# EXTRACTION OF CLOUD HEIGHTS FROM SENTINEL-2 MULTISPECTRAL IMAGES

Thomas Krauß

DLR – German Aerospace Center, Münchener Str. 20, 82234 Weßling/Oberpfaffenhofen, Germany

**Commission I, WG I/1**

**KEY WORDS:** Sentinel-2, Inter-Band-Parallax, Clouds, DEM generation, Semi-Global-Matching

**ABSTRACT:**

Investigation of the focal plane assembly of the Sentinel-2 satellites show slight delays in the acquisition time of different bands on different CCD lines of about 0.5 to 1 second. This effect was already exploited in the detection of moving objects in very high resolution imagery as from WorldView-2 or -3 and also already for Sentinel-2 imagery. In our study we use the four 10-m-bands 2, 3, 4 and 8 (blue, green, red and near infrared) of Sentinel-2. In the level 1C processing each spectral band gets orthorectified separately on the same digital elevation model. So on the one hand moving objects on the ground experience a shift between the spectral bands. On the other hand objects not on the ground also show a slight shift between the spectral bands depending on the height of the object above ground. In this work we use this second effect. Analysis of cloudy Sentinel-2 scenes show small shifts of only one to two pixels depending on the height of the clouds above ground. So a new method based on algorithms for deriving dense digital elevation models from stereo imagery was developed to derive the cloud heights in Sentinel-2 images from the parallax from the 10-m-bands. After detailed description of the developed method it is applied to different cloudy Sentinel-2 images and the results are cross-checked using the shadows of the clouds together with the position of the sun at acquisition time.

## INTRODUCTION

Actual high and very-high resolution Satellites like Sentinel-2 or WorldView-2/-3 acquire images scanning the ground line by line while traveling along their orbit. Such a push-broom-scanner consists of different CCD lines for each panchromatic and multispectral band acquiring one line at one time. In the processing of the raw imagery the individual multispectral and panchromatic bands get coregistered using a common digital elevation model (DEM). So moving objects on ground show a shift of the objects between the different bands as shown in fig. 1, left.



Figure 1. Examples from a Sentinel-2 image, left: color shifts due to moving cars, right: color shifts due to non-ground objects (clouds), sections $2.5 \times 2.5$ km$^2$

This effect was already exploited in the detection of moving objects in very high resolution imagery as from WorldView-2 or -3 as shown for example in (Krauß et al., 2013) or (Kääb, 2011). For Sentinel-2 the time shift was used for detection of aircrafts and ships by (Heiselberg, 2019) or for trucks in (Fisser, 2020). On the other hand objects not sticked to the ground also show a slight shift between the spectral bands depending on the height of the object above ground as shown in fig. 1, right. In the study we present in this publication we use this second effect.

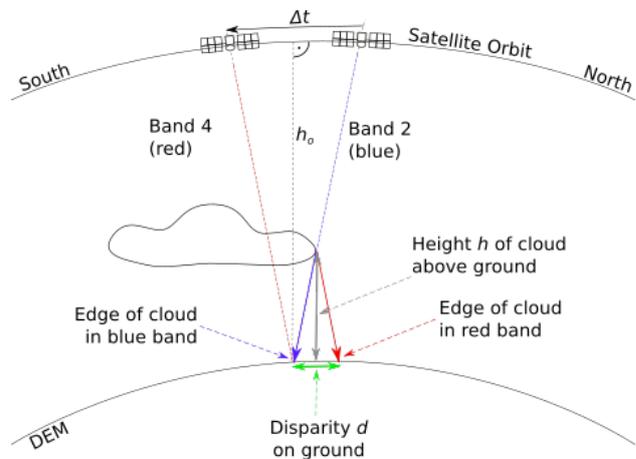The principle of the spectral shift due to non-ground objects is depicted in fig. 2.



Figure 2. Principle of acquisition geometry of image bands separated in a FPA; $h_o$: height of the orbit, $\Delta t$: time lag as listed in table 1.

For exploiting the spectral shift of moving or non-ground objects we need to know the details of the focal plane assembly (FPA) of the sensor. Fig. 3 shows the CCD lines in the FPA of Sentinel-2 (ESA, 2021). The FPA for the visual/near-infrared bands (VNIR) of the multispectral sensor (MSI) is build as a staggered array of 12 CCD elements covering the whole 290-km-swath of Sentinel-2. But the order of the CCD lines is inverted in neighbouring CCD elements. The blue bands of the odd/even sensors are nearest with a time lag of 2.32 s.

