Atmos. Meas. Tech. Discuss., 7, 2783–2825, 2014
www.atmos-meas-tech-discuss.net/7/2783/2014/
doi:10.5194/amtd-7-2783-2014

Atmospheric Measurement Techniques Discussions
Open Access

Automatic cloud top height determination using a cost-effective time-lapse camera system

H. M. Schulz[1], **S.-C. Chang**[2], **B. Thies**[1], **and J. Bendix**[1]

[1]Laboratory for Climatology and Remote Sensing, Faculty of Geography, Philipps-University, Marburg, Germany
[2]Department of Natural Resources and Environmental Studies, National Dong Hwa University, Hualien, Taiwan

Received: 17 February 2014 – Accepted: 5 March 2014 – Published: 21 March 2014

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

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

A new method for the determination of cloud top heights from the footage of a time-lapse camera is presented. 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 the ecology of which 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) has shown 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 calculate 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 below 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

AMTD
7, 2783–2825, 2014

Organization, 1992). z_{CB} is calculated by subtracting the cloud geometrical thickness (Δz) 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 away from perfect. To overcome these problems an extensive validation of z_{CT} and z_{CB} values used in future fog detection algorithms will be helpful.

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

This lack of cloud height data in remote regions is also an issue if it is needed for preliminary studies conducted to design a ground fog detection scheme that shall be used to map fog frequencies. If the inter-diurnal dynamic of cloud heights is high, the ground fog detection should be based on 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 is necessary as a base for solid decision making, 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 MODIS data and used them to map cloud immersion frequencies in their study area in the Costa Rican mountains. Because of the mentioned 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

AMTD
7, 2783–2825, 2014

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 of manual evaluation of the photos makes the approach hardly applicable for any comprehensive statistical investigation as the mentioned preliminary studies for the design of ground fog mapping techniques. The authors have, 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 digital elevation model (DEM) projected to the camera view so that z_{CT} and z_{CB} can easily be determined. This simple approach is not suited 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 interference, it shows the potential for further automation.

The aim of this paper is to develop and validate a cost-effective, fully automated method for z_{CT} and z_{CB} determination from camera footage in a cloud forest area of Taiwan. The method (cf. Sect. 4) has been developed to be in principle suited 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 for the cloud top height. Results will be presented in Sect. 5 and discussed in Sect. 6.

AMTD
7, 2783–2825, 2014

2 Study area and camera setup

2.1 Study area

The Taroko Gorge located in Eastern Taiwan is famous for a frequently (almost daily) occurring sea of clouds, which can be observed from higher terrain and is therefore well suited for cloud top height determination. 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. Therefore the area is ideally suited to test 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 hardly accessible, only some places (mostly near roads) can be considered as suited 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. Fig. 2). 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 the 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.

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°10′42.44″ N, 121°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

AMTD
7, 2783–2825, 2014

it oversees the gorge (cf. Fig. 3). The clouds usually form in different heights below the camera location, but sometimes the camera itself is immersed into clouds.

Another camera (further referred to as validation cam) is installed at 24°10′46.64″ N, 121°21′33.95″ E on the northern slope of the gorge. The validation cam is located at a height of 2377 m, that is much more often cloud immersed than the main cam height. Its location can be seen from the main 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 in 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 × 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 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 is the ASTER Global Digital Elevation Model (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 20 m (Tachikawa and Hato, 2011).

AMTD
7, 2783–2825, 2014

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 in an image editing software. Although the image distortion never exceeded about 2 pixels, a camera calibration as described in Zhang (2009) 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 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 on the base of these mean images instead of single frames. Therefore the output of the presented workflow are z_{CT} values in a temporal resolution of 5 min.

Scenes are discarded if the location of the main cam itself is cloudy (this results in a undifferentiated image. Therefore the coefficient of variation of the brightness of all pixels below the horizon of the projected DEM is below 0.8 for each color channel) or

AMTD
7, 2783–2825, 2014

the image is too dark to be analyzed (mean brightness of all pixels below the horizon is below 25). Otherwise 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 are 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 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 parameters position (longitude, latitude and height), orientation (rotation around the x, y and z axis) and its 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 to initially fine adjust the rotation of the virtual camera manually by using the main cam footage as an overlay over the reprojected DEM.

The rotation as well as 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 sun position. 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 whether it is usable for camera adjustment. This is the case if the horizon is visible. To check for this a buffer of 20 pixels is placed around the virtual horizon. Due to the initial adjustment and the

AMTD
7, 2783–2825, 2014

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 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 suited scene is used.

To adjust the camera the fit between the virtual horizon and the horizon in the mean image is calculated. This is done based on an edge image that is generated from the mean image by calculating the sum of the euclidean distances in RGB space for each pixel to its neighboring pixels. For all N pixels p_i of the virtual horizon the sum s_i of all edge image pixel values ep_{xy} in a 10×10 pixels window surrounding p_i (with x, y ranging from -5, -5 to 5, 5) weighted by the reciprocal value of the distance to p_i is calculated. The fit is then calculated as the average of all s_i.

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

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°. Afterwards also the focal length is 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.

