Customer:	JRC	Document Ref.:	S2GM-ATBD-BC
Framework contract No:	933113	Issue Date: Issue:	06-November-2020 1.3.3

Algorithm Theoretical Basis Document

Sentinel 2 Global Mosaics

Copernicus Sentinel-2 Global Mosaic (S2GM) within the Global Land Component of the Copernicus Land Service

Author(s):

Grit Kirches, BC
(Science Lead)

Reviewer(s):

C. Brockmann, BC
M. Riffler, GeoVille

Approver(s):

Nadine Gobron, JRC
European Commission

Amendment Record

Releasing a new edition of the document in its entirety shall amend this document. The Amendment Record Sheet below records the history and issue status of this document.

Amendment Record Sheet

Issue	Date	Reason for change
V1.0	08-June-2018	Version delivered to JRC as Annex SC2-D1
V1.1	05-July-2018	Stand-alone version of the ATBD; major revision of all chapters
V1.2	08-January-2019	Updating of the ATBD regarding to STC (sec. 2.4.2)
V1.3	22-March-2019	Updating of the ATBD regarding to STC (sec. 2.4.2) – in line with mosaic version V1.1.0
V1.3.1	16-October-2019	Updating of the ATBD regarding to STC (sec. 2.4.2) regarding the LOT2 review – in line with mosaic version V1.1.0
V1.3.2	29-October-2019	Updating of the ATBD regarding to STC (sec. 2.4.2) regarding the second LOT2 review – in line with mosaic version V1.1.0
V1.3.3	06-November-2020	Removing of all references regarding the Specific Contract 2 “Delivery of Mosaics over Europe for 2017 -2018”

Applicable documents

AD-1	S2GM Framework Contract, Contract 933113
AD-2	Technical Proposal, Brockmann Consult, reference 32412 of 28.04.17
AD-3	Technical Specifications JRC/IPR/2017/D.6/0004/OC COPERNICUS SENTINEL-2 GLOBAL MOSAIC

Scope

This document is the Algorithm Theoretical Basis Document (ATBD) related to the Sentinel-2 L2A composite/mosaic products processed within the Framework Contract No 933113. This ATBD is written through successive increments. This version (V1.3) focuses on the updates of the composite/mosaic algorithms with respect to V1.2.

Table of Contents

List of Figures

List of Tables

1 Overview and Background Information

1.1 Introduction

The objective of the S2GM service is the provision of mosaic surface reflectance products derived from Sentinel-2 A and B platforms. Input to the processing are Level 2A products provided by the Copernicus Ground Segment, i.e. ESA Sentinel 2 core products. The S2GM service will generate regional and temporal composites at global scale and produced on-demand over specific areas of interest. The generated products will directly support international policy agreements to which the EU committed, e.g., the Paris agreement of the United Nations Framework Convention on Climate Change (UNFCC) and its activities for Reducing Emissions from Deforestation and forest Degradation (REDD+). These mosaicked surface reflectance products shall serve as input to further thematic processing, i.e. the generation of higher-level products, and analyses. Thus, the overarching requirement for the mosaicking is the selection of the most representative spectrum for a given pixel, selected from the set of observations made during the temporal compositing period and to provide this information as Analysis Ready Dataset (ARD).

The Sentinel-2 mission includes a twin-satellite constellation covering all Earth's land surfaces, large islands, inland and coastal waters every five days at the equator with even higher observation frequencies in mid- and high-latitude regions with the primary aim to support the monitoring of vegetation, land cover, and the environment in general. The Sentinel-2 Multispectral Instrument (MSI) provides multi-spectral information from 13 spectral bands ranging from visible and near-infrared to shortwave infrared wavelengths along a 290-km orbital swath. The MSI sensor data are complementary to data acquired by the U.S. Geological Survey Landsat 8 Operational Land Imager and Landsat 7 Enhanced Thematic Mapper Plus. Sentinel-2A was launched in June 2015 and Sentinel-2B in April 2017.

Input to the mosaicking process are surface reflectance values, from the so-called Level 2A product. Level 2As are operationally produced by the Copernicus (ESA) ground segment. Currently ESA is using the Sen2Cor atmospheric correction processor for the generation of Level 2A products. These Level 2A contain directional surface reflectances in 10 spectral bands (i.e. not BRDF corrected), a scene classification layer (SCL) providing information on cloudiness, snow and other pixel classification information, as well as aerosol and water vapour used during the atmospheric correction process. The S2GM mosaicking algorithm has to rely on this information for its processing.

The S2GM processing chain to calculate the mosaic image products is fully automated and is based on a modular design - see Figure 1-1. The three following main modules form the basis of the chain.

1. Quality assurance/ quality check (QA/QC) of the input products
2. Composite/Mosaic algorithm
3. Quality assurance/ quality check (QA/QC) of the input products

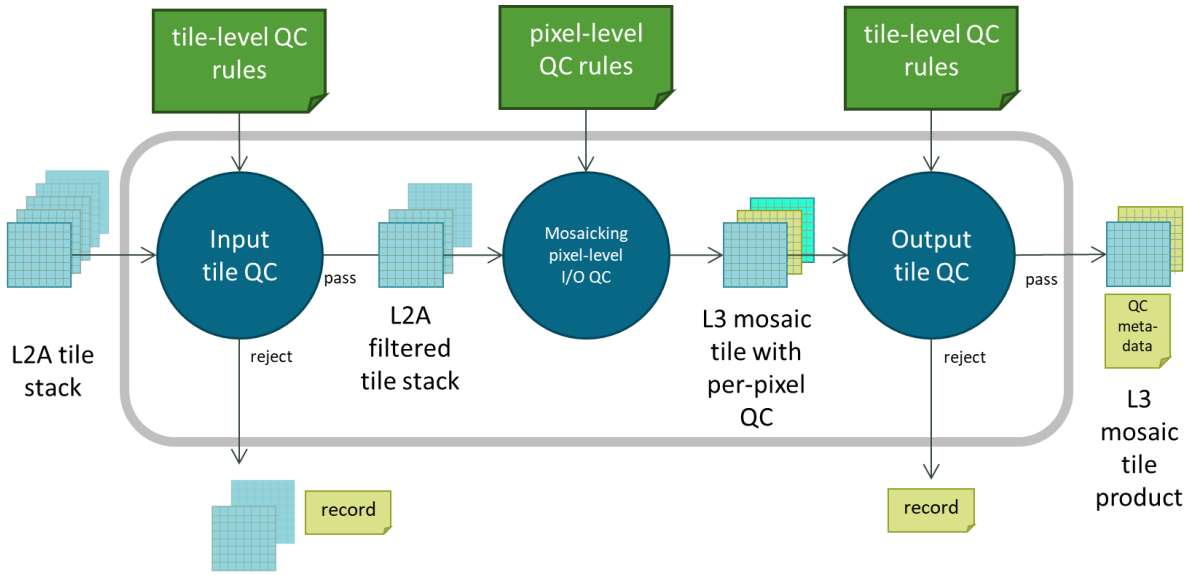


Figure 1-1: Processing chain – main steps

1.2 Review of Mosaicking Algorithms

1.2.1 Temporal Resampling

Image compositing aims at identifying the best suited observation in a given period of time on the basis of pre-defined criteria at the pixel-level or image-level [Franz, 2017]. The following section provides an overview of the most known and commonly applied image composing algorithms based on best-pixel approaches. The different algorithms are compared with respect to selection criteria that are taking into account the ITT requirements, but also aspects like processing time, storage and others. Furthermore, the advantages and disadvantages of the considered algorithms are presented and summarized in Table 1-1. The following pixel-based image compositing algorithms are assessed here [White et al., 2014]:

- WELD [Roy et al., 2010 & 2011]
- Short Term composite - STC
- Medoid [Flood 2013]
- Median NIR Composite [Potapov et al., 2011 & 2012]
- Weighted pixel-based scoring system [Griffiths et al., 2013; Franz et al., 2017]
- Cosine similarity [Nelson & Steinwand, 2015]
- Sen2Three [ESA, 2017A]

Web-enabled Landsat Data –WELD [Roy et al., 2011]

According to Roy et al. 2011, the composited mosaics generated on a monthly, seasonal, and annual basis have been designed to provide consistent Landsat data that can be used to derive land cover and geo-physical and bio-physical products for detailed regional assessments of land-cover dynamics and to study Earth System functioning. The compositing approach has been designed to preferentially select valid land surface observations with minimum cloud, snow, and atmospheric contamination. Therefore, the composited mosaics are not for Landsat studies of cloud, snow or the atmosphere. The selection criteria of the best pixel observation based on the analysis of the surface reflectance value, pixel classification, flags, normalized difference vegetation index (NDVI) and the brightness temperature. These WELD compositing criteria are used to compare two acquisitions of a pixel [Roy et al. 2011].

Short Term Composite - STC – adaption of the WELD algorithm regarding Sentinel-2

The STC approach has been motivated by the WELD method and is, like WELD, based on a decision tree regarding the surface reflectance values, the scene classification, and the different indices. Compared to WELD, the STC has to work without the thermal bands available on Landsat 8, and is adapted to the spectral characteristics, as well as the Scene Classification available in the Sentinel 2 Level 2A product. STC is part of the S2GM processing chain. Section 2.4.2 provides a detailed description of this algorithm.

Medoid Composite [Flood, 2013]

The Medoid composite is part of the combined mosaicking algorithm to produce the composites in the S2GM service. The approach determines the medoid of a set of observations which can be considered as a representative value in a period. The algorithm is described in detail in Section 2.4.2 [Flood 2013].

Median Near Infrared (NIR) Composite [Potapov et al., 2011 & 2012]

A median NIR composite has been used by Potapov et al. (2011) for a regional-scale study of boreal forest cover and change mapping using Landsat data in Russia. The image data compositing process at the pixel-level uses a pre-filtered and quality assessed data pool that represents all selected image dates. Each selected image additionally includes the information about cloud, water, and shadow likelihood, image date, and different quality measures on a per-pixel basis. For each pixel from the filtered image dataset, the median NIR value was used to determine the finally selected date. A test of the median band value compositing approach on boreal Landsat data revealed that median NIR composites produce the least noisy outputs over forested areas.

Weighted pixel-based scoring system [Griffiths et al.,2013 and Franz et al., 2017]

Griffiths et al. (2013) presented a compositing algorithm using a weighted pixel-based scoring system to create cloud free, seasonally and radiometrically consistent datasets. The parametric weighting scheme builds on score functions considering acquisition year, day of year, and distance to clouds to select the best suitable observation in the time interval [Franz et al. (2017)].

Cosine similarity [Nelson & Steinwand, 2015]

A cosine similarity approach has been applied by Nelson and Steinwand (2015) to produce bi-annual images optimized for change detection. On pixel-level the approach selects the observation in the time interval with greatest similarity to others in case more than three or more observations are available. If less than three observations are available, a rule-based selection is applied. The cosine similarity is defined as the difference between 1 and the cosine of the angle between the two vectors containing the surface reflectance values from the selected observation dates. The image data preparation for the compositing process includes the analysis of the image quality, e.g., cloud coverage. A comparison of this approach with the WELD methodology regarding the change detection potential has indicated that the Cosine similarity approach is better suited to detect and classify landscape change [Nelson & Steinwand, 2015].

Sen2Three [ESA, 2017A]

Sen2Three has been developed for the spatio-temporal synthesis of bottom-of-atmosphere corrected Sentinel-2 L2A images, as they are generated by the Sen2Cor atmospheric correction [ESA, 2018]. The idea is to generate a synthetic output image from a time-series of L2A images by replacing step by step all contaminated or erroneous pixels of the L2A image that is regarded as best. Pixels used for replacement are taken from other scenes of the time series. The following three criteria can

be applied to rank the scenes in the time interval: (i) newer acquisition date is ranked higher, (ii) ranking according to the sum of good pixels (greater is better), (iii) ranking according to average aerosol optical thickness (smaller is better) or the average of the solar zenith angle (higher is better). Additionally, Sen3Three offers the possibility to calculate the output scene as an average of the good pixels of all scenes in the time series.

Table 1-1: Comparison of the different pixel-based image compositing approaches regarding different criteria

Criterion / **Method**	**Best pixel approach**	**Applicability to S2 L2A**	**Universality (specific tuning in location and time)**	**Consistency proof results available**
WELD (Roy et al., 2011)	Yes, decision tree, subsequent computation, TOA-Reflectance	No; requires brightness temperature	universal	Yes; see Roy et al. (2011)
STC	Yes, decision tree, subsequent computation	Yes	universal	in progress
Medoid Composite (Flood et al., 2013)	Yes, multidimensional median, minimum sum of distances	Yes	universal	Yes; see Flood (2013)
Median NIR Composite (Potapov et al., 2011)	Yes, per pixel QA assessment to identify data pool for best pixel selection, Median of NIR band	Yes	universal	Yes; see Potapov (2011)
Parametric weighting scheme based on score functions and adaptations like phenology adaptive best pixel scoring (Griffiths et al. 2013, Franz et al. 2017)	Yes, Weighted pixel-based scoring system based on acquisition year, acquisition day of year, and distance of a given pixel to cloud	Yes	requires background info for best day of year, weights optimized to local seasonality, requires weights	Yes; see Griffiths et al. (2013) and Franz et al. (2017)
Cosine similarity (Nelson & Steinwand 2015)	Yes; selects observation with greatest similarity to others; rule-based selection if less than 3 observations	Yes	requires desired DOY (day of the year) definition	Yes; see Nelson & Steinwand paper

Criterion / Method	Best pixel approach	Applicability to S2 L2A	Universality (specific tuning in location and time)	Consistency proof results available
Sen2Three (ESA, 2017A)	Yes/No depends on the selection of the method; selects observation or aggregated observations based on the scene statistics, scene classification or other criteria	Yes	universal	Yes; see Sen2Three documentation

Criterion / Method	Sensitivity to non-clear sky misclassification	Common artefacts	Robustness against (other) contingencies	Visual quality
WELD (Roy et al., 2011)	not much affected in case of enough clear observations	averaged clouds in case of few observations	not very robust if spectrum passes QC	risk of salt-pepper @ daily
STC	not much affected in case of enough clear observations	none	not very robust if spectrum passes QC	risk of salt-pepper @ daily
Medoid Composite (Flood et al., 2013)	not much affected in case of enough clear observations	none	robust due to median criteria if sufficient clear sky observations exist	risk of salt-pepper @ monthly
Median NIR Composite (Potapov et al., 2011)	not much affected in case of enough clear observations	none	robust due to median criteria if sufficient clear sky observations exist	risk of salt-pepper @ monthly
Parametric weighting scheme based on score functions and adaptations like phenology adaptive best pixel scoring (Griffiths et al. 2013,	sensitive	cloud/shadow mask ghosts	not very robust because 2 criteria do not use the measurement but just the date	risk of salt-pepper @ seasonal

Franz et al. 2017)				
Cosine similarity (Nelson & Steinwand 2015)	not much affected in case of enough clear observations	cloud/shadow mask ghosts	robust if the contingency is an exceptional case	risk of salt-pepper @ seasonal
Sen2Three (ESA, 2017A)	sensitive	cloud/shadow mask ghosts	not very robust if spectrum passes QC	risk of salt-pepper decreased

