



# Automatic cloud top height determination in mountainous areas using a cost-effective time-lapse camera system

H. M. Schulz<sup>1</sup>, S.-C. Chang<sup>2</sup>, B. Thies<sup>1</sup>, and J. Bendix<sup>1</sup>

<sup>1</sup>Laboratory for Climatology and Remote Sensing, Faculty of Geography, Philipps University, Marburg, Germany

<sup>2</sup>Department of Natural Resources and Environmental Studies, National Dong Hwa University, Hualien, Taiwan

Correspondence to: H. M. Schulz (martin.schulz@geo.uni-marburg.de)

Received: 17 February 2014 – Published in Atmos. Meas. Tech. Discuss.: 21 March 2014

Revised: 29 September 2014 – Accepted: 5 November 2014 – Published: 5 December 2014

**Abstract.** A new method is presented for the determination of cloud top heights using the footage of a time-lapse camera that is placed above a frequently occurring cloud layer in a mountain valley. Contact points between cloud tops and underlying terrain are automatically detected in the camera image based on differences in the brightness, texture and movement of cloudy and non-cloudy areas. The height of the detected cloud top positions is determined by comparison with a digital elevation model projected to the view of the camera. The technique has been validated using data about the cloud immersion of a second camera as well as via visual assessment. The validation shows a high detection quality, especially regarding the requirements for the validation of satellite cloud top retrieval algorithms.

## 1 Introduction

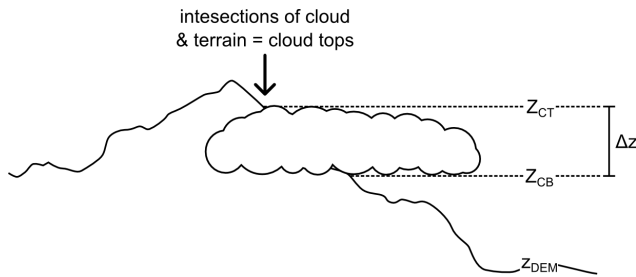
Cloud forests are forest types whose ecology is strongly affected by the frequent occurrence of ground fog (Bruijnzeel et al., 2011). Since they are often hard to access due to their location in mountainous areas, the mapping of cloud forest is difficult in many aspects. An appropriate method for mapping efforts is, however, satellite-borne remote sensing. Mulligan (2010) showed that satellite-derived ground fog frequency maps can be used to distinguish between cloud forest and other forest types by the simple application of a frequency threshold.

Different existing approaches for ground fog detection from the imagery of geostationary (GEO) and low Earth orbit (LEO) satellite platforms do exist (cf. Cermak and Bendix, 2011, for a GEO approach and Welch et al., 2008, for a LEO

approach). In most cases they estimate the cloud base height ( $z_{CB}$ ) for each pixel of an image and compare this value to the terrain height of the pixel ( $z_{DEM}$ ) taken from a digital elevation model (DEM). If  $z_{DEM}$  is equal to or above  $z_{CB}$ , a pixel is considered as foggy because the cloud is assumed to touch the ground and consequently reduces the horizontal visibility to below 1 km (which is the meteorological definition of fog; World Meteorological Organization, 1992). The estimated value for  $z_{CB}$  is calculated by subtracting the cloud geometrical thickness ( $\Delta z$ ) from the cloud top height ( $z_{CT}$ ) (cf. Fig. 1). However, the calculation of  $z_{CT}$ , which is usually done based on the cloud top temperature under the assumption of a certain atmospheric profile, is still causing some problems regarding its precision (cf. Marchand et al., 2010, for a review). State-of-the-art methods for the retrieval of the cloud thickness and therefore of the cloud base height, for instance the one proposed by Cermak and Bendix (2011), are imposing even more uncertainties. This results in a distinction of foggy pixels that is far from perfect. To overcome these problems, an extensive validation of  $z_{CT}$  and  $z_{CB}$  values used in future fog detection algorithms would be helpful.

Typical data that can be used for this purpose are those of visibility sensors, ceilometers and cloud radars. The data of visibility sensors, however, are of low informative value since they can only be used to decide whether a distinct point in space is immersed in clouds or not. Ceilometers and radar devices are expensive and thus generally not available in remote areas like mountain cloud forests.

The lack of cloud height data in remote regions also impedes the design of ground fog detection schemes for mapping fog frequencies. If the intradiurnal dynamic of cloud heights is high, the ground fog detection should be based on



**Figure 1.** Some definitions used in this paper.

GEO data to prevent the resulting cloud frequency map from being heavily biased by the low temporal sampling rate of LEO satellites. Especially in complex terrain, however, the maps could strongly benefit from the high spatial resolution of LEO data. Therefore cloud top and base height data measured with a high temporal sampling rate directly in an investigation area are necessary as a basis for solid decision making regarding whether to use GEO or LEO data for mapping purposes.

Nair et al. (2008) and Welch et al. (2008) compared different approaches for the computation of cloud base heights from Moderate-Resolution Imaging Spectroradiometer (MODIS) data and used them to map cloud immersion frequencies in their study area in the Costa Rican mountains. Because of the aforementioned problems, a comprehensive validation of the different approaches was performed using ceilometer data from US airports located in relatively flat terrain only. It is obvious that the validity of such an approach that transfers the validation results from flat terrain in the US to complex terrain in tropical regions is limited. A validation in the study area itself was only possible on a much smaller scale: photos from cloud-immersed mountains in the study area were taken and used to estimate the height of the intersection between the cloud base and the terrain based on expert knowledge about the area as well as GPS positions of visually prominent features. Due to this non-automatic approach, only a small number of photos could be incorporated into the validation, the significance of which is therefore decreased. The necessity for manual evaluation of the photos makes the approach hardly applicable for any comprehensive statistical investigation such as the analysis of the intradiurnal variability of cloud heights. The above authors, however, stressed the suitability of camera footage for cloud height determination.

A semi-automated approach presented by Bendix et al. (2008) detects clouds in webcam footage by applying manually defined brightness thresholds on the images. The resulting cloud masks are compared to a DEM projected to the camera view so that  $z_{CT}$  and  $z_{CB}$  can easily be determined. This simple approach is not suitable for complex lighting conditions and viewing geometries that entail greatly varying viewing distances to the different image

pixels (cf. Sect. 4.2.3). Although it still needs human intervention, it shows the potential for further automation.

The aim of this paper is to develop and validate a cost-effective method for  $z_{CT}$  determination from camera footage in a cloud forest area of Taiwan with a much higher degree of automation. The method (cf. Sect. 4) has been developed to be in principle suitable for  $z_{CT}$  as well as  $z_{CB}$  depending on the position of the camera. For the current study the camera was placed above a frequently occurring cloud layer (cf. Sect. 2). Therefore the method will be demonstrated and validated for the cloud top height. Results will be presented in Sect. 5 and discussed in Sect. 6. The suitability of the method for  $z_{CB}$  will be addressed in future studies.

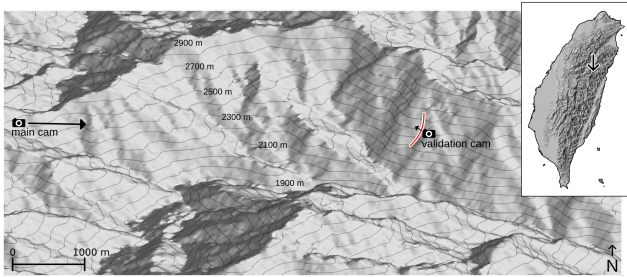
## 2 Study area and camera setup

### 2.1 Study area

Taroko Gorge, located in eastern Taiwan, is famous for a frequently (almost daily) occurring sea of clouds, which can be observed from higher terrain. Since cloud forest is present on the slopes of the gorge, the frequency of ground fog will be mapped using satellite data in a future study. The width of the gorge is between about 3 and 7 km for most of its length. Many smaller side valleys incised into its slope form a complex topography that would be barely recognizable in the low spatial resolution (e.g., up to 1 km per pixel for MTSAT) of GEO satellites. The shape could be much better reproduced from LEO data (e.g., Terra/Aqua imagery with a resolution of up to 250 per pixel). It is, however, not known whether the temporal dynamic of the cloud occurrence can be captured unbiasedly by the sampling rate of polar-orbiting satellites. Therefore the area is ideally suitable for testing a technique that can be used to design and validate methods for ground fog retrieval from satellite data. The area is accessible via roads but due to its sparse population electric power is not available in most parts. Since the terrain is mostly steep and therefore difficult to access, only some places (mostly near roads) can be considered as suitable for the installation of the cameras used for our cloud top height determination approach.

### 2.2 Camera setup

Two cameras of the type PlotWatcher Pro (Day 6 Outdoors, LLC, USA) were installed near the western end of the Taroko Gorge (cf. Figs. 2 and 3). The PlotWatcher Pro is usually used to observe game animals but it is also well suited for cloud top height determination purposes. The waterproof housing and a battery-powered operation mode allow for installation independent of any infrastructure. Due to its construction, the camera can easily be mounted on any pole. In our case, traffic sign posts were used. After setup the camera automatically takes images in an adjustable time interval for a given time span each day. The footage is saved to a SD card.



