θ_v = OZA; ϕ_s = SAA; ϕ_v = OAA)

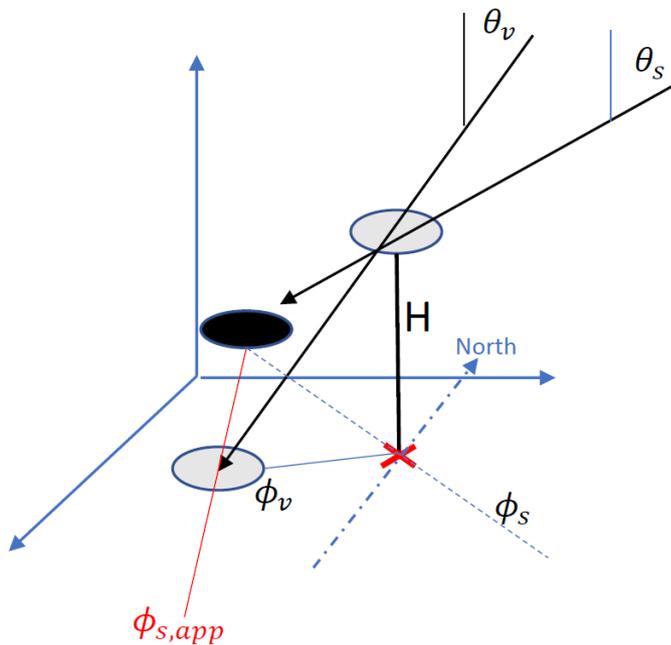### 3.6.2 More details - Geometric considerations for apparent sun azimuth

Figure 5 Position of cloud (grey, actual position above ground at red cross, apparent position in plane) and cloud shadow (black). Azimuth angles are counted counterclockwise from North. The correction of the sun azimuth angle to the apparent sun azimuth angle (red) is given by the angle between the red line and the sun azimuth line through the actual position.

Direction of the cloud projected onto the surface (origin at actual position of the cloud at the ground) p_{cloud}, direction of sun projected on the surface p_{sun}:

$$\vec{p}_{cloud} = H \tan(\theta_v) \ (\cos\phi_v, -\sin\phi_v)$$

$$\vec{p}_{sun} = H \tan(\theta_s) \ (\cos\phi_s, \sin\phi_s)$$

Angle between $\vec{p}_{cloud} - \vec{p}_{sun}$ and $\vec{p}_{sun}$ gives the needed correction for SAA:

$$\cos\Delta = \frac{\vec{p}_{sun}}{\|\vec{p}_{sun}\|} * \frac{\vec{p}_{cloud} - \vec{p}_{sun}}{\|\vec{p}_{cloud} - \vec{p}_{sun}\|}$$

This correction can be large particularly in the west of the swath, where OZA for OLCI get large.

3.6.3   Determine the maximum of search distance along illumination direction

The search algorithm starts at a cloud free pixel, finds the direction towards the sun, and investigates, whether there is a cloud at the necessary height along this path, which would cast a shadow. The angle towards the sun against North is the corrected sun azimuth angle.

A first guess for the maximum length of the search path (in meter) is defined by the maximum cloud height, a function of latitude, and the sun zenith angle:

$$\text{maxDist} = \text{maxCloudTopHeight} \times \tan(\theta_s)$$

with maximum cloud top height of 12000 m.

If the sun zenith angle is large, the distance can become very long and the end point at this distance may fall outside the scene.

In this case, the end position of the search path is derived as follows: the minimum distance between a cloud and its shadow is defined by the spatial resolution of the OLCI grid (here: 300 m) and the sun zenith angle.

$$\text{cloudDistanceMin} = \frac{300 \text{ m}}{\tan(\theta_s)}$$

The search path length starts at a length of maxDist, and will be reduced step by step by cloudDistanceMin, until the end point lies within the raster, or the distance has become smaller than twice cloudDistanceMin.

3.6.4   Cloud top pressure

To derive cloud top pressure an artificial neural network has been trained from a simulated data set.