Criterion / Method	**Time compositing applicability**	**Minimum time series**	**Thematic up-sampling daily->10day-> … -> yearly**	**Methodological uncertainties**
WELD (Roy et al., 2011)	yearly, seasonal, monthly, 10day, daily	2 observations	Yes	Sensitivity regarding the geolocation uncertainty, most of the implementations of the method do not take into account the uncertainty of the surface reflectance and of the other input parameter
STC	10day	2 observations	Yes	
Medoid Composite (Flood et al., 2013)	yearly, seasonal, monthly	3 observations	No; median calculation requires full time series at each compositing period	
Median NIR Composite (Potapov et al., 2011)	yearly, seasonal, monthly	3 observations	No; Medoid calculation requires full time series at each compositing period	

Parametric weighting scheme based on score functions and adaptations like phenology adaptive best pixel scoring (Griffiths et al. 2013, Franz et al. 2017)	yearly, seasonal	1 observation	Yes (requires adaptation)	
Cosine similarity (Nelson & Steinwand 2015)	yearly, seasonal	1 observation	Yes	
Sen2Three (ESA, 2017A)	yearly, seasonal, monthly, 10day, daily	1 observation	Yes	

Criterion / **Method**	**Neighbourhood requirements**	**Auxiliary data requirements**	**Computational costs**	**Storage requirements**
WELD (Roy et al., 2011)	none	none	low; compositing criteria applied to 2 spectra at a time acquisitions of a pixel	low; only best pixel product persistent
STC	none	temporal cloud filter threshold	low; N criteria applied to 2 spectra at a time	low; only best pixel product persistent
Medoid Composite (Flood et al., 2013)	none	none	middle; time series	middle; L2A access regarding composite interval
Median NIR Composite (Potapov et al., 2011)	none	none	middle; time series	middle-high; L2A access regarding composite interval

Parametric weighting scheme based on score functions and adaptations like phenology adaptive best pixel scoring (Griffiths et al. 2013, Franz et al. 2017)	yes; calculation of distance to next cloud	L2A from previous years; best DOY per pixel per compositing interval or for interval from phenometrics	high; distance to cloud, potentially long time series; phenometrics	high; L2A multiple years access, phenometrics;
Cosine similarity (Nelson & Steinwand 2015)	none	definition of targeted DOY	middle; time series	middle; L2A access regarding composite interval
Sen2Three (ESA, 2017A)	yes - scene analysis for best pixel selection	none	low; compositing criteria applied to 2 spectra at a time acquisitions of a pixel	low; only best pixel product persistent

1.2.2 Spatial resampling

The S2GM service has the objective to deliver Sentinel-2 surface reflectance mosaics at global scale at 10m, 20m and 60m of spatial resolution and over arbitrary Areas of Interest (AOI). This requires spatial resampling since the bands are measured different spatial resolutions. Additional information such as illumination and viewing geometry as well as uncertainties are included on user request. Also, this information needs to be properly treated when resampled. In the following sub-chapters, we discuss the issues to be taken into account and pros and cons of different approaches. It should be remembered here that S2GM will deliver the selected best pixel (spectrum) which poses constraints on the final selection of the resampling method.

Terminology:

Up-sampling is used when measurements with a larger spatial resolution (e.g. S2 band 1 with 60m) are resampled onto a grid with higher spatial resolution grid (e.g. to a grid at 10m resolution).

Down-sampling is used when measurements with a higher spatial resolution (e.g. S2 band 2 with 10m) are resampled onto a grid with lower spatial resolution (e.g. to a grid with 60m resolution).

1.2.2.1 Up-sampling methods

Resampling entitles the process of determination and interpolation of pixels in the source product for computation of the pixel values in the target product. We evaluate three different resampling methods for the up-sampling here.

Nearest neighbour

Every pixel value in the output product is set to the nearest input pixel value.

Table 1-2: Advantages and disadvantages of the nearest neighbour.

Pros	Cons
Very simple, fast	Some pixels get lost and others are duplicated
No new values are calculated by interpolation	Loss of sharpness
Fast, compared to Cubic Convolution resampling	

Figure 1-2 illustrates the calculation of the new pixel value.

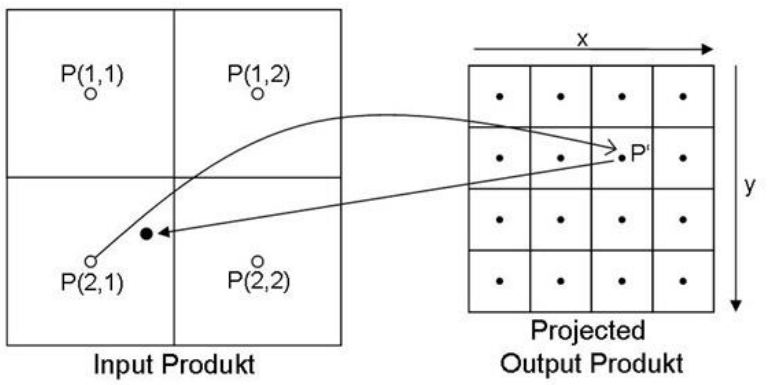


Figure 1-2: Nearest Neighbour

Each output cell value in the nearest neighbour method is the unmodified value from the closest input cell. Less computation is involved than in other methods, thus, large input raster files are processed faster than with other methods. Preserving the original cell values can also be an advantage if the resampled raster will be used for quantitative analyses, such as automatic classification. However, nearest neighbour resampling can cause feature edges to be offset by distances up to half of the input cell size. If the raster is resampled to a different cell size, a blocky appearance can result from the duplication (smaller output cell size) or dropping (large cell size) of input cell values.

Bi-linear interpolation

The new pixel value is calculated from the weighted average of the four surrounding pixels.

Table 1-3: Advantages and disadvantages of the bilinear interpolation.

Pros	Cons
Extrema are balanced	Less contrast compared to Nearest Neighbour
Image losses sharpness compared to Nearest Neighbour	New values are calculated which are not present in the input product

Figure 1-3 illustrates the calculation of the new pixel value.

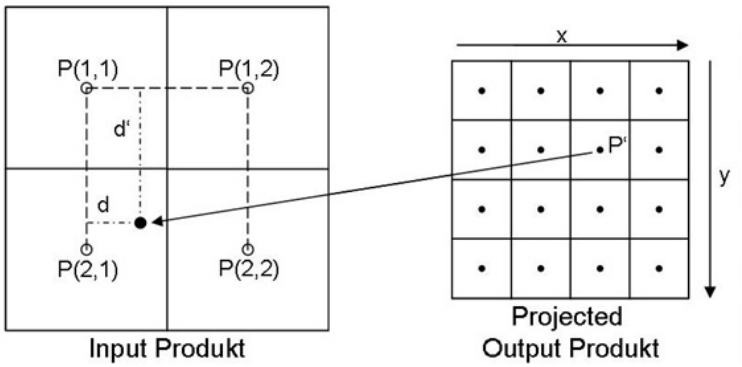


Figure 1-3: Bilinear Interpolation.

The bilinear interpolation is performed by the following equation:

$$P'(x,y) = P(1,1)(1-d)(1-d') + P(1,2)(d)(1-d') + P(2,1)(d')(1-d) + P(2,2)d\,d'$$

An output cell value in the bilinear interpolation method is the weighted average of the four closest input cell values, with weighting factors determined by the linear distance between output and input cells. This method produces a smoother appearance than the nearest neighbour approach, but it can diminish the contrast and sharpness of feature edges.

Cubic convolution

Calculation of the new pixel value is performed by weighting the 16 surrounding pixels.

Table 1-4: Advantages and disadvantages of the cubic convolution.

Pros	**Cons**
Extremes are balanced	Less contrast compared to Nearest Neighbour
Image is sharper compared to bi-linear Interpolation	New values are calculated which are not present in the input product
	Computational demanding, compared to Nearest Neighbour resampling

Figure 1-4 illustrates the calculation of the new pixel value.

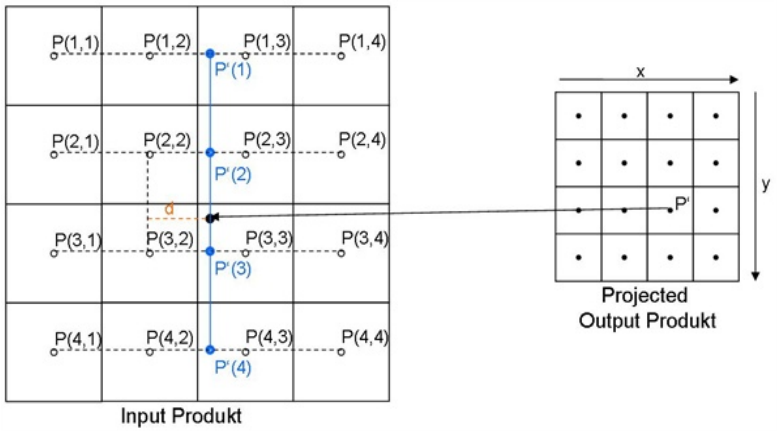


Figure 1-4: Cubic Convolution.

The cubic interpolation is performed by the following equation:

$$P'(k) = P(k,1)[4 - 8(1+d) + 5(1+d)^2 - (1+d)^3] + P(k,2)[1 - 2d^2 + d^3] + P(k,3)[1 - 2(1-d)^2 + (1-d)^3] + P(k,4)[4 - 8(2-d) + 5(2-d)^2 - (2-d)^3]$$

In the first step, the average value for each line is calculated, afterwards the new pixel value is calculated with the four, new average values P'(1) - P'(4) similar to the preceding calculation.

The cubic convolution method calculates an output cell value from a 4 X 4 block of surrounding input cells. The output value is a distance-weighted average, but the weight values vary as a nonlinear function of distance. This method produces sharper images than bilinear interpolation, but it is the most computationally intensive resampling method.

Summary

Figure 1-5 shows the visual comparison between the three algorithms under consideration. The nearest neighbour can have a blocky appearance, whereas the other approaches produce smoothed images, with reduced contrast and sharpness of feature edges in most but not all cases.

Nearest Neighbour | Bi-linear Interpolation | Cubic Convolution

Figure 1-5: Visual comparison - up-sampling

1.2.2.2 Down-sampling methods

As explained before, resampling entitles the process of determination and interpolation of pixels in the source product for computation of the pixel values in the target product. Three different resampling methods for the down-sampling proposed by Bian & Butler (1999) are considered:

- Averaging aggregation method
- Central pixel aggregation method
- Median aggregation method

Averaging aggregation method

Calculation of the new pixel value is performed by the mean computation of the input pixel values. Figure 1-6 illustrates the calculation of the new pixel value.

Table 1-5: Advantages and disadvantages of the averaging aggregation.

Pros	Cons
Very simple, fast	New values are calculated which are not present in the input product
	Loss of sharpness

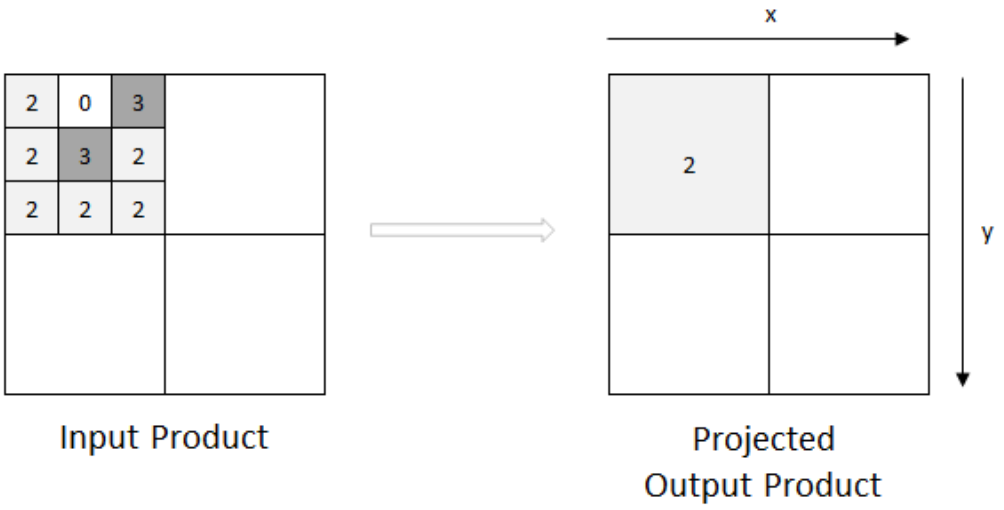


Figure 1-6: Aggregation method.

An output cell value in the averaging aggregation method is the average value from all corresponding input cells regarding the output cell. Thus, new values are calculated which are not present in the input product. This type of data aggregation, however, resembles the physical behaviour of sensors capturing data with different spatial resolutions.

Central pixel aggregation method

Every pixel value in the output product is set to the value of the central input pixel value.

Table 1-6: Advantages and disadvantages of the central pixel aggregation

Pros	Cons
Very simple, fast	Loss of sharpness
No new values are calculated	Strong dependency of the results from aggregation window

Figure 1-7 illustrates the calculation of the new pixel value.

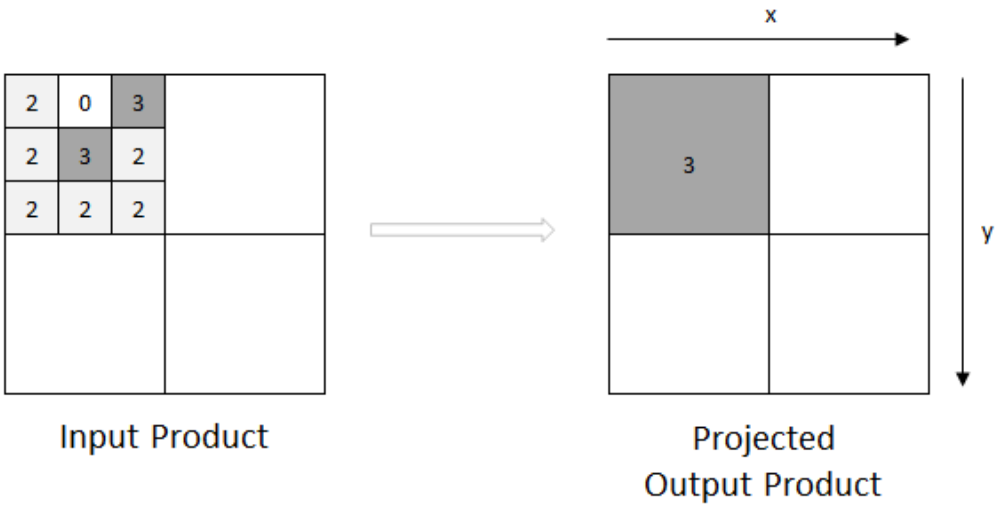


Figure 1-7: Central pixel aggregation.

Each output cell value in the central pixel aggregation method is the unmodified value from the central input cell. Less computation is involved than in other methods, leading to a speed advantage for large input raster files. But the resultant image is highly dependent on the resampling output cell size.

Median aggregation method

Calculation of the new pixel value is performed by the median computation of the input pixel values.

Table 1-7: Advantages and disadvantages of the median aggregation.

Pros	Cons
No new values are calculated	Synthetic resultant spectra
	Slow, compared to both other resampling

Figure 1-8 illustrates the calculation of the new pixel value.

