RESEARCH PAPER

Cloud shadow removal for optical satellite data

Viktoria Zekoll[ID]*

DLR, German Aerospace Center, Wessling, Germany

ABSTRACT. An improved cloud shadow removal algorithm for high spatial resolution optical satellite data over land is presented. The method is based on the matched filter method, which consists of the calculation of a covariance matrix and the corresponding zero-reflectance matched filter vector and the computation of the shadow function. The new additions consist of the usage of an improved cloud shadow map and further evaluations performed on the shadow function. The performance of the cloud shadow removal algorithm incorporated in the software package Python-based atmospheric correction (PACO) is compared to the deshadowing algorithm in atmospheric correction on a set of 25 Sentinel-2 scenes distributed over the globe covering a wide variety of environments and climates. Furthermore, an evaluation of the relative ratio between clear and shadow pixels with and without deshadowing is performed. The visual, spectral, and statistical results show that the new additions performed on the deshadowing algorithm can improve the cloud shadow removal performance used so far.

1 Introduction

For optical remote sensing of the Earth's surface, clouds and their shadows have always been a major disadvantage, since a lot of remote sensing applications are impacted by their presence. These applications involve, for example, radiation, image classification, the calculation of surface reflectance or land surface temperature, vegetation indices, etc.[1,2]

The annual cloud coverage of the Earth lies around ∼70%.[3] Therefore, it is inevitable that no observations of a specific location on Earth will be continuously cloud- and cloud shadow-free, and the information that can be extracted from a scene will have a high percentage of degradation.[4] This means that scientists will have to find ways to work around or with the presence of clouds and cloud shadows. This is especially crucial for land applications for which the amount of usable data per scene and specific timing is of high importance, for example for crop yield estimation.[5] Furthermore, the use of cloud and cloud shadow free images enables to determine ground properties of the Earth's surface[6–9] and facilitates crop monitoring tasks.[2] Even geological applications[10] are disturbed if clouds and their shadows cover parts of high spatial resolution optical satellite data. This proves how important it is nowadays to have a correct and exact masking of clouds and cloud shadows as preprocessing step for atmospheric correction (ATCOR) and shadow removal of multi-spectral imagery. Hence, in the past years, more and more cloud and cloud shadow detection and removal approaches have been developed[11–18] and used to enable various applications.

---

*Address all correspondence to Viktoria Zekoll, vz13110@my.bristol.ac.uk

To undergo a cloud shadow removal algorithm, the cloud and their shadows have to be detected and mapped. The detection of clouds can be done by studying each satellite scene separately using a mono-temporal approach[19–26] or through a multi-temporal methodology[27,28] and hence studying a time series of images. For the detection of the correct location and geometry of a cloud shadow, the direction of observation is crucial since they represent projections of clouds in an image.[1] In this paper, a mono-temporal cloud shadow detection approach is used as preprocessing step for cloud shadow removal, called thresholds, indices, projections (TIP) method.[29]

Due to remotely sensed optical imagery of the Earth's surface being contaminated by cloud and cloud shadows, the surface information underneath a cloud covered region cannot be retrieved with optical sensors. The surface information underneath a cloud shadow, on the other hand, can be retrieved since the ground reflected solar radiance is a small non-zero signal. Now the total radiation signal that is measured at the sensor is composed out of a direct beam and a diffuse, reflected skylight component. This means that even if there is no direct solar beam arriving at the sensor from the shadow region, there will still be some information arriving from the reflected diffuse flux.[30]

The proposed shadow removal method works with this knowledge and uses the estimate of the fraction of direct solar irradiance for a fully or partially shadowed pixel as basis for the removal algorithm. The aim is to provide an improved cloud shadow removal algorithm based on the current version of the matched filter proposed by ATCOR.[30] As opposed to the IDL ATCOR algorithm, the new cloud shadow removal algorithm is implemented into the Python-based atmospheric correction (PACO) software.

In this paper, the multispectral instruments (MSIs) of the Copernicus Sentinel-2 satellites are used.[31] The MSIs are sensors on-board of the satellite, which allow free access to the data and a high revisit time.[32] The 13 spectral bands of a Sentinel-2 scene are composed of 4 bands at 10 m, 6 bands at 20 m and 3 bands at 60 m spatial resolution. Furthermore, the PACO atmospheric processor is used. This is the python-based version of ATCOR. For PACO, the input data are in L1C radiances in units of $[\mathrm{mW}/(\mathrm{cm}^2 * \mathrm{sr} * \mu\mathrm{m})]$. If the input scene is given in terms of top of atmosphere (TOA) reflectance, it has to be converted into TOA radiance. PACO performs the ATCOR using the spectral information in all bands by resampling them to a 20-m cube, yielding an image cube of 13 bands with a size of 5490×5490 pixels. This so called "merged cube" will be a Sentinel-2 TOA cube considered in the rest of this study.

2 Basic Concept of Atmospheric Correction

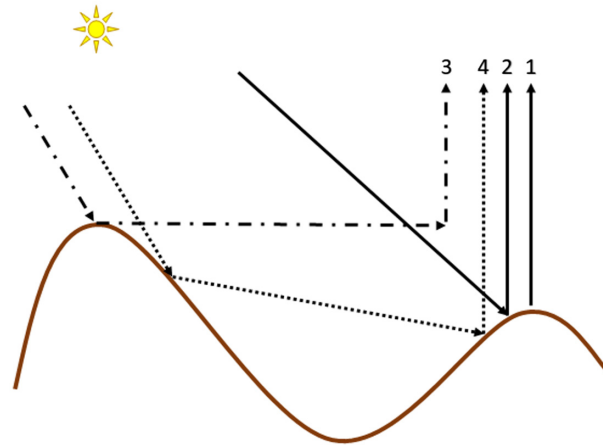### 2.1 Radiation Components and Surface Reflectance

The radiation signal in the solar region (0.35 to 2.5 μm) arriving at the sensor is due to four different components.[33,34]

- Path radiance, L_{path}: from photons that did not have contact with the ground and are scattered into the field-of-view of the sensor.
- Ground reflected radiation from a pixel, L_{ground}: the fraction of the diffuse and direct solar radiation incident on the pixel that is getting reflected from the surface.
- Reflected radiation from the surrounding, L_{adj}: the fraction of the solar radiation reflected from the neighborhood and scattered by the air volume into the field-of-view of the sensor. This radiation is also called adjacency radiation.
- Reflected terrain radiance from opposite mountains, L_{terrain}.

Figure 1 shows the four different components arriving at the sensor. For a full evaluation of the radiation component in rugged terrain, please refer to Ref. 34 and Sec. 6.2 in Ref. 30.

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

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

Fig. 1 Radiation components arriving at the sensor. The path radiance is represented by component number 1. The ground reflected radiation is represented by component number 2. The reflected radiation from the surrounding is represented by component number 3. The reflected terrain radiance is represented by component number 4.

$$L = L_{\text{path}} + L_{\text{ground}} + L_{\text{adj}} + L_{\text{terrain}}. \tag{1}$$

3 Methods

3.1 Previous Method: Matched Filter

3.1.1 *MF method: surface reflectance and covariance matrix*

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

The MF method needs at least one channel in visible and at least one spectral band in the NIR. The bands used in the MF are: blue, green, red, near-infrared (NIR), short-wave infrared 1 (SWIR1), and short-wave infrared 2 (SWIR2), if existing.