AMTD
7, 2783–2825, 2014

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:

- A mask (white area 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 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 distance of a pixel of the reprojected DEM to its upper neighboring pixel exceeds a threshold of 400 (northern slope) or 200 (southern slope) meters (both thresholds are empirically determined). These edges could adulterate values for 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. Also 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 left of the lowest height pixel is marked as northern slope, every pixel 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 30×30 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 they 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

AMTD
7, 2783–2825, 2014

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 in some heights by terrain that is nearer to the camera. A fine segmentation as shown in Fig. 6 as well as a coarse segmentation with a drastically decreased number of classes (6 per slope) are performed to obtain both segment images.

- 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 is for the sake of computational time calculated in a resolution that is degenerated by the factor of 4 in each dimension) 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.

AMTD
7, 2783–2825, 2014

The SD images contain for each pixel p_i the standard deviation of the pixel values in a vertical 1 × 40 pixel window surrounding p_i.

4.2.3 Image analysis for the derivation of z_{CT}

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

A simple threshold approach as used by Bendix et al. (2008) is not suited for this purpose. Despite 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 2.5 km to the main cam while the furthest pixels are about 10 km away, the atmospheric signal that influences the pixel colors is greatly varying. 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 by 1 has a color value of (R: 137, G: 140, B: 149) while pixel 2, which is much nearer to the camera, has a color value of (R: 121, G: 126, B: 129). Obviously the distance depended 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 often is 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 distant dependent. 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

AMTD
7, 2783–2825, 2014

clouds (e.g. white if the sun is high, reddish if the sun is low) are greatly varying over time, which would also complicate the application of thresholds.

The algorithm used instead is based on the idea that, in spite of small scale variations caused by lighting conditions etc., the mean pixel value of different input images (mean image (R, G and B channel), movement image and SD images of those images; cf. Section 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:

- 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.

- $\mathrm{diff}_{SD(R)}$, $\mathrm{diff}_{SD(G)}$ and $\mathrm{diff}_{SD(B)}$ 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.

- $\mathrm{diff}_{movement}$ and $\mathrm{diff}_{SD(movement)}$ 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:

- mov_{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.

AMTD
7, 2783–2825, 2014

– $\mathrm{diff}_{\mathrm{local}}$ is the color difference between a DEM section and the next highest DEM section of the same terrain segment. It is is calculated as

$$\mathrm{diff}_{\mathrm{local}} = |R_{\mathrm{DEM\ section}} - R_{\mathrm{DEM\ section}+5\,\mathrm{m}}| \times |G_{\mathrm{DEM\ section}} - G_{\mathrm{DEM\ section}+5\,\mathrm{m}}| \times |B_{\mathrm{DEM\ section}} - B_{\mathrm{DEM\ section}+5\,\mathrm{m}}| \quad (2)$$

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 in the height of z_{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_{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 is calculated for different distance classes. The result (further referred to as edge value_z) is shown in Fig. 7 for a scene that includes the frame shown in Fig. 4. The standardization is done on base of the coarse segments as many fine segments do no cover the whole height of a slope, and hence some of them would be entirely situated above or below z_{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. 7) can be considered as the mean z_{CT} for each slope (further referred to as $z_{\mathrm{CT\ slope}}$). This is, however, only true if clouds are present and if the maximum of the edge value in the corresponding height is actually caused

AMTD
7, 2783–2825, 2014

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 diff_R, diff_G and diff_B. If more than 50 % of the fine segments of one slope are brighter (this means at least 2 of 3 color channels are brighter) below $z_{\mathrm{CT\ slope}}$ than above this height, it is actually accepted as the height of the cloud top surface. Otherwise $z_{\mathrm{CT\ slope}}$ 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 segments the cloud top should be near to $z_{\mathrm{CT\ slope}}$. Therefore z_{CT} for each fine segment (further referred to as $z_{\mathrm{CT\ segment}}$) is assumed to be located inside a height interval around $z_{\mathrm{CT\ slope}}$ (dashed line in Fig. 7). The more even the cloud surface is the smaller is this interval. For each slope its size is calculated from the root-mean-square deviation (RMSD) of the height of all N DEM sections to $z_{\mathrm{CT\ slope}}$. In the calculation of the RMSD the height distance of each DEM section to $z_{\mathrm{CT\ slope}}$ is weighted by its edge value$_z$:

$$\text{weighted RMSD} = \sqrt{\frac{\sum_{i=1}^{N}\left(\left(\text{height}_i - z_{\mathrm{CT\ slope}}\right)^2 \times (\text{edge value}_z)_i\right)}{\sum_{i=1}^{N}(\text{edge value}_z)_i}} \tag{3}$$

The weighted RMSD times 1.2 has proven to be a good size for the height interval in which $z_{\mathrm{CT\ segment}}$ is determined.

Fine segments that are not reaching $z_{\mathrm{CT\ slope}}$ may also not reach their $z_{\mathrm{CT\ segment}}$. Therefore they are excluded from further analysis. Also the 1.2 RMSD interval may not necessarily contain the cloud top of each fine segment. $z_{\mathrm{CT\ segment}}$ would be detected in a wrong height in these cases. Therefore the validity of each $z_{\mathrm{CT\ segment}}$ is checked individually by using a linear regression between the height distance of the DEM sections of the associated fine segment to $z_{\mathrm{CT\ segment}}$ and the image brightness (mean value of the R, G and B channel standardized on a fine segments base). This

AMTD
7, 2783–2825, 2014

regression is carried out for DEM sections that are located in a RMSD times 1.2 height interval around $z_{\text{CT segment}}$ only. Since clouds are overall brighter than non cloud 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 height meter most probably do not contain cloud tops inside the 1.2 RMSD interval. 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 $z_{\text{CT segment}}$.