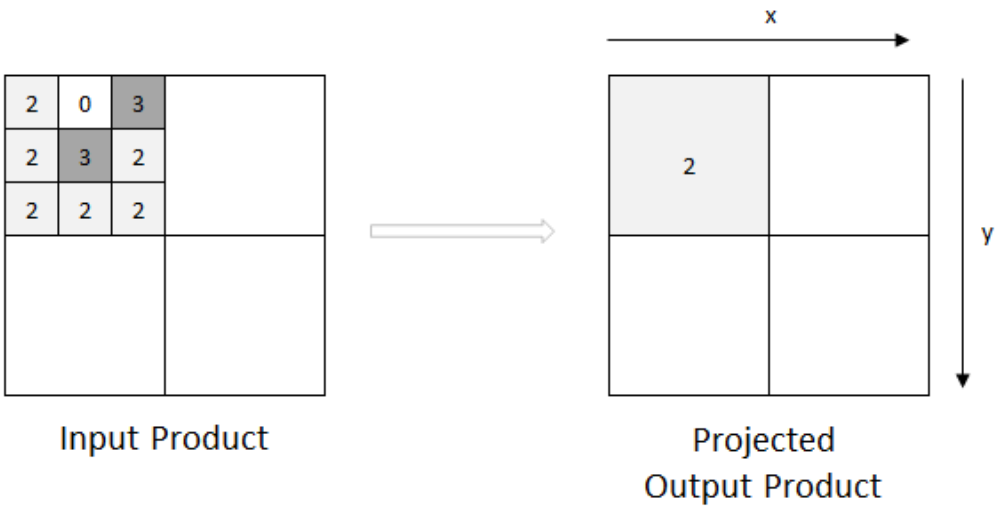


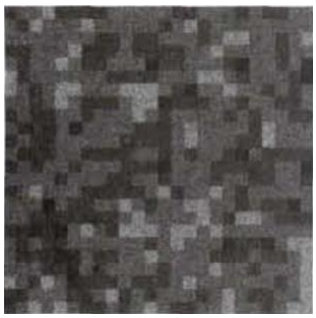
Figure 1-8: Median aggregation.

Each output cell value in the median aggregation method is the unmodified median value from all corresponding input cells regarding the output cell. More computations are involved than in both

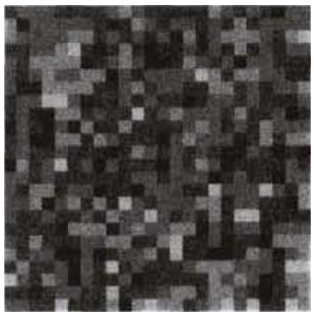
other methods. Furthermore, the resultant spectrum is composed from the different median values and is not a selected observation. The adaptation to the Medoid aggregation would solve this issue.

Summary

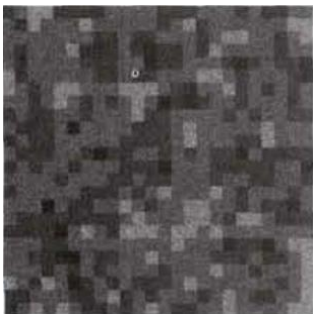
Figure 1-9 shows the visual comparison between the three algorithms under consideration. All three methods produce smoothed images, with reduced contrast and sharpness of feature edges in most but not all cases. Some areas reveal sharp contrasts.



Averaging aggregation



Central pixel aggregation



Median aggregation

Figure 1-9: Visual comparison - down-sampling (images taken from Bian & Butler, 1999].

1.2.3 Algorithm Selection

The alternative temporal mosaicking algorithms are introduced in chapter 1.2.1 and the advantages and disadvantages under different criteria relevant for the S2GM service are listed in Table 1-1. It shall be recalled here that the different criteria include quality, documentation, applicability, robustness and implementation aspects.

The decision matrix shows that the MEDOID algorithm is the most favourable algorithm. However, it requires a minimum number of good observations before it can be applied successfully. In actual value of the minimum number depends of the quality of the filtering of the data, i.e. removal of unsuitable spectral (clouds, haze, …). In case of S2GM the processing currently starts from sen2cor processed Level 2A products and the Scene Classification Layer of sen2cor is the main input to the filtering. Experimentally we found that 4 or more spectra are required to run the MEDOID algorithm.

In those cases where 3 or less spectra are retained in the compositing period an alternative method has to be chosen. According to the decision table and motivated by the argument to be comparable with the USGS Weld, the STC-S2 approach has been selected.

Both, the MEDOID and the STC-S2 will be described in detail in the following sections. The finally selected resampling algorithms are also described in the following sections.

2 Algorithm description

2.1 General overview

This section describes in detail the compositing and mosaicking approach, which is the basis of the main module of the processing chain. This approach has been adopted to generate the Mosaics V1.1.0. The main steps of the processing are:

1. Selection of the region of interest or the whole globe (typically a user option)
2. Selection of the spatial resolution and the aggregation period of the composite (typically a user option)
3. The processing system identifies all observations in space and time which lie within the region and aggregation period definitions of the user. These are then submitted to the composite algorithm, which also includes the pre-processing.
4. The compositing algorithm selects the best pixel observation in time, depending on statistical characteristics of the spectral measurements and other corresponding information like geometry; pixel identification is included in the composite regarding the chosen spatial resolution
5. Mosaicking in the region of interest

The main module performs the following operations (Figure 2-1): pre-processing and compositing/mosaicking. There are two distinct methods used for selection of the best pixel, namely the MEDOID and the Short-Term Composite (STC). They are briefly introduced below, and all steps are described in full detail in Section 2.4

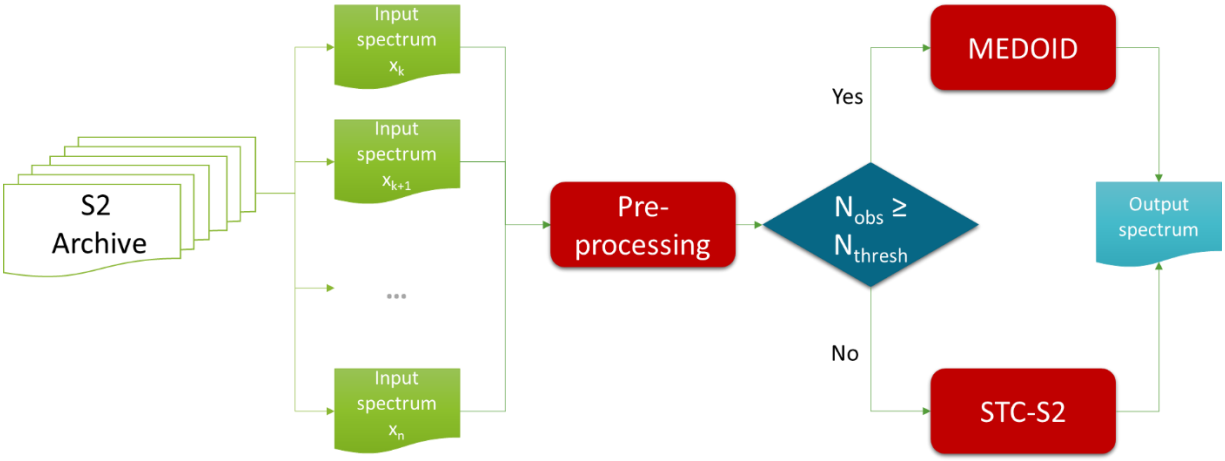


Figure 2-1: Compositing/mosaicking approach

ALGORITHM CONFIGURATION AND MOSAIC ASSESSMENT

The algorithm used here for the generation of Sentinel-2 mosaics selects the original pixel from the Sentinel-2 L2A product that is supposed to be most representative for the mosaicking period. The only manipulation of actual measured values may come from the resampling required to arrive at a common spatial resolution of the output format. Two different methods, STC and the Medoid, are applied to make the selection for an individual pixel for few or numerous observations, respectively.

While both methods are sufficiently simple and robust for automated large-scale application, the quality of the resulting mosaics are sensitive to errors in Sen2Cor's scene classification and to an optional pre-filtering of input products based on the scene classification. Finally, the threshold of

valid observations per pixel to decide which of the algorithms to be used influences the resulting mosaic. The consortium thoroughly and systematically assessed the effects of these issues at different test sites and for different compositing periods.

For the Mosaics V1.1.0, a threshold of four valid observations has been set for the usage of the Medoid. If there are less than four observations, the STC algorithm has been applied. By this means, a sufficiently large sample size for the Medoid to work properly is ensured. It is worth noting that the ability of the method to eliminate outliers such as erroneously classified surfaces including clouds and cloud borders is proportional to the number of observations. On the contrary, the decision-tree approach of the STC produces useful results even for very few observations and, thus, maximises the number of pixels in the mosaics for the user. Although relying on a mixture of the radiometric data and the scene classification, it is, however, critically depending on the quality of the scene classification. Consequently, the resulting mosaics may contain pixels selected by either of the two methods, which select for different criteria. Depending on the specific application, this may not be desirable, and instead, a mosaic generated by only one method may be the better choice.

2.2 Input Data- Sentinel-2 L2A processed with Sen2Cor

The Sentinel-2 L2A products produced with Sen2Cor and delivered by ESA via the two Copernicus hubs serve as input for the mosaic service. The Sen2Cor processor [ESA 2018], which generates the L2A products has been analysed in the Atmospheric Correction Inter-Comparison Exercise (ACIX) exercise and was found to provide reasonable results regarding the aerosol optical depth, water vapour and surface reflectance values. The scene classification was not part of the assessment in the ACIX exercise [Doxani et al., 2018].

The processing methodology for the mosaicking algorithm relies on quality of the input data in terms of surface reflectance values, on the absence of artefacts and on a correct pixel classification, in particular for clouds and cloud shadows. Despite the high accuracy and quality of the L2A images reported in the sen2cor documentation, we encountered several issues concerning remaining haze, cloud omission errors and commission errors for bright surfaces. In particular, urban areas are systematically flagged with the low probability flag. This has a significant influence on the quality of the mosaic, see also section 2.3.

The Sentinel-2 L2A products generated by Sen2Cor include different bands at different spatial resolutions. Table 2-1 shows the different Sen2Cor bands related to the provided spatial resolution.

Table 2-1: Used Sen2Cor bands regarding the spatial resolution

R10m	R20m	R60m
		B01_60m
B02_10m	B02_20m	B02_60m
B03_10m	B03_20m	B03_60m
B04_10m	B04_20m	B04_60m
	B05_20m	B05_60m
	B06_20m	B06_60m
	B07_20m	B07_60m
B08_10m		
	B8A_20m	B8A_60m
	B11_20m	B11_60m
	B12_20m	B12_60m

R10m	R20m	R60m
	SCL_20m	SCL_60m
AOT_10m	AOT_20m	AOT_60m

2.2.1 Harmonization of Sentinel-2A and Sentinel-2B

With both S2 satellites in orbit, ESA routinely provides high spatial resolution optical images covering the entire globe. More explicitly, high spatial resolution data (10-60 m) in combination with short revisiting times (in the order of days for most locations on Earth with S2A and S2B being both operational) is realised.

Such synergistic use of the S2A and S2B data results, however, in large data volume to be processed. The differences in the spectral response functions (SRF) between these two sensors can be seen in Figure 2-2. These different spectral response functions [ESA, 2017] may require an adaptation of the mosaicking processor software to ensure that the two sensors are aligned in terms of radiometry in the final time series of the surface reflectance. As it stands, however, this adaptation has not been realised due to comparable small differences between the two sensors. We plan to implement them after more important issues (namely scene classification) have been resolved.

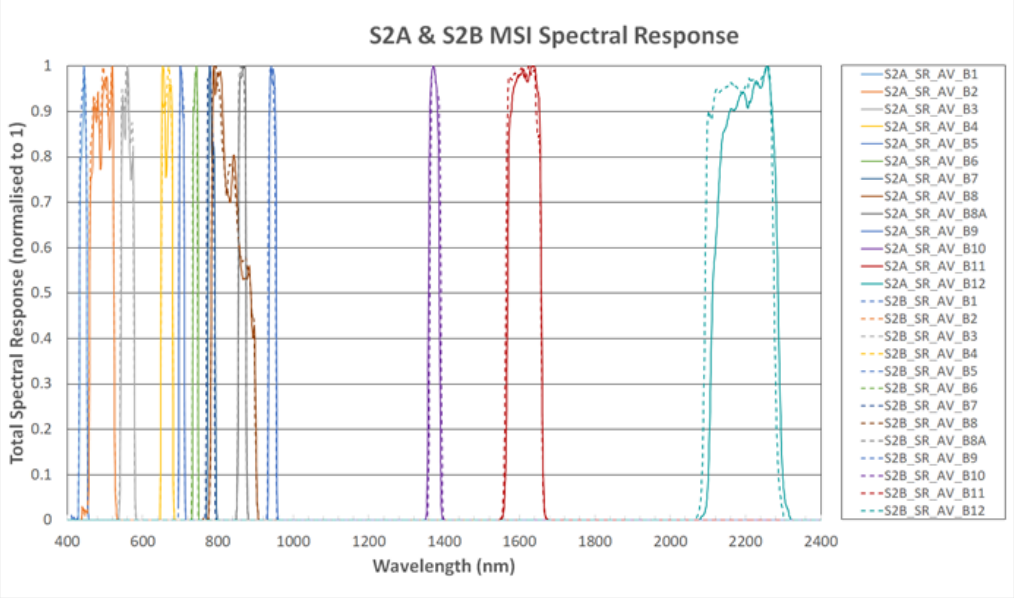


Figure 2-2: Sentinel-2A and Sentinel-2B spectral response functions (SRF)

2.2.1.1 Data and method

The approach to consider the harmonization factor is based on

- simulation of S2A and S2B observations using spectral response curves of both Sentinel-2 sensors
- linear regression between the simulation results for corresponding bands of the two sensors to derive band harmonization factors

The generation of simulated S2A and S2B observation data is based on the data of the USGS Spectral Library [Kokaly et al, 2017]. The USGS Spectral Library is a reference data base that covers the wavelength range from the ultraviolet to far infrared along with sample documentation has been assembled over many years. The library includes samples of minerals, rocks, soils, physically constructed as well as mathematically computed mixtures, plants, vegetation communities, microorganisms, and man-made materials. The collected in situ samples and spectra have been

assembled for the remote detection of these and similar materials by means of spectra and their features from space.

Materials contained in the Spectral Library

- Minerals
- Elements
- Soils, Rocks, Mixtures, and Coatings
- Liquids, Liquid Mixtures, Water, and Other Volatiles Including Frozen Volatiles
- Artificial (Man-Made) Materials Including Manufactured Chemicals
- Plants, Vegetation Communities, and Mixtures with Vegetation
- Micro Organisms

2.2.1.2 Harmonization Factors

The following harmonization factors (see Table 2-2) have to be applied to align the S2B surface reflectance ($sr_{S2B\_aligned}$) to the S2A surface reflectance in terms of radiometry in the final time series of the surface reflectance from S2A and S2B.

$$y = f(x) = a_1 x + a_0 \; with \; y = sr_{S2B\_aligned} \; and \; x = sr_{S2B}$$

Table 2-2: Harmonization factors for Sentinel-2B

S2 Band	a_1	a_0	rmse
B1	1.001E+00	1.696E-04	2.593E-05
B2	1.000E+00	8.478E-05	6.344E-06
B3	1.001E+00	1.640E-04	1.378E-04
B4	9.998E-01	3.546E-05	1.598E-05
B5	9.990E-01	8.812E-04	2.398E-04
B6	1.001E+00	1.523E-03	3.354E-03
B7	9.995E-01	8.810E-04	4.600E-04
B8	1.000E+00	-3.302E-05	2.072E-06
B8A	9.999E-01	1.126E-04	3.053E-05
B9	9.992E-01	-9.443E-05	6.193E-04
B10	1.003E+00	4.080E-03	1.080E-02
B11	9.994E-01	1.180E-03	3.145E-04
B12	9.861E-01	1.622E-03	8.759E-03

Because of the small deviations in the spectral responses between S2A and S2B, the Mosaics V1.1.0 have been produced without applying harmonization factors.