Images acquired by push-broom-scanners with focal plane assemblies like those shown in fig. 3 cause small time gaps between image bands clearly separated in the focal plane assembly – e.g. about one second between the blue and red CCD lines – as listed in table 1 from (ESA, 2021).
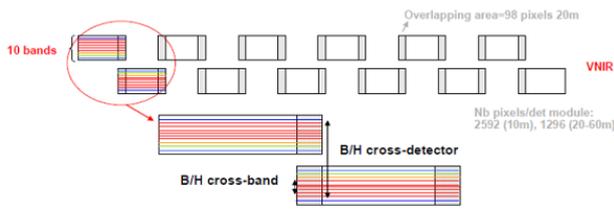
Figure 3. Focal plane assembly of Sentinel-2

Table 1. Sentinel-2 multispectral VNIR bands

| Band | Name | $\Delta t$ [s] | GSD [m] | Sentinel-2A $\lambda$ [nm] | $\beta$ [nm] | Sentinel-2B $\lambda$ [nm] | $\beta$ [nm] |
|------|------|------|------|------|------|------|------|
| B01 | coastal | 2.314 | 60 | 442.7 | 21.0 | 442.3 | 21.0 |
| B02 | blue | 0.000 | 10 | 492.4 | 66.0 | 492.1 | 66.0 |
| B03 | green | 0.527 | 10 | 559.8 | 36.0 | 559.0 | 36.0 |
| B04 | red | 1.005 | 10 | 664.6 | 31.0 | 665.0 | 31.0 |
| B05 | red-edge | 1.269 | 20 | 704.1 | 15.0 | 703.8 | 16.0 |
| B06 | red-edge | 1.525 | 20 | 740.5 | 15.0 | 739.1 | 15.0 |
| B07 | red-edge | 1.790 | 20 | 782.8 | 20.0 | 779.7 | 20.0 |
| B08 | NIR | 0.264 | 10 | 832.8 | 106.0 | 833.0 | 106.0 |
| B08a | narrow-NIR | 2.055 | 20 | 864.7 | 21.0 | 864.0 | 22.0 |
| B09 | water-vapour | 2.586 | 60 | 945.1 | 20.0 | 943.2 | 21.0 |

GSD: Ground sampling distance, $\lambda$: central wavelength, $\beta$: bandwidth, FWHM (full width at half max)

For our study we use only the bands with the highest resolution of 10 m: blue, green and red, but not the NIR band (B08, nead infrared) since the spectral differences are mostly too large of the NIR band to the others (see fig. 9).

Table 2. Sentinel-2 orbit parameters

| Parameter | Symbol | Value |
|-----------|--------|-------|
| Orbit height | $h_o$ | 786 km |
| Period | $t_o$ | 100.7 min |
| Descending node | $N$ | 10:30 |
| Inclination angle | $\alpha$ | $-98.62°$ |
| Earth radius | $R_E$ | 6357.5 km |

Using the orbit parameters for Sentinel-2 as shown in table 2 we can calculate the height over ground $h$ for a given disparity $d$ in metres to:

$$h = d \cdot \frac{h_o}{v_o \cdot \Delta t} \qquad (1)$$

using the orbital speed $v_o$ with orbit period $t_o$ and orbit length $l_o$ calculated from the mean earth radius $R_E$ and the orbit height $h_o$

$$v_o = \frac{l_o}{t_o} = \frac{2 \cdot \pi \cdot (R_E + h_o)}{t_o} = 7.429 \frac{\text{km}}{\text{s}} \qquad (2)$$

A shift of one pixel (10 m) between the blue and the red band corresponds to a height above ground of 1053 m using $\Delta t = 1.005$ s and eq. 1. For normal clouds we will thus expect a shift of about one to three pixels. However cirrus clouds may form in heights of 5000 to 13700 m, so shifts of about five to 14 pixel may occur. Fig. 4 shows a stereo image of an area $8.7 \times 6.5$ km$^2$ containing different types of clouds at different heights. The 3D effect of the different heights can only be observed if the image is viewed using appropriate red-cyan-glasses.



Figure 4. Example stereo image with different cloud heights

After the detailed description of the developed method in section Method the method is applied with different parameter settings to different Sentinel-2 images in section Experiments and Discussion. The results are also discussed and cross-checked using casted shadows of the clouds and knowledge of the position of the sun at acqustion time in Evaluation. Finally a resumé of the work is given in the last section Conclusion and Outlook.