In the calculation of $z_{\text{CT slope}}$ and the preliminary results for $z_{\text{CT segment}}$ the absolute (not standardized) edge value has not been respected. Therefore it is unclear if 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 a RMSD times 1.2 interval around the preliminary $z_{\text{CT segment}}$ of each associated fine segment a linear regression between the height distance of each DEM segment to the corresponding preliminary $z_{\text{CT segment}}$ 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 height meter, 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 do 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 images (cf. Fig. 8) and/or written to a CSV file for each scene of a video file. These CSV

AMTD
7, 2783–2825, 2014

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).

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:

- 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.
- $\mathrm{diff}_{\mathrm{local}}$ needs to be calculated as follows:

$$\mathrm{diff}_{\mathrm{local}} = |R_{\mathrm{DEM\ section}} - R_{\mathrm{DEM\ section-5\,m}}| \times |G_{\mathrm{DEM\ section}} - G_{\mathrm{DEM\ section-5\,m}}| \times |B_{\mathrm{DEM\ section}} - B_{\mathrm{DEM\ section-5\,m}}| \quad (4)$$

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

AMTD
7, 2783–2825, 2014

these scenes the cloud top positions were calculated and validated using two different approaches. This was done by summarizing the validation results in a confusion matrix for each validation approach. From these matrices the following statistical measures (Jollife and Stephenson, 2003) 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)

4.3.1 Validation 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. 9. 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 if cloud tops in a horizontal radius of 600 m around the validation cam do 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. Also 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 excluded from the validation. Thus the validation results show to what extent the presented method is suited to determine z_{CT} 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

AMTD
7, 2783–2825, 2014

correspond to the current scene of the main cam is analyzed. In a first step it is checked if at least one frame is too dark for further analysis (mean brightness of the mean of all 3 color channels < 40). In these cases the whole scene is excluded from the validation. For usable scenes it is checked if the validation cam is cloud immersed in at least 3 of 5 frames of the scene. If this is the case, it 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 too ensure that the validation cam is cloud immersed if it is located below z_{CT}.

The test if the validation cam is immersed into clouds is based on the contrast in its image. As Fig. 10 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 tests 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 × 30 pixels instead (two of which are exemplarily marked in Fig. 10). 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_{vmax} of these three values is determined and the sum

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

is calculated for all N boxes for which c_{vmax} is below an empirical threshold of 0.03. As c_{vmax} is low for boxes with a cloud obstructed view due to their low image contrast, s is

AMTD
7, 2783–2825, 2014

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.

The validation cam based approach does consider only one point in space, namely the validation cam position. Therefore it would not be known if 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 be 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 can be interpreted as the answer to the question if the top height of clouds is derived correctly if they have been detected.

4.3.2 Visual validation

Since the automatized validation has been performed under the assumption of correctly determined cloud presence a second, visual, approach has been conducted. 10 % of the 8400 main cam scenes were randomly chosen and detected cloud top heights that were marked as shown in Fig. 8 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.

AMTD
7, 2783–2825, 2014

– Despite the absence of cloud tops in the detection area, cloud tops were detected. Or: cloud tops are present and were detected, but they were detected in the wrong height: false positives.

– Cloud tops are present in the detection area but they were not detected: false negative.

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 on the base of two-dimensional images. Therefore it does not provide any information whether the height determination of the cloud tops detected in 2-D images using the three-dimensional terrain is valid.

5 Results

An example output of the cloud top height determination algorithm is shown in Fig. 8. Interpolated cloud tops heights are shown in Fig. 11. Both figures are derived from the 5 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 examples for the whole day can be downloaded from http://lcrs.geographie.uni-marburg.de/uploads/media/3_15_2013_CTH_camView.zip (camera view as shown in Fig. 8, digital object identifier (DOI): 10.5678/LCRS/MS.2) and 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 (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. 11) vs. time is shown in Fig. 12 for the same date. The increase of z_{CT} over time that can be observed in the original footage of the main cam is well reproduced and also times in which the view of the camera was obstructed

AMTD
7, 2783–2825, 2014

by clouds (red areas) were correctly detected. The data gap between about 06:30 to 08:15 is caused by cloud tops that are too low to be detected.

The algorithm also works for more complex lighting conditions (cf. Fig. 13) and – as long as the clouds are touching the terrain – in situations with a shattered cloud cover (cf. Fig. 14).

Validation results

Table 1 shows the result of the validation based on the validation cam footage. The PC calculated from this matrix is 0.9544. 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 cloud tops in the two-dimensional camera footage was correctly determined in 87.56 % of the scenes (PC) while the frequency of cloud tops in the area of detection was slightly underestimated in general (bias). Cloud tops were detected in 82.02 % of the scenes in which cloud tops were present (POD) and in 8.14 % of the cloudless scenes. This corresponds to about 11.34 % of the determined cloud tops being mistakenly detected (FAR).

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 POD and FAR are low. The Bias shows that the cloud frequency is only slightly underestimated. A method that is designed to provide validation data for another method should, however, be as near to perfect as possible. Due to the fact that the intended use of the presented method is the validation of a satellite algorithm this is not problematic at all. The presence or absence of clouds can already be determined from satellite data with a high degree of certainty (Reuter et al., 2009). For a fog detection algorithm thus only the cloud height retrieval and not

AMTD
7, 2783–2825, 2014

the cloud detection needs to be validated. This can be done for periods of time in which clouds were detected by the satellite algorithm as well as the camera approach. False positives and false negatives of the camera approach would not be used.