2.2.2 Geolocation of Sentinel-2A and Sentinel-2B

For the assessment of the geo-location accuracy of Sentinel-2A and Sentinel-2B image data, the GeoTool developed in the Land Cover CCI project phase I has been adapted and applied to the Sentinel-2 L1C. The input here is as well the collocated image of Sentinel-2A and Sentinel-2B. Sentinel-2A (B8) is here used as master and Sentinel-2B (B8) as slave.

2.2.2.1 Methodology of GeoTool

The GeoTool moves the slave image over the master image, i.e. in this case the Sentinel-2B image is being moved over the Sentinel-2A. The principle of GeoTool is to estimate shifts in line and column between two images on each pixel. For each pixel, a similarity measurement between a 3x3 mask of the master image (or reference image) and a 3x3 mask of the slave image is computed. The slave is then successively shifted with respect to the master image. In total 9 directions of translation are examined (illustrated in Figure 2-3 and Figure 2-4).

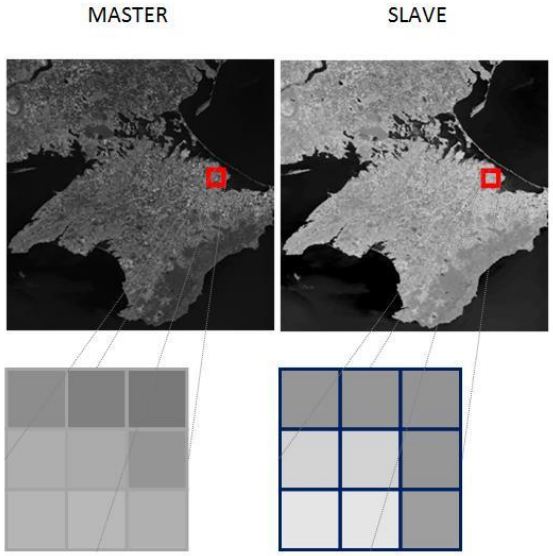


Figure 2-3: Principle of similarity measurements between two images

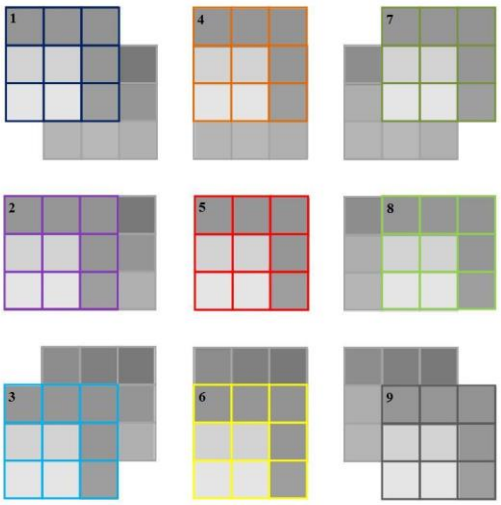


Figure 2-4: Scheme of applied GeoTool - 9 directions of translation (master: grey 3x3 window, slave: coloured 3x3 window)

Shifts in line and column are estimated in a local 3x3 window centred on each point by assessing the translation that maximizes the similarity feature between the master image and the slave image. The similarity measurement between two images is based on Pearson's correlation coefficient, which is a measure of the strength of the association between two quantitative, continuous variables. It is defined as the covariance of the two variables divided by the product of their standard deviations and is calculated as follows:

$$\rho_{X_M,X_S} = \frac{cov(X_M, X_S)}{\sigma_{X_M}\sigma_{X_S}} = \frac{E\left[\left(X_M - \mu_{X_M}\right)\left(X_S - \mu_{X_S}\right)\right]}{\sigma_{X_M}\sigma_{X_S}}$$

$$\rho_{X_M,X_S} = \frac{\sum_{i=1}^{n}\left(X_{M_i} - \mu_{X_M}\right)\left(X_{S_i} - \mu_{X_S}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_{M_i} - \mu_{X_M}\right)^2}\sqrt{\sum_{i=1}^{n}\left(X_{S_i} - \mu_{X_S}\right)^2}}$$

$$with$$

$$\mu_{X_M} = \sum_{i=1}^{n} X_{M_i} \quad and \quad \mu_{X_S} = \sum_{i=1}^{n} X_{S_i}$$

$$\sigma_{X_M} = \sqrt{\sum_{i=1}^{n}\left(X_{M_i} - \mu_{X_M}\right)^2} \quad and \quad \sigma_{X_S} = \sqrt{\sum_{i=1}^{n}\left(X_{S_i} - \mu_{X_S}\right)^2}$$

Additionally, the ‘Maximum of correlation coefficient drift’ is calculated for each position (Figure 2-5). The colours in the generated output image are corresponding to the colours in the scheme of Figure 2-6. Value “5” indicates no drift of the image data and hence a perfect match of master and slave. Based on the ‘Maximum of correlation coefficient drift’ image, the corresponding histogram is produced and reveals the frequency of the individual window positions. To make the different output data comparable, the colours used in the histogram are corresponding to the colours in the source image.

A thorough examination of the produced images and the respective histograms (Figure 2-7), both illustrating the maximum of the correlation coefficient drift, reveals that value “4” dominates the output data set. Thus, the Sentinel-2B image tends to shift towards northward by one pixel.

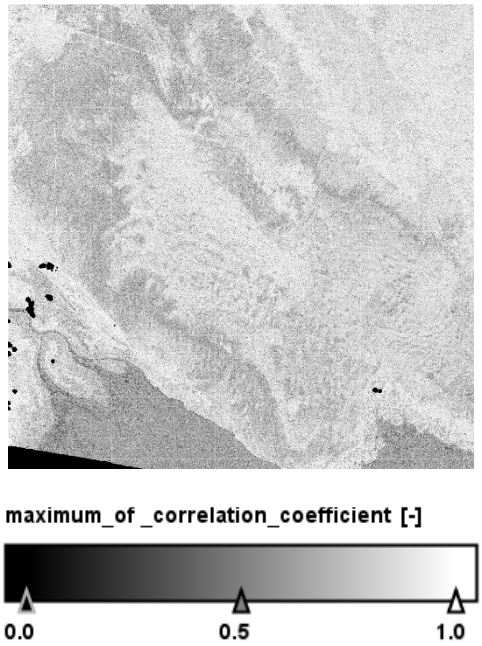


Figure 2-5: Maximum of correlation coefficient of a collocated tile generated using Sentinel-2A and Sentinel-2B data from August 6th and 11th 2017.

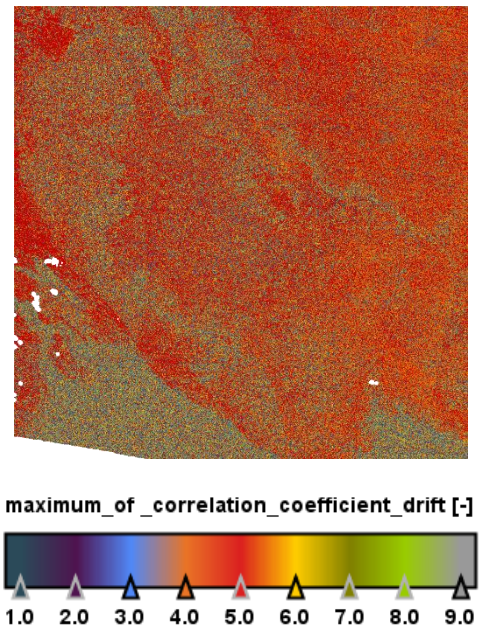


Figure 2-6: Maximum of correlation coefficient drift of a collocated tile generated using Sentinel-2A and Sentinel-2B data from August 6th and 11th 2017.

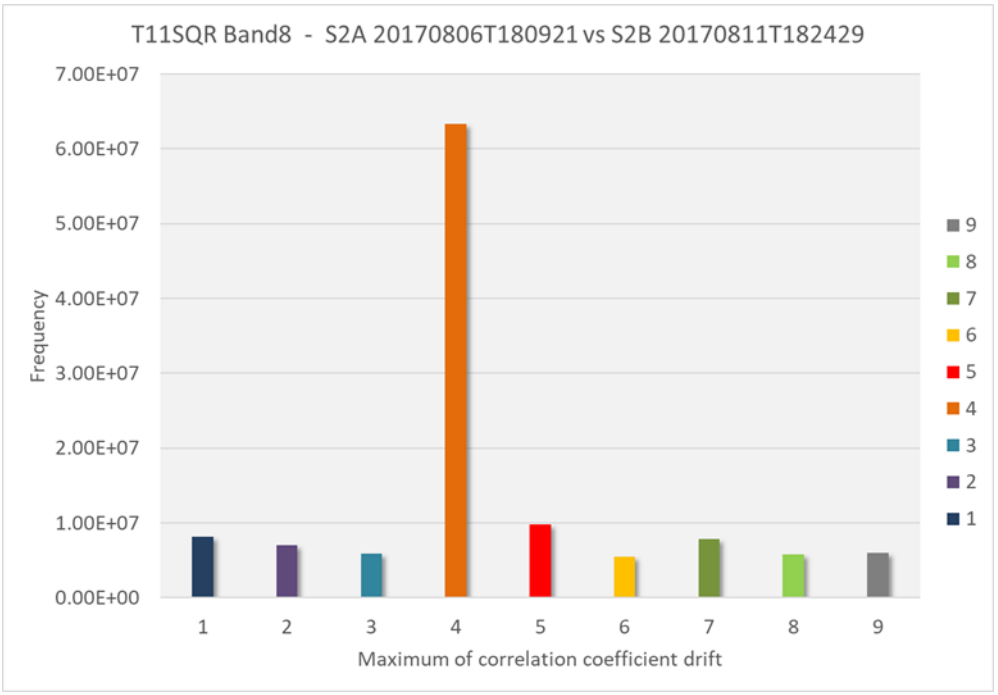


Figure 2-7: Histogram of the maximum of correlation coefficient drift of a collocated tile generated using Sentinel-2A and Sentinel-2B data from August 6th and 11th 2017.

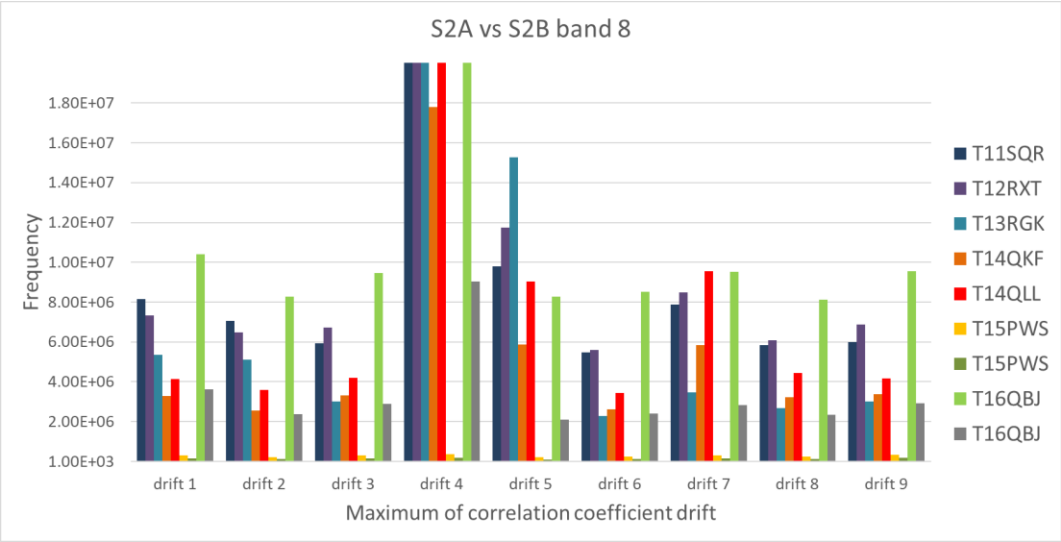


Figure 2-8: Histogram of the maximum of correlation coefficient drift of a collocated tiles generated using Sentinel-2A and Sentinel-2B data from August 2017.

The geolocation correction should be applied at L1C generation. It is known to ESA and we expect a fix in due time. It is not part of Mosaics processing chain.

2.3 Scene Classification analysis

2.3.1 General overview and findings

Sen2cor includes a scene classification (SCL) algorithm, which detects clouds, their shadows, and snow and generates a classification map. The SCL labels comprise 12 classes, which are compiled in Table 2-3. The algorithm is based on threshold tests on the top-of-atmosphere spectral reflectance, band ratios, and indices. For each of these threshold tests, a level of confidence is associated. Furthermore, the processing chain delivers products with probabilistic cloud and snow mask [Main-Korn et al., 2017].

Table 2-3: Sen2Cor scene classification specification

Label	Value	Description
NODATA	0	No data
SATURATED_DEFECTIVE	1	Saturated or defective
DARK_FEATURE_SHADOW	2	Dark feature shadow
CLOUD_SHADOW	3	Cloud shadow
VEGETATION	4	Vegetation
BARE_SOIL_DESERT	5	Bare soil/ desert
WATER	6	Water
CLOUD_LOW_PROBA	7	Cloud (low probability)
CLOUD_MEDIUM_PROBA	8	Cloud (medium probability)
CLOUD_HIGH_PROBA	9	Cloud (high probability)
THIN_CIRRUS	10	Thin cirrus
SNOW_ICE	11	Snow or Ice

Visual analysis of several randomly selected L2A data by experts has revealed that the S2 L2A scene classification contains numerous insufficiencies.

There is a substantial amount of confusions between pixel classes, in particular between the classes for cloud and urban. Part of these confusions are probably caused by the Sentinel-2 sensor design, which makes it difficult in some cases to distinguish between clouds and urban areas applying a pixel-wise classification on a single scene. Others may be rooted in the Sen2Cor's scene classification methodology itself. Besides these confusions, some errors seem to be systematic. Topographic shadow flagging for example appear to be based only on geometries and terrain without additional spectral tests. This leads to the flagging of bright snow on sun averted hill sides as dark surface even though there is no recognizable shadow present. Another systematic error is the flagging of shallow water areas as dark areas. These issues have direct and adverse effects on the results of both mosaicking methods applied here, the Medoid as well as the STC.