The method starts with the calculation of the surface reflectance image cube, ρ. The surface reflectance, ρ, is computed with the assumption of fully solar illumination, excluding water and clouds. Then the covariance matrix, $C(\rho)$, is calculated where ρ represents the surface reflectance vector of the three selected bands [see Eq. (2)]

$$V_{\text{MF}} = \frac{C^{-1}(\rho_t - \overline{\rho})}{(\rho_t - \overline{\rho})^T C^{-1}(\rho_t - \overline{\rho})}. \tag{2}$$

The MF vector, V_{MF}, is tuned to a certain target reflectance spectrum, ρ_t, to be detected. $\overline{\rho}$ is the scene-average spectrum without water and cloud pixels. For the shadow target, a target reflectance spectrum of $\rho_t = 0$ is selected, which will give the simplified version of the shadow MF vector, V_{sh}, as follows:[35]

$$V_{\text{sh}} = -\frac{C^{-1}\overline{\rho}}{\overline{\rho}^T C^{-1}\overline{\rho}}. \tag{3}$$

3.1.2 *MF method: unscaled shadow function*

The MF shadow vector, V_{sh}, can be applied to the non-water and non-cloud part of the scene to give the un-normalized values, Φ, also called unscaled shadow function. $\Phi(x, y)$ gives a relation measure of the fractional direct illumination for each pixel x, y [see Eq. (4)]

$$\Phi(x, y) = V_{\text{sh}}^T(\rho(x, y) - \overline{\rho}). \tag{4}$$

3.1.3 *MF method: rescaling and scaled shadow function*

The MF calculates a minimum RMS shadow target abundance for the entire scene. The values of Φ can be both, positive and negative. Therefore, Φ is rescaled into the physical range from 0 (full shadow) to 1 (full direct illumination). The histogram of the unscaled shadow function is used for rescaling and an illustration can be found in Fig. 3 of Ref. 36.

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

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

$$\Phi^* = \frac{\Phi - \Phi_{min}}{\Phi_{max} - \Phi_{min}} \quad \text{if } \Phi \le \Phi_{max} \quad \Phi = 1 \quad \text{if } \Phi > \Phi_{max}. \tag{5}$$

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

- min and max direct sun fraction in the shadow regions (a_{min} and a_{max}; the defaults are $a_{min} = 0.20$ and $a_{max} = 0.95$).[36]
- Corresponding shadow thresholds Φ_{min} and Φ_{max} obtained from the normalized histogram of Φ.

The normalized and scaled shadow function is given as

$$\Phi_n = a_{min} + \frac{(\Phi - \Phi_{min})(a_{max} - a_{min})}{\Phi_{max} - \Phi_{min}}. \tag{6}$$

3.1.4 *MF method: iteration*

The potential shadow pixels are those which satisfy $\phi_n < 1$, but as can be seen from Eq. (6), the value strongly depends on Φ_{max}. Therefore, an iterative strategy is applied and the exact steps of the ATCOR MF iteration method can be found in Ref. 36.

3.1.5 *Deshadowing reflectance equation*

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

The final deshadowing is performed by multiplying the direct illumination, E_{dir}, with the pixel-dependent Φ_n. This reduces the direct solar term and increases the brightness of a shadow pixel, since it is located in the denominator of the deshadowing equation [see Eq. (7)]

$$\rho_i(x, y) = \frac{\pi(d^2(c_0(i) + c_1(i)\text{DN}_i(x, y)) - L_{p,i})}{\tau_i(E_{dir,i}\Phi_n(x, y) + E_{dif,i})}. \tag{7}$$

3.2 Proposed Method: Cloud Shadow Removal MF Method with New Additions

The proposed cloud shadow removal algorithm was created for data acquired by satellite/airborne sensors in multispectral and hyperspectral imagery. It is computed to create a fully automated shadow removal algorithm and based on the main concept of the MF method.[30,36]

The MF method was first implemented from the IDL ATCOR version into the python-based PACO software. The new cloud shadow removal algorithm incorporates the main equations from the MF method used by the ATCOR deshadowing algorithm after.[30] To improve the results, additions have been added and are proposed in this paper.
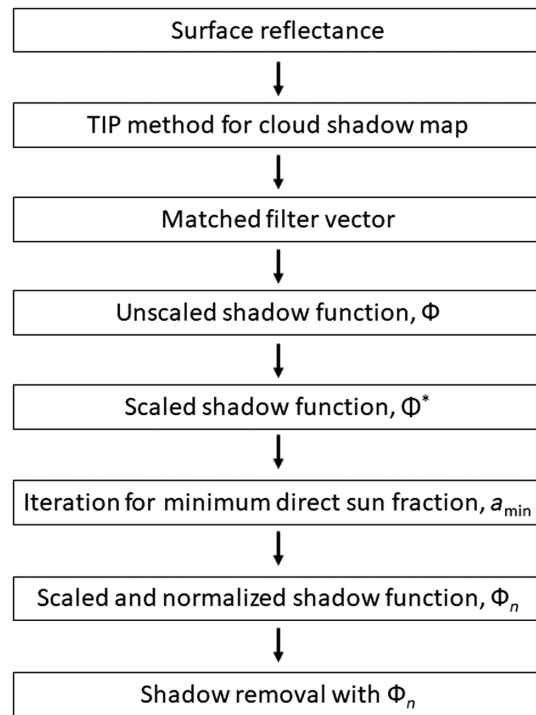
Fig. 2 Flowchart of the eight main steps of the proposed cloud shadow removal method.

Figure 2 shows a flow chart of the eight main steps performed during the new cloud shadow removal algorithm. The steps that have to be performed are: the calculation of the surface reflectance; the evaluation of the cloud shadow map with the TIP method after;[29] the calculation of the MF vector, which gives the unscaled shadow function, Φ; the calculation of the scaled shadow function, Φ^*; the calculation of the minimum direct sun fraction, a_{min} and the normalized scaled shadow function, Φ_n; and the final step of the shadow removal.

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

The new additions are the cloud shadow map from the TIP method developed in Ref. 29 and a newly defined iteration over the scaled minimum shadow function, a_{min}.

The TIP cloud shadow detection is based on thresholds, indices and projections and has been able to improve the results of the previous ATCOR cloud shadow map calculation. The detailed evaluation of the TIP cloud shadow map calculation can be found in Ref. 29. Having included a better cloud shadow map will improve the results of the shadow removal algorithm.

The second newly implemented addition is performed within the computation of the matched filter vector Φ. The normalization of Φ into the physical range between 0 and 1 is evaluated with the minimum and maximum direct sun fraction in the cloud shadow area a_{min} and a_{max} and their corresponding shadow thresholds Φ_{min} and Φ_{max}. Φ_{min} and Φ_{max} are obtained from the normalized histogram of Φ. The default values of a_{max} is set to 0.95, and as opposed to the previous method, the starting default value of a_{min} is set to 0.01. Then the normalized scaled shadow function is calculated using Eq. (6).