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. The detection of cloud top positions in the 2-D image as well as the projection onto the three-dimensional DEM work well for the camera location in the Taroko Gorge.

Since a valid cloud height determination depends on clouds touching the terrain, the approach does only work for selected locations, ideally with frequently occurring sea of cloud conditions. Since several thresholds used in the cloud height determination process have been derived empirically for the location in the Taroko Gorge and also the segmentation of the terrain into slopes and segments has 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 are used to derive z_{CB} (cf. Sect. 4.2.4) could cause problems since cloud bases are often blurry.

For suited locations above the cloud top height, however, the presented method supplies data that is applicable for the design and validation of satellite algorithms for ground fog detection. The method is cheap and easy to setup, even in terrain that is difficult to access as it often is the case for cloud forest areas. It provides continuous daytime data in a temporal resolution of 5 min over several months without the need of maintenance. 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 does provide data for a whole valley. This gives the opportunity to validate satellite derived cloud tops 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.

AMTD
7, 2783–2825, 2014

Appendix A

Formulas used in the validation

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

$$\mathrm{PC} = \frac{n_{11} + n_{00}}{n_{11} + n_{10} + n_{01} + n_{00}} \tag{A1}$$

$$\mathrm{Bias} = \frac{n_{11} + n_{01}}{n_{11} + n_{10}} \tag{A2}$$

$$\mathrm{POD} = \frac{n_{11}}{n_{11} + n_{00}} \tag{A3}$$

$$\mathrm{POFD} = \frac{n_{01}}{n_{01} + n_{00}} \tag{A4}$$

$$\mathrm{FAR} = \frac{n_{01}}{n_{11} + n_{01}} \tag{A5}$$

$$\mathrm{HKD} = \mathrm{POD} - \mathrm{POFD} \tag{A6}$$

Acknowledgements. This research was funded by 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 in the field work.

AMTD
7, 2783–2825, 2014

References

Abeles, P.: BoofCV, available at: http://boofcv.org/ (last access: 11 February 2014), 2012.

Bendix, J., Rollenbeck, R., Göttlicher, 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.

Jolliffe, I. T. and Stephenson, D. B.: Forecast Verification – a Practitioner's Guide in Atmospheric Science, John Wiley & Sons, Chicester, 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, 1–25, doi:10.1029/2009JD013422, 2010.

Mulligan, M.: Modeling the tropics-wide extent and distribution of cloud forest and cloud forest loss, with implications for conversation 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.

Reuter, M., Thomas, W., Albert, P., Lockhoff, M., Weber, R., Karlsson, K.-G., and Fischer, J.: The CM-SAF and FUB cloud detection schemes for SEVIRI: validation with synoptic data and initial comparison with MODIS and CALIPSO, J. Appl. Meteorol. Clim., 48, 301–316, 2009.

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: 11 February 2014), 2006.

AMTD
7, 2783–2825, 2014

Tachikawa, T., Hato, M., Kaku, M., and Iwasaki, A.: Characteristics of ASTER GDEM version 2, in: Proceedings of the IGARSS 2011, Vancouver, 24–29 July 2011, 3657–3660, 2011.

United States Geological Survey: Global data explorer, available at: http://gdex.cr.usgs.gov/gdex/ (last access: 11 February 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.

AMTD

7, 2783–2825, 2014

Table 1. Confusion matrix for the automated validation.

Validation cam cloud immersed?	Validation cam below interpolated z_{CT}? True/1	False/0
True/1	$n_{11} = 240$	$n_{10} = 7$
False/0	$n_{01} = 43$	$n_{00} = 807$

AMTD
7, 2783–2825, 2014

Table 2. Confusion matrix for the visual validation.

Cloud tops present in detection area?	Cloud tops detected at correct position? True/1	False/0
True/1	$n_{11} = 219$	$n_{10} = 48$
False/0	$n_{01} = 28$	$n_{00} = 316$

AMTD
7, 2783–2825, 2014

Table 3. Results of the visual validation.

PC	Bias	POD	POFD	FAR	HKD
0.8756	0.9251	0.8202	0.0814	0.1134	0.7388

AMTD
7, 2783–2825, 2014

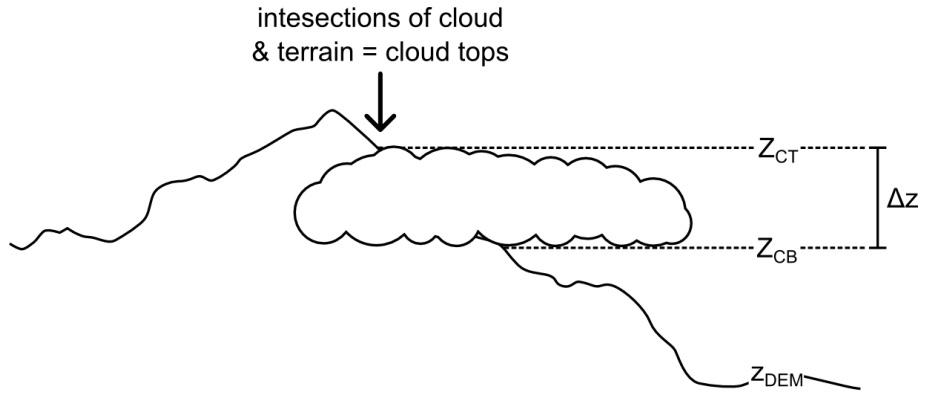

Fig. 1. Some definitions used in this paper.