2.3.2 Influence of different flag combinations on the outputs

In case all cloud flags in the L2A product were correct, a mosaic generate by the STC method would have no clouds and thus no commissioning errors due to the cloud flag. Unfortunately, as we have explained in the previous section, the scene classification routinely contains errors. To illustrate the impact of erroneous flags on mosaic results, tests with different pre-filtering set-ups, i.e. with different combinations of flags, have been conducted. We have tested five different combinations of flags, from very strict to very tolerant flagging. The following configurations have been applied:

- Strongest : scl == VEGETATION || scl == BARE_SOIL_DESERT || scl == WATER || scl == SNOW_ICE;
- Weak_1 : scl == DARK_FEATURE_SHADOW || scl == VEGETATION || scl == BARE_SOIL_DESERT || scl == WATER || scl == SNOW_ICE;
- Weak_2 : scl == DARK_FEATURE_SHADOW || scl == VEGETATION || scl == BARE_SOIL_DESERT || scl == WATER || scl == SNOW_ICE || scl==CLOUD_LOW_PROBA;
- Weak_3 : scl != CLOUD_HIGH_PROBA && scl != SATURATED_DEFECTIVE && scl != NODATA && scl != CLOUD_MEDIUM_PROBA && scl != CLOUD_SHADOW;
- Weak_4 : scl != CLOUD_HIGH_PROBA && scl != SATURATED_DEFECTIVE && scl != NODATA;

2.3.3 Influence of flag combinations on Medoid

Figure 2-9 shows the effect of the different flag combinations for filtering at the example of an annual product over the Kapuvar test site. Weakening the flag conditions has different effects over different surface types:

- Agriculture areas: Bare surfaces are chosen over vegetation. Which surface type should be preferred should be judged by the user.
- Vegetation: Haze gets a stronger influence on the best pixel
- General: Higher salt & pepper effect

Very strict filtering leads to:

- Single pixels are missing even over one year. Shorter periods will lead to even more lost pixels.

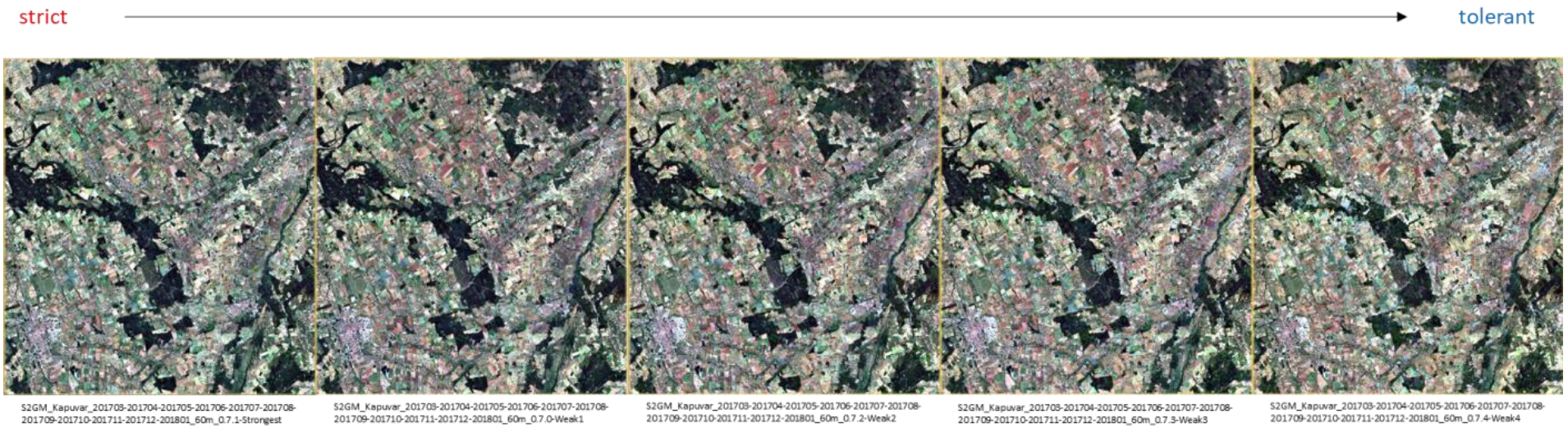


Figure 2-9: Annual mosaics over Kapuvar test site with different pre-filtering configurations, from strict to weak, definitions of configurations are explained in 2.3.2.

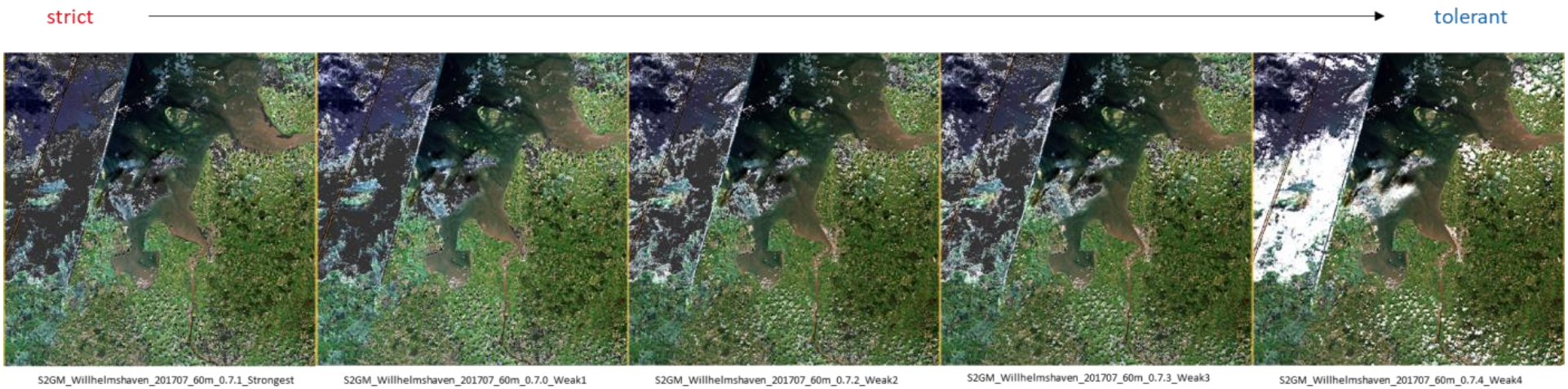


Figure 2-10: Annual mosaics over Wilhelmshaven test site with different pre-filtering configurations, from strict to weak, definitions of configurations are explained in 2.3.2.

2.3.4 Influence of flag combinations on STC

The influence of the different flag combinations on the STC algorithms is demonstrated at a monthly product over the Wilhelmshaven test site (Figure 2-10).

Weakening the flag conditions leads to:

- Increase in residual clouds and in the worst case, massive clouds remain in the product

Very strict filtering leads to:

- Immense loss in good pixels
- Loss of shallow water (caused by the L2A dark area flag)
- Loss of urban areas (caused by medium and low cloud probability flags)

2.3.5 Conclusion on pre-filtering based on the scene classification

The erroneous S2 L2A scene classification has a direct impact on the quality of the resulting mosaics. The configuration of the pre-filtering based on the L2A flags has different effects on the resulting mosaics depending on the compositing method (Medoid or STC) applied. In summary, a very strict pre-filtering leads to systematic removal of numerous valid observations while undetected cloud pixels are still present. In turn, a weak pre-filtering results in numerous occurrences of cloud spectra (Medoid) or even remaining clouds in the final mosaic products (STC). The second strongest filtering (called weak_1 in the examples above) has been found as best compromise.

2.4 Mosaicking

The mosaic processing is organized in two distinct modules (see Figure 2-11):

- the pre-processing module
- and a combined mosaicking module based on Medoid and Short-Term Composite approaches.

A threshold related to the number of valid observations is defined and applied in the processing chain for the selection of the mosaicking approach. In case of sufficient valid observations, the Medoid approach is selected otherwise the Short-Term composite is chosen. The threshold has been set to 4 for Mosaics V1.1.0.

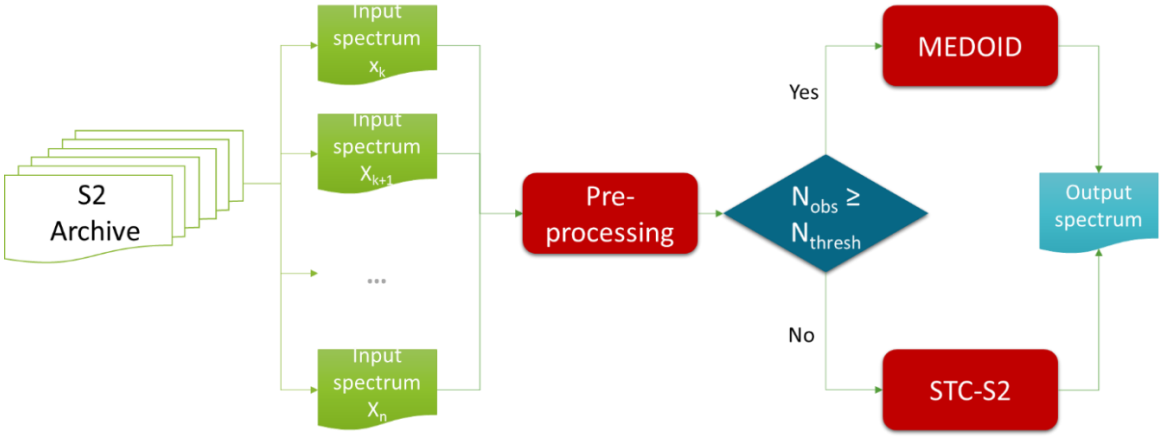


Figure 2-11: Scheme of the processing chain – mosaicking

2.4.1 Pre-processing - filtering

The pre-processing module analyses all input spectra and identified all pixel which should be used in the compositing/mosaicking approach., which are defined as valid observation. As explained in section 2.3.5, the unreliability of the S2 L2A scene classification requires this pre-processing of the Sen2Cor data and the second strongest filtering has been found as best compromise regarding the quality of the resulting mosaics. Furthermore, unflagged artefacts on the swath border in the Sen2Cor data have been also filtered out by the pre-processing to ensure the quality of the mosaics - see Figure 2-12. This filtering is done by using the view zenith, because the swath border can be identified through the view zenith angle. Additionally, all input spectral bands containing any Not-a-Number (NaN) or infinite value have been also identified and filtered out by the pre-processing regarding the quality of the mosaics.

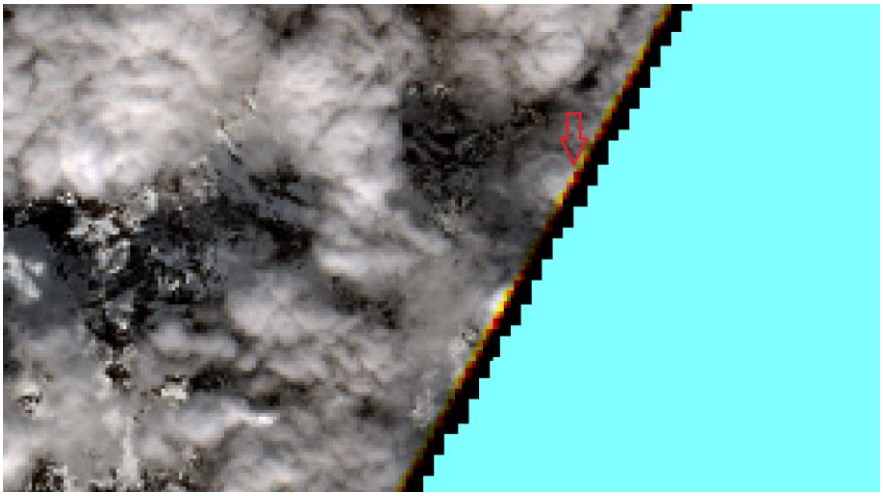


Figure 2-12: RGB with artefacts on swath border (arrow) in Sen2Cor product – blue area is flagged as invalid by scene classification (S2A_MSIL2A_20170526T105031_N0205_R051_T34WDA_20170526T105029)

The pre-processing of the input spectra has been applied to perform the mosaicking only for valid observations. The definition of a valid pixel is based on the spectra, the viewing geometry, and also of the Sen2Cor scene classification layer (SCL). In the Mosaics V1.1.0, an observation is classified as valid in the pre-processing, if the following conditions are fulfilled:

- The spectrum does not contain any Not-a-Number (NaN) or infinite value
- The mean view zenith value of all bands is less than 11.0°
- The pixel is classified as DARK_FEATURE_SHADOW or VEGETATION or BARE_SOIL_DESERT or WATER or SNOW_ICE (weak 1 filtering in section 2.3.2)

Furthermore, an L2A snow refinement test has to be applied caused by misclassified clouds as snow in the original S2 L2A scene classification. This adjustment of the snow detection has been necessary to prevent that cloudy pixels classified as snow have a negative effect on the composites. The following pseudo code (Figure 2-13) describes the applied the L2A snow refinement test and the Figure 2-14 presents the comparison of the original and adjusted snow mask for an Sentinel L2A product.

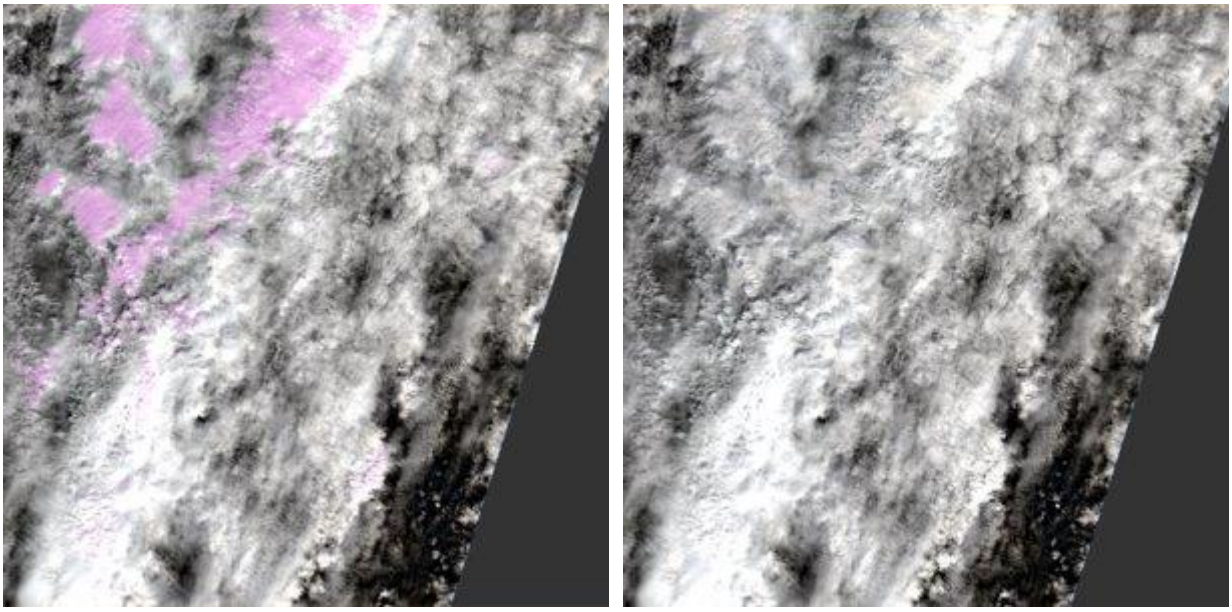
After pre-processing, all valid pixels are identified and used in following processing steps.

```
TCB = 0.3029·B02 + 0.2786·B03 + 0.4733·B04 + 0.5599·B8A + 0.508·B11 + 0.1872·B12
NDSI = (B03 - B11) / (B03 + B11)

function isSnow(sample) {
   let scl_snow = sample.SCL === 11;
   let tcb = computeTCB(sample); //

   let ndsi = computeNdsi(sample); //
   let s2gm_snow = ndsi > 0.6 && tcb > 0.36;
   return scl_snow && s2gm_snow;
}
```

Figure 2-13: Pseudo-code of the L2A snow refinement test



Original L2A snow mask

L2A snow mask cleaned by the L2A snow refinement test

Figure 2-14: Snow masks (pink) for the Sentinel 2 product over Ireland - 2018/07/23 - S2B_MSIL2A_20180723T115359_N0208_R023_T29UNV_20180723T192112

2.4.2 Temporal Resampling

The temporal sample aggregation is an integration over all valid input data in a given grid cell providing the best pixel value of surface directional reflectance (SDR) and SDR uncertainties, and additional bands derived from the pixel identification, the method itself, and input products and quality assessment. The temporal sample aggregation selects the best pixel from all valid measurements as well as the corresponding uncertainty and status flag for a pixel. The 10-day compositing period is adapted on a fixed basis: January, 1st is set as the starting point, the last composite of each year and the last February composite of a leap year comprise 11 instead of 10 days.