The proposed iteration calculates the mean reflectance of all shadow pixels (Φ_{sh}) and all sun-lit pixels (Φ_{sun}) for all bands, j, and terminates when the overall absolute difference of these two values, $\sum D(j)$, is less than the difference of the previous iteration, $\sum D(j)_{previous}$ [see Eq. (8)]

$$\sum D(j) = \mathrm{abs}(\Phi_{\mathrm{sh}}(j) - \Phi_{\mathrm{sun}}(j)) < \sum D(j)_{\mathrm{previous}}. \qquad (8)$$

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

The final calculation of the reflectance is done using Eq. (7) with the corresponding scaled shadow function.

4 Results

The results of the deshadowing algorithm presented in this paper are first shown through three selected scenes from the data set where a visual and spectral comparison is performed (see Sec. 4.2). Additionally, a metric quantitative comparison for all evaluated scenes is shown and discussed in Sec. 4.3.

4.1 Data and Material for Training Set

To test the new cloud shadow removal method on a set of data, 25 Sentinel-2 (S2) scenes are chosen. A list of the investigated Sentinel-2 A and B scenes is given in Table 1. The scenes were selected to cover a wide variety of regions over the entire globe (see Fig. 3). This enables to validate the shadow removal algorithm for different continents, climates, seasons, weather conditions, and land cover classes. Furthermore, they have been selected to represent flat and mountainous sites with a cloud cover from 3% to 80% and they include the presence of cumulus, thin and thick cirrus clouds. The land cover types represented are: desert, urban, cropland, grass, forest, wetlands, and sand and coastal areas. The range of solar zenith angles is from 27 deg to 67 deg.

4.2 Cloud Shadow Removal Results

In the following section, three scenes (scene ID 18, 16, and 10 from Table 1) are chosen to show the results of the cloud shadow removal algorithm. For each scene, the deshadowing results are given for a subset to better evaluate the results spectrally and visually. In each subset, a clear pixel and a shadow pixel are selected that are located close by and represent the same ground properties. The visual and spectral comparison is done between the original scene, the new presented cloud shadow removal algorithm from PACO and the cloud shadow removal algorithm as given by ATCOR.

4.2.1 Netherlands, Amsterdam (scene ID 18)

Figure 4 shows the first scene results to be analyzed in this paper. It is scene number 18 of Table 1, located in Amsterdam, Netherlands, and taken on the 13th of September 2018 with a zenith angle of 49.7 deg. To show the visual results of the deshadowing algorithm, a subset from scene ID 18 was chosen. The exact location of the subset is shown in the top row of Fig. 4. The left image of Fig. 4 represents the original Amsterdam scene stretched into the RGB = 665/560/490 nm bands for better optical comparison. The right image shows the chosen subset.

To compare the new method visually with the previous version, Fig. 4 provides the subset from the original scene and the two deshadowed subsets from PACO (new version) and ATCOR (old version) in the middle row, respectively. For each subset the same zoom is provided (see Fig. 4 bottom row). In this image zoom, a clear pixel and a cloud shadow pixel are chosen. To not only provide visual results, the spectra of the clear pixel, the cloud shadow pixel, the PACO deshadowed cloud shadow pixel, and the ATCOR deshadowed cloud shadow pixel are presented in Fig. 5.

The black curve of Fig. 5 represents the reflectance spectrum obtained from the cloud shadow pixel without correction. The pink curve represents the chosen clear pixel close by. The orange curve represents the reflectance spectrum of the cloud shadow pixel after deshadowing with the new PACO version. Finally the blue curve represents the cloud shadow pixel after being deshadowed with ATCOR. As can be seen from Fig. 5 both methods nicely deshadow the

Table 1 Sentinel-2 level L1C test scenes. Information on scene climate, main surface cover, and rural/urban. (SZA = solar zenith angle).

Scene	Location	Date	Tile	SZA	Desert	Ice/snow	Nonveg	Veg	Water	Mountains	Rural	Urban
1	Africa, Gobabeb	06/03/2019	T33KWP	32.5 deg	X					X	X	
2	Africa, Namibia			deg	X					X	X	
3	Antarctic-1	03/02/2019	T21EVK	54.9 deg		X	X		X			
4	Argentina, Buenos Aires	27/08/2018	T21HUC	51.5 deg					X		X	X
5	Australia, Lake Lefroy	19/08/2018	T51JUF	51.5 deg			X					
6	Bolivia, Puerto Siles	06/09/2018	T19LHF	30.6 deg				X	X		X	
7	China, Dunhuang	22/01/2018	T46TFK	62.3 deg	X	X					X	
8	France, Arcachon-1	15/11/2017	T30TXQ	63.9 deg			X	X	X		X	X
9	France, Arcachon-2	18/02/2017		deg								
10	France-3	16/01/2016	T31TFJ	66.8 deg			X	X	X	X	X	X
11	Estonia, Tallin	14/07/2018	T35VLG	39.0 deg					X		X	X
12	Germany, Berlin	04/05/2018	T33UUU	38.0 deg			X	X	X		X	X
13	Italy, Etna	09/03/2017	T33UUU	45.1 deg		X			X		X	X
14	Kazakhstan, Balkhash	30/07/2018	T43TFM	30.7 deg	X		X		X		X	
15	Mexico, Cancun	27/05/2018	T16QDJ	18.4 deg				X	X		X	
16	Morocco, Quarzazate	30/08/2018	T29RPQ	27.2 deg	X		X	X	X			X
17	Mosambique, Maputo	13/07/2018	T36JVS	54.4 deg					X		X	X
18	Netherlands, Amsterdam	13/09/2018	T31UFU	49.7 deg				X	X		X	X
19	Philipines, Manila	19/03/2018	T51PTS	27.4 deg					X		X	X

Table 1 (*Continued*).

Scene	Location	Date	Tile	SZA	Desert	Ice/snow	Nonveg	Veg	Water	Mountains	Rural	Urban
20	Russia, Sachalin	09/05/2018	T54UVC	35.5 deg		X	X		X			
21	Russia, Yakutsk	08/08/2017	T52VEP	45.9 deg			X	X	X			
22	Spain, Barrax-1	09/05/2017	T30SWH	24.1 deg				X		X	X	
23	Spain, Barrax-2	19/05/2017	T30SWH	22.0 deg				X		X	X	
24	Switzerland, Davos	17/04/2019	T32TNS	37.7 deg		X	X	X	X	X		
25	United States of America, Rimrock	12/05/2018	T11TMM	30.4 deg		X			X		X	X
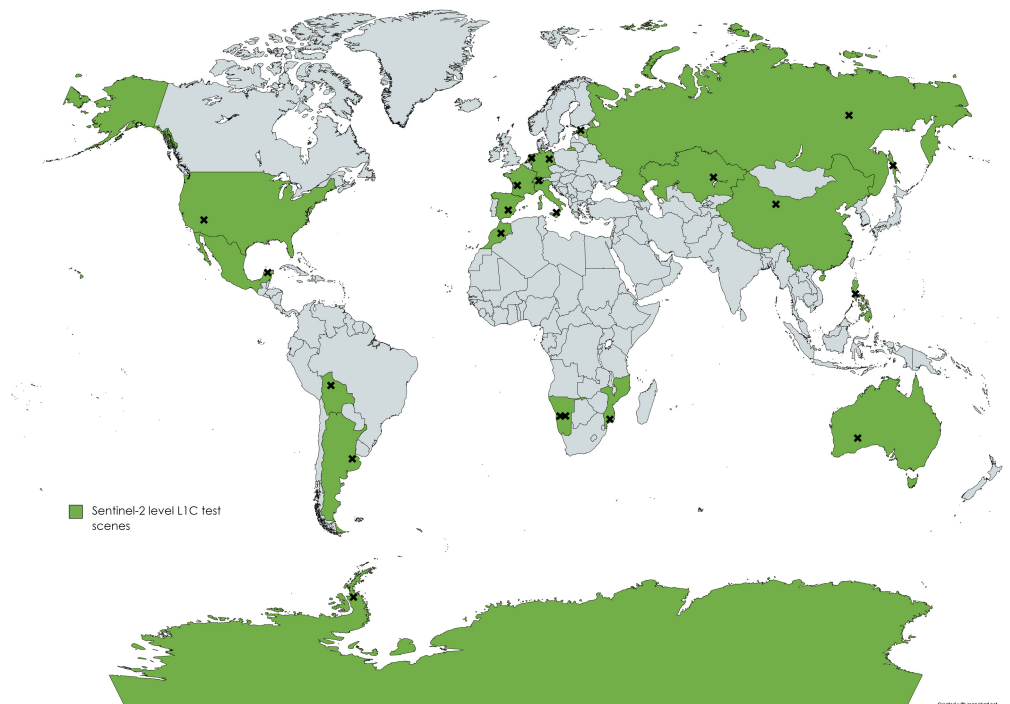
Fig. 3 Geographical distribution of the Sentinel-2 selected test sites. Black crosses represent the exact location of the test sites.