AMTD
7, 2783–2825, 2014

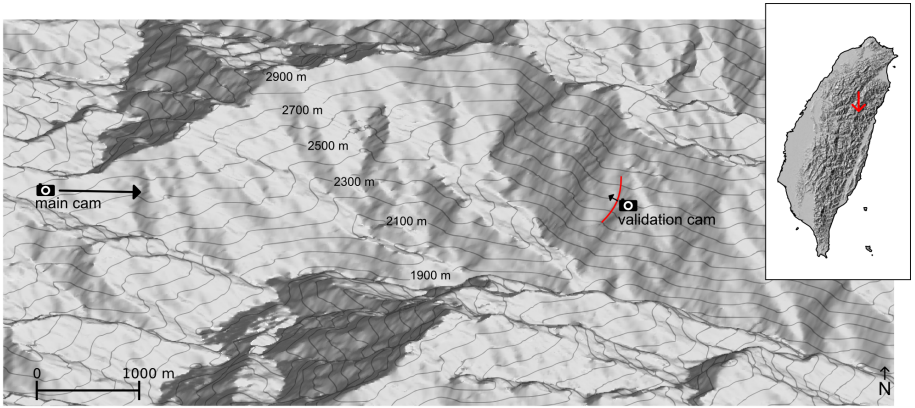

Fig. 2. Isometric view of the western end of the Taroko Gorge with the positions of the cameras. The black arrows denote the cameras' viewing directions. The red line marks a small recess in the slope. The location of the area in Taiwan is marked in the upper right overview map.

AMTD
7, 2783–2825, 2014

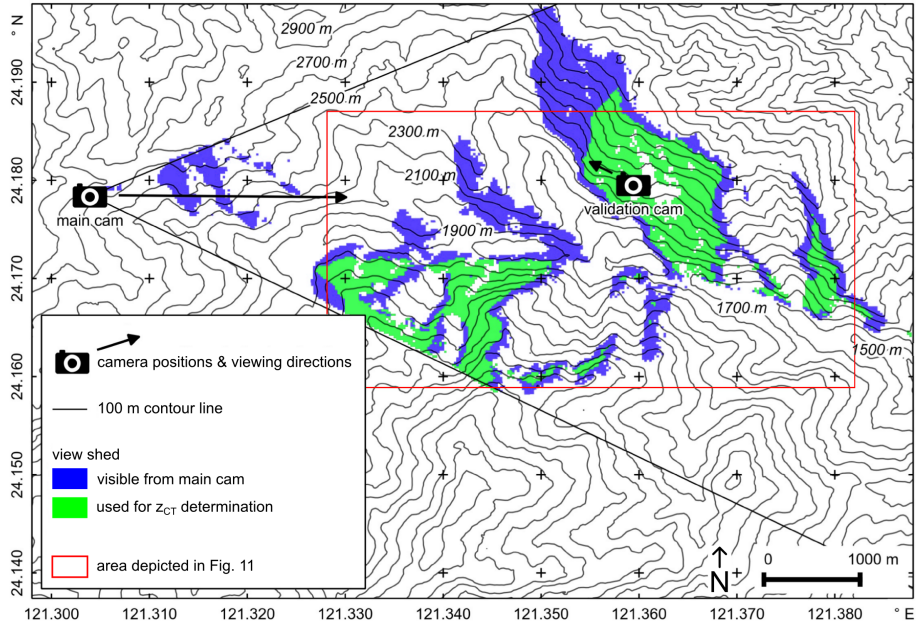

Fig. 3. Camera positions and the view shed of the main cam.

AMTD
7, 2783–2825, 2014

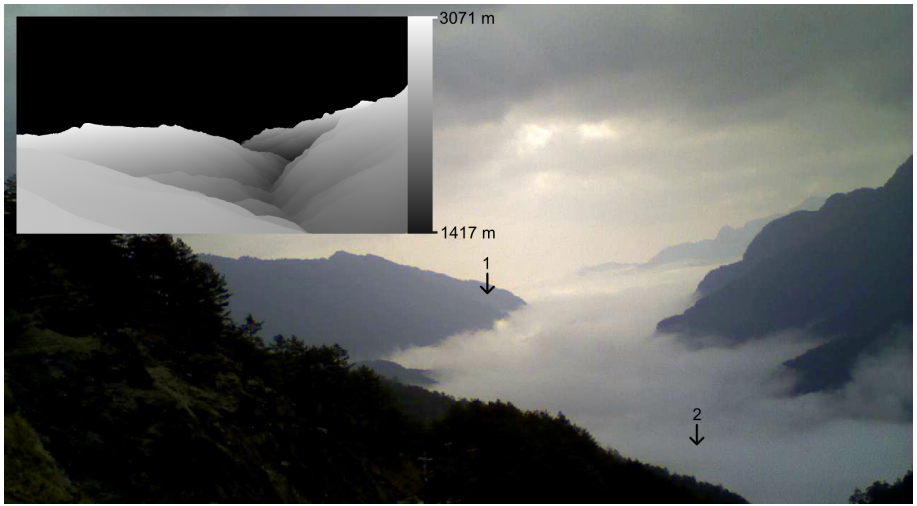

Fig. 4. View of the main cam on 15 March 2013, 10:14 UTC + 8 (big picture) and the DEM projected to the view of the main cam (upper left). See text for explanation of the arrows.