**Figure 2.** Isometric view of the western end of the Taroko Gorge with the positions of the cameras. The black arrows denote the cameras' viewing directions. The red line marks a small recess in the slope. The location of the area in Taiwan is marked in the upper right overview map.

The positions of the cameras were taken from differential GPS measurements. Their viewing directions were roughly determined using a compass.

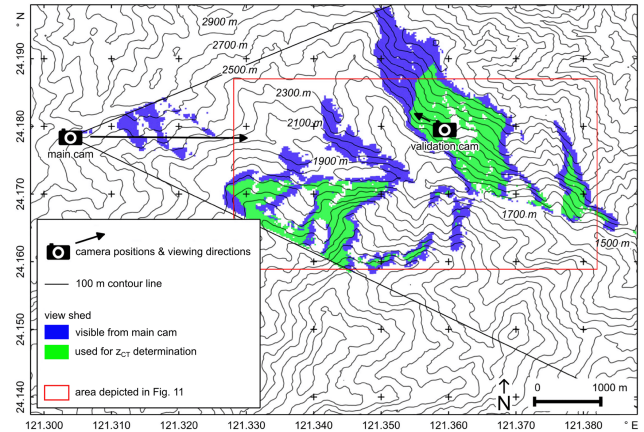
The camera used for the determination of  $z_{CT}$  (further referred to as main cam) was installed at  $24^{\circ}10'42.44''$  N,  $121^{\circ}18'14.18''$  E at a height of about 2681 m on the eastern slope of the North Peak of Hehuan Mountain. The camera is facing eastwards so that it oversees the gorge (cf. Fig. 3). The clouds usually form at different heights below the camera location, but on some days the camera itself is immersed in clouds for approximately 1 h in the afternoon. As the height of the camera sets an upper limit, cloud top heights cannot be detected in these cases.

Another camera (further referred to as validation cam) is installed at  $24^{\circ}10'46.64''$  N,  $121^{\circ}21'33.95''$  E on the northern slope of the gorge. The validation cam is located at a height of 2377 m, which is cloud immersed much more often than the main cam height. Its location can be seen from the validation cam viewpoint. The camera is facing northwestwards onto the valley slope. Because of a small recess in the slope (red line in Fig. 2), its viewing axis intersects the terrain at a distance of about 200 m. The validation cam is used to check whether its position is cloud immersed or not. The footage is needed to validate the cloud top heights derived from the main cam footage (cf. Sect. 4.3.1).

Both cameras in the Taroko Gorge were set up to take pictures from 05:00 to 19:00 UTC + 8 each day in 1 min intervals (resulting in 840 captured frames each day). With these settings they can operate autonomously for several months until the batteries and SD cards need to be replaced.

### 3 Data

In this study, camera footage from 14 March 2013 to 3 May 2013 has been used. For each day the cameras create a separate video file in a resolution of  $1280 \times 720$  pixel. The video files are saved in 8 bit RGB color space ranging from 0 to 255. The footage is largely free of image distortions and



**Figure 3.** Camera positions and the view shed of the main cam.

can therefore be used without further image calibration (cf. Sect. 4.1). An example image of the main cam is shown in Fig. 4.

The only data used in addition to the camera footage are from the Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model 2 (ASTER GDEM 2, property of METI and NASA) as distributed via the USGS global data explorer (United States Geological Survey, 2013). The DEM has a horizontal resolution of 30 m per pixel and the 95 % confidence interval for the vertical accuracy is 17.01 m. The root-mean-square deviation (RMSD) of the height is 8.68 m (Tachikawa et al., 2011).

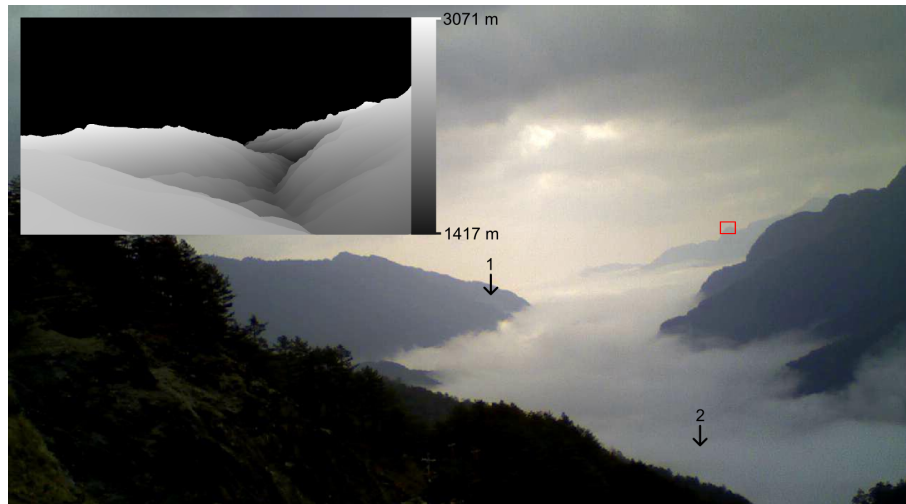
## 4 Methodology

### 4.1 Check for image distortions

The existence of image distortions in the PlotWatcher Pro footage was tested by taking pictures of checkerboard patterns and measuring the straightness of the lines between the rows and columns with an image editing program. Although the image distortion never exceeded about 2 pixels, a camera calibration as described in Zhang (1998) and implemented by Abeles (2012) has been performed for testing purposes. Since the calibration did not further enhance the straightness of the lines, the image is used without any image calibration.

### 4.2 Workflow for the retrieval of $z_{CT}$

The basic idea of Bendix et al. (2008) is adopted in our approach: the DEM is reprojected to the view and resolution of the main cam (cf. Fig. 4, upper left) so that the height of the contact points between the cloud surface and the terrain (further referred to as cloud tops, cf. Fig. 1) that have been detected in the camera image can be read from the corresponding pixels of the DEM. The new approach should, unlike the method by Bendix et al. (2008), be suitable for



**Figure 4.** View of the main cam on 15 March 2013, 10:14 UTC + 8 (main picture) and the DEM projected to the view of the main cam (inset). The mountain top marked by a red square is used in Sect. 4.3.3 for the assessment of errors caused by a misfit between the real camera and the virtual camera. See Sect. 4.2.3 for explanation of the arrows.

batch processing. Therefore several video files (each consisting of 840 frames) need to be processed without human intervention. This higher degree of automation entails a more complex workflow (cf. Fig. 5). It also increases the computation time. Therefore each PlotWatcher Pro video that is used as input is split into groups of five successive frames (further referred to as “scene”). For each group, the mean image is calculated. The analysis of each video is performed based on these mean images instead of single frames. Therefore the output of the presented workflow are  $z_{CT}$  values at a temporal resolution of 5 min.

Scenes are discarded if the location of the main cam itself is cloud immersed. Cloud immersion results in an undifferentiated mean image. Therefore the coefficient of variation of the brightness of all image pixels that are below the horizon of the reprojected DEM can be used to detect that condition. If it is below 0.8 for each color channel, the camera is cloud immersed. Scenes that are too dark to be analyzed (mean brightness of all pixels below the horizon is below 25) are also excluded from further analysis. For all other scenes the image analysis algorithm delineated in Sect. 4.2.3 is used to detect cloud tops in the main cam imagery. The projection of the DEM to the camera perspective and other input derived from the DEM that is used for the image analysis as well as input that is derived from the mean image is described in Sects. 4.2.1 and 4.2.2, respectively.

#### 4.2.1 Input derived from the DEM

The projection of the DEM to the main cam view accounts for the curvature of the Earth. It is conducted under the assumption of a perfectly round geoid with a radius of 6370 km. The reprojected DEM contains information about the distance to the camera calculated using Jcoord (Stott,

2006) and the terrain height for each pixel. The projection is performed using a virtual camera that is defined by the three parameters position (longitude, latitude and height), orientation (rotation around the  $x$ ,  $y$  and  $z$  axis) and focal length. Ideally the parameters of the virtual camera would match those of the real main cam. While the position of the real camera is known from the GPS measurements (cf. Sect. 2.2) with high precision, the rotation is based on the much more imprecise compass readings in the field. Therefore an interactive tool was written that allows for the virtual camera to initially be fine-adjusted manually by using the main cam footage as an overlay over the reprojected DEM.