cloud shadow pixel and get very close to the true clear pixel reflectance spectrum. ATCOR slightly overshoots the compensation whereas the new PACO method closely follows the pink curve but slightly lower.

4.2.2 Morocco, Quarzazate (scene ID 16)

In order to prove the promising results of the cloud shadow removal algorithm presented in this paper, a second scene is illustrated in Fig. 6. Figure 6 shows scene number 16 located in Quarzazate, Morocco. The scene was taken on the 30'th of August and has a zenith angle of 27.2 deg. The top left image of Fig. 6 shows the RGB = 665/560/490 nm true color composite of Quarzazate and the top right image the chosen subset from this scene.

The same evaluation as done in Sec. 4.2.1 for scene ID 18 is performed on the Morocco example. Figure 6 shows the scene subsets of the original, PACO deshadowed and ATCOR deshadowed scene in the middle row and a zoom of the subset in the bottom row.

In Fig. 7, the reflectance spectra for the clear and cloud shadow pixel of the original scene, the PACO deshadowed scene, and the ATCOR deshadowed scene are given. The reflectance spectra again show how the ATCOR deshadowing algorithm rather overcompensates the cloud shadow pixel, whereas the PACO deshadowing algorithm follows the reflectance spectrum of the clear pixel very nicely for low wavelengths and then slowly becomes smaller.

4.2.3 France (scene ID 10)

As a final example, scene ID 10 is chosen. It represents a scene from France taken on the 16'th January 2016 and has a zenith angle of 66.8 deg. The top row of Fig. 8 shows on the left the original scene stretched into the RGB = 665/560/490 nm bands for better optical comparison and the chosen subset on the right. For this example, two image zooms of the subset were taken.

As performed in Secs. 4.2.1 and 4.2.2, Figs. 8 and 10 show the scene subsets of the original, PACO deshadowed and ATCOR deshadowed scene in the middle row and the first zoom of the subset in the bottom row. In Fig. 9, the reflectance spectra for the clear and cloud shadow pixel of the original scene, the PACO deshadowed scene and the ATCOR deshadowed scene are given for
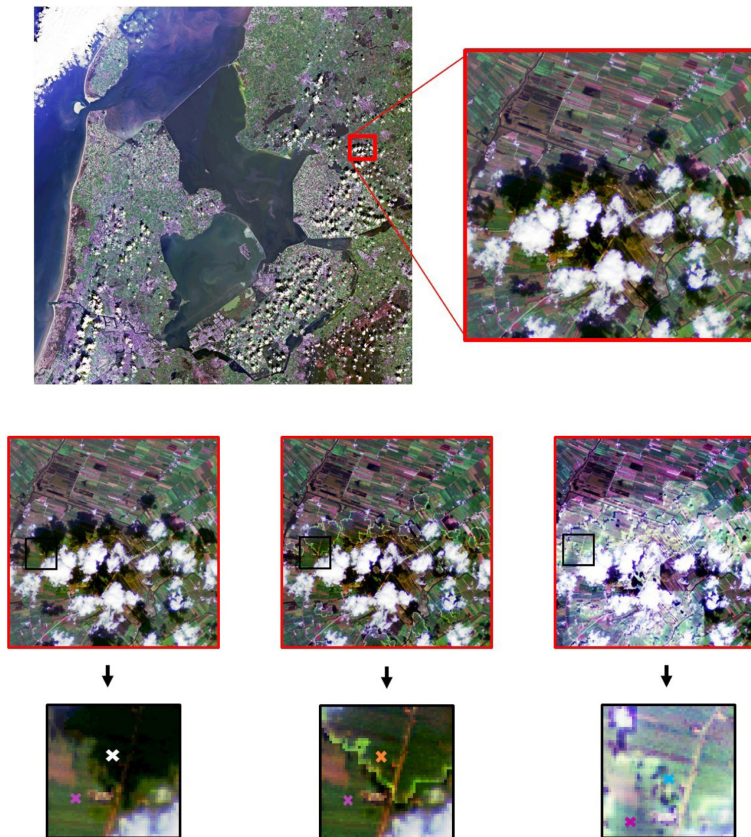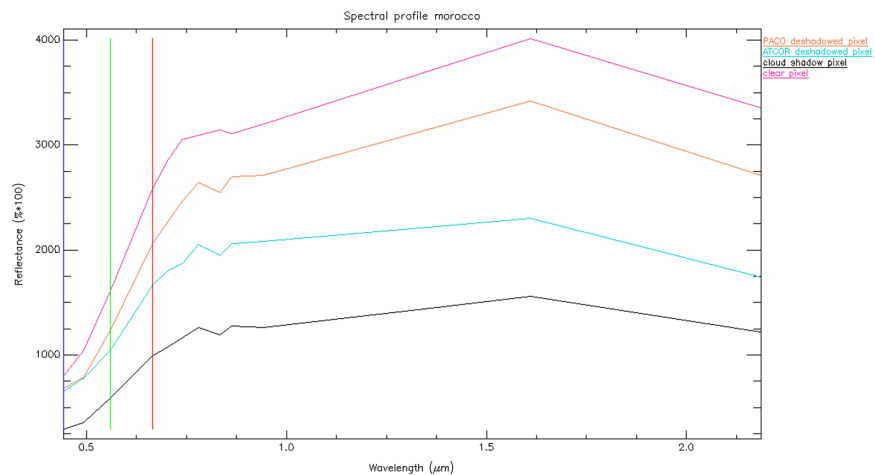
Fig. 4 Top row (left): Amsterdam (Netherlands) RGB = 665/560/490 nm with equalization stretching. Top row (right): subset of scene ID 18 with equalization stretching. Middle row (left to right): subset of Amsterdam (Netherlands) RGB = 665/560/490 nm with equalization stretching; PACO deshadowed subset of Amsterdam with equalization stretching; and ATCOR deshadowed subset of Amsterdam with equalization stretching. Bottom row (left to right): zoom of subset of Amsterdam (Netherlands) RGB = 665/560/490 nm with equalization stretching (white cross = cloud shadow pixel; pink cross = clear pixel); zoom of PACO deshadowed subset of Amsterdam with equalization stretching (orange cross = PACO deshadowed cloud shadow pixel; pink cross = clear pixel); and zoom of ATCOR deshadowed subset of Amsterdam with equalization stretching (blue cross = ATCOR deshadowed cloud shadow pixel; pink cross = clear pixel).
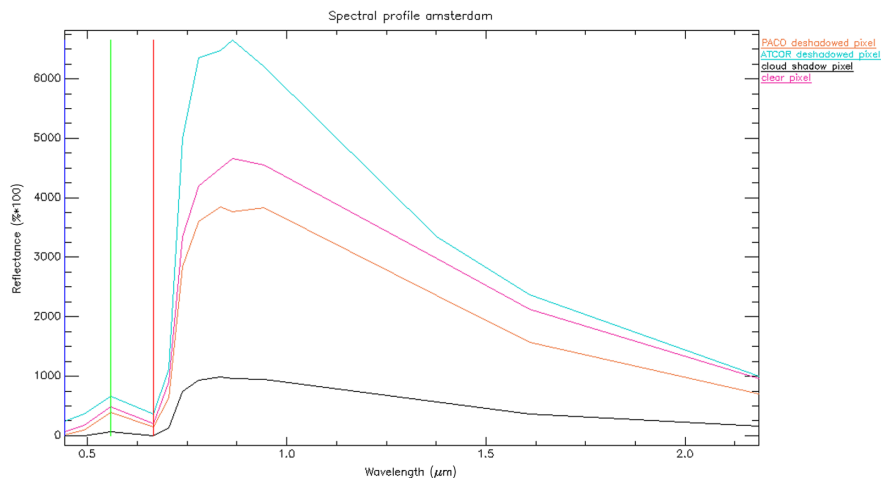