## METHOD

### Background

The proposed method is based on small shifts between spectral bands of Sentinel-2 scenes for non-ground objects as described in section Introduction and shown in fig. 2. If no objects above ground exist there should be no shifts between the bands since they were all coregistered and orthorectified to the same ground DEM. The shifts occur in scanning direction of the satellite. For Sentinel-2 there are only Level-1C and Level-2A data available. Since both are orthorectified the knowledge of the scanning direction and also which pixel belongs to which of the staggered CCD arrays (cf. fig. 3) is lost in the data. In the proposed workflow we have to determine the scanning direction, calculate the shifts and create a digital elevation model (DEM) of the clouds (or any other non-ground objects).

Since a delivered Sentinel-2 scene has normally a size of about $10980 \times 10980$ pixels the scene is split to tiles of size $1000 \times 1000$ pixels and the processing is done on each tile separately. The processing of a tile needs approximately three minutes on a Dell Latitude E7270 laptop under Ubuntu 16.04.

### Workflow

The proposed workflow for deriving cloud heights from parallaxes between different spectral bands of the VNIR Sentinel-2 MSI instrument is shown in fig. 5.

The core of the processing is the derivation of disparities between spectral bands using the well known Semi-Global-Matching (SGM) algorithm (Hirschmüller, 2005) extended and optimized for satellite imagery (dAngelo and Reinartz, 2011). SGM needs the input stereo image pair in epipolar direction with disparities occuring only in one direction – horizontal or vertical. In our case we selected the vertical direction since the Sentinel-2 scenes are acquired roughly from north to south.
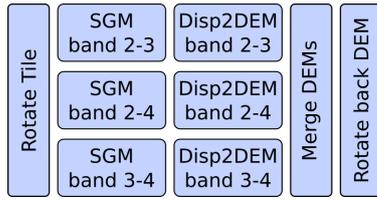
Figure 5. Workflow of proposed method

After determining the scanning direction the input tile is rotated to vertical epipolar direction. Next for a statistically more reliable result for each of the band combinations (2,3), (2,4) and (3,4) (blue-green $\Delta t = 0.527$ s, blue-red $\Delta t = 1.005$ s and green-red $\Delta t = 0.478$ s according to table 1) the shifts or disparities between the spectral bands are calculated using SGM. In the next step these disparities are converted to heights above ground. Finally the three independently derived DEMs are merged and rotated back to the tile geometry. Each step is described in detail in the following subsections.

**Rotating to scan-direction**

For calculating the disparities epipolar images are needed. But as stated in subsection Background the available Sentinel-2 images of processing level 1C or 2A are already rectified and have no more any information on the scanning direction. Luckily Sentinel-2 always scans nadir and in orbit-direction. So knowing the direction of the orbit at the acquisition position we can calculate the scan-angle. Fig. 6 shows the Sentinel-2-orbit and the definition of the inclination angle $\alpha$.
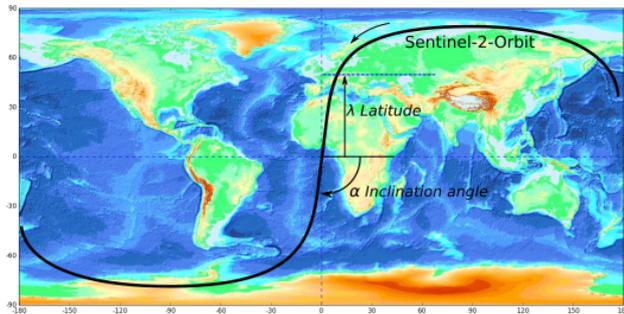


Figure 6. Orbit of Sentinel-2 on day-side from north to south, descending node with Inclination angle $\alpha = -98.62°$