The rotation and the focal length of the real-world main cam have, however, proven not to be stable over time. The footage shows that they change over the day depending on the position of the sun. This is most probably caused by thermal extension of the camera body. Additionally, the camera could be slightly rotated when the SD card or batteries are replaced. To overcome these problems the virtual camera is automatically readjusted every 24th scene of the camera footage (this corresponds to a time span of 120 min). For this purpose the current mean image is checked for whether it is usable for camera adjustment. This is the case if the horizon is visible (i.e., the view to the horizon is obstructed by neither clouds nor mist). To check for this a buffer of 20 pixels is placed around the virtual horizon. Due to the initial adjustment and the regularly performed automatic adjustments the horizon in the real-world camera image should always be located in the buffer area. For each column of pixels in the buffer, the differences (Euclidean distances in RGB color space) between every two pixels with a vertical distance of 3 pixels to each other is calculated. The ratio between the maximum value (which is high if the horizon is visible) and

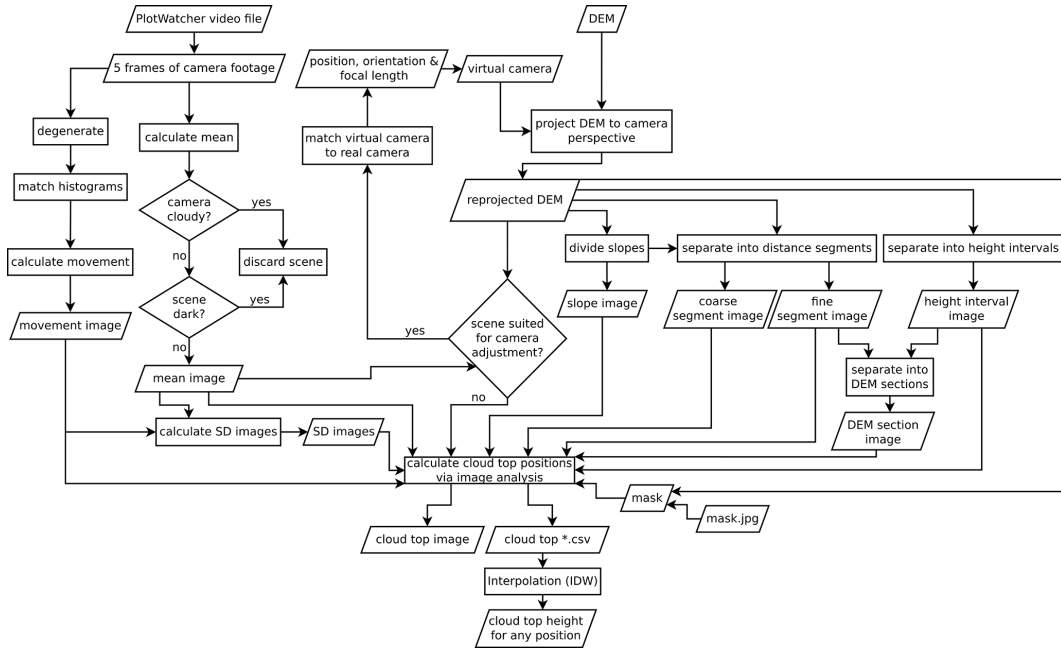


Figure 5. Workflow for the retrieval of  $z_{CT}$ .

the mean value (which accounts for image characteristics as image noise and texture that are also affecting the maximum value) of these differences is calculated for each column and averaged over the whole buffer. If this value exceeds an empirically derived threshold of 6.4, the scene is regarded as usable for camera adjustment. Otherwise, the next suitable scene is used.

To adjust the camera, the fit between the virtual horizon and the horizon in the mean image is calculated. As the horizon can (if it is not obstructed by clouds or mist) be seen as an edge in the camera footage, a simple edge detection is applied: for each pixel of the mean image the sum of the Euclidean distances in RGB space to its neighboring pixels is calculated and written to a new image further referred to as edge image. For each pixel  $p_i$  of the scene that touches the virtual horizon, the sum  $s_i$  of all edge image pixel values  $ep_{xy}$  in a  $10 \times 10$  pixel window surrounding  $p_i$  (with  $x, y$  ranging from  $-5, -5$  to  $5, 5$ ) weighted by the reciprocal value of their distance in pixels to  $p_i$  is calculated. The fit, which quantifies how well the virtual horizon matches with edges in the mean image, is then calculated as the average of all  $N$  values of  $s_i$ .

$$\text{fit} = \frac{\sum_{i=1}^N \left( \sum_{x,y=-5}^5 \frac{ep_{xy}}{\max(\sqrt{x^2+y^2}, 0.5)} \right)}{N} \quad (1)$$

The closer the horizon pixels (i.e., pixels with high edge image values) of the mean image are to the virtual horizon, the higher the fit value.

The actual camera adjustment is done using an iterative algorithm: the virtual camera is rotated clockwise as well as

counterclockwise around each of its three axes one after another by  $0.05^\circ$ . Afterwards, the focal length is also increased and decreased. After each of these eight possible parameter changes the fit is calculated. If it has been increased, the parameter change is maintained and the whole procedure is done from the beginning. Otherwise the change is undone and the algorithm continues with the next parameter change. If every possible parameter change has been conducted without the fit being improved, the camera adjustment is done.

The DEM reprojected to the view of the correctly adjusted virtual camera serves as an input for the image analysis that is used to determine  $z_{CT}$ . Other inputs entirely or partially derived from the reprojected DEM are as follows:

- A mask (white areas in Fig. 6, blue areas in Fig. 3) marks the areas of the camera footage that are not used for  $z_{CT}$  determination. These areas include sky, foreground objects that are included in a manually created JPEG image, terrain with a distance of more than 10 km to the main cam (this corresponds to a natural segmentation of the camera’s view shed: every visible pixel with a distance of more than 10 km to the camera is also at least 13 km away; at such a distance the accuracy with which  $z_{CT}$  can be determined is significantly lowered; cf. Sect. 4.3.3), and areas that are in a vertical buffer of 10 pixels around distinctive edges in the terrain. Edges in the terrain are defined as areas where the difference between the distance from a pixel of the reprojected DEM to the main cam and the distance of its upper neighboring pixel to the main cam exceeds a threshold of 400 (northern slope) or 200 (southern slope) meters (both thresholds are empirically determined and are



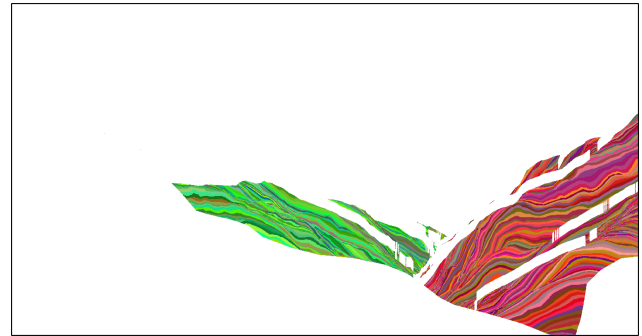
dependent on the slopes' topography and the viewing angle). Near those edges the presented method could provide fundamentally incorrect values of  $z_{CT}$  since a small misfit of the virtual camera parameters could drastically influence the height that is attached to a main cam pixel in these areas. Furthermore they could be mistakenly considered as cloud edges in the image analysis.

All other input images are calculated for areas that are not masked out only.

- The slope image is used to separate between the northern (greenish in Fig. 6) and the southern slope of the valley (reddish in Fig. 6). It is calculated by identifying the pixel with the lowest height of each row of pixels in the reprojected DEM. Every visible DEM pixel to the left of the lowest height pixel is marked as northern slope, and every pixel to the right of it is marked as southern slope. This very simple approach needs to be corrected in order to achieve valid results by removing isolated islands with a bounding box size of less than  $30 \times 30$  pixels of southern slope pixels surrounded by northern slope pixels and vice versa.
- The segment images separate each slope into distance classes. These classes are not based on fixed intervals but are the result of a  $k$ -means clustering of the distance of each pixel of the reprojected DEM to the main cam. This approach was chosen to ensure that each segment corresponds to the natural segmentation of the slopes as seen from the camera as far as possible. In other words, it helps to prevent the view to each segment from being cutoff at some heights by terrain that is nearer to the camera. A fine segmentation (using 400 initial centroids for each slope) as shown in Fig. 6 and a coarse segmentation with a drastically decreased number of classes (six per slope) are performed to obtain both segment images. The numbers of centroids have proven to be suitable for the location in the Taroko Gorge and may need to be adapted for different locations.
- The height interval image is created by subdividing the DEM into several height classes of each 5 m.
- The combination of the height interval image and the fine segment image is further referred to as the DEM section image. It separates the DEM into several small sections, each of which is the intersection of a fine segment and a height interval.

#### 4.2.2 Input derived from the main cam footage

Inputs for the image analysis algorithm derived from the main cam footage are the mean image for the current scene as well as a movement image. The movement image (which, for the sake of computational time, is calculated in a resolution that is decreased by a factor of 4 in each dimension)



**Figure 6.** Segmentation of the terrain into northern (greenish area) and southern (reddish area) slope as well as into fine segments (color shades). The white areas correspond to parts of the image that are masked out.