Fig. 5 Spectral profile of scene ID 18 (Amsterdam). Orange = PACO deshadowed cloud shadow pixel; blue = ATCOR deshadowed cloud shadow pixel; black = original cloud shadow pixel spectrum; and pink = clear pixel spectrum.
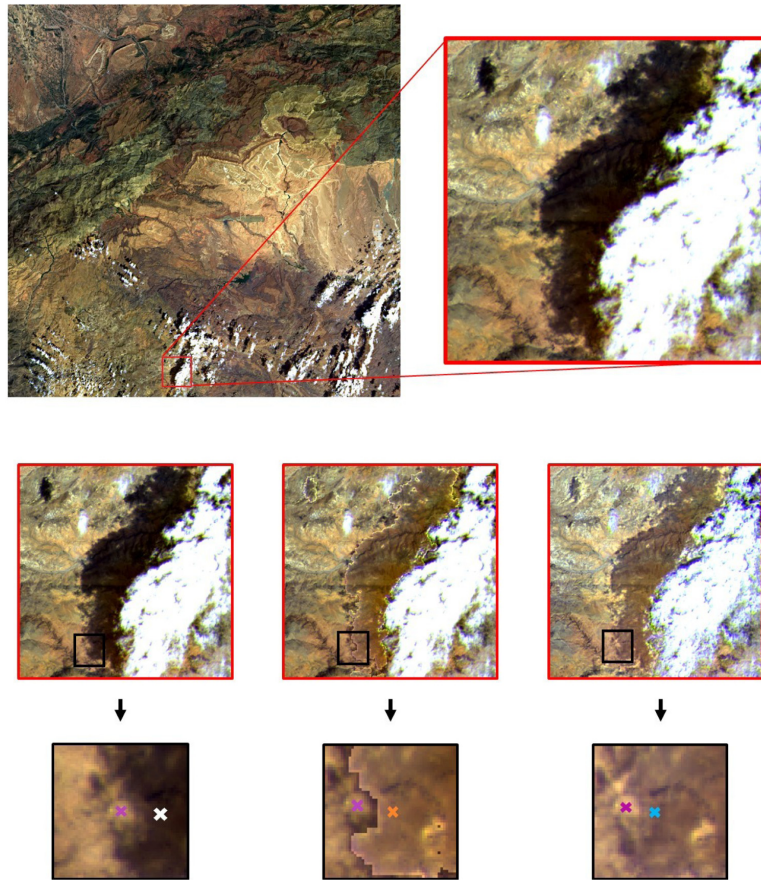
Fig. 6 Top row (left): Quarzazate (Morocco) RGB = 665/560/490 nm true color composite. Top row (right): RGB = 665/560/490 nm true color composite subset of scene ID 16. Middle row (left to right): subset of Quarzazate (Morocco) RGB = 665/560/490 nm true color composite; PACO deshadowed subset of Quarzazate; and ATCOR deshadowed subset of Quarzazate. Bottom row (left to right): zoom of subset of Quarzazate (Morocco) RGB = 665/560/490 nm true color composite (white cross = cloud shadow pixel; pink cross = clear pixel); zoom of PACO deshadowed subset of Quarzazate (orange cross = PACO deshadowed cloud shadow pixel; pink cross = clear pixel); and zoom of ATCOR deshadowed subset of Quarzazate (blue cross = ATCOR deshadowed cloud shadow pixel; pink cross = clear pixel).