Considering the number of valid observations in a period, two different approaches have been used. In case of sufficient number of valid observations, the Medoid algorithm [Flood, 2013] has been applied to select the most representative pixel of the period. Otherwise, the Short-Term Composite approach (STC) has been selected, which is based on the LC-CCI WELD S2 algorithm and the WELD algorithm [Roy et al., 2011]. The latter has been developed in the course of the ESA LC-CCI project and is an adaptation of the WELD algorithm for Landsat [Roy et al., 2010 & 2011]. As described above, the selection of the mosaicking approach is based on the threshold with respect to the number of valid observations. Both approaches are described in detail in the following.

Short-Term Composite approach

The Short-Term-Composite (STC) approach is based on the LC-CCI WELD S2 algorithm and the WELD algorithm [Roy et al., 2011], which have been developed to reduce residual cloud and aerosol contamination in time series. WELD has been adapted for Sentinel-2 and the Sen2Cor L2A product in particular.

The STC approach identifies from a time series of observations the pixel that best satisfies compositing criteria. For the calculation of the surface reflectance composites value, a decision tree on the surface reflectance values, the cloud test and different indices and their mean values (brightness, mean value of bands, NDVI, mNDWI and TCB) is applied. Ideally, the criteria should select from the time series only near-nadir observations that have reduced cloud and atmospheric contamination. The different indices and the mean of the indices are defined as follows:

Modified normalized difference water index (identical to the definition of normalized difference snow index, NDSI) [Du et al., 2016 and Xu, 2006]

$$mNDWI = \frac{\rho_{B3} - \rho_{B11}}{\rho_{B3} + \rho_{B11}}$$

Normalized difference vegetation index

$$NDVI = \frac{\rho_{B8} - \rho_{B4}}{\rho_{B8} + \rho_{B4}}$$

Tasseled cap transformation brightness

$$TCB = 0.3029\,\rho_{B2} + 0.2786\,\rho_{B3} + 0.4733\,\rho_{B4} + 0.5599\,\rho_{B8A} + 0.508\,\rho_{B11} + 0.1872\,\rho_{B12}$$

Brightness

$$Brigthness = \rho_{B2} + \rho_{B3} + \rho_{B4}$$

Mean of B11 and B12 values

$$Mean\ of\ B11\ and\ B12 = \frac{\rho_{B11} + \rho_{B12}}{2}$$

Mean of index X (X = NDVI, mNDWI …)

$$\bar{X} = \sum_{i=1}^{N} \frac{X_i}{N} \; with\; N = number\; of\; valid\; observations$$

Decision logic and thresholds

Consequently, the STC algorithm is based on the selection of a "best" pixel over the compositing period. The decision tree is based on the comparison of all acquisitions of the same pixel. For each pixel in the composite, the product index is stored. This product index can be used to identify the product from which the selected pixel has been taken. Figure 2-15 presents the pseudo-code of the STC and

Table 2-4 summaries the STC compositing logic.

The STC compositing approach as well as the WELD approach [Roy et al., 2011] preferably select observations over vegetation from the entirety of cloud-free observations, because a considerable number of applications focus on vegetation. Both approaches are based on the assumption that clouds and aerosols typically decrease NDVI over land surfaces. In case of the STC approach, the WELD criterion regarding the brightness temperature has been compensated by the proposed indices [Roy et al., 2011]. From the WELD experience, it is known that the TOA NDVI of a cloud can be higher than the TOA NDVI of cloud free surfaces like surfaces with low vegetation coverage, including certain dark and bright soils, water and snow [Roy et al., 2011]. The sensitivity of NDVI regarding the brightness of the underlying soil in vegetation canopies [Huete 1988] and regarding atmospheric effects [Liu and Huete, 1995] is well understood and is investigated in various studies. These considered aspects have been taken into account for the comparison of pixels covered by vegetation and/or bare soil and/or snow/ice and/or water and/or dark feature surfaces, the corresponding compositing criteria are used to provide a differentiation between undetected clouds and the underlying surface.

In two of the decision criteria, additional refinement tests regarding the cloud and snow identification have to be applied caused by undetected clouds or snow covered surfaces. The adjustments have been necessary to prevent that undetected cloudy or snow covered pixels are selected as the preferred pixel in case of more than one observation. The additional refinement tests are based on thresholds regarding different spectral indices. The following pseudo code (Figure 2-15) describes the applied additional refinement tests.

```
private static boolean isCloudOrSnow(double[] spectrum) {
   double ratioB3B11 = B3 / B11;
   double ratioB11B3 = B11 / B3;
   double rgbMean = (B02 + B03 + B04) / 3;
   double tcHaze = -0.8239 * B02 + 0.0849 * B03 + 0.4396 * B04 - 0.058 * B8A + 0.2013 * B11 - 0.2773 * B12;
    double normDiffB8B11 = (B08 - B11) / (B08 + B011);

   boolean isSnow = (((B3 - B11) / (B3 + B11)) > 0.7) && !((ratioB3B11 > 1) &&
                        ((0.3029 * B2 + 0.2786 * B3 + 0.4733 * B4 + 0.5599 * B8 + 0.508 * B11 + 0.1872 * B12) < 0.36));
   if (isSnow) {
      return true;
   }
   boolean isHighProbCloud = ((((ratioB3B11 > 1) && (rgbMean > 0.3)) && ((tcHaze < -0.1) || (tcHaze > -0.08 &&
                              normDiffB8B11 < 0.4))) || (tcHaze < -0.2) || ((ratioB3B11 > 1) && (rgbMean < 0.3)) &&
                              ((tcHaze < -0.055) && (rgbMean > 0.12)) || (!((ratioB3B11 > 1) &&
                              (rgbMean < 0.3)) && ((tcHaze < -0.09) && (rgbMean > 0.12))));
   if (isHighProbCloud && !isSnow) {
      return true;
   }
   boolean isLowProbCloud = ((((ratioB11B3 > 1) && (rgbMean < 0.2)) && ((tcHaze < -0.1) || (tcHaze < -0.08
                              && (normDiffB8B11) < 0.4))) || (tcHaze < -0.2) || ((ratioB3B11 > 1) && ((rgbMean < 0.2)) &&
                              ((tcHaze < -0.055) && (rgbMean > 0.12)) || (!((ratioB3B11 > 1) && (rgbMean < 0.2)) &&
                             ((tcHaze < -0.02) && (rgbMean) > 0.12))));
   if (isLowProbCloud && !isSnow && !isHighProbCloud) {

      return true;
   }
   return false;
}
```

Figure 2-15: Pseudo-code of the additional refinement tests[1]

[1] Note: In the implementation code on GitLab the variable tcHaze is defined as tcBright

Table 2-4: STC compositing logic – V1.1.0[2]

Number of observations	Compositing Criteria
1	Select the valid pixel if it has passed the filtering Note the identification of a valid observation is done during the pre-processing step.
2-n	Select one or none of the both, three or n-valid pixels: Select the pixel 1. with the maximum NDVI if it has meanmNDWI < -0.55 and maxNDVI – meanNDVI < 0.05 else 2. with the maximum mNDWI if it has meanNDVI < - 0.3 and meanmNDWI – minNDVI < 0.05 else 3. with the maximum NDVI if it has meanNDVI > 0.6 and meanTCB < 0.45 else 4. with the minimum TCB if the cloudTest is false else 5. if the snowTest is false a. with the minimum TCB if minTCB < 1.0 b. none of the valid pixels if minTCB > 1.0 else 6. with the maximum mNDWI if the meanNDVI < -0.2 else 7. with the minimum NDVI if the meanTCB > 0.45 else 8. with the maximum NDVI

```
samples.forEach(function (element, i, samples) {
    cloudTest[i] = isCloudOrSnow(samples[i]);
    snowTest[i] = isSnow(samples[i]);
    tcb[i] = computeTCB(samples[i]);
    tcbSum = tcbSum + tcbi];
    mndwi[i] = computeMndwi(samples[i]);
    mndwiMean = mndwiSum + mndwi[i];
    ndvi[i] = computeNdvi(samples[i]);
    ndviSum = ndviSum + ndvi[i];
});
……
```

[2] Note: In the implementation code on GitLab the variable tcb is defined as tc4, and the variable TCB is defined as TC4

```
switch (n) {
    //one valid observation
    case 1:
      index = PRIMARY_IDX;
      break;
    //two and three valid observations
    case 2:
    case 3:
      if (mndwiMean < -0.55 && ndvi[ndviMaxIndex] - ndviMean < 0.05) {
        index = ndviMaxIndex;
      } else {
        if (ndviMean < -0.3 && mndwiMean - mndwi[mndwiMinIndex] < 0.05) {
          index = mndwiMaxIndex;
        } else {
          if (ndviMean > 0.6 && tcbMean < 0.45) {
            index = ndviMaxIndex;
          } else {
            if (!cloudTest[tcbMinIndex]) {
              index = tcbMinIndex;
            } else {
              if (!snowTest[tcbMinIndex]){
                if (tcb[tcbMinIndex] > 1.0) {
                  index = undefined;
                } else {
                  index = tcbMinIndex;
                }
              } else {
                if (ndviMean < -0.2) {
                  index = mndwiMaxIndex;
                } else {
                  if (tcbMean > 0.45){
                    index = ndviMinIndex;
                  } else {
                    index = ndviMaxIndex;
                  }
                }
              }
            }
          }
        }
      }
      break;
    default:
    throw Error("Number of observations in STC not handled: " + n)
  }
```

Figure 2-16: Pseudo-code of the STC[3]

[3] Note: In the implementation code on GitLab the variable tcb is defined as tc4, and the variable TCB is defined as TC4

	
previous version STC based on scene classification and indices	current version STC based on indices and applied L2A snow refinement test

Figure 2-17: Ireland – quarterly composite – July – September 2018

Medoid composite approach

For the calculation of the surface reflectance composites with sufficient valid pixels, the Medoid in reflectance space and time is applied [Flood, 2013]. The Medoid is the representative object of a data set whose average dissimilarity to all the objects in the data set is at a minimum [Struyf et al., 1996]. The Medoid is robust against extreme values in the data set. Flood (2013) has shown that the Medoid composite over a season is representative for that period and that the contamination by clouds and other features has been reduced using this technique.

The Medoid is defined as follows:

$$medoid(X) = \underset{x_i \in X}{\arg min} \sum_{x_j \in X} \|x_j - x_i\|$$

With the Euclidian distance

$$\|x_j - x_i\| = \sqrt{\sum_{b=1}^{n_{bands}} \left(x_{j,b} - x_{i,b}\right)^2}$$

With the normalized difference

$$\|x_j - x_i\| = \sum_{b=1}^{n_{bands}} \left|\frac{x_{j,b} - x_{i,b}}{x_{j,b} + x_{i,b}}\right|$$

Given a set of points in an n-dimensional space, one of those points can be defined as representative for the group. For a time-series of satellite images, the dimensions are the spectral bands of the image, and the observations are the dates. The following spectral bands have been used: B02, B03, B04, B06, B08, B11, B12. The proposed Medoid method selects a single date per pixel for the resulting composite. The reflectance values for that date are taken as the reflectance values for the entire period. The result is thus always a genuine observation at the location of the pixel, therefore preserving relationships between bands. The Medoid implicitly removes outliers, because it represents the observation that has the minimum distance to all considered observations. An explicit prior outlier removal by means of a strict pre-filtering based on the scene classification is applied to account for issues in the Sentinel-2 scene classification – see section 2.4.1.

2.4.3 Spatial resampling – up-sampling and down-sampling

In section 1.2.2 we discussed the necessity for spatial resampling and the impact of different methods. The following paragraph explain the solutions finally chosen. It should be recalled here that:

Up-sampling is used when measurements with a larger spatial resolution (e.g. S2 band 1 with 60m) are resampled onto a grid with higher spatial resolution grid (e.g. to a grid at 10m resolution).

Down-sampling is used when measurements with a higher spatial resolution (e.g. S2 band 2 with 10m) are resampled onto a grid with lower spatial resolution (e.g. to a grid with 60m resolution).

The S2GM service produces Sentinel-2 surface reflectance composites at global/regional scale at spatial resolutions of 10m, 20m and 60m, which include all bands except B9 and B10. The Sentinel L2A input products do not include all bands in all three spatial target resolutions; a spatial resampling is thus necessary prior to the production of composites. The following list summarizes the different approaches to spatial re-sampling and compositing in the different spatial resolutions:

- for the 10m composite (see Figure 2-18):
  - up-sampling to 10m via nearest neighbour method for B01_60m, B05_20m, B06_20m, B07_20m, B8A_20m, B11_20m, B12_20m and SCL_20m
  - selection of the best representative spectra based on all original and up-sampled bands in 10m
  - Consequently, all bands of lower spatial resolution, may exhibit spatial (artificial) variability below the spatial resolution of the detector, because several values from different observation times may be used to generate the spatial composite in the higher resolution.
- for the 20m composite (see Figure 2-19)
  - down-sampling via mean aggregation method required for B08_10m
  - up-sampling via nearest neighbour method required for B01_60m
  - selection of the best representative spectra based on all original and down- and up-sampled bands in 20m
- for the 60m composite (see Figure 2-20)
  - down-sampling via mean aggregation method required for B08_10m
  - selection of the best representative spectra based on all original and down-sampled bands in 60m