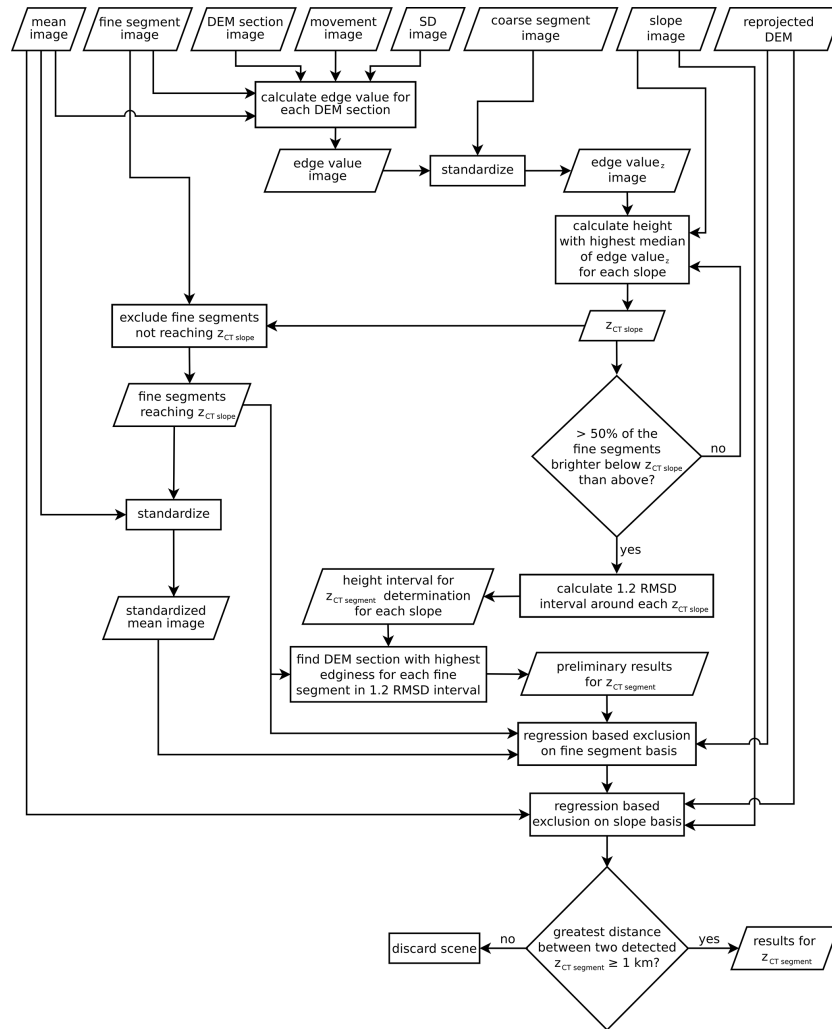
shows the difference between the five frames a scene consists of. For the latter four of these five images the difference measured in terms of the Euclidean distance in RGB space to the previous frame is calculated for each pixel. The movement image is the mean of the four difference images. Before the difference image can be calculated, the latter four frames are histogram-matched (Burger and Burge, 2008) to the first frame of the scene. This helps to overcome the problem of exaggerated movement values that are caused by changes in the brightness of the whole images. The histogram matching is performed for the sky and the terrain separately since the brightness of both areas reacts differently to illumination changes.

For the R, G and B channel of the mean image, as well as for the movement image, an image of the standard deviation (further referred to as “SD image”) is calculated. For each pixel  $p_i$  the SD images contain the standard deviation of the pixel values in a vertical  $40 \times 1$  pixel window surrounding  $p_i$ .

#### 4.2.3 Image analysis for the derivation of $z_{CT}$

The image analysis algorithm for the derivation of  $z_{CT}$  is driven with the input described in Sects. 4.2.1 and 4.2.2.

A simple threshold approach as used by Bendix et al. (2008) is not suitable for this algorithm. In addition to the fact that it is only partially automated and therefore could not handle the amount of scenes the main cam produces over time, the viewing geometry in the Taroko Gorge is too complex to detect clouds by simply applying a brightness threshold. Since the nearest pixels of the mean image that are not masked out have a distance of about 2 km to the main cam while the furthest pixels are about 10 km away, the atmospheric signal that influences the pixel colors varies greatly. As seen in Fig. 4, terrain generally appears to be darker than clouds. If the two pixels marked with the arrows are compared to each other, it becomes apparent that this is an optical illusion. The pixel denoted as 1 has a color value of (R: 137, G: 140, B: 149), while pixel 2, which is much



**Figure 7.** Simplified flowchart for the image analysis in order to derive  $z_{CT}$ . For the sake of simplicity, the mask as an input is not shown as every processing step is restricted to the pixels that are not masked out. In addition, the height extracted from the reprojected DEM is used in a lot more steps than shown (see text).

nearer to the camera, has a color value of (R: 121, G: 126, B: 129). Obviously the distance-dependent influence of mist prevents a simple separation of fog and terrain based on one global brightness threshold for the whole image. This is the reason why the segmentation of the terrain into several distance classes (cf. Sect. 4.2.1) is conducted. A valid separation based on thresholds, however, is not possible, even if in each segment an individual threshold is applied. This is caused by the fact that the sea of clouds is often illuminated irregularly due to its complex surface structure (this is especially of importance if the sun is low) as well as shadows of overlying cloud layers. These local differentiations in the brightness are not dependent on distance. Thus, the problems they cause for a threshold approach cannot be overcome by segmenting the image into distance classes. Additionally the pixel colors of the terrain (e.g., green under sunny, conditions, dark gray under overcast conditions) and the clouds (e.g., white if the

sun is high, reddish if the sun is low) vary greatly over time, which would also complicate the application of thresholds.

The algorithm used instead (cf. Fig. 7) is based on the idea that, in spite of small-scale variations caused, for example, by lighting conditions, the mean pixel value of different input images (mean image (R, G and B channel), movement image and SD images of those images; cf. Sect. 4.2.1) differs between the cloudy part of a fine segment and the non-cloudy terrain. Therefore, for each DEM section and each mentioned input, image the difference between all pixels above and all pixels below or in the DEM section is calculated. The resulting differences are as follows:

- $diff_R$ ,  $diff_G$  and  $diff_B$ . These values contain the color difference between the pixels above and below the DEM section. They are high for cloud tops since clouds are overall brighter than terrain.

- $\text{diff}_{\text{SD}(\text{R})}$ ,  $\text{diff}_{\text{SD}(\text{G})}$  and  $\text{diff}_{\text{SD}(\text{B})}$  can be considered as simple measures for texture differences between the area above and below a DEM section. Their absolute values are high for cloud tops due to texture differences between clouds and terrain.
- $\text{diff}_{\text{movement}}$  and  $\text{diff}_{\text{SD}(\text{movement})}$  account for the degree and structure of movement that is much higher in cloudy areas than for the terrain. Therefore these values are also high in the peripheral area of clouds.

All these differences are non-local and therefore hardly disturbed by small-scale variations in the input images. In addition the following local variables are calculated for each DEM section:

- $\text{mov}_{\text{local}}$  is the average movement at a distinct height level. It is high where the cloud surface is touching the terrain due to the billowing movement of clouds.
- $\text{diff}_{\text{local}}$  is the color difference between a DEM section and the next highest DEM section of the same terrain segment. It is calculated as

$$\begin{aligned} \text{diff}_{\text{local}} = & |\text{R}_{\text{DEMsection}} - \text{R}_{\text{DEMsection}+5\text{ m}}| & (2) \\ & \times |\text{G}_{\text{DEMsection}} - \text{G}_{\text{DEMsection}+5\text{ m}}| \\ & \times |\text{B}_{\text{DEMsection}} - \text{B}_{\text{DEMsection}+5\text{ m}}| \end{aligned}$$

and is high for height intervals that contain a cloud edge, as well as for other edges in the image.

The local variables are high at the height of  $z_{\text{CT}}$  as well as in some other areas of the image. Also, the non-local differences are not necessarily high in cloud top areas only but all local and non-local variables have in common that their absolute value is high at least in DEM sections at  $z_{\text{CT}}$ . Therefore the absolute values of all variables are combined to their geometric mean for each DEM section. Since its calculation involves multiplying all input variables, a high geometric mean is particularly obtained for the edges between clouds and terrain where each of the mentioned variables is high. Therefore the geometric mean is further referred to as edge value. It is calculated for each pixel that is not masked out.

Since cloud tops are less distinct in distant terrain and therefore most edge values cannot be compared to each other, the standard score of the edge value was calculated for different distance classes. The result (further referred to as edge value $_z$ ) is shown in Fig. 8 for a scene that includes the frame shown in Fig. 4. The distance classes used for the standardization are the coarse segments as many fine segments do not cover the whole height of a slope, and hence some of them would be entirely situated above or below  $z_{\text{CT}}$ . Low edge values that are not caused by cloud tops would be exaggerated in these segments after standardization.

For each slope, the height interval with the highest median of the edge value $_z$  is calculated. This height (red line in Fig. 8) can be considered as the mean  $z_{\text{CT}}$  for each slope

(further referred to as  $z_{\text{CT slope}}$ ). This is, however, only true if clouds are present and if the maximum of the edge value at the corresponding height is actually caused by the cloud top instead of the cloud base. The latter is tested for each fine segment of the image using the absolute values of  $\text{diff}_{\text{R}}$ ,  $\text{diff}_{\text{G}}$  and  $\text{diff}_{\text{B}}$ . If more than 50 % of the fine segments of one slope are brighter (this means at least two of three color channels are brighter) below  $z_{\text{CT slope}}$  than above this height, it is actually accepted as the height of the cloud top surface. Otherwise  $z_{\text{CT slope}}$  is the height with the next highest median of the edge value for which more than 50 % of the fine segments are brighter below. If such a height does not exist, the slope is marked as cloudless.