Fig. 7 Spectral profile of scene ID 16 (Quarzazate). Orange = PACO deshadowed cloud shadow pixel; blue = ATCOR deshadowed cloud shadow pixel; black = original cloud shadow pixel spectrum; and pink = clear pixel spectrum.
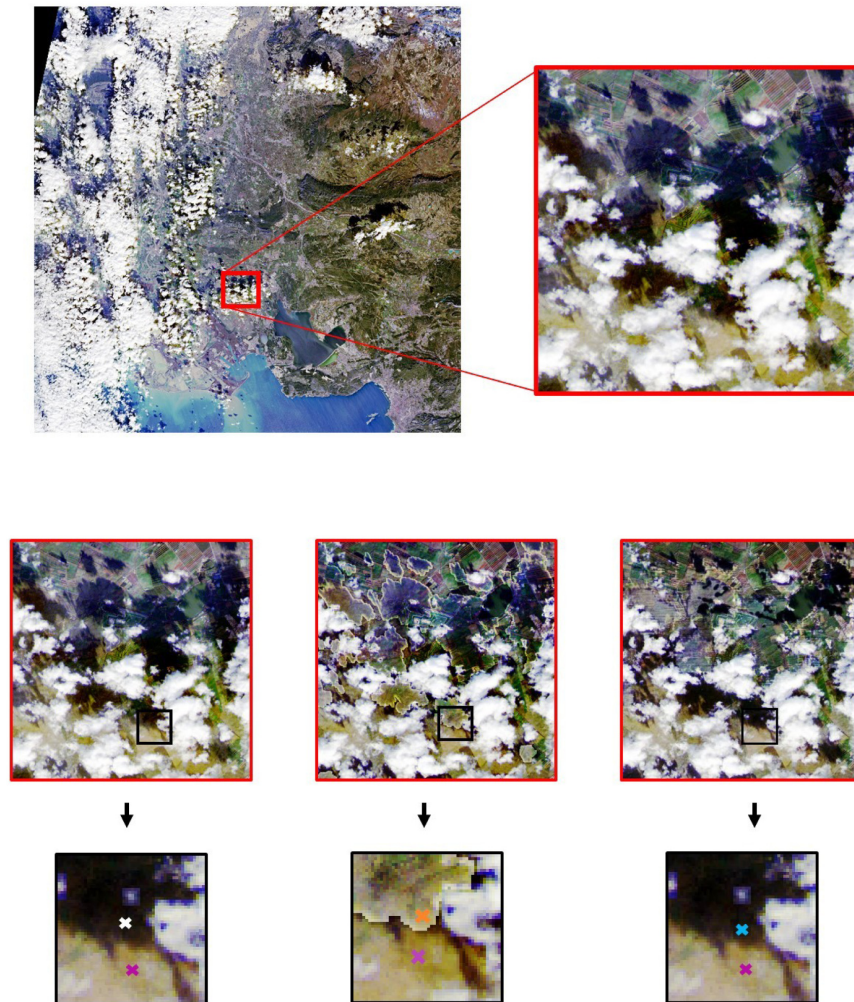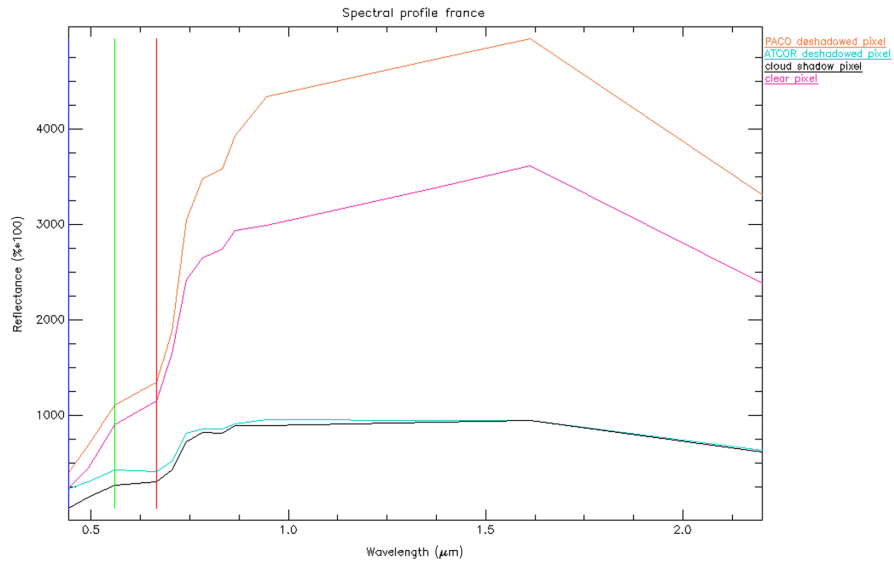
Fig. 8 Top row (left): France RGB=665/560/490 nm with equalization stretching. Top row (right): subset of scene ID 10 with equalization stretching. Middle row (left to right): subset 1 of France RGB = 665/560/490 nm with equalization stretching; PACO deshadowed subset 1 of France with equalization stretching; and ATCOR deshadowed subset 1 of France with equalization stretching. Bottom row (left to right): zoom of subset 1 of France RGB = 665/560/490 nm with equalization stretching (white cross = cloud shadow pixel; pink cross = clear pixel); zoom of PACO deshadowed subset 1 of France with equalization stretching (orange cross = PACO deshadowed cloud shadow pixel; pink cross = clear pixel); and zoom of ATCOR deshadowed subset 1 of France with equalization stretching (blue cross = ATCOR deshadowed cloud shadow pixel; pink cross = clear pixel).

zoom number 1. As can be seen from the visual zoom in Fig. 8 and the reflectance spectra given in Fig. 9, ATCOR is not able to deshadow this area of the cloud shadow. The new version on the other hand performs similar to the previous examples and nicely recovers most of the reflectance spectrum.

Figure 10 shows the scene subsets of the original, PACO deshadowed and ATCOR deshadowed scene in the middle row and the second zoom of the subset in the bottom row. In Fig. 11 the reflectance spectra for the clear and cloud shadow pixel of the original scene, the PACO deshadowed scene and the ATCOR deshadowed scene are given for zoom number 2. In this example of France, ATCOR performs better visually as seen in Fig. 11, but looking at the reflectance spectras, the new deshadowing algorithm gives better results.

4.3 Validation of Dataset

To evaluate the data set, a metric for the comparison of the surface reflectance retrieval without deshadowing and the deshadowed reflectance is calculated. This metric is represented by the