From the inclination angle $\alpha$ at the equator the maximum polar latitude of the orbit can be calculated as

$$\lambda_{max} = 180° - \mid \alpha \mid = 180° - 98.62° = 81.38° \quad (3)$$

At the equator the scan-direction is simply $\alpha$. At the maximum polar latitude in the north and south, $\lambda_{max}$ and $-\lambda_{max}$, the orbit is straight west or the scan-angle is $180°$. So following (Heiselberg, 2019) we can calculate the scan-angle $\phi$ as function of the inclination angle $\alpha$ and the center latitude $\lambda$ of the scene as

$$\cos \phi = \frac{\cos \alpha}{\cos \lambda} \quad (4)$$

Since we want to rotate the scene in vertical epipolar direction the rotation angle of the scene is $90° - \phi$ in clockwise direction. For Frankfurt am Main ($50.11°$ north) $\phi$ is $103.515°$ and thus the images will be rotated by $13.515°$ to the left. For rotating

the tile of size $w_t \times h_t$ pixels a larger tile with additional borders $b$ of

$$b = \frac{\max(w_t, h_t)}{2} \tan(\min(\alpha, 90 - \alpha)) \quad (5)$$

in each direction will be cut from the original image and rotated around the center of the tile.

**Creating dense stereo disparity images**

Since the expected shifts of normal clouds with heights of about 1000 m above ground are only about one pixel for the maximum time distance of $\Delta t = 1.005$ s (band combination blue-red) and only about half a pixel for the other band combinations (blue-green and green-red) the resulting disparity images will be very noisy. To overcome this we create disparity images for all possible band combinations and try to merge them to a statistically more stable result.

Using the rotated VNIR spectral bands 2, 3 and 4 (blue, green and red) we generate DEMs from the stereo-pairs (2,3), (2,4) and (3,4) using the Semi-Global-Matching (SGM) algorithm (Hirschmüller, 2005) with the Census cost function (dAngelo and Reinartz, 2011) and disparities in y-direction from -20 to 20 pixel – corresponding to heights above ground of about $\pm 21055$ m using eq. 6.

**Converting disparities to heights above ground DEM**

Depending if a pixel belongs to an odd or even CCD array the stereo direction is inverted. In one case the red band is acquired 1.005 s before the blue band, in the other it is acquired 1.005 s after the blue band. When using the red and blue bands as stereo pairs for the DEM generation the disparities get inverted or the height of a cloud is in the first case above ground, in the second case below ground. If we assume heights below ground can not occur we can simply take the absolute value of our disparities.

The disparity maps from the SGM processing are floating point images with values $d$ from -20.0 to 20.0 pixels or no-data values for areas where no disparities can be calculated or the forward-backward-test failed. These disparity images are converted to "heights above ground" $h$ using the time lag $\Delta t$ between the spectral bands, the image resolution $\rho = 10$ m/px and eqs. 1 and 2 as

$$h = \mid d \mid \cdot \rho \cdot \frac{h_o}{v_o \, \Delta t} \quad (6)$$

**Merging band-stereo DEMs**

In this processing step the three resulting DEMs for the three band combinations get merged by averaging all non-no-data values. All pixels with DEM-values exceeding $0.25\sigma$ with $\sigma$ as the standard deviation of the DEM-values of this pixel in the three input DEMs will be set to zero height as well as all remaining no-data values. Finally a small morphological opening with an element size of (5,5) and a Gaussian filter of $\sigma = 1.0$ will be applied for smoothing the results.

**Rotating back to tile**

The final step is rotating the resulting DEM back by $90° - \phi$ in counterclockwise direction and cutting the added borders $b$ to fit the result on the input tile.

## EXPERIMENTS AND DISCUSSION

For the experiments we selected two scenes over Germany and one over Austria as listed in table 3.

Table 3. Sentinel-2 scenes used

| ID | Region | Date | Scene-ID |
|---|---|---|---|
| A | Frankfurt am Main | 23.06.2020 | S2A_MSIL1C_20200623T103031 _N0209_R108_T32UMA_ 20200623T124659 |
| B | München | 16.05.2020 | S2B_MSIL1C_20200516T101559 _N0209_R065_T32UPU_ 20200516T123335 |
| C | Innsbruck | 03.05.2020 | S2B_MSIL1C_20200503T100549 _N0209_R022_T32TPT_ 20200503T130430 |

The test regions listed in table 4 are used for the experiments. The Scene-ID refers to the ID given in table 3, the section is given in pixels as left, top, width and height.

Table 4. Test regions

| ID | Purpose | Scene | Section |
|---|---|---|---|
| Fch | Example for scattered clouds | A | 450 200 300 300 |
| Ffm | City center of Frankfurt, can high buildings be derived | A | 7500 4700 500 500 |
| Mch | Clouds in different layers | B | 2000 6800 500 500 |
| Inb | Snow and clouds over mountains | C | 7500 5500 1000 1000 |

### Fch – Estimating processing parameters from a cloudy area north of Frankfurt

The first test case "Fch" uses an area of $3 \times 3$ km$^2$ north of Frankfurt am Main containing some scattered clouds as shown in fig. 7.