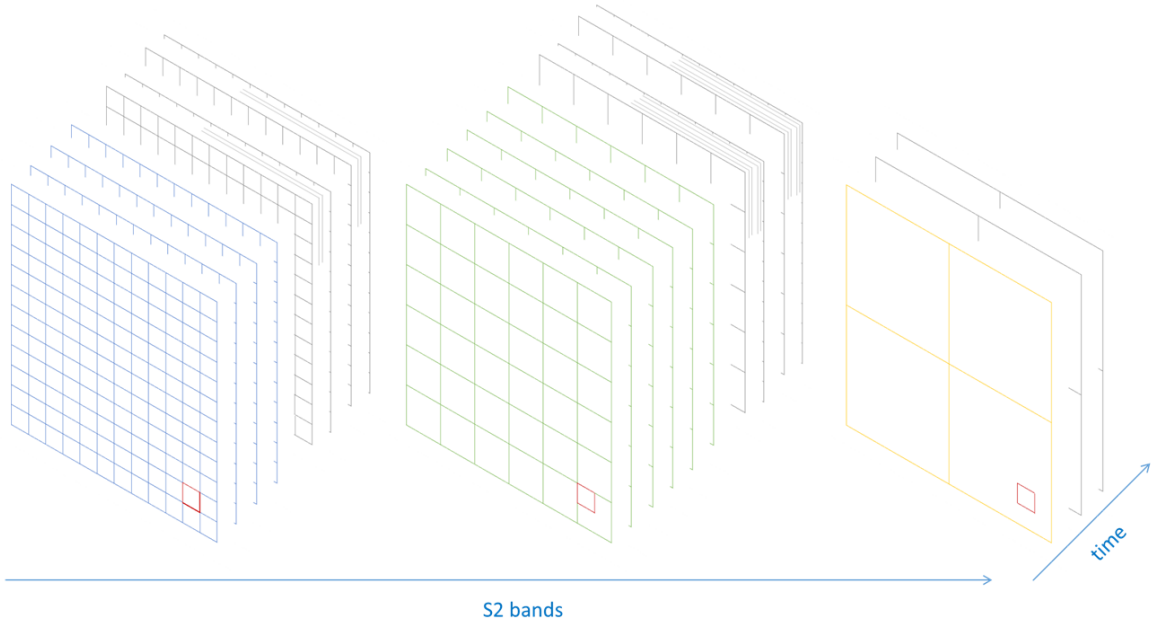


Figure 2-18: Spatial aggregation of the bands for the 10m composite

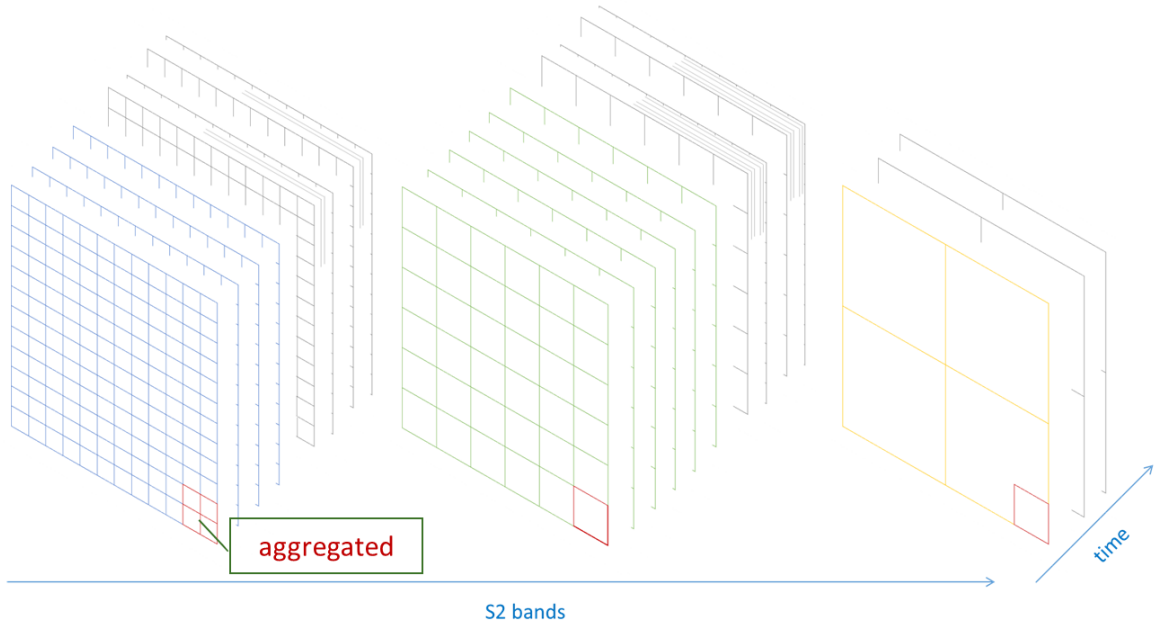


Figure 2-19: Spatial aggregation of the bands for the 20m composite

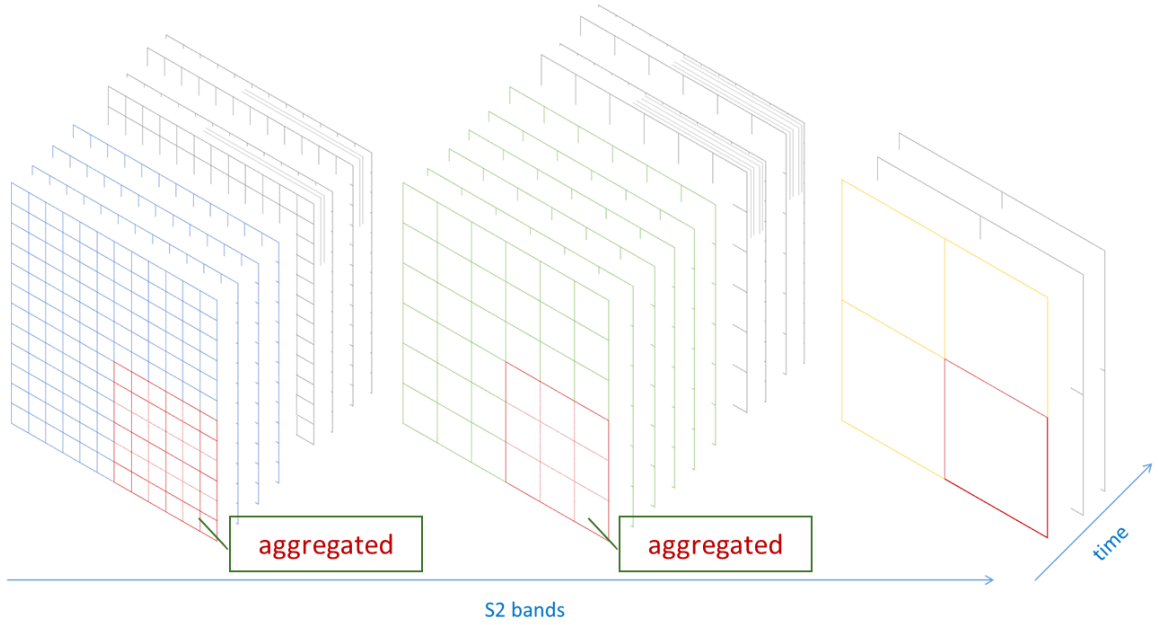


Figure 2-20: Spatial aggregation of the bands for the 60m composite

The method delivers the requested mosaic in the desired spatial resolution as a composite of genuine observations within the aggregation period, albeit at (potentially) different observation times for each pixel. As a consequence, a later spatial aggregation, in particular a down-sampling to lower resolution is not advisable, because of the different selected observation time in the spatial grid.

2.4.4 Spatial resampling – mosaicking

The user can select a region of interest and the corresponding Sentinel-2 granules have been processed with the compositing approach and mosaicked to the region of interest afterwards. The "filling" of the mosaic product follows the principle of fetching pixels from source products instead of putting all source product pixels into the target bins. The mosaic processor loops over all cells in the target grid and determines a pixel from a source product may be suitable to read into it. This is done by looking-up the nearest neighbour pixel in the input product that contains the geographical centre coordinate of the current output pixel being handled. In case two pixels have the same coordinates; the pixel with the smallest AOD is selected.

For the Mosaics V1.1.0, the standard SNAP mosaicking approach is applied. The SNAP Mosaic Operator combines overlapping products into a single composite product. The mosaicking is achieved based on the geocoding of the source products therefore the geocoding needs to be very accurate.

3 Product demonstration and initial assessment

The following chapter is a brief demonstration of the mosaic products generated and a discussion of results and limitations.

Figure 3-1 is one example of a 3-monthly composite over the Sevilla area, Spain. Between 2 and 11 valid observation were retained after filtering and used for the temporal compositing (Figure 3-2); i.e., for most pixels the medoid method could be applied. This is a typical example for a product under good conditions.

Figure 3-3 shows an example from northern Europe, Aberdeen area in Scotland. Although this is an annual composite, there are issues visible in the image: remaining haze (blueish colour) and partially remaining snow cover. In some areas a salt-and-pepper effect is visible. These artefacts are caused by erroneous scene classification: semi-transparent clouds are not flagged as low-probability clouds but are also not correctly treated (de-hazed) in the atmospheric correction. On the contrary, the necessity for a strict filtering, including the low probability clouds, leads partly to low number of valid observations and thus selection of snow as most representative spectrum.

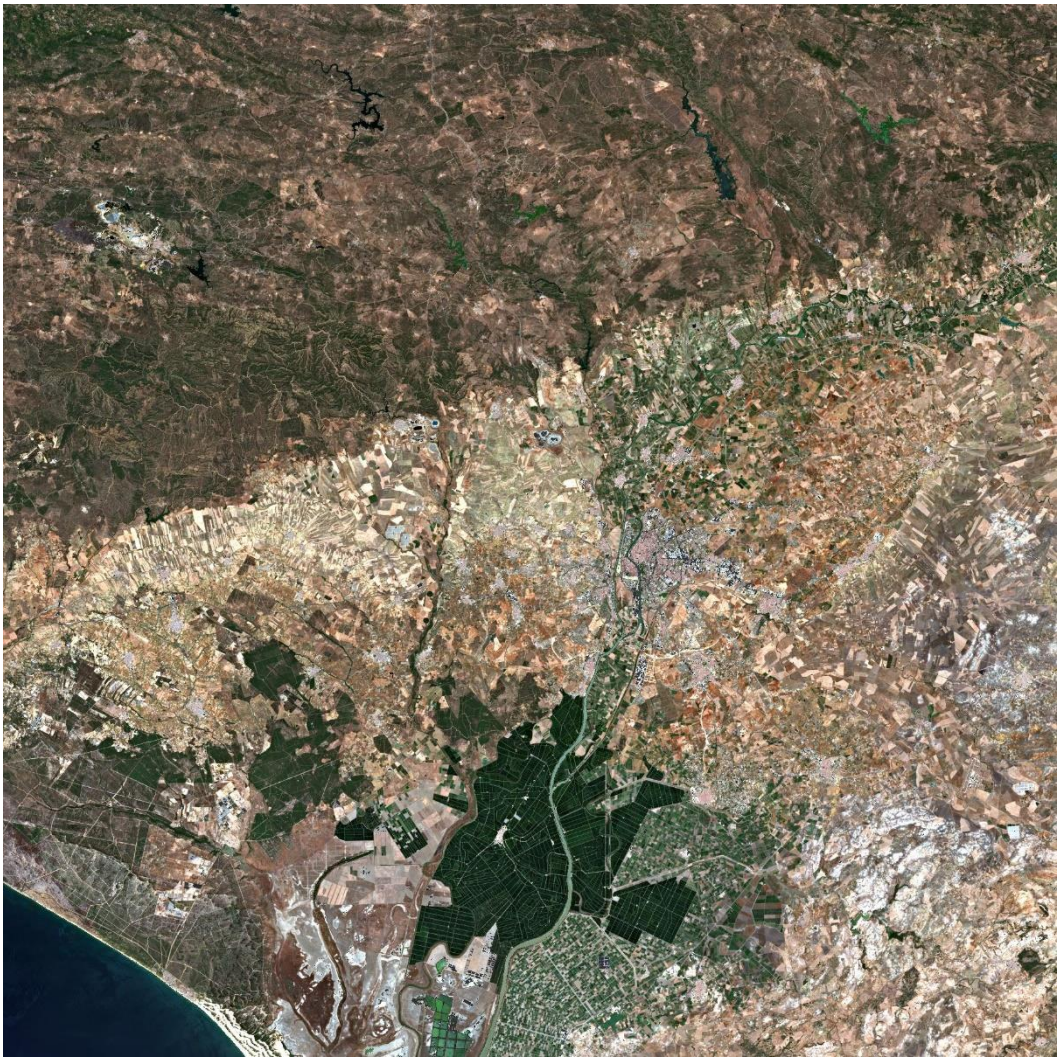


Figure 3-1: July – August – September composite of Sevilla area, 60m spatial resolution

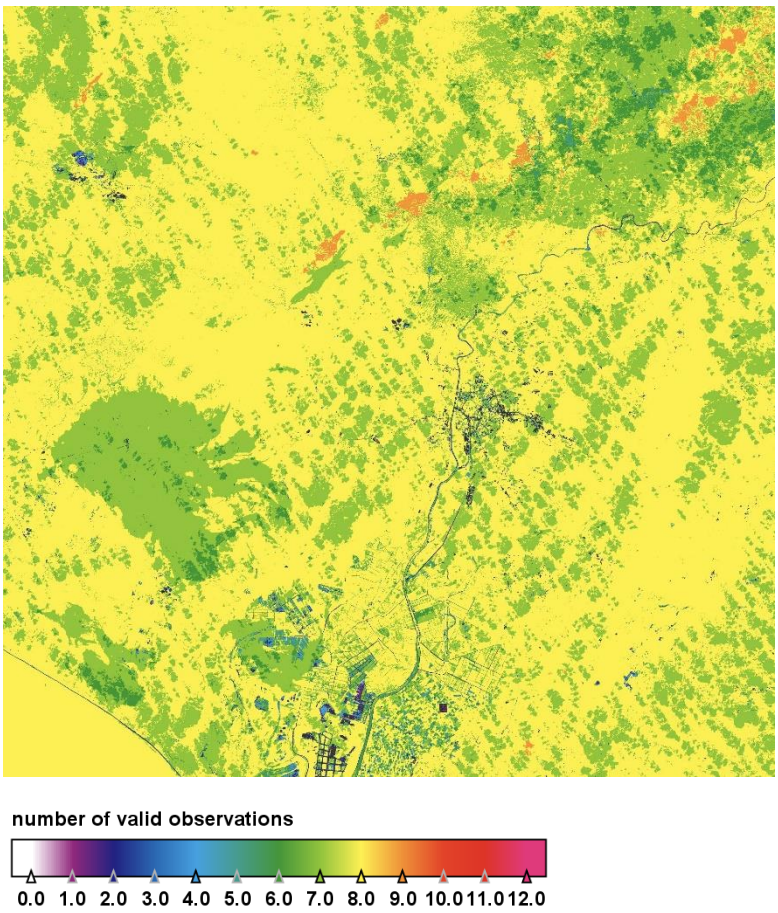


Figure 3-2: Number of valid observations for the Sevilla product of Figure 3-1.