Fig. 9 Spectral profile of scene ID 10 (France). Orange = PACO deshadowed cloud shadow pixel; blue = ATCOR deshadowed cloud shadow pixel; black = original cloud shadow pixel spectrum; and pink = clear pixel spectrum.
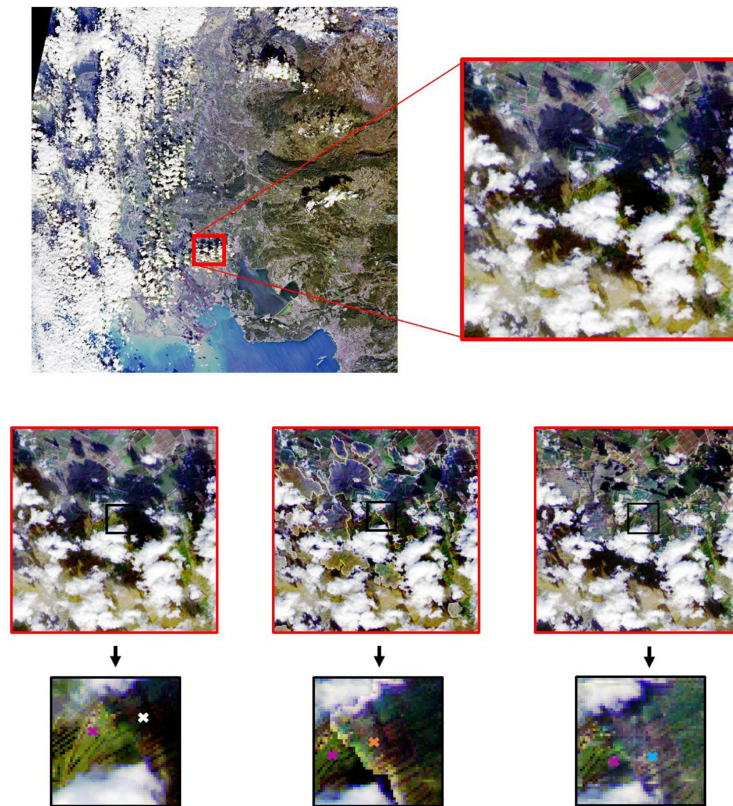
Fig. 10 Top row (left): France RGB = 665/560/490 nm with equalization stretching. Top row (right): subset of scene ID 10 with equalization stretching. Middle row (left to right): subset 2 of France RGB = 665/560/490 nm with equalization stretching; PACO deshadowed subset 2 of France with equalization stretching; and ATCOR deshadowed subset 2 of France with equalization stretching. Bottom row (left to right): zoom of subset 2 of France RGB = 665/560/490 nm with equalization stretching (white cross = cloud shadow pixel; pink cross = clear pixel); zoom of PACO deshadowed subset 2 of France with equalization stretching (orange cross = PACO deshadowed cloud shadow pixel; pink cross = clear pixel); and zoom of ATCOR deshadowed subset 2 of France with equalization stretching (blue cross = ATCOR deshadowed cloud shadow pixel; pink cross = clear pixel).
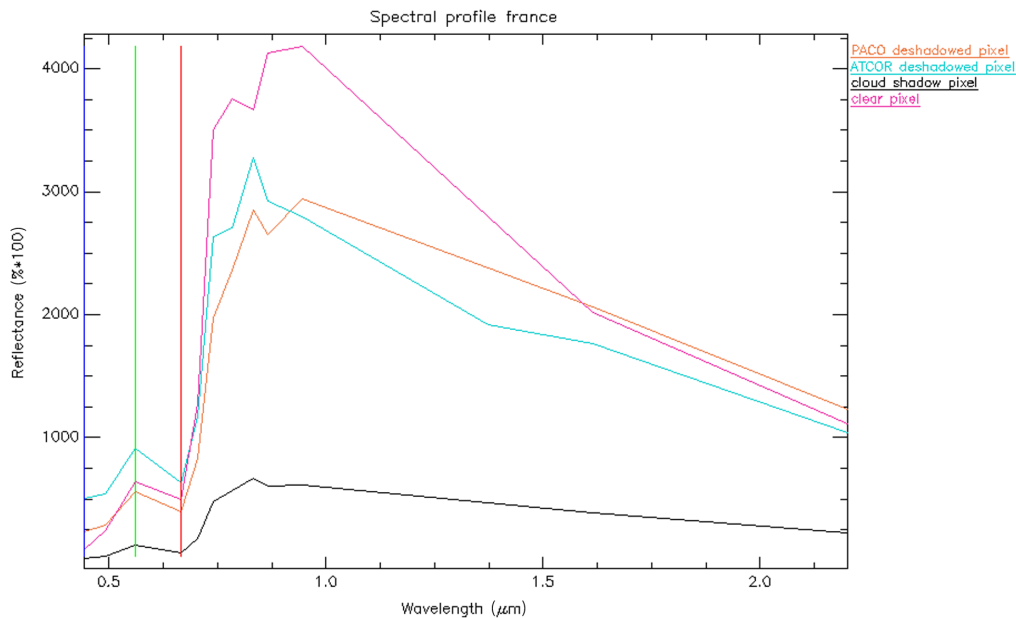
Fig. 11 Spectral profile of scene ID 10 (France). Orange = PACO deshadowed cloud shadow pixel; blue = ATCOR deshadowed cloud shadow pixel; black = original cloud shadow pixel spectrum; and pink = clear pixel spectrum.

relative ratio of the mean reflectance vectors over all spectral bands with deshadowing and without deshadowing. Hence, the mean reflectance of all the clear pixels is divided by the mean reflectance of all cloud shadow pixels. For perfect deshadowing, the value of the relative ratio should lie as close as possible to the perfect value of +1. Depending on the disagreement between clear pixels and cloud shadow pixels, the relative ratio will deviate from +1.

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

- Calculation of the relative ratio R1: ratio of the mean reflectance vector of the clear scene pixels with ATCOR versus the mean reflectance of the shadow pixels but without deshadowing.
- Calculation of the relative ratio R2 for PACO and ATCOR: ratio of the mean reflectance vector of the clear scene pixels versus the mean reflectance of the shadow pixels after deshadowing.
- Comparison of R1 and R2 (see Table 2).

For an improvement in the reflectance vector after deshadowing, the relative ratio R2 should be closer to the value +1 than the relative ratio R1. Table 2 gives the values of R1 and R2 for each scene. R2 was evaluated for the new method, PACO, and for the previous ATCOR method. The bold face numbers of Table 2 indicate the correlation coefficient with a value closest to +1 and hence with the best correlation. No cloud shadows are present in scene ID 5 and 15, hence no values for R1 and R2 are obtained.

As can be deduced from Table 2, all of the relative ratios are improved by the deshadowing algorithm for PACO apart from the case of scene ID 8. The best performance of PACO is obtained with scene ID 1 located in Gobabeb (Africa) with a value of R2 = 0.999. The worst performance of PACO is obtained with scene ID 23 located in Barrax (Spain) with a value of R2 = 0.596. This is due to the value of R1 for this scene to be the worst outlier and hence represents a hard scene to be deshadowed.

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

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

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

spectrum of this cloud shadow will appear lower. This has as a consequence, that the value of R2 is a bit less than the value of R1.

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

5 Discussion

The improved deshadowing algorithm implemented in the python ATCOR (PACO) has shown very promising results and was able to improve the previous matched filter version implemented in ATCOR through its additions. Relative ratios close to the value of +1 are reached as shown in Table 2 with a range between 0.596 and 1.245 with deshadowing, R2 (PACO) and 0.469 and

12.434 without deshadowing, R1. The high values that are reached without going through the cloud shadow removal algorithm can be explained by the presence of haze covering the scene (scene ID 8) or when the scene has a lot of water bodies which are part of the clear pixels (scene ID 17). The cloud shadow removal algorithm performed within PACO is done without any ATCOR on the haze particles. This is one additional step that can be taken into account to further improve the deshadowing algorithm. Hence, to fully correct the scene for the visibility and then perform the cloud shadow removal.

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

The advantages of the deshadowing method presented in this paper are that it is performed through a fully automatic algorithm which is based on the matched filter method, the TIP cloud shadow masking and a iterative process of the scaled shadow function. It works for multispectral and hyperspectral imagery over land acquired by satellite/airborne sensors. Since the deshadowing algorithm relies on the TIP method after,[29] it must be noted that it is applicable for VNIR-SWIR sensors, such as EnMAP,[37] PRISMA,[38] Landsat-8,[39] and Landsat-9,[40] but for VNIR sensors, such as for example the DESIS sensor,[41] the method can only be implemented partially.

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

6 Conclusions

An improved cloud shadow removal algorithm for high spatial resolution optical satellite data was presented. It is based on the matched filter method with the addition of an improved cloud shadow masking and an iterative process for the final reflectance value calculation. Through visual and spectral inspection it was shown that the new method improves the previous de-shadowing algorithm. This was furthermore presented through an evaluation of the relative ratio between the reflectance of clear and cloud shadow pixel and showed promising values for the new method. Future work will have to include the cloud shadow detection improvements of the TIP method and the evaluation of the cloud shadow border correction.