AMTD
7, 2783–2825, 2014

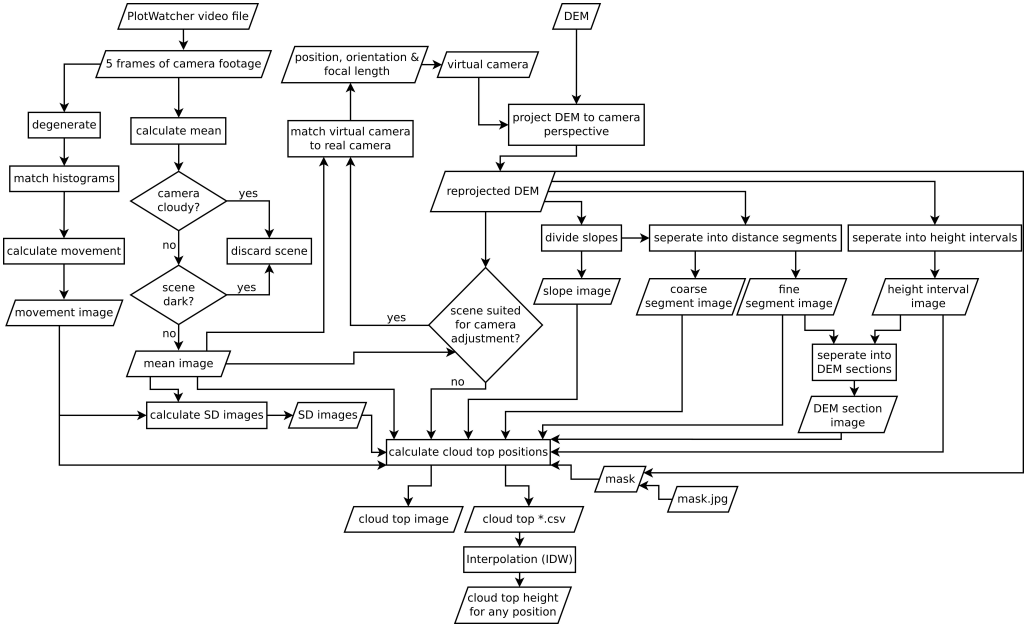

Fig. 5. Workflow for the retrieval of z_{CT}.

AMTD
7, 2783–2825, 2014

Automatic cloud top height determination

H. M. Schulz et al.

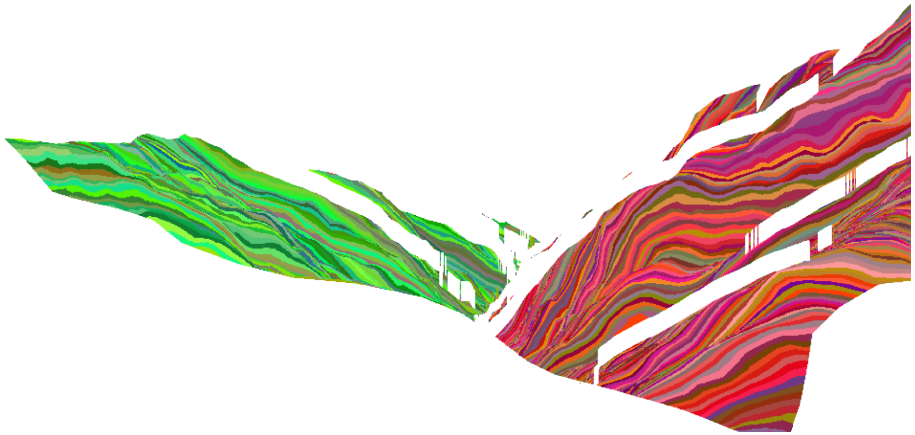

Fig. 6. Segmetation of the terrain in slopes (reddish/greenish areas) and fines segments. The white areas are masked out.

AMTD
7, 2783–2825, 2014

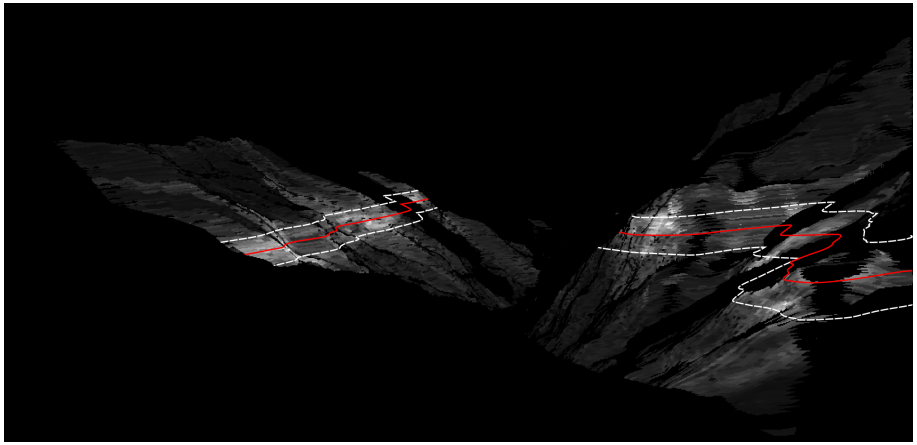

Fig. 7. Standardized edge value for the scene from on 19 March 2013 from 10:12 to 10:16 UTC + 8. (see also Fig. 4). The higher the edge value$_z$, the brighter the shade. $z_{\text{CT slope}}$ is marked with the red line, while the 1.2 times RMSD interval is the area between the dashed lines.