Figure 7. Left: Test area Fch with scattered clouds north of Frankfurt, $3 \times 3$ km$^2$, right: Test area Ffm over the city of Frankfurt am Main, $5 \times 5$ km$^2$

Fig. 8 shows the influence of an additional (bicubic) scaling of the multispectral image by factors of 1, 2 and 4. The DEMs in this subsection are scaled from black for ground to white at 2000 m above ground. It can clearly be seen that the band-combinations with $\Delta t$ of only half a second ((2,3) and (3,4)) show much more noise than (2,4) with a $\Delta t$ of 1.005 s. The pre-scaling of the bands before applying the SGM algorithm on the other hand has no real effect. So we can stick to only using scale 1 for further processing.

Fig. 9 shows the stereo pairs containing the near infrared band (NIR, band 8). We see the results are really worse due to the
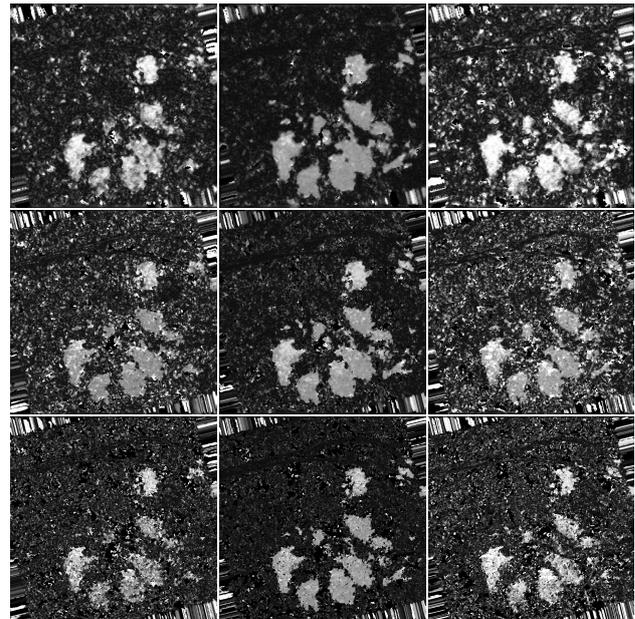


Figure 8. Resulting DEMs, rows: scale 1, 2, 4, columns: DEMs from bands (2,3), (2,4) and (3,4) respectively
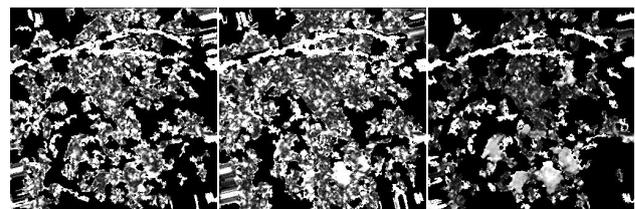


Figure 9. Resulting DEMs containg the NIR band, l.t.r. bands (2,8), (3,8) and (4,8)

bad correlation of the NIR band with the blue, green or red band.

An other effect may be the selection of the cost-function in the SGM algorithm. Beside testing absolute gray differences and many more the best remaining cost functions were MI (mutual information) and Census transform. Operating experience of processing DEMs from normal very high resolution satellite stereo imagery shows also best results for a weighted sum of MI and Census as cost function. Fig. 10 shows the resulting DEMs using different weights of the MI (mutual information) and Census cost functions.
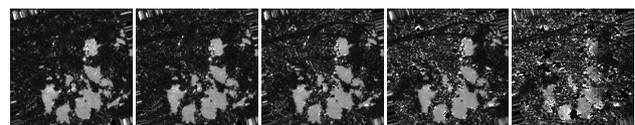


Figure 10. Resulting DEMs using weights of 0, 0.25, 0.5, 0.75 and 1 (l.t.r.) of the MI cost function (Census is one minus the weight of MI) for bands (2,4) and scale 1

As can be seen in our case the simple usage of Census as cost function for the SGM algorithm shows the best results. Beneath this also the processing using MI needs about 22 s while using Census only needs 4 s.

**Ffm – Can we derive building heights?**

Fig. 11 shows a small section of the epipolar image containing the city center of Frankfurt am Main with skyscrapers together with the final merged cloud-height DEM (heights 0–500 m from black to white).



