Manuscript prepared for Atmos. Meas. Tech.
with version 2015/09/17 7.94 Copernicus papers of the LaTeX class copernicus.cls.
Date: 8 May 2017

Using CALIOP to estimate the base height of optically thick clouds
Used to be: "Optimized Detection by Radar/lidAr for base of Nuages (ODR/AN)"

Johannes Mülmenstädt[1], Odran Sourdeval[1], and Johannes Quaas[1]

[1]Institute of Meteorology, Universität Leipzig, Leipzig, Germany
[2]Anyone else?

Correspondence to: Johannes Mülmenstädt (johannes.muelmenstaedt@uni-leipzig.de)

Abstract. A measurement technique is presented that uses CALIOP lidar profiles to estimate cloud base heights. The technique provides cloud base heights even when clouds are thick enough to attenuate the lidar beam (optical thickness $\tau \gtrsim 5$) by treating the cloud base height of nearby thinner clouds as representative of the entire cloud field. Using ground-based ceilometer data, uncertainty estimates are derived as a function of various properties of the CALIOP lidar profiles. Evaluation of the predicted cloud base heights and their predicted uncertainty using a second, statistically independent, ceilometer data set shows that cloud base heights and uncertainties are biased by less than 10%.

{ j_μ: General comments:

– switch to heights above ground $\longrightarrow$ done

}

1 Introduction

Cloud base height is one of the most important parameters of a cloud. It controls how much downwelling longwave radiation the cloud emits. Aerosol concentration and updraft speed at that level control the microphysics of the cloud. It is one of the parameters that is necessary to calculate the subadiabaticity of the cloud. However, it is also one of the most difficult parameters to retrieve from satellite. Oxygen absorption bands: planning to evaluate these too. VIIRS cloud-base temperature method: Zhu et al. (2014, doi 10.1002/2013GL058970). MISR stereoscopic imager method: working on it. However, these are all not on the A-Train. CloudSat misses the small droplets at the base and cannot retrieve in the ground clutter region. Calipso detects the bases of only the thinnest clouds ($\tau < 5$, according to Mace and Zhang, 10.1002/2013JD021374); frequently, these are not the clouds you are looking for.

Because the lifting condensation level is approximately homogeneous within an airmass, the cloud bases retried by CALIOP for thin clouds may be a good proxy for the cloud base heights of the entire cloud cluster, including the optically thicker clouds within the cluster. We have designed an algorithm that extrapolates the CALIOP cloud-base measurements into locations where CALIOP attenuates before reaching cloud base. This algorithm is called CBASE (Cloud Base Altitude Spatial Extrapolator). In this paper we evaluate its performance by comparing CBASE cloud base heights against cloud base heights observed by ground-based ceilometers.

Section 2 describes the data sources used in determining and evaluating the cloud base height. In Section 3 we describe the algorithm and evaluate its performance, including error statistics. We conclude in Section 5 with an outlook on the longstanding questions that this retrieval can help address.

Literature: Meerkötter and Zinner (2007) (10.1029/2007GL030347) retrievals using cloud top properties and adiabatic assumption; presumably there are others. Zhu et al. (2014) for cloud base temperature using the thinnest cloud in a cloud cluster. Mace and Zhang (2014) using CPR and CALIOP. At least one more using CloudSat; check Stephens et al (2012) energy-balance paper.

2 Data

Description of the CALIOP VFM data.

Description of the airport ceilometers. METARs: `https://library.wmo.int/pmb_ged/wmo_49-v2_2013_en.pdf`; ASOS: `http://www.nws.noaa.gov/asos/pdfs/aum-toc.pdf`

3 Algorithm and evaluation

The algorithm and evaluation proceed in four steps:

1. We determine the cloud base height from all CALIOP profiles where the surface generates a return ($\longrightarrow$ lidar is not attenuated above cloud base). We refer to these cloud base heights as *local cloud base heights* in the sense that they are local to the CALIOP profile.

2. Using ground-based ceilometer data, we determine quality of cloud base height depending on a number of factors.

3. Based on the predicted quality of each local cloud base, we either reject the local cloud base or combine it with other local cloud bases within a distance $D_{\max}$ of the point of interest (POI) $\longrightarrow$ estimate of cloud base height, estimate of cloud base height uncertainty

4. Using a statistically independent validation dataset, we verify that the predicted cloud base height and uncertainty are correct.

This section is divided into four subsections, one for each step enumerated above. Figure 1 illustrates the method.

3.1 Determination of local cloud base height

Source of local cloud base from the CALIOP VFM: any profile with a surface return.

3.2 Determination of local cloud base quality

We assess the quality of the CALIOP cloud base height z using the ceilometer-observed cloud base height $\hat{z}$ using the following metrics:

Correlation coefficient between the satellite retrieval and ground-based observation of the cloud base. We use the Pearson correlation coefficient. Ideally the correlation coefficient would be unity.

Linear regression slope and intercept (ideally 1 and 0, respectively).