The TensorFlow library is needed to interpret these NN files. There is an installation guide for TensorFlow in C, see https://www.tensorflow.org/install/lang_c

The license terms for TensorFlow can be found here: https://github.com/tensorflow/tensorflow/blob/master/LICENSE

An example, how the TensorFlow library is used in a JAVA implementation of the cloud shadow algorithm, can be found here: https://github.com/bcdev/snap-idepix/blob/master/idepix-olci/src/main/java/org/esa/snap/idepix/olci/TensorflowNNCalculator.java

The neural network takes the following input variables (in this order):

- Sun zenith SZA transformed into its cosine: $\cos\left(\text{SZA} \times \frac{\pi}{180°}\right)$
- Observation zenith OZA, transformed into its cosine: $\cos\left(\text{OZA} \times \frac{\pi}{180°}\right)$
- Azimuth difference, transformed by multiplication with the sine of the observation zenith angle to prevent discontinuities at the nadir: $(\text{SAA} - \text{OAA}) \times \frac{\pi}{180°} \times \sin\left(\text{OZA} \times \frac{\pi}{180°}\right)$
- Top of atmosphere reflectance at 754 nm. The simulated data uses the following definition for this reflectance: Oa12_reflectance = Oa12_radiance/ solar_flux_band_12, so it has been trained and applied in this fashion in the training and the application of the neural network.

- During the O_2 Harmonization, the corrected transmittances for O_2 absorption bands (bands 13, 14, 15) are calculated. The transmittances are transformed by taking the negative value of the natural logarithm, i.e., $-\ln(T_{13}), -\ln(T_{14}), -\ln(T_{15})$.

The neural network calculates one output, the normalized cloud top pressure. To transform the output value back into a meaningful range, it has to be converted in the following manner:

$$CTP = outputValue \times 228.03508502 \text{ hPa} + 590 \text{ hPa}.$$

3.6.5 Convert cloud top pressure to cloud top height

With an estimation of the cloud top height CTH the cloud shadow can be calculated geometrically based on the cloud mask.

In the MERIS algorithm, the conversion of CTP to CTH has been calculated by:

$$CTH = -8000 \text{ m} \times \ln\left(\frac{CTP}{1013 \text{ hPa}}\right)$$

It has been decided to use the international barometric height formula, which considers the air pressure at sea level p_0 and the temperature T_s at pressure CTP derived from the temperature profiles given in the OLCI product.

$$CTH = -\frac{T_s}{0.0065 \text{ Km}^{-1}} \times \left(\left(\frac{CTP}{p_0}\right)^{\frac{1}{5.255}} - 1\right)$$

This approach reduces the cloud top height, which becomes rather large in the MERIS algorithm; often above 10 km. If validated by means of the cloud shadow positions, the lower values lead to more realistic and accurate coverage.

3.7   Flag definitions

Table 3 list all flags produced by the Sentinel-3 OLCI IdePix. The methodological basis for the flags was described in the previous sections. Some of the listed flags in Table 3, like IDEPIX_CLOUD_SURE, are common flags, available for each Satellite sensor processable with IdePix (e.g. Sentinel-2 MSI, Envisat MERIS, or Proba-V), while other flags are sensor specific and do not exist for all sensors, like IDEPIX_CIRRUS_AMBIGUOUS.

Table 3: IdePix flagging

Bit	Integer value	Flag name	Description
1	1	IDEPIX_INVALID	Invalid pixels
2	2	IDEPIX_CLOUD	Pixels which are either CLOUD_SURE or CLOUD_AMBIGUOUS
3	4	IDEPIX_CLOUD_AMBIGUOUS	Semi-transparent clouds, or clouds where the detection level is uncertain
4	8	IDEPIX_CLOUD_SURE	Fully opaque clouds with full confidence of their detection
5	16	IDEPIX_CLOUD_BUFFER	A buffer of n pixels around a cloud. n is a user supplied parameter. Applied to pixels masked as 'cloud'
6	32	IDEPIX_CLOUD_SHADOW	Pixel is affected by a cloud shadow
7	64	IDEPIX_SNOW_ICE	Clear snow/ice pixels
8	128	IDEPIX_BRIGHT	Bright pixels
9	256	IDEPIX_WHITE	Currently not defined
10	512	IDEPIX_COASTLINE	Pixels at a coastline
11	1024	IDEPIX_LAND	Land pixels identified by the used land/water mask; L1b or SRTM (default L1b)
12	2048	IDEPIX_MOUNTAIN_SHADOW	Pixel is affected by mountain shadow