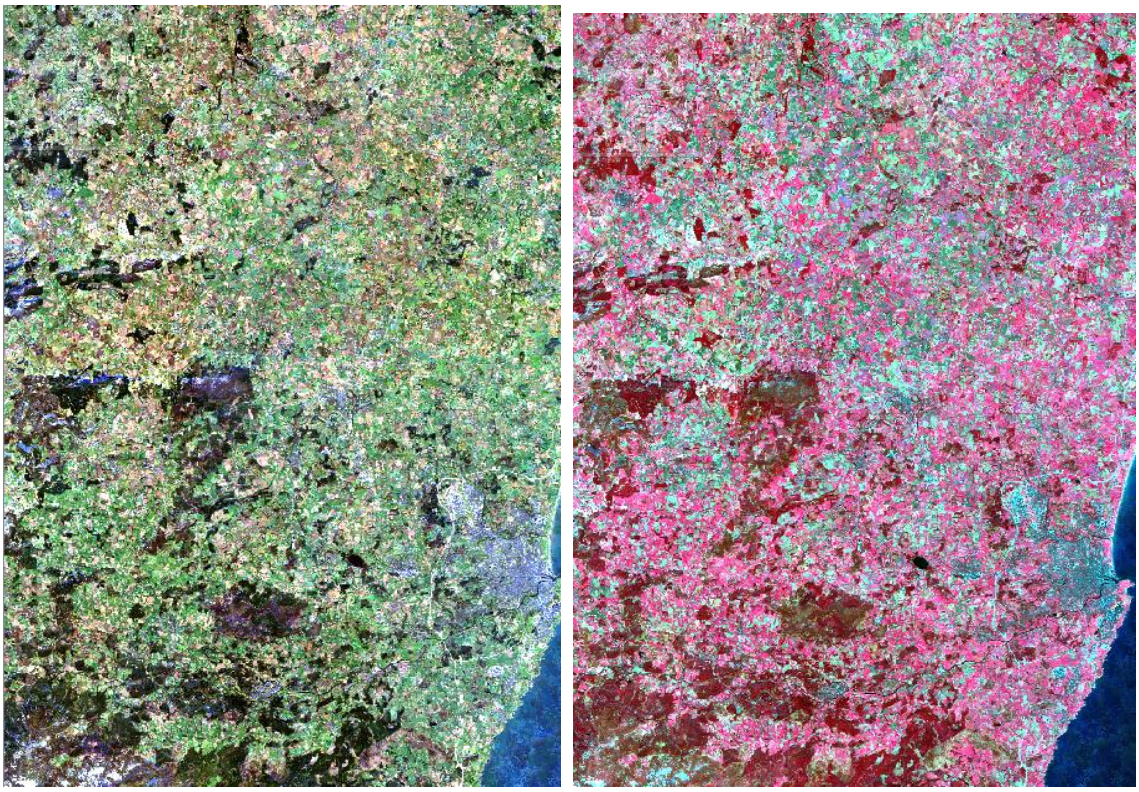


Figure 3-3: Aberdeen (Scotland) annual product of 2017. True colour RGB (left) and false colour RGB (right)

The process to select the most representative spectrum is demonstrated in Figure 3-4. All products acquired during the year 2017 went into the analysis. The left figure shows the L2A reflectances of all observations before filtering, and the right figure shows those spectra which were retained after filtering and used in the compositing algorithm. The spectra which were selected in different compositing periods are identified by colour coded dots (A … annual, M … monthly, S … seasonal. "60" indicates that a 60m spatial aggregation was applied. The two dates indicate the dates of the first and last product used in the compositing interval.)

The filtering removes largely those spectra which are bright and flat over the whole spectral range and which are obviously clouds. Two groups of spectra remain after filtering: snow spectra which are bright in the VIS and darker in the SWIR, and a group of vegetation spectra. The snow spectra are selected by the monthly and seasonal products from the spring and autumn periods, while the vegetation spectra prevail in summer. A snow spectrum has been selected for the annual product.

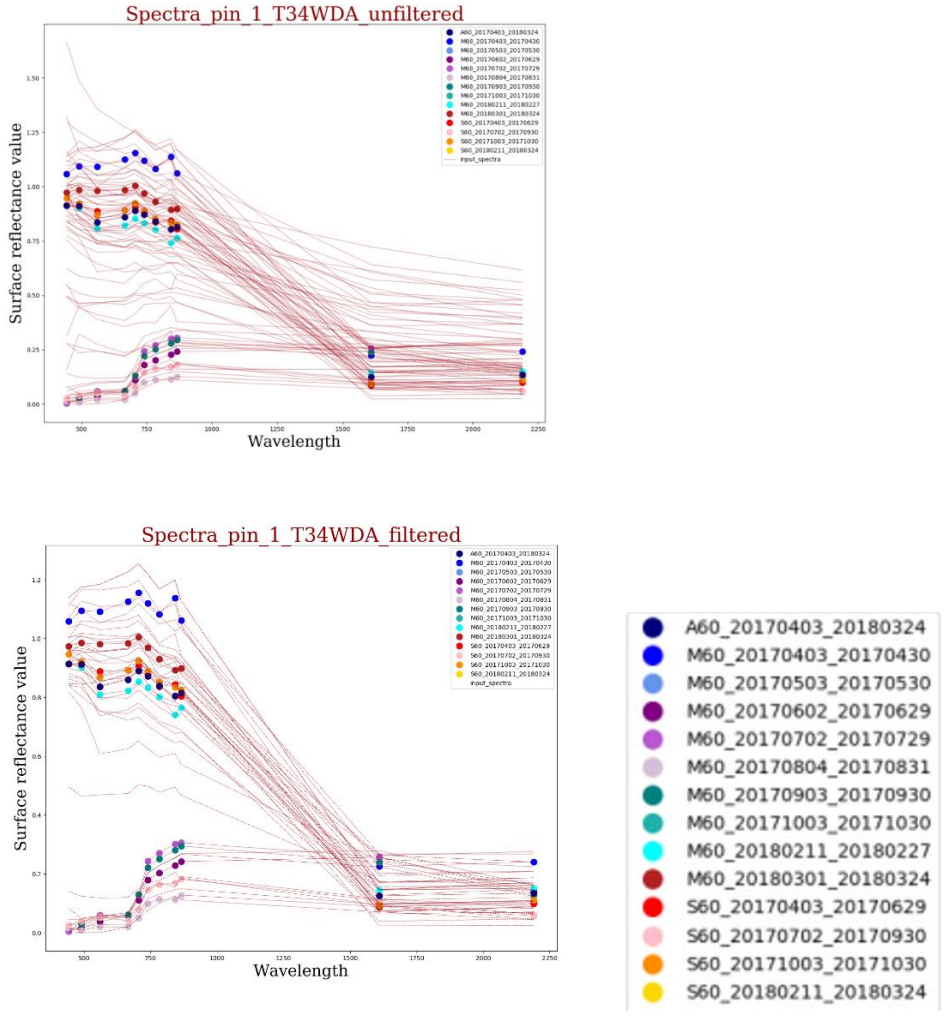


Figure 3-4: Example of the selection process of an arbitrary pixel over Kiruna, Sweden. Top: all spectra acquired during the year 2017, bottom: spectra retained after applying the filtering. Left: Colour coding used to label the spectra which were selected for the different compositing periods.

4 References

BIAN & BUTLER, 1999: BIAN, L., & BUTLER, R. (1999). COMPARING EFFECTS OF AGGREGATION METHODS ON STATISTICAL AND SPATIAL PROPERTIES OF SIMULATED SPATIAL DATA. PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING, 65(1), 73–84.

DOXANI ET AL., 2018: DOXANI, G.; VERMOTE, E.; ROGER, J.-C.; GASCON, F.; ADRIAENSEN, S.; FRANTZ, D.; HAGOLLE, O.; HOLLSTEIN, A.; KIRCHES, G.; LI, F.; LOUIS, J.; MANGIN, A.; PAHLEVAN, N.; PFLUG, B.; VANHELLEMONT, Q. ATMOSPHERIC CORRECTION INTER-COMPARISON EXERCISE. REMOTE SENS. 2018, 10, 352.

DU, Y., ZHANG, Y., LING, F., WANG, Q., LI, W., & LI, X. , 2016: 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, 8(4), 354. HTTPS://DOI.ORG/10.3390/RS8040354

ESA, 2017: TECHNICAL DOCUMENT SENTINEL-2 SPECTRAL RESPONSE FUNCTIONS, COPE-GSEG-EOPG-TN-15-0007, ISSUE 3.0, DATE 19 DECEMBER 2017 - HTTPS://EARTH.ESA.INT/WEB/SENTINEL/USER-GUIDES/SENTINEL-2-MSI/DOCUMENT-LIBRARY/-/ASSET_PUBLISHER/WK0TKAJIISAR/CONTENT/SENTINEL-2A-SPECTRAL-RESPONSES.

ESA, 2017A: ESA ,2017: SEN2THREE: SENTINEL-2 LEVEL 3 SPATIO-TEMPORAL SYNTHESIS PROCESSOR, VERSION 1.0.0.- DOCUMENTATION. HTTP://STEP.ESA.INT/THIRDPARTIES/SEN2THREE/1.0.0/SEN2THREE-1.0.0_DOC/# + HTTP://STEP.ESA.INT/THIRDPARTIES/SEN2THREE/1.0.0/SEN2THREE-1.0.0_DOC/SUM.HTML#ALGORITHMS

ESA, 2018: SEN2COR CONFIGURATION AND USER MANUAL. HTTP://STEP.ESA.INT/THIRDPARTIES/SEN2COR/2.5.5/DOCS/S2-PDGS-MPC-L2A-SUM-V2.5.5_V2.PDF

FLOOD 2013: FLOOD, N. ,2013: SEASONAL COMPOSITE LANDSAT TM/ETM+ IMAGES USING THE MEDOID (A MULTI-DIMENSIONAL MEDIAN). REMOTE SENSING, 5(12), 6481–6500. HTTP://DOI.ORG/10.3390/RS5126481

FRANTZ ET AL., 2017: FRANTZ, D., ET AL. (2017). "PHENOLOGY-ADAPTIVE PIXEL-BASED COMPOSITING USING OPTICAL EARTH OBSERVATION IMAGERY." REMOTE SENSING OF ENVIRONMENT 190: 331-347.

GRIFFITHS ET AL.,2013: GRIFFITHS, P., VAN DER LINDEN, S., KUEMMERLE, T., & HOSTERT, P. (2013). A PIXEL-BASED LANDSAT COMPOSITING ALGORITHM FOR LARGE AREA LAND COVER MAPPING. IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, 6(5), 2088–2101. HTTP://DOI.ORG/10.1109/JSTARS.2012.2228167

HUETE, 1988: HUETE, A., 1988: A SOIL ADJUSTED VEGETATION INDEX (SAVI). REMOTE SENSING OF ENVIRONMENT, 25, 295–309.

KOKALY ET AL, 2017: KOKALY, R.F., CLARK, R.N., SWAYZE, G.A., LIVO, K.E., HOEFEN, T.M., PEARSON, N.C., WISE, R.A., BENZEL, W.M., LOWERS, H.A., DRISCOLL, R.L., AND KLEIN, A.J., 2017, USGS SPECTRAL LIBRARY VERSION 7: U.S. GEOLOGICAL SURVEY DATA SERIES 1035, 61 P., DOI: 10.3133/DS1035.

LIU AND HUETE, 1995: LIU, H. Q., AND HUETE, A. R (1995). A FEEDBACK BASED MODIFICATION OF THE NDVI TO MINIMIZE CANOPY BACKGROUND AND ATMOSPHERIC NOISE, IEEE TRANS. GEOSCI.REMOTE SENS., VOL. 33, PP. 457-465.

MAIN-KORN ET AL., 2017: MAIN-KNORN, M. UND PFLUG, B. UND LOUIS, J. UND DEBAECKER, V. UND GASCON, F. UND MÜLLER-WILMS, U. (2017) SEN2COR FOR SENTINEL-2. IN: PROCEEDINGS OF SPIE, 10427, SEITEN 1-13. SPIE REMOTE SENSING, 11.-14.SEP. 2017, WARSCHAU. DOI: 10.1: 117/12.2278218

NELSON& STEINWAND, 2015: NELSON, K.J., & STEINWAND, D. (2015). A LANDSAT DATA TILING AND COMPOSITING APPROACH OPTIMIZED FOR CHANGE DETECTION IN THE CONTERMINOUS UNITED STATES. PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING, 81, 573-586.

POTAPOV ET AL., 2011: POTAPOV, P., TURUBANOVA, S., & HANSEN, M. C. (2011). REGIONAL-SCALE BOREAL FOREST COVER AND CHANGE MAPPING USING LANDSAT DATA COMPOSITES FOR EUROPEAN RUSSIA. REMOTE SENSING OF ENVIRONMENT, 115(2), 548–561. HTTP://DOI.ORG/10.1016/J.RSE.2010.10.001

POTAPOV ET AL., 2012: POTAPOV, P. V., TURUBANOVA, S. A., HANSEN, M. C., ADUSEI, B., BROICH, M., ALTSTATT, A., JUSTICE, C. O. (2012). QUANTIFYING FOREST COVER LOSS IN DEMOCRATIC REPUBLIC OF THE CONGO, 2000-2010, WITH LANDSAT ETM+ DATA. REMOTE SENSING OF ENVIRONMENT, 122, 106–116. HTTP://DOI.ORG/10.1016/J.RSE.2011.08.027

ROY ET AL., 2010: ROY, D. P., JU, J., KLINE, K., SCARAMUZZA, P. L., KOVALSKY, V., HANSEN, M., …, ZHANG, C. (2010). WEB-ENABLED LANDSAT DATA (WELD): LANDSAT ETM+ COMPOSITED MOSAICS OF THE CONTERMINOUS UNITED STATES. REMOTE SENSING OF ENVIRONMENT, 114(1), 35-49 DOI: 10.1016/J.RSE.2009.08.011

ROY ET AL., 2011: ROY, D. P., JU, J., KOMMAREDDY, I., HANSEN, M., VERMOTE, E., ZHANG, C., KOMMAREDDY, A. (2011). WEB-ENABLED LANDSAT DATA (WELD) PRODUCTS – ALGORITH THEORETICAL BASIS DOCUMENT, FEBRUARY 2011, 63 PP

STRUYF ET AL., 1996: STRUYF, A., HUBERT, M., & ROUSSEEUW, P. (1996). CLUSTERING IN AN OBJECT-ORIENTED ENVIRONMENT, JOURNAL OF STATISTICAL SOFTWARE, 1(4), 1–30, DOI:10.18637/JSS.V001.I04

XU, H. ,2006: MODIFICATION OF NORMALISED DIFFERENCE WATER INDEX (NDWI) TO ENHANCE OPEN WATER FEATURES IN REMOTELY SENSED IMAGERY. INTERNATIONAL JOURNAL OF REMOTE SENSING, 27(14), 3025–3033.

5 List of abbreviations

Abbreviation	Long name
AC	Atmospheric Correction
AOD	Aerosol Optical Depth
AOI	Area of Interest
ASCII	American Standard Code for Information Interchange
AWS	Amazon Web Services
B1 – B12	Sentinel-2 Band 1 - 12
BRDF	Bidirectional reflectance distribution function
CCI	ESA' Climate Change Initiative
CEOS	Committee on Earth Observation Satellites
CEOS	Committee on earth observation satellites
CSR	Coordinate Reference System
CSV	Comma Separated Values
DIAS	Copernicus Data and Information Access Services
DLR	German Aerospace Centre
DWH	Data Warehouse phase
EO	Earth Observation
ESA	European Space Agency
EU	European Union
GAC	GMES Advisory Committee
GIO	GMES Initial Operations
GMES	Global Monitoring for Environment and Security
GROW	DG Internal Market, Industry, Entrepreneurship and SMEs
ISRSE33	International Conference on Remote Sensing of Environment
ITT	Invitation To Tender
JRC	DG Joint Research Centre, European Commission
JSON	JavaScript Object Notation
L1	Level 1
L2	Level 2
L2A	Level 2A
LC	Land Cover
LUT	Look-up table
MSI	Multispectral Instrument
NaN	Not-a-Number
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized difference water index
NIR	Near Infrared
QA	Quality assurance
QC	Quality control
OGC	Open Geospatial Consortium
QA4EO	Quality assurance framework for earth observation
REDD	Reducing Emissions from Deforestation and forest Degradation
RGB	

RMSE	Root Mean Square Error
S2	Sentinel-2
S2A	Sentinel-2- A
S2B	Sentinel-2- B
S2GM	Copernicus Sentinel-2 Global Mosaic
SCL	Scene Classification Layer
Sen2Cor	Atmospheric Correction Processor
Sen2Three	Level 3 processor
SNAP	Sentinel Application Platform
SR	Surface Reflectance
SRF	Spectral Response Functions
STC	Short-Term Composite
TOA	Top of Atmosphere
USGS	United States Geological Survey
UTC	Universal Time Coordinated
VM	Virtual Machine
WELD	Web-Enabled Landsat Data
WMS	Web Mapping services
WP	Work Package
XML	Extensible Mark-up Language