Response to Anonymous Referee #1

We would like to thank Anonymous Referee #1 for his / her critical comments helping us to improve our paper.

The referee's original comments are written in italics, while our response is intended to the right.

The abstract (and title) would be clearer if it was mentioned that the technique is meant for mountain areas and a camera placed above the cloud layer.

=> We agree. The title will be changed to "Automatic cloud top height determination **in mountainous areas** using a cost-effective time-lapse camera system". The beginning of the abstract will be changed to "A new method for the determination of cloud top heights ~~from~~ **using** the footage of a time-lapse camera, **that is placed above a frequently occurring cloud layer in a mountain valley**, is presented."

As the aim of the technique is to provide data for the validation of algorithms using imagery from GEO or LEO satellites, it would be illustrative to give the horizontal resolution of such imagery (especially current GEO satellites viewing the area) and compare the pixel size to the geographical extension of the Taroko Gorge.

We agree. We will add the following lines to p. 2787, l. 5-8:

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 hardly 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 suited to test a technique that can be used to design and validate methods for ground fog retrieval from satellite data.

The study focuses on cloud top height (ZCT), and the paper actually gives only an indication of how the technique described for ZCT could be extended to ZCB (cloud bottom height). However, sometimes the paper gives the impression that the study is equally concerned with ZCT and ZCB. For instance, the sentence in P3 L9-10 is misleading, as the technique is not validated for ZCB.

=> We agree. p. 2785, l. 8-9 will be changed to "To overcome these problems an extensive validation of zCT and zCB values used in future fog detection algorithms ~~will~~ **would** be helpful. " since the word "will" suggest that the method suggested in our paper would deliver validation data (which is only true for zCT).

Further p. 2786 will be changed as follows:

"The aim of this paper is to develop and validate a cost-effective, fully automated method

for zCT ~~and zCB~~ 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 zCT as well as zCB 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 only. Results will be presented in Sect. 5 and discussed in Sect. 6. **The suitability of the method for zCB will be addressed in future studies.**"

P1 L21 (and others). It sounds more appropriate to use "estimate" instead of "calculate", both for ZCB and ZCT, considering the uncertainties involved (and mentioned by the authors in L29-32).

=> We agree. p. 2784, l. 24 – p. 2785, l. 2 will be changed as follows:

In most cases they ~~calculate~~ **estimate** the cloud base height (zCB) for each pixel of an image and compare this value to the terrain height of the pixel (zDEM) taken from a digital elevation model (DEM). If zDEM is equal to or below zCB , 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** zCB is calculated by subtracting the cloud geometrical thickness (Δz) from the cloud top height (zCT) (cf. Fig. 1).

P1 L23. Please check "If ZDEM is equal to or below ZCB".

=> Thanks for the correction. The sentence will be changed to: "If zCB is equal to or below zDEM "

P2 L11. The expression "inter-diurnal dynamic" sounds somewhat odd. Do authors mean "day-to-day variation"?

=> Thanks for noticing. "Inter-diurnal" is actually a typo. It should be "intradiurnal".

P3 L9-10. Please rephrase, as the study does not attempt to validate the technique for ZCB.

We agree. See above.

P4 L12. It would be interesting to know how often the main camera is immersed in clouds, roughly. The height of its location sets an artificial upper limit to the values of ZCT that can be provided by the system in that environment.

Clouds in the height of the camera occur clearly less than every second day. Typically they form in the afternoon and stay for around one hour. P. 2788, l. 1 – 2 will be reformulated as follows:

"The clouds usually form in different heights below the camera location, but ~~sometimes~~ **on**

some days the camera itself is immersed into clouds **for approximately one hour in the afternoon. As the height of the camera sets an upper limit, cloud top heights can not be detected in these cases.**"

P5 L29-31. Please rephrase, as it is not quite clear. Presumably it refers to those times when the main camera is immersed in clouds. How is that condition detected?

=> You are right. We will reformulate the sentence:

Scenes are discarded if the location of the main cam itself is ~~cloudy~~ **cloud immersed (This results in an undifferentiated image. Therefore the coefficient of variation of the brightness of all image pixels below the horizon of the reprojected DEM can be used to detect that condition. If it is below 0.8 for each color channels the camera is cloud immersed.)** (~~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~~) [...]

P6 L25. What is the meaning of "the horizon is visible"? Does it mean not obscured by clouds?

=> We agree that the phrase is a little bit unclear. We will change it to:
"This is the case if the horizon is visible **(i.e. the view to the horizon is obstructed neither by clouds nor mist**.)"

P6 L25. This is not quite clear. Consider rephrasing to explain how the edge image is obtained and how the fit between horizons is calculated.

=> We reformulated p. 2791, l. 11 – 19. Feel free to contact us if anything is still unclear.

"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 touch 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 their distance in pixels to p_i is calculated. The fit is then calculated as the average of all N values for s_i.

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

The more pixels of the edge image have high values near to the virtual horizon, the better is the fit."

P14, 4.3.1 Validation using the validation cam. The number of cases in the confusion matrix is a very small number compared with the number of scenes (8400). It would be interesting to know why only such a small fraction of the number of scenes is suitable for validation with the validation cam.

=> Only scenes with cloud top occurrences in a radius of 600 m around the validation cam are incorporated here. This is already mentioned in the text. If you think the formulation is unclear, please let us know.

“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, zCT 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.”

P15, L3. Please check "vertical distance of less than 50 m".

=> “vertical distance of 50 m” is correct. Since the precision of our system is, of course, restricted, it is validated with a vertical tolerance of 50 m. This is mentioned in p. 2800, l. 21-23: “Thus the validation results show to what extent the presented method is suited to determine zCT with a precision of 50 m.”

P18 L8. Please check "POD and FAR are low". Presumably it is POFD.

=> Exactly. Thanks a lot for the correction.

P18, L23-24. It would be helpful if this useful information (or similar) is included in the abstract.

=> Since you don't give a direct quote here and you are not using the page and line numbers from the discussion article, we are not sure which information is meant. Most probably you are referring to the sentence “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. ” (p. 2805, l. 9-11)? Please let us know if this assumption is not correct.

We agree that this information will be useful in the abstract. Since we will change the abstract according to you first request (see above), we think it is now obvious that the method works only for certain situations (namely a camera in a mountain valley above a frequently occurring cloud layer).

P19, L5. Actually not for the whole valley, but for the part of the valley seen by the main camera, which may be a fraction of it.

You are right. We will change p. 2805, l. 23-25 to

“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~~ **much wider area**. ”

Appendix A. Please check POD formula.

You are right. The formula should be $POD = \frac{n_{11}}{n_{11} + n_{10}}$. Luckily this was just a typo and the POD calculation in the paper was done using the correct formula.

Fig. 5. There is a line from "mean image" to "match virtual camera to real camera". Is that correct? From the description, it looks that this is only done if the scene is suitable for the adjustment.

=> Yes, the virtual camera is only matched to the real camera when the scene is suitable for camera adjustment. The mean image is necessary as an input for the camera matching. The arrow from "mean image" to "match virtual camera to real camera" was meant to illustrate this. But you are right, this could be confusing. We will remove the arrow.

Fig. 5. Please check all instances of "seperate"

=> Thanks for noticing. We will change that to "separate".