RMS retrieval error, defined as

$$\mathrm{RMSE} = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(z_i - \hat{z})^2}, \tag{1}$$

(ideally 0) {j_μ: This is the only metric that is actually used, so maybe we should just get rid of the rest.}

Retrieval bias, defined as

$$\mathrm{bias} = \frac{1}{N}\sum_{i=1}^{N}(z_i - \hat{z}), \tag{2}$$

(ideally 0)

Efficiency, i.e., probability that a retrieval is available at the desired location (ideally 1).

These metrics are calculated based on the set of all ground-based observations for which a Calipso overpass is available and which meet the following additional conditions:

- the ground-based cloud base height is below 3 km above ground level (AGL); this is the height to which the ceilometer-based aviation cloud base reports are reliable
- for the evaluation of low-cloud retrievals, the ground-based cloud base height is below 3 km above mean sea level; this threshold height emulates the ISCCP definition of low cloud (cloud top pressure below 690 hPa)

The sources of ground-based observations are the following:

METARs aviation routine and special weather reports (METARs) where the cloud base height is measured by ceilometer; to ensure that this is the case, we restrict ourselves to the contiguous continental United States, where the cloud base height is mostly derived automatically by laser ceilometers that form part of Automated Surface Observing Stations (ASOS)

HD(CP)^2 ceilometers {j_μ: Unless there is an urgent reason, these will have to wait for a later paper}

ICOADS {j_μ: Unless there is an urgent reason, these will have to wait for a later paper}

R/V *Polarstern* ceilometer {j_μ: Unfortunately, there is only a minute number of overpasses within 100 km, mostly at high latitudes}

MAGIC ceilometer {j_μ: Unfortunately, the number of overpasses here is also very small; need to check that I have all the MAGIC data}

3.3 Combination of local cloud bases

These cloud base retrievals only exist sporadically (on average x% of columns), when CALIOP happens to hit a sufficiently thin cloud. To infer the cloud base height z at a point of interest (POI) that does not necessarily coincide with the location of a CALIOP profile, we proceed as follows. We first select all local CALIOP cloud base heights within a horizontal distance $D_{\max} = 100$ km of the POI. On the basis of Section 3.2, we discard cloud base heights unless

- VFM quality flag high confidence in the lowest cloud feature
- lowest cloud feature is liquid (required because not enough data is available for reliable uncertainty prediction)
- minimum horizontal averaging distance required for detection of the lowest cloud layer is 1 km or 1/3 km
- geometric thickness of the lowest cloud layer is less than 1 km
- the feature immediately above the surface is neither "invalid" nor "no signal" (indicating that although the surface may generate a detectable return, the lidar is sufficiently attenuated that the cloud base, which scatters less strongly than the surface, is unreliable)

For each remaining local cloud base height z_i, we determine the predicted uncertainty σ_i based on the categories established in the previous section. We determine a combined cloud base height

$$z = \frac{\sum_i^n w_i z_i}{\sum_i^n w_i} \tag{3}$$

with weights

$$w_i = \frac{1}{\sigma_i^2} \tag{4}$$

(optimal weights for uncorrelated least-squares). In practice, the individual measurements of cloud base are highly correlated with fairly similar σ_i. The cloud base estimate by Eq. (3) with weights given by Eq. (4) remains unbiased even in the presence of correlations. However, for the combined cloud base uncertainty, the uncorrelated weights would yield a biased estimate in the presence of correlations. The expression

$$\sigma^2 = \frac{1}{n}\sum_i^n \sigma_i^2 \tag{5}$$

yields acceptable results, as would be expected for highly correlated and fairly similar σ_i.

3.4 Evaluation of cloud base heights and cloud base height errors

Having tuned the algorithm on data from the year 2008, we evaluate it using a statistically independent data set from the year 2007. In the evaluation data set, the "true" (i.e., measured by the ceilometer) cloud base height $\hat{z}$ is known in addition to the estimated cloud base height z and the estimated cloud base height uncertainty σ, determined according to the procedure described in the previous section.

For satellite-derived measurements of the cloud base height z that are unbiased with respect to the ceilometer-observed cloud base heights $\hat{z}$ and have correctly estimated uncertainties σ, the pdf of the quantity $(z-\hat{z})/\sigma$ has zero mean and unit standard deviation. In our evaluation data set, we find a mean of 0.03 and a standard deviation of 1.05 – so both the cloud base estimate and the uncertainty estimate are unbiased at the better than 10% level.

Discussion of representativeness of continental clouds over North America for the remainder of the globe. A marine validation dataset would be very welcome.

4 Results and data product availability

Geographic distributions of mean or median cloud base and thickness. Annual cycle (if it's interesting).

Comparison with 2B-GEOPROF-LIDAR cloud bases: distinguish between radar-only and lidar-only bases (both is rare for warm cloud). Radar-only mean error is large because the radar cloud base height predominantly clusters around the top of the ground clutter region with little dependence on the actual cloud base height. Lidar-only 2B-GEOPROF-LIDAR cloud base performs about as well as the CBASE cloud base on average, but does not provide an uncertainty estimate (and therefore no way to select only low-uncertainty cloud base estimates).

Data is available at DKRZ or PANGEA (need to look into which one is more appropriate). {j_μ: Obtain DOI, put it here.}