Since the surface of the sea of clouds is more or less plain, for each fine segment the cloud top should be near to  $z_{\text{CT slope}}$ . Therefore  $z_{\text{CT}}$  for each fine segment (further referred to as  $z_{\text{CT segment}}$ ) is assumed to be located inside a height interval around  $z_{\text{CT slope}}$  (area between the dashed lines in Fig. 8). The more even the cloud surface, the smaller the interval. For each slope, its size is calculated from the RMSD of the height of all  $N$  DEM sections to  $z_{\text{CT slope}}$ . In the calculation of the RMSD, the height distance of each DEM section to  $z_{\text{CT slope}}$  is weighted by its edge value $_z$ :

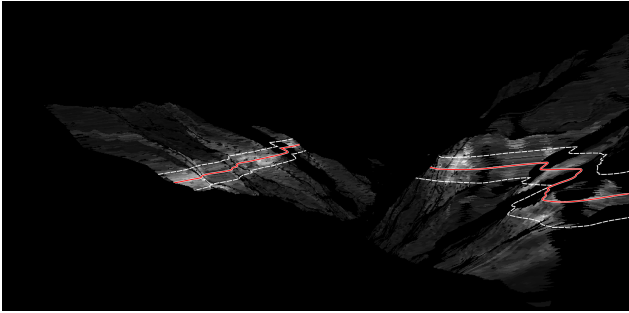
weighted RMSD = (3)

$$\sqrt{\frac{\sum_{i=1}^N \left( (\text{height}_i - z_{\text{CT slope}})^2 \times (\text{edge value}_z)_i \right)}{\sum_{i=1}^N (\text{edge value}_z)_i}}$$

Visual evaluation has shown that the weighted RMSD times 1.2 is a reasonable size for the height interval in which  $z_{\text{CT segment}}$  is determined.

Fine segments that do not reach  $z_{\text{CT slope}}$  may also not reach their  $z_{\text{CT segment}}$ . Therefore they are excluded from further analysis. For all other fine segments the height inside the interval around  $z_{\text{CT slope}}$  with the highest edge value $_z$  is marked as the preliminary result for  $z_{\text{CT segment}}$ . The 1.2 RMSD interval may not necessarily contain the cloud top of each fine segment.  $z_{\text{CT segment}}$  would be detected at a wrong height in these cases. Therefore the validity of each  $z_{\text{CT segment}}$  is checked individually by using a linear regression between the height distance of the DEM sections of the associated fine segment to  $z_{\text{CT segment}}$  and the image brightness (mean value of the R, G and B channel standardized on a fine segment basis). This regression is carried out for DEM sections that are located in a height interval of RMSD times 1.2 around  $z_{\text{CT segment}}$  only. Since clouds are overall brighter than non-cloud- (and non-snow-) covered terrain (if analyzed for each fine segment separately), fine segments for which the slope of the regression is above an empirically derived value of  $-0.01$  standard deviations per meter of height most probably do not contain cloud tops inside the 1.2 RMSD interval.





**Figure 8.** Standardized edge value for the scene from on 19 March 2013 from 10:12 to 10:16 UTC + 8 (see also Fig. 4). The higher the edge value  $z$ , the brighter the shade.  $z_{CT\ slope}$  is marked with the red line, while the interval of 1.2 times RMSD is the area between the dashed lines.

In the calculation of  $z_{CT\ slope}$  and the preliminary results for  $z_{CT\ segment}$ , the absolute (not standardized) edge value has not been respected. Therefore it is unclear whether the detected edges are distinct enough to be cloud tops. In the absence of clouds they could also be any edge in the terrain. Consequently, for each valley slope, the validity of the detected edges as cloud top indicators is tested. For heights that are located in an interval of RMSD times 1.2 around the preliminary  $z_{CT\ segment}$  of each associated fine segment, a linear regression between the height distance of each DEM segment to the corresponding preliminary  $z_{CT\ segment}$  and the image brightness (mean value of the non-standardized R, G and B channel) is performed. If the absolute value of the regression slope is above an empirically determined threshold of 0.1 per meter of height, the edge can be considered as distinct enough to be a cloud top. Otherwise the corresponding valley slope is most probably not cloudy and therefore excluded from further analysis.

Every test described above is performed under the assumption that clouds in the Taroko Gorge occur as a more-or-less stratiform sea of clouds or at least touch the terrain. In these cases it is valid to read the height and position of cloud tops that have been identified in the two-dimensional mean image from the three-dimensional reprojected DEM. In other cases this method would cause wrong results for  $z_{CT}$ . Therefore scenes with a horizontal distance of less than 1 km between the two detected cloud tops that have the greatest distance to each other are classified as not applicable for further use.

For usable scenes the detected cloud top positions are marked in the mean image (cf. Fig. 9) and/or written to an output file for each scene of a video file. These files contain the heights as well as the geographical location of all detected cloud top positions and may be further processed in different ways. In this study the  $z_{CT}$  values were spatially interpolated using inverse distance weighting (IDW) with an exponent of 2 (Shepard, 1968). The interpolation was conducted for the area of the bounding box of all detected cloud top positions (cf. Sect. 5).



**Figure 9.** Cloud tops (green) detected in the scene taken on 19 March 2013 from 10:12 to 10:16 UTC + 8.

#### 4.2.4 Adaptation of the method for the derivation of $z_{CB}$

For the detection of  $z_{CB}$  instead of  $z_{CT}$  some minor changes in the method presented in Sect. 4.2.3 are necessary. In detail these adaptations are follows:

- The non-local difference values for each DEM section need to be calculated between all pixels below the section and all pixel in or above the section.
- $diff_{local}$  needs to be calculated as follows:

$$diff_{local} = |R_{DEM\ section} - R_{DEM\ section-5m}| \quad (4)$$

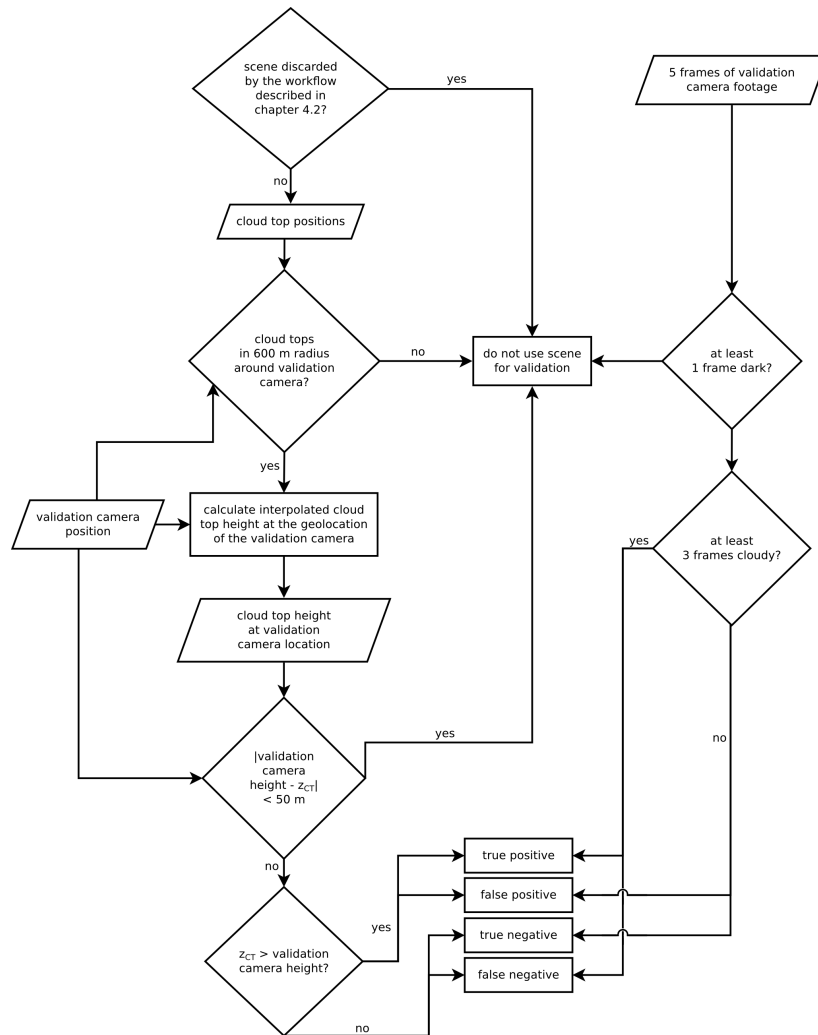
$$\times |G_{DEM\ section} - G_{DEM\ section-5m}|$$

$$\times |B_{DEM\ section} - B_{DEM\ section-5m}|.$$

- While a test is performed for each detected  $z_{CT\ slope}$  to check whether it is actually a cloud top height, for each  $z_{CB\ slope}$  it needs to be tested whether it is the cloud base height. This is done by checking whether more than 50 % of the fine segments of the corresponding slope are brighter above  $z_{CT\ slope}$  than below this height.
- For the validity check of each  $z_{CB\ segment}$ , a threshold of 0.01 per meter of height (instead of  $-0.01$  per meter of height for  $z_{CT\ segment}$ ) has to be used on the slope of the regression.