AMTD
7, 2783–2825, 2014

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

AMTD
7, 2783–2825, 2014

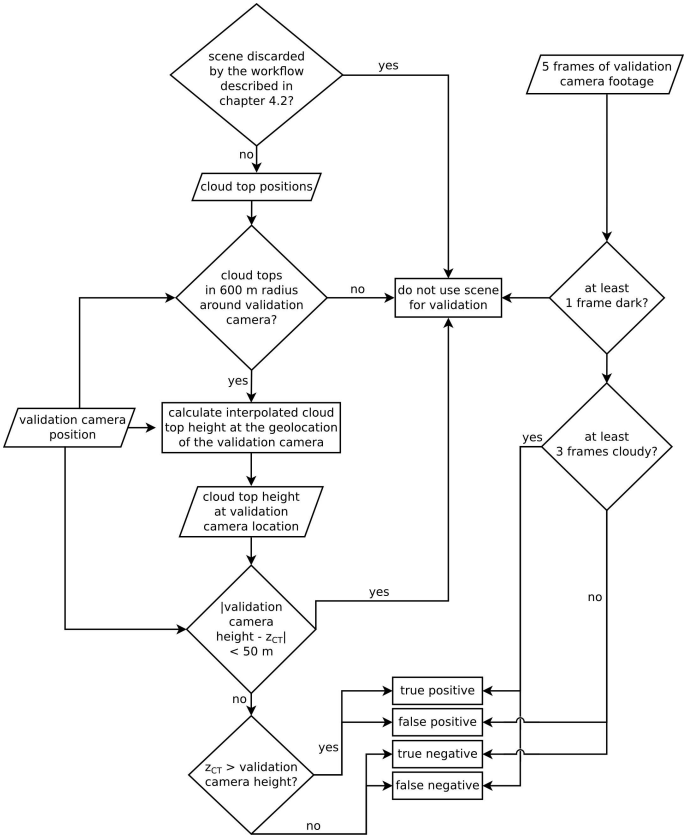

Fig. 9. Workflow used for the automated validation.

AMTD
7, 2783–2825, 2014

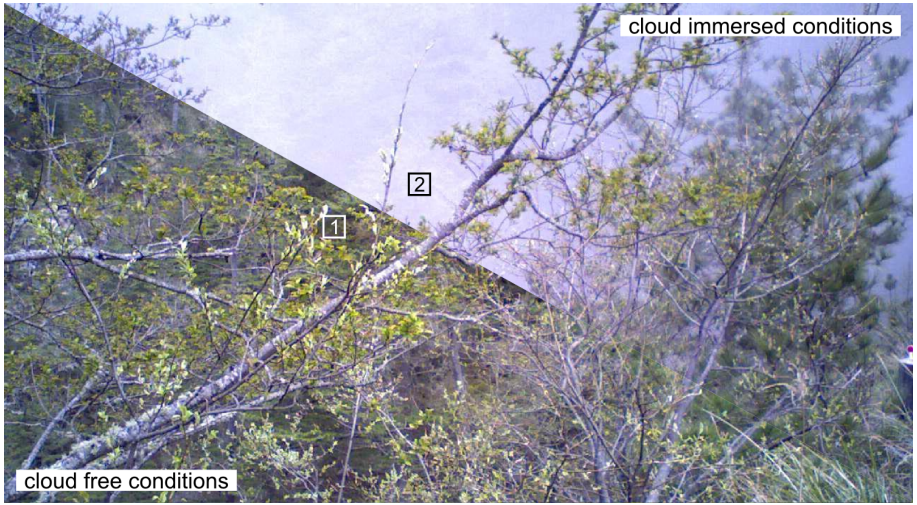

Fig. 10. Image of the validation cam under cloudy and cloud free conditions. c_{vmax} is 0.230 for the 30 × 30 pixels box denoted with "1" and 0.020 for the box denoted with "2" (see text for further explanation).

AMTD
7, 2783–2825, 2014

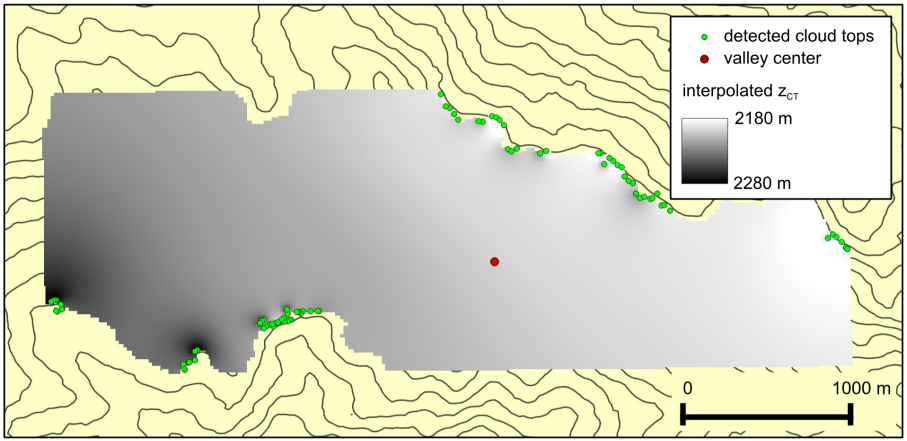

Fig. 11. Cloud tops heights derived from the main cam footage taken on 15 March 2013 from 10:12 to 10:16 UTC + 8. Cf. Fig. 3 for orientation.