Figure 11. City center of Frankfurt am Main, $1500 \times 1200$ m$^2$, left: multispectral epipolar image, right: merged heights above ground

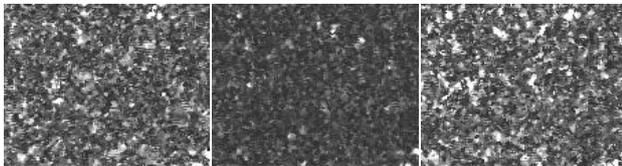Fig. 12 shows the three single DEMs of this area before merging. The noise in the DEMs can clearly be seen.



Figure 12. Single DEMs from bands (2,3), (2,4) and (3,4) for the City center of Frankfurt am Main, $1500 \times 1200$ m$^2$

Looking on the noisy DEMs of fig. 12 and comparing the mutispectral image and the resulting merged DEM in fig. 11 visually we can clearly say, that building heights can not be extracted using the described method.

**Mch – Clouds in different heights**

The test area Mch contains clouds in different heights as shown already in fig. 4. Fig. 13 shows the multispectral image and fig. 14 three DEMs of the clouds calculated using a bicubic scaling of the bands before SGM with factors one, two and four (heights 0–10000 m from black to white).



Figure 13. Test area Mch west of Munich, $5 \times 5$ km$^2$

As can clearly be seen, the thin high clouds at heights of about 6000 to 7000 m above ground vanish in higher scalings.

**Inb – Clouds and Snow over mountains**

The last test case is a very mountainous area around Innsbruck in Austria. Fig. 15 shows the epipolar scene together with the
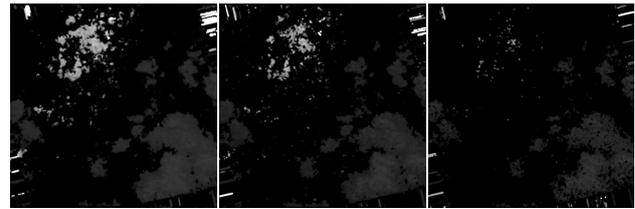


Figure 14. Derived cloud heights of test area Mch, $5 \times 5$ km$^2$, scales 1, 2 and 4

resulting cloud heights above ground (heights 0–3000 m from black to white). The cloud heights seem to go down and up again following the depicted profile line from top left to bottom right.
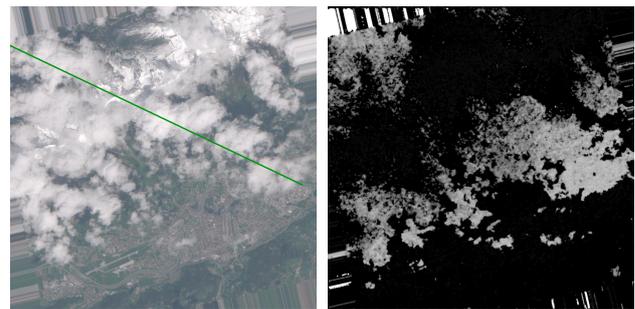


Figure 15. Test area Inb at Innsbruck, $10 \times 10$ km$^2$, the green line represents the position of the profile line used in fig. 16

Fig. 16 shows the profiles of the resulting cloud DEM with heights above ground in green, the SRTM DEM in red and the sum of both – the absolute height of the clouds – in blue. It can clearly be seen, that the derived cloud-heights are really heights above ground since adding the ground-DEM gives a nearly constant cloud-height at 2000 to 2500 m with no clouds at the tops of higher mountains.
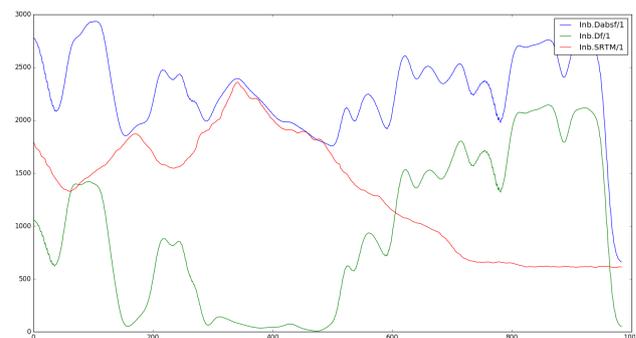