#### 4.3 Validation

Cloud top heights derived from daily camera footage taken between 14 March and 3 May 2013 were used to validate the method for cloud top height determination presented in Sect. 4.2. This corresponds to a total number of 8400 scenes. For these scenes the cloud top positions were calculated and validated using two different approaches. This was done by summarizing the validation results in three confusion matrices for each validation approach. One of these matrices contains all incorporated scenes. Another matrix contains only scenes that were taken under complex lighting conditions



**Figure 10.** Workflow used for the automated validation.

defined by a sun elevation between  $5$  and  $15^\circ$  (calculated using code from FMet (Cermak et al., 2008), additionally taking into account atmospheric refractions as described by Sæmundsson, 1986). Below a sun elevation of  $5^\circ$  the valley is located in the shadows of the surrounding mountains, which would result in an indirect illumination that could not be considered as a complex lighting condition. A third matrix contains only scenes in which the detection area was visibly affected by shadows of overlying broken cloud layers. From these matrices the following statistical measures (Jolliffe and Stephenson, 2003; Matthews, 1975) were calculated (see Appendix A for formulas):

- proportion correct (PC),
- bias,
- probability of detection (POD),
- probability of false detection (POFD),

- false alarm rate (FAR),
- Hanssen–Kuipers discriminant (HKD),
- Matthews correlation coefficient (MCC).

#### 4.3.1 Validation of retrieved cloud top heights using the validation cam

An automated validation of the cloud top positions that were calculated for each of the 8400 scenes has been performed using the approach shown in Fig. 10. Scenes that have been discarded as not usable for cloud top determination by the method described in Sect. 4.2 are also excluded from the validation. For the remaining scenes it is checked whether cloud tops in a horizontal radius of 600 m around the validation cam exist. If that is the case,  $z_{CT}$  for the location of the validation cam is calculated via IDW interpolation. Otherwise the scene is not used in the validation since the interpolation over long distances would cause too much error for accurate

validation. Scenes for which the interpolated height at the cloud cam position has a vertical distance of less than 50 m to the validation cam are also excluded from the validation. Thus the validation results show to what extent the presented method is suitable for determining whether  $z_{CT}$  is above or below the height of the validation cam with a precision of 50 m.

The interpolated value for  $z_{CT}$  at the validation cam position is compared with the validation cam footage. For this purpose, each of the five validation cam frames that correspond to the current scene of the main cam is analyzed. As a first step it is checked whether at least one frame is too dark for further analysis (mean brightness of the mean of all three color channels  $< 40$ ). In these cases the whole scene is excluded from the validation. For usable scenes it is checked whether the validation cam is cloud immersed in at least three of five frames of the scene. If this is the case, the validation cam is regarded as cloud immersed for the whole scene.

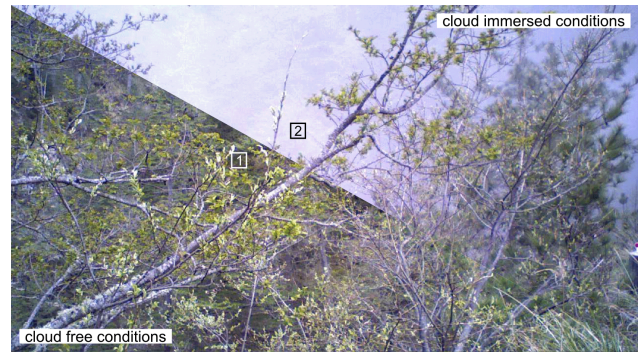
Depending on the height of the interpolated value for  $z_{CT}$  (below or above the validation cam height?) and the information from the validation cam footage (cloud immersed or not?) each scene is registered as a “true positive”, “true negative”, “false positive” or “false negative” in the confusion matrix. This is done under the assumption (which is based on the manual analysis of several scenes as well as field studies) that the sea of clouds is thick enough to ensure that the validation cam is cloud immersed if it is located below  $z_{CT}$ .

The test for whether the validation cam is immersed in clouds is based on the contrast in its image. As Fig. 11 shows, the contrast is low in parts where the view is obstructed by clouds, while it is high in areas where the image shows vegetation on the valley slope. Since nearby vegetation causing high contrasts is visible even under cloud-immersed conditions, a test for low contrast cannot be performed on the whole image. The vegetation cannot be masked out as it grows over the year and moves in the wind. The image is separated into boxes of  $30 \times 30$  pixels instead (two of which are marked as examples in Fig. 11). For each of these boxes the coefficient of variation ( $c_v$ ) is calculated for the red, the green and the blue channel. The maximum  $c_{v \max}$  of these three values is determined and the sum

$$s = \sum_{i=1}^N (0.03 - c_{v \max}) \quad (5)$$

is calculated for all  $N$  boxes for which  $c_{v \max}$  is below an empirical threshold of 0.03. As  $c_{v \max}$  is low for boxes with a cloud-obstructed view due to their low image contrast,  $s$  is high under cloudy conditions. Above a threshold of 0.5 the validation cam is regarded to be cloud immersed.

The results of this analysis were manually verified. For a time span of 10 days (14 March to 3 May 2013), each scene of the validation cam footage was visually checked for cloud immersion. Discrepancies between the visual analysis and the algorithm’s result did not occur.



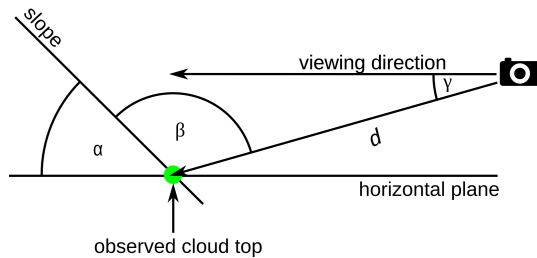
**Figure 11.** Image of the validation cam under cloudy and cloud-free conditions.  $c_{v \max}$  is 0.230 for the  $30 \times 30$  pixels box denoted “1” and 0.020 for the box denoted “2” (see text for further explanation).

The validation-cam-based approach considers only one point in space, namely the validation cam position. Therefore it would not be known whether clouds are completely absent or merely below the camera for scenes in which the validation cam image does not show cloud immersion. Hence, the over- and underestimation of cloud occurrence cannot be reasonably derived from the validation cam footage. This is the reason why all scenes that were excluded from the cloud top height determination need to be excluded from the validation too. After this exclusion, the validation results indicate the fraction of detected cloud tops with correctly retrieved heights.

#### 4.3.2 Visual validation of cloud top detection

Since the automatized validation has been performed under the assumption of correctly determined cloud presence, a second, visual approach has been conducted. Twenty percent of the 8400 main cam scenes were randomly chosen, and detected cloud top heights that were marked as shown in Fig. 9 were manually assessed. Only night scenes and scenes in which the main cam was cloud immersed were excluded from the analysis. The scenes were added to the confusion matrix as follows:

- Cloud tops are present in the detection area (green area in Fig. 3). They were detected at the correct position: true positives.
- No cloud tops in the detection area. Therefore no cloud tops were detected: true negatives.
- Despite the absence of cloud tops in the detection area, cloud tops were detected. Alternatively, cloud tops are present and were detected, but they were detected at the wrong height: false positives.
- Cloud tops are present in the detection area but they were not detected: false negatives.



**Figure 12.** Definitions used in Eqs. (6) and (7).

The disadvantages of this manual validation approach are the decreased number of scenes that could be included as well the fact that it has been carried out based on two-dimensional images. Therefore it does not provide any information regarding whether the height determination of the cloud tops detected in 2-D images using the three-dimensional terrain is valid.

#### 4.3.3 Basic error analysis

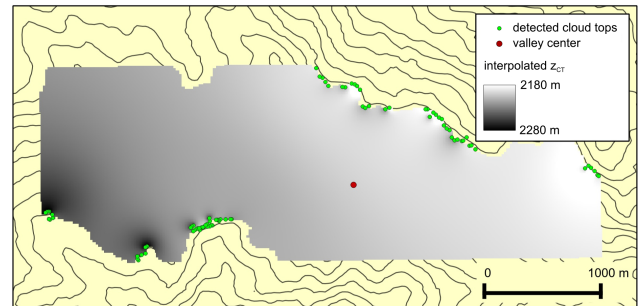
For heights of cloud tops considered as true positives in the visual validation, a simple error analysis was carried out. Three main error sources were identified:

1. Misfits between the reprojected DEM and the real landscape as seen from the main cam. They are caused by an imperfect automatic adjustment of the virtual camera (cf. Sect. 4.2.1) as well as small image distortions (cf. Sect. 4.1). To quantify the size of this error, 20 randomly chosen scenes with a correctly detected appearance of cloud tops and a visible horizon were analyzed. For each scene the deviation in pixels between the position of a mountain top (in a distance of  $\sim 15\,300$  m to the main cam, cf. Fig. 4) in the mean image and in the reprojected DEM was calculated. The RMSD calculated from these deviations ( $\text{RMSD}_1$ ) is 2.73 pixels. These 2.73 pixels correspond to a height difference of

$$\Delta h = 2.73 \text{ px} \cdot (|d| \cdot \cos(\gamma) / 1333 \text{ px}) \cdot (1 / \cos(\beta - 90^\circ)) \cdot \sin(\alpha), \quad (6)$$

with  $|d|$  being the length of the vector  $d$  from the camera to the observed cloud top and  $\gamma$  being the angle between the camera's viewing direction and  $d$ .  $\beta$  is the angle between  $d$  and the slope on which the cloud heights are measured.  $\alpha$  is the steepness of the terrain measured as an angle (cf. Fig. 12).