References

1. T. Wang et al., "Effect of solar-cloud-satellite geometry on land surface shortwave radiation derived from remotely sensed data," *Remote Sens.* **9**(7), 690 (2017).
2. M. Segal-Rozenhaimer et al., "Cloud detection algorithm for multi-modal satellite imagery using Convolutional Neural-Networks (CNN)," *Remote Sens. Environ.* **237**, 111446 (2020).
3. B. Dunbar, "NASA's ICESat satellite sees changing world affecting many," https://www.nasa.gov/vision/earth/lookingatearth/icesat_light.html (2004).
4. Z. Li et al., "Multi-feature combined cloud and cloud shadow detection in GaoFen-1 wide field of view imagery," *Remote Sens. Environ.* **191**, 342–358 (2017).
5. M. Marshall et al., "Field-level crop yield estimation with prisma and sentinel-2," *ISPRS J. Photogramm. Remote Sens.* **187**, 191–210 (2022).
6. C. B. Pande et al., "Prediction of soil chemical properties using multispectral satellite images and wavelet transforms methods," *J. Saudi Soc. Agric. Sci.* **21**(1), 21–28 (2022).
7. J. Hu et al., "Quantitative estimation of soil salinity using UAV-borne hyperspectral and satellite multispectral images," *Remote Sens.* **11**, 736 (2019).
8. P. A. Wiefels and C. Baroja, "Red edge detects vegetative stress earlier in plant growth cycle," https://up42.com/blog/tech/red-edge-detects-vegetative-stress-earlier-in-plant-growth-cycle#:~:text=The%20Red%20Edge%20Difference&text=%22Vegetative%20stress%20will%20show%20itself,nm%20sliver%20of%20the%20spectrum (2022).

9. M. Dalponte et al., "Mapping forest windthrows using high spatial resolution multispectral satellite images," *Int. J. Appl. Earth Obs. Geoinf.* **93**, 102206 (2020).

10. F. van der Meer, H. van der Werff, and F. van Ruitenbeek, "Potential of ESA'S sentinel-2 for geological applications," *Remote Sens. Environ.* **148**, 124–133 (2014).

11. H. Lv, Y. Wang, and Y. Shen, "An empirical and radiative transfer model based algorithm to remove thin clouds in visible bands," *Remote Sens. Environ.* **179**, 183–195 (2016).

12. E. Menaka, S. S. Kumar, and M. Bharathi, "Cloud removal using efficient cloud detection and removal algorithm for high–resolution satellite imagery," *Int. J. Comput. Appl. Technol.* **51**(1), 54–61 (2015).

13. R. Eckardt et al., "Removal of optically thick clouds from multi-spectral satellite images using multi-frequency SAR data," *Remote Sens.* **5**(6), 2973–3006 (2013).

14. K. P. Paudel and P. Andersen, "Monitoring snow cover variability in an agropastoral area in the trans Himalayan region of Nepal using MODIS data with improved cloud removal methodology," *Remote Sens. Environ.* **115**(5), 1234–1246 (2011).

15. J. Li et al., "Optimal cloud-clearing for airs radiances using MODIS," *IEEE Trans. Geosci. Remote Sens.* **43**(6), 1266–1278 (2005).

16. A. Gafurov and A. Bárdossy, "Cloud removal methodology from MODIS snow cover product," *Hydrol. Earth Syst. Sci.* **13**(7), 1361–1373 (2009).

17. D.-C. Tseng, H.-T. Tseng, and C.-L. Chien, "Automatic cloud removal from multi-temporal spot images," *Appl. Math. Comput.* **205**(2), 584–600 (2008).

18. F. Díaz et al., "Improving shadows detection for solar radiation numerical models," *Appl. Math. Comput.* **319**, 71–85 (2018).

19. Z. Zhu, S. Wang, and C. E. Woodcock, "Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and sentinel 2 images," *Remote Sens. Environ.* **159**, 269–277 (2015).

20. Z. Zhu and C. E. Woodcock, "Object-based cloud and cloud shadow detection in Landsat imagery," *Remote Sens. Environ.* **118**, 83–94 (2012).

21. S. A. Ackerman et al., "Discriminating clear sky from clouds with MODIS," *J. Geophys. Res.: Atmos.* **103**(D24), 32141–32157 (1998).

22. J. Louis et al., "Sentinel-2 sen2cor: L2a processor for users," in *Proc. Living Planet Symp. 2016*, pp. 1–8, Spacebooks Online (2016).

23. A. Hollstein et al., "Ready-to-use methods for the detection of clouds, cirrus, snow, shadow, water and clear sky pixels in sentinel-2 MSI images," *Remote Sens.* **8**(8), 666 (2016).

24. A. Baraldi and D. Tiede, "Autocloud+, a "universal" physical and statistical model-based 2D spatial topology-preserving software for cloud/cloud–shadow detection in multi-sensor single-date earth observation multi-spectral imagery—part 1: systematic ESA EO level 2 product generation at the ground segment as broad context," *ISPRS Int. J. Geo-Inf.* **7**(12), 457 (2018).

25. H. Zhai et al., "Cloud/shadow detection based on spectral indices for multi/hyperspectral optical remote sensing imagery," *ISPRS J. Photogramm. Remote Sens.* **144**, 235–253 (2018).

26. T. Nazarova, P. Martin, and G. Giuliani, "Monitoring vegetation change in the presence of high cloud cover with Sentinel-2 in a lowland tropical forest region in Brazil," *Remote Sens.* **12**(11), 1829 (2020).

27. A. H. Sanchez et al., "Comparison of cloud cover detection algorithms on Sentinel–2 images of the Amazon tropical forest," *Remote Sens.* **12**(8), 1284 (2020).

28. O. Hagolle et al., "MAJA algorithm theoretical basis document (version 1.0)," *Zenodo* (2017).

29. V. Zekoll, R. de los Reyes, and R. Richter, "A newly developed algorithm for cloud shadow detection–tip method," *Remote Sens.* **14**(12), 2922 (2022).

30. R. Richter and D. Schläpfer, "Atmospheric and topographic correction (atcor theoretical background document)," 2017, https://www.rese-apps.com/pdf/atcor_ATBD.pdf.

31. European Space Agency, "Level-1c," 2022, https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/level-1c/algorithm.

32. E. Ibrahim et al., "Cloud and cloud-shadow detection for applications in mapping small-scale mining in Colombia using sentinel-2 imagery," *Remote Sens.* **13**(4) (2021).

33. R. Richter, "A fast atmospheric correction algorithm applied to Landsat TM images," *Int. J. Remote Sens.* **11**, 159–166 (1990.

34. R. Richter, "Correction of satellite imagery over mountainous terrain," *Appl. Opt.* **37**, 4004–4015 (1998).

35. S. Adler-Golden et al., "An algorithm for de-shadowing spectral imagery," *Proc. SPIE* **4816**, 10 (2002).

36. R. Richter and A. Müller, "De-shadowing of satellite/airborne imagery," *Int. J. Remote Sens.* **26**(15), 3137–3148 (2005).

37. L. Guanter et al., "The enmap spaceborne imaging spectroscopy mission for Earth observation," *Remote Sens.* **7**(7), 8830–8857 (2015).

38. N. Acito et al., "Atmospheric compensation of prisma data by means of a learning based approach," *Remote Sens.* **13**(15), 2967 (2021).

39. B. Markham et al., "Landsat-8 operational land imager radiometric calibration and stability," *Remote Sens.* **6**(12), 12275–12308 (2014).
40. J. Masek et al., "Landsat 9: empowering open science and applications through continuity," *Remote Sens. Environ.* **248**, 111968 (2020).
41. K. Alonso et al., "Data products, quality and validation of the DLR Earth Sensing Imaging Spectrometer (DESIS)," *Sensors* **19**(20), 4471 (2019).

Viktoria Zekoll received her MS degree in physics from the University of Bristol in 2017 and her MS degree in atsrophysics from the AstroMundus program in 2019. In 2023, she received her PhD in mathematics and informatics from the University of Osnabrück. Currently, she is working at Munich Public Observatory.