Figure 16. Profile test case Inb, green: cloud heights above ground, red: ground DEM (SRTM 1 arcsec), blue: absolute cloud heights

**EVALUATION**

For evaluation of the results we measure the distance of the cloud positions to the corresponding positions of the cloud shadow as shown in fig. 17. $h_c$ is the cloud height above ground which was derived using our described method. $d$ is the measured distance of a cloud-point with the corresponding shadow-point. $H_c$ and $H_s$ are the heights of the surface model on which the Sentinel-2 scene was orthorectified. $s_z$ is the sun zenith angle for the acquisition time.
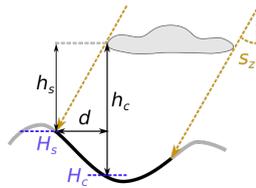
Figure 17. Principle of the measurement of cloud- and shadow-borders

Since we have no reliable azimuth and incidence angles of the looking direction of the satellite we assume nadir looking (which could add a shift the cloud position in the ortho image). For the validation we measure $d$ and take $h$ from our results and $H_c$ and $H_s$ from the fitting SRTM-C-1-arcsec surface model as listed in table 5 for test case Fch. Fig. 18 shows the Fch ortho image together with our derived cloud heights above ground and the positions of the measurements.
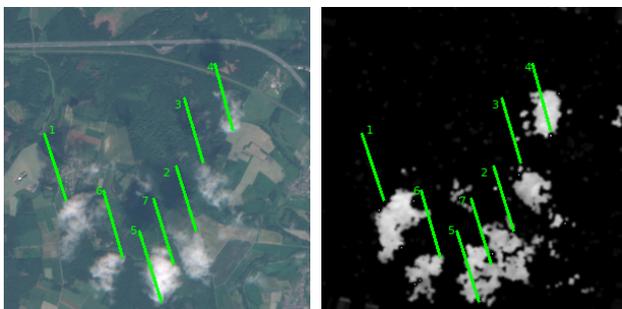


Figure 18. Left: Ortho image Fch, right: cloud heights above ground Fch (black: 0 m, white: 2048 m), green lines: manually measured cloud/shadow borders

Table 5 shows the measured distances and derived cloud-heights. The measured mean azimuth angle is $s_a = 163.46°$ which corresponds to an acquisition time of 11:00:09 UTC and for the center of the test-area at 7.6737°E, 50.5132°N a solar zenith angle of $s_z = 27.76°$.

Table 5. Measured cloud-shadow distances, azimuth angles and heights of the ground at shadow position ($H_s$) and cloud corner ($H_c$) and cloud height above ground from the processing $h$

| ID | $\Delta x$ [m] | $\Delta y$ [m] | $d$ [m] | Azimuth | $H_s$ [m] | $H_c$ [m] | $h$ [m] |
|----|------|------|--------|----------|-----|-----|------|
| 1 | -220 | -670 | 705.20 | 161.822° | 305 | 285 | 1360 |
| 2 | -195 | -650 | 678.62 | 163.301° | 343 | 306 | 1360 |
| 3 | -190 | -650 | 677.20 | 163.706° | 337 | 332 | 1430 |
| 4 | -180 | -670 | 693.76 | 164.962° | 325 | 291 | 1550 |
| 5 | -215 | -705 | 737.06 | 163.040° | 331 | 296 | 1206 |
| 6 | -185 | -675 | 699.89 | 164.673° | 328 | 291 | 1455 |
| 7 | -205 | -660 | 691.10 | 162.745° | 333 | 307 | 1327 |

From the measurement of $d$ in the ortho image and the knowledge of the sun zenith angle $s_z$ we can calculate the height of the cloud above the position of the shadow edge $h_s$

$$\tan s_z = \frac{d}{h_s} \quad (7)$$

Using the difference of the ground heights we get the cloud

height above ground $h_c$ as

$$h_c + H_c = h_s + H_s \quad \text{thus} \quad h_c = \frac{d}{\tan s_z} + H_s - H_c \quad (8)$$

Calculating the differences of the measured $h_c$ (from $d$, cf. eq. 8) versus the automatically derived $h$ (see table 5, right-most column) we get a mean error of 31.03 m with a standard deviation of 137.39 m. Taking into account measurement errors of about one pixel for each end-point we get an error in $h_c$ of about 30 m and errors in deriving the correct $h_c$ in the cloud heights of about 100 m (steep variations at the cloud borders) and errors in the surface model from incorrect endpoints of also about 20 m we estimate an error of about 105 m alone based on the manual measurement.