2. Blurriness of the cloud tops. In Sect. 4.3.1, cloud tops were considered as true positives if they had been correctly detected at the transition between cloud and non-cloud areas. This transition, however, is not a distinct edge in all cases and may extend over altitude differences of several meters. The “real” cloud top height –



**Figure 13.** Cloud tops heights derived from the main cam footage taken on 15 March 2013 from 10:12 to 10:16 UTC + 8. See Fig. 3 for orientation.

for instance defined as the height at which the World Meteorological Organization's definition of fog (cf. Sect. 1) is fulfilled – is located somewhere in this transition zone. For 20 randomly chosen scenes with a correctly detected cloud top appearance, these transition zones were visually identified in the two-dimensional mean image and their height was calculated using the reprojected DEM. It was assumed that the deviation between real cloud top heights as well as the detected cloud top heights would be typically half the height of the transition zone. The RMSD calculated from these deviations ( $\text{RMSD}_2$ ) is 41.93 m.

3. Imprecisions of the ASTER GDEM 2. The RMSD for the ASTER GDEM 2 ( $\text{RMSD}_3$ ) calculated using GPS measurements is given as 8.68 m by Tachikawa et al. (2011).

If the RMSD is assumed to be the typical uncertainty caused by each of these error sources, according to the Gaussian law of error propagation for additive magnitudes, the uncertainty  $u_h$  of the retrieved cloud top heights can be calculated as

$$u_h = \sqrt{\text{RMSD}_1^2 + \text{RMSD}_2^2 + \text{RMSD}_3^2} \quad (7)$$

$$= \sqrt{(2.73 \text{ px} \cdot (|d| \cdot \cos(\gamma) / 1333 \text{ px}) \cdot (1 / \cos(\beta - 90^\circ)) \cdot \sin(\alpha))^2 + 1833.47}$$

For a vertical slope, directly facing orthogonally towards the viewing direction of the camera at a distance of 2500 m and with the camera being orientated horizontally, this would result in  $u_h = 43.12$  m.

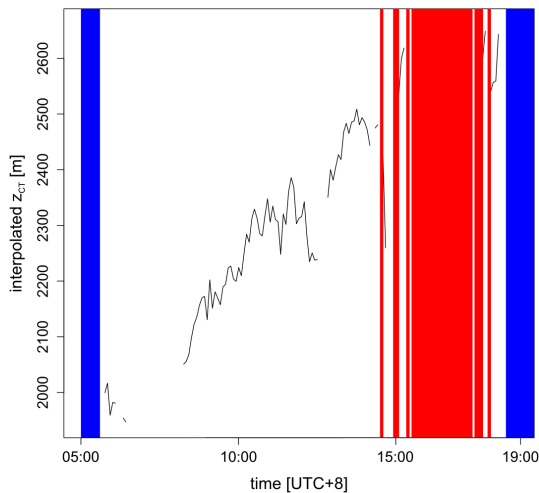
## 5 Results

An example output of the cloud top height determination algorithm is shown in Fig. 9. Interpolated cloud tops heights are shown in Fig. 13. Both figures are derived from the five frames taken on 15 March 2013 from 10:12 to 10:16 UTC + 8 and show a sea-of-cloud situation as it typically occurs in the Taroko Gorge. Video



**Table 1.** Confusion matrices for the automated validation.

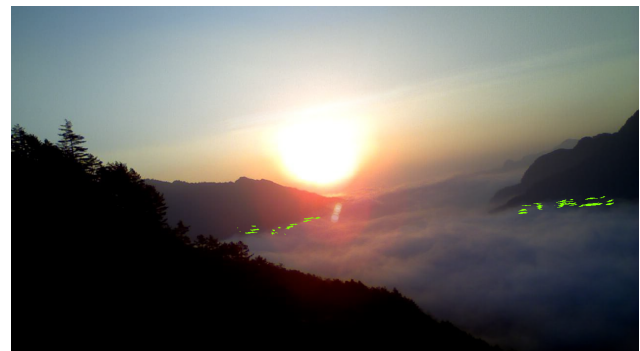
Validation cam cloud immersed?	Validation cam below interpolated $z_{CT}$ ?		
	True/1	False/0	
True/1	$n_{11} = 262$	$n_{10} = 9$	all lighting conditions
False/0	$n_{01} = 49$	$n_{00} = 835$	
True/1	$n_{11} = 41$	$n_{10} = 1$	complex illumination
False/0	$n_{01} = 10$	$n_{00} = 154$	
True/1	$n_{11} = 34$	$n_{10} = 2$	cloud shadow affected
False/0	$n_{01} = 1$	$n_{00} = 144$	



**Figure 14.** Increase in the cloud top height on 15 March 2013 for the position marked as “valley center” in Fig. 13.

examples for the whole day can be downloaded from [http://lcrs.geographie.uni-marburg.de/uploads/media/3\\_15\\_2013\\_CTH\\_camView.zip](http://lcrs.geographie.uni-marburg.de/uploads/media/3_15_2013_CTH_camView.zip) (camera view as shown in Fig. 9, digital object identifier (DOI): doi:10.5678/LCRS/MS.2) and [http://lcrs.geographie.uni-marburg.de/uploads/media/3\\_15\\_2013\\_CTH\\_isoView\\_01.zip](http://lcrs.geographie.uni-marburg.de/uploads/media/3_15_2013_CTH_isoView_01.zip) (isometric view as shown in Fig. 2, doi:10.5678/LCRS/MS.3). The original footage of the main cam is downloadable from [http://lcrs.geographie.uni-marburg.de/uploads/media/3\\_15\\_2013\\_rawVideo.zip](http://lcrs.geographie.uni-marburg.de/uploads/media/3_15_2013_rawVideo.zip) (doi:10.5678/LCRS/MS.1). A plot of  $z_{CT}$  (interpolated to the position in the valley center marked with a red dot in Fig. 13) versus time is shown in Fig. 14 for the same day. The increase in  $z_{CT}$  over time that can be observed in the original footage of the main cam is well reproduced, and times in which the view of the camera was obstructed by clouds (red areas) were also correctly detected. The data gap between about 06:30 and 08:15 UTC + 8 is caused by cloud tops that are too low to be detected.

The algorithm also works for more complex lighting conditions (cf. Fig. 15) and for scenes affected by the shadows of overlying broken cloud cover (cf. Fig. 16).



**Figure 15.** Cloud tops (green) detected in the scene taken on 19 March 2013 from 06:12 to 06:16 UTC + 8.

### Validation results

Table 1 shows the result of the validation based on the validation cam footage. The PC calculated from this matrix is 0.9498 (complex illumination: 0.9466; cloud shadow affected: 0.9889). It is the only statistical measure described in Sect. 4.3 that can be interpreted in a meaningful way for the validation cam approach.

The results of the visual validation are shown in Tables 2 and 3. These results show that the presence and absence as well as the height of cloud tops in the two-dimensional camera footage were correctly determined in 84.73 % (complex illumination: 83.23 %; cloud shadow affected: 90.00 %) of the scenes (PC) while the frequency of cloud tops in the area of detection was slightly underestimated in general (bias: 0.8418; complex illumination: 0.8438; cloud shadow affected: 0.9927). Cloud tops were detected in 77.21 % (complex illumination: 78.13 %; cloud shadow affected: 91.97 %) of the scenes in which cloud top were present (POD) and in 7.35 % (complex illumination: 9.23 %; cloud shadow affected: 13.70 %) of the scenes where no cloud tops were present in the detection area (POFD). This corresponds to about 8.28 % (complex illumination: 7.41 %; cloud shadow affected: 7.35 %) of the determined cloud tops being mistakenly detected (FAR).

**Table 2.** Confusion matrices for the visual validation.

Cloud tops present in detection area?	Cloud tops detected at correct position?		
	True/1	False/0	
True/1	$n_{11} = 454$	$n_{10} = 134$	all lighting conditions
False/0	$n_{01} = 41$	$n_{00} = 517$	
True/1	$n_{11} = 75$	$n_{10} = 21$	complex illumination
False/0	$n_{01} = 6$	$n_{00} = 59$	
True/1	$n_{11} = 126$	$n_{10} = 11$	cloud shadow affected
False/0	$n_{01} = 10$	$n_{00} = 63$	

**Table 3.** Results of the visual validation.

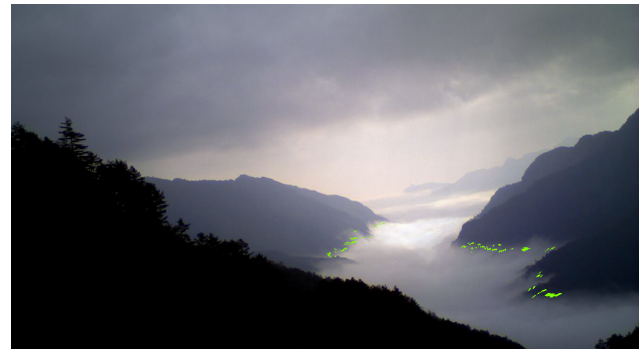
PC	Bias	POD	POFD	FAR	HKD	MCC	
0.8473	0.8418	0.7721	0.0735	0.0828	0.6986	0.7050	all lighting conditions
0.8323	0.8438	0.7813	0.0923	0.0741	0.6889	0.6760	complex illumination
0.9000	0.9927	0.9197	0.1370	0.0735	0.7827	0.7803	cloud shadow affected

## 6 Discussion and conclusion

The results of the visual validation can be regarded as promising. The accuracy, the HKD and the POD are quite high and the POFD and FAR are low. The bias shows that the cloud frequency is slightly underestimated. A method that is designed to provide validation data for another method should, however, be as near to perfect as possible. For this reason the presented method should only be used to validate satellite-derived cloud heights in scenes where there is no doubt about the presence of clouds. False positives and false negatives of the camera approach would be ignored in that way.