4 References

Aktaruzzaman, M. D. (2008). Simulation and Correction of Spectral Smile Effect and Its Influence on Hyperspectral Mapping. ITC. https://books.google.de/books?id=F5teAQAACAAJ

Baig, M. H. A., Zhang, L. , Shuai, T. & Tong, Q. (2014) Derivation of a tasselled cap transformation based on Landsat 8 at-satellite reflectance, Remote Sensing Letters, 5:5, 423-431, DOI: 10.1080/2150704X.2014.915434

ESRI 2016. How hillshade works. https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/how-hillshade-works.htm, last visited: 14.12.2021

Farr, T. G., et al. (2007). The Shuttle Radar Topography Mission, Rev. Geophys., 45, RG2004, doi:10.1029/2005RG000183.

Frantz, D., Haß, E., Uhl, A., Stoffels, J., & Hill, J. (2018). Improvement of the Fmask algorithm for Sentinel-2 images: Separating clouds from bright surfaces based on parallax effects. *Remote sensing of environment*, *215*, 471-481.

Gao, B-C (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sensing of Environment, 58(3), 257–266. doi:10.1016/s0034-4257(96)00067-3

Kauth, R.J., & Thomas, G.S. (1976). The tasselled cap - A graphic description of the spectral-temporal development of agricultural crops as seen by Landsat.

Liou, K.N. (1992). Radiation and Cloud Processes in the Atmosphere. Oxford University Press, Oxford, 487pp., 1992.

Luo, Y., Trishchenko, A. and Khlopenkov, K. (2008). Developing clear-sky, cloud and cloud shadow mask for producing clear-sky composites at 250-meter spatial resolution for the seven MODIS land bands over Canada and North America', Remote Sensing of Environment, 112(12), pp. 4167–4185. doi: 10.1016/j.rse.2008.06.010

Preusker, R. & Fischer, J. (2021): Algorithm Theoretical Basis Document (ATBD) including Requirements Baseline (RB) Product Validation Plan (PVP) Input/Output Data Definition (IODD), EUMETSAT, https://www.eumetsat.int/media/48528

Preusker, R., Fischer, J., Bozzo, A., Chimot, J., Kwiatkowska, E. (2021): Zoom on the OLCI O2-A harmonization as done within OCTPO2. EUMETSAT, https://www.eumetsat.int/media/48768

Rossow, W.B., and Schiffer, R.A. (1999). Advances in Understanding Clouds from ISCCP. Bull. Amer. Meteor. Soc., 80, 2261-2288, https://doi.org/10.1175/1520-0477(1999)080<2261:AIUCFI>2.0.CO;2.

Schuckman, K. 2020, "GEOG 480, Exploring Imagery and Elevation Data in GIS Applications, Lesson 8: Terrain Modeling and Analysis, Slope, Aspect, and Hillshade", John A. Dutton e-Education Institute, College of Earth and Mineral Sciences, The Pennsylvania State University. https://www.e-education.psu.edu/geog480/node/490, last visited 14.12.2021

Upadhyay, A. 2015. Formula to Find Bearing or Heading angle between two points: Latitude Longitude. https://www.igismap.com/formula-to-find-bearing-or-heading-angle-between-two-points-latitude-longitude/, last visited: 14.12.2021

Wylie, D., 1998: Cirrus and Weather: A satellite perspective. OSA Meeting, pp.66-68, Baltimore, U.S.A., 6-8, Oct. 1998