So the results of the described method seem to be in very good correlation with the manual evaluation. Other errors influencing the result of the proposed algorithm are:

- Terrain height: Since we use the orbit height above ground as $h_o$ in eq. 1 we get a maximum error in $h_o$ of about 8 km or 1 % which results in a $h_c$ also 1 % lower or typical 10 to 20 m – neglectable with general errors in height of about 150 m

- Scan direction: using a wrong scan direction $\phi$ in eq. 6 gives an error of $\cos(\Delta \phi)$ for the disparity $d$, for a maxium error of about 10° this means also derived heights $h_c$ are about 1.5 % too small – also neglectable

- Orbit height: if the orbit height is not always 786 km above the ellipsoid the influence is similar to the one of the terrain height above

- Movement of clouds: for average not too cloudy scenes at most a strong breeze (Beaufort 6, 25 knots or 13 m/s) can be assumed. This means an additional 1.5 pixel disparity in any direction. So the derived cloud heights may vary by about ±1500 m. But those shifts originating by the movement of clouds also influence the position of the shadows and in this case also the borders of the cloud-shadows should show colored borders or additional disparities. In our examples this effect cannot be observed. So the wind speeds at acquisition time were significantly below 10 m/s. This possible error is not neglectable – cross-checking the shadow borders and areas should be done if proper scenes can be found.

Fig. 19 shows positive (green) and negative (red) heights derived from the ortho-image not using the absolute value of $d$ in eq. 6. The five CCD elements of the staggered array can clearly be seen. The full swath of Sentinel-2 is 290 km covered by 12 CCDs so the 109.8 km of scene B only cover 4.5 CCD elements.



Figure 19. Test image B, lowest 500 rows, 10980 columns, top: ortho image, bottom: cloud-heights, green: positive heights above ground, red: negative heights above ground

## CONCLUSION AND OUTLOOK

In this paper we presented a new method for deriving cloud heights from spectral shifts in the visual 10-m-bands of the Sentinel-2 MSI instrument. Typical clouds show a shift of about one pixel between the blue and red band but only about half a pixel between the blue and green or green and red band. Combining derived disparity maps from all three band combinations give a relative consistent and good result. Cross-checking the derived heights with manual measurements of the cloud shadows show an error of about 150 m at cloud heights of 1500 m above ground.

The resulting digital elevation models of the clouds cover mostly a larger area than clouds are visible in the ortho imagery. The method can often also derive heights for very thin, nearly invisible clouds. So also high cirrus clouds can be detected, but since they are mostly very thin their DEM is very perforated and only small parts can be seen in the DEM. Due to the very small shifts between the spectral bands the method is also prone to generate noise in areas not covered by clouds.

In future work the cloud height may be fused with spectral cloud masks and a better filtering may be implemented to get rid of noise in non-cloud regions. Also a point to be solved in subsequent works is the detection of moving clouds. This can be done by detecting the same additional spectral shifts in cloud shadow areas. Since already a moderate wind speed of 10 m/s means one more pixel shift between the blue and red band the heights may be wrong by $\pm 1000$ m.

## REFERENCES

dAngelo, P., Reinartz, P., 2011. Semiglobal Matching Results on the ISPRS Stereo Matching Benchmark. *International Archives of Photogrammetry and Remote Sensing*, XXXVIII-4/W19, 79–84.

ESA, 2021. Sentinel-2 MSI. https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/msi-instrument. (accessed 1/2021).

Fisser, H., 2020. Truck detection using Sentinel-2 data. http://students.eagle-science.org/detect-trucks-using-sentinel-2-data/. (accessed 1/2021).

Heiselberg, H., 2019. Aircraft and Ship Velocity Determination in Sentinel-2 Multispectral Images. *Sensors*, 19(13), 2873. http://dx.doi.org/10.3390/s19132873.

Hirschmüller, H., 2005. Accurate and efficient stereo processing by semi-global matching and mutual information. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2, 807–814.

Krauß, T., Stätter, R., Philipp, R., Bräuninger, S., 2013. Traffic Flow Estimation from Single Satellite Images. *ISPRS Archives*, XL-1/W, 241-246.

Kääb, A., 2011. Vehicle velocity from WorldView-2 satellite imagery. *IEEE Data Fusion Contest*, 2011.