The results for complex lighting conditions only differ slightly from those for all scenes incorporated in the visual validation. The results for scenes affected by the shadows of overlying cloud layers are generally better (even if the POFD is slightly higher) than those for the other classes. This finding, however, might be strongly biased since the cloud-shadow-affected scenes are not randomly distributed over the investigation period but the appearance of cloud shadows is temporally clustered.

The ability of the presented approach to determine the cloud top height for cases in which the presence of clouds is known was shown using the validation cam approach and can be regarded as very good for all scenes incorporated in the validation as well as for scenes under complex lighting conditions and for scenes that are affected by cloud shadows. The detection of cloud top positions in the 2-D image, as well as the projection onto the three-dimensional DEM, works well for the camera location in the Taroko Gorge, although the derived heights are afflicted with uncertainties of above 40 m that are mostly caused by the blurriness of cloud tops (cf. Sect. 4.3.3).

**Figure 16.** Cloud tops (green) detected in the scene taken on 15 March 2013 from 09:02 to 09:06 UTC + 8.

Since a valid cloud height determination depends on clouds touching the terrain, the presented approach only works for selected locations, ideally with frequently occurring sea-of-cloud conditions. Furthermore, the occurrence of snow cover, which is unlikely for the area used for  $z_{CT}$  determination in this study, might cause problems as the presented method relies on differences in the brightness between clouds and terrain. Since several thresholds used in the cloud height determination process have been derived empirically for the location in the Taroko Gorge and the segmentation of the terrain into slopes and segments has also been custom-tailored to the view of the main cam, an adaptation of the algorithm to the footage of cameras on other positions would be necessary. The adaptation to the footage of cameras below the cloud base that would be used to derive  $z_{CB}$  (cf. Sect. 4.2.4) could cause problems since cloud bases are generally more blurry than cloud tops.

For suitable locations above the cloud top height, however, the presented method supplies data that are applicable for the design and validation of satellite algorithms for ground fog detection. The method is cheap and easy to set up, even



in terrain that is difficult to access, as is often the case for cloud forest areas. It provides continuous daytime data at a temporal resolution of 5 min over several months without the need for maintenance. Besides the necessary adaptations of the algorithm to different camera locations, the data analysis is fully automated. While other techniques observe only one column of the atmosphere or one point in space, the newly developed method provides data for a much wider area. This

provides the opportunity to validate satellite-derived cloud top heights using interpolated  $z_{CT}$  values for several positions all over a pixel of the satellite image. An extrapolation of one single point in space to a whole pixel is less valid than this procedure, especially for the low-resolution imagery of GEO satellites.

**Appendix A: Formulas used in the validation**

The following formulas were used for the calculation of the statistical measure used in Sect. 4.3. See Tables 1 and 2 for explanations of  $n_{11}$ ,  $n_{10}$ ,  $n_{01}$  and  $n_{00}$ .

$$PC = \frac{n_{11} + n_{00}}{n_{11} + n_{10} + n_{01} + n_{00}}, \quad (A1)$$

$$\text{Bias} = \frac{n_{11} + n_{01}}{n_{11} + n_{10}}, \quad (A2)$$

$$POD = \frac{n_{11}}{n_{11} + n_{10}}, \quad (A3)$$

$$POFD = \frac{n_{01}}{n_{01} + n_{00}}, \quad (A4)$$

$$FAR = \frac{n_{01}}{n_{11} + n_{01}}, \quad (A5)$$

$$HKD = POD - POFD, \quad (A6)$$

$$MCC = \quad (A7)$$

$$\frac{n_{11} \cdot n_{00} - n_{01} \cdot n_{10}}{\sqrt{(n_{11} + n_{01}) \cdot (n_{11} + n_{10}) \cdot (n_{00} + n_{01}) \cdot (n_{00} + n_{10})}}.$$

*Acknowledgements.* This research was funded through a grant (TH 1531/2-1) from the German Research Council (DFG) in cooperation with the Taiwanese National Science Council (NSC). The authors would also like to thank Hsiao-Fan Chiou, Chih-Wei Hui, Wei Hao Ho and Sebastian Achilles for their help with the field work.

Edited by: A. Stoffelen

## References

- Abeles, P.: BoofCV, available at: <http://boofcv.org/> (last access: 1 December 2014), 2012.
- Bendix, J., Rollenbeck, R., Göttlischer, D., Nauss, T., and Fabian, P.: Seasonality and diurnal pattern of very low clouds in a deeply incised valley of the eastern tropical Andes (South Ecuador) as observed by a cost-effective webcam system, *Meteorol. Appl.*, 15, 281–291, 2008.
- Bruijnzeel, L. A., Mulligan, M., and Scatena, F. N.: Hydrometeorology of tropical montane cloud forests: Emerging patterns, *Hydrol. Process.*, 25, 465–498, 2011.
- Burger, W. and Burge, M. J.: *Digital Image Processing: An algorithmic introduction using Java*. Springer Science + Business Media, New York, USA, 2008.
- Cermak, J. and Bendix, J.: Detecting ground fog from space – a microphysics-based approach, *Int. J. Remote Sens.*, 32, 3345–3371, 2011.
- Cermak, J., Bendix, J., and Dobbermann, M.: FMet – an integrated framework for Meteosat data processing for operational scientific applications, *Comput. Geosci.*, 34, 1638–1644, 2008.
- Jolliffe, I. T. and Stephenson, D. B.: *Forecast Verification – A Practitioner’s Guide in Atmospheric Science*, John Wiley & Sons, Chichester, West Sussex, UK, 2003.
- Marchand, R., Ackerman, T., Smyth, M., and Rossow, W. B.: A review of cloud top height and optical depth histograms from MISR, ISCCP, and MODIS, *J. Geophys. Res.*, 115, D16206, doi:10.1029/2009JD013422, 2010.
- Matthews, B. W.: Comparison of the predicted and observed secondary structure of T4 phage lysozyme, *Biochim. Biophys. Acta*, 405, 442–451, 1975.
- Mulligan, M.: Modeling the tropics-wide extent and distribution of cloud forest and cloud forest loss, with implications for conservation priority, in: *Tropical Montane Cloud Forests: Science for Conservation and Management*, edited by: Bruijnzel, L. A., Scatena, F. N., and Hamilton, L. S., Cambridge University Press, Cambridge, UK, 14–38, 2010.
- Nair, U. S., Asefi, S., Welch, R. M., Ray, D. K., Lawton, R. O., Manoharan, V. S., Mulligan, M., Sever, T. L., Irvin, D., and Pounds, J. A.: Biogeography of tropical montane cloud forests. Part II: Mapping of orographic cloud immersion, *J. Appl. Meteorol. Clim.*, 47, 2183–2197, 2008.
- Sæmundsson, T.: Atmospheric Refraction, *Sky Telescope*, 72, p. 70, 1986.
- Shepard, D.: A two-dimensional interpolation function for irregularly-spaced data, in: *Proceedings of the 23rd ACM National Conference*, Princeton, 27–29 August 1968, 517–524, 1968.
- Stott, J.: Jcoord, available at: <http://www.jstott.me.uk/jcoord/> (last access: 1 December 2014), 2006.
- Tachikawa, T., Kaku, M., Iwasaki, A., Gesch, D., Oimoen, M., Zhang, Z., Danielson, J., Krieger, T., Curtis, B., Haase, J., Abrams, M., Crippen, R., and Carabajal, C.: *ASTER Global Digital Elevation Model Version 2 – Summary of Validation Results*, NASA Land Processes Distributed Active Archive Center and the Joint Japan-US ASTER Science Team, 2011.
- United States Geological Survey: Global data explorer, available at: <http://gdex.cr.usgs.gov/gdex/> (last access: 1 December 2014), 2013.
- Welch, R. M., Asefi, S., Zeng, J., Nair, U. S., Han, Q., Lawton, R. O., Deepak, K. R., and Manoharan, V. S.: Biogeography of tropical montane cloud forests. Part I: Remote sensing of cloud-base heights, *J. Appl. Meteorol. Clim.*, 47, 960–975, 2008.
- World Meteorological Organization: *International Meteorological Vocabulary*, 2nd Edn., Secretariat of the World Meteorological Organization, Geneva, 1992.
- Zhang, Z.: A flexible new technique for camera calibration, *IEEE T. Pattern Anal.*, 22, 1330–1334, 1998.