AMTD
7, 2783–2825, 2014

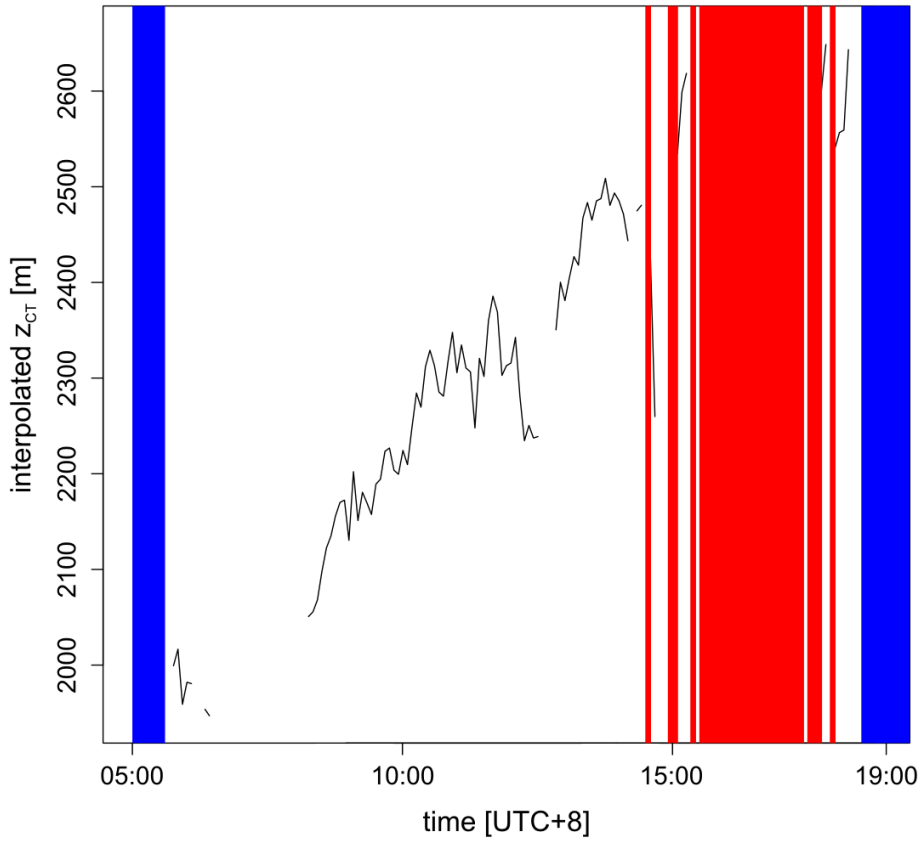

Fig. 12. Increase of the cloud top height on 15 March 2013 for the position marked as "valley center" in Fig. 11.

AMTD
7, 2783–2825, 2014

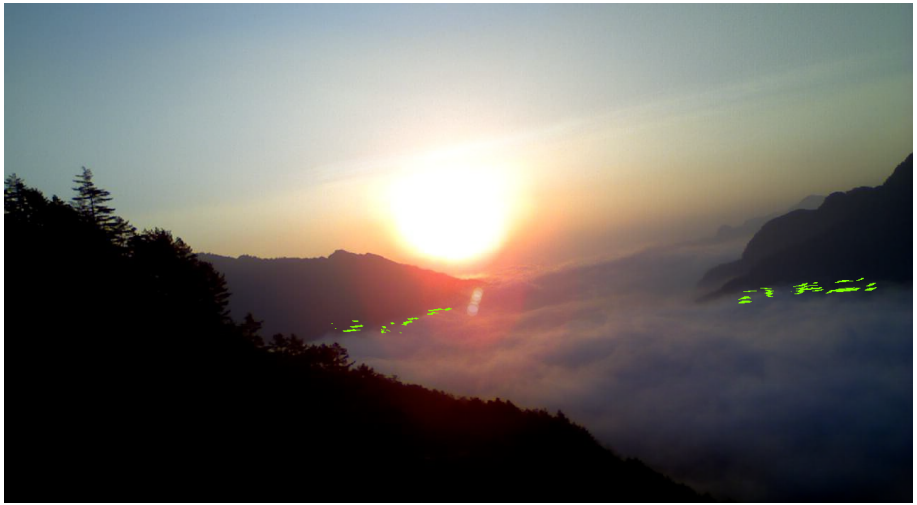

Fig. 13. Cloud tops (green) detected in the scene taken on 19 March 2013 from 06:12 to 06:16 UTC + 8.

AMTD
7, 2783–2825, 2014

Fig. 14. Cloud tops (green) detected in the scene taken on 4 April 2013 from 07:04 to 07:08 UTC + 8.

AMTD
7, 2783–2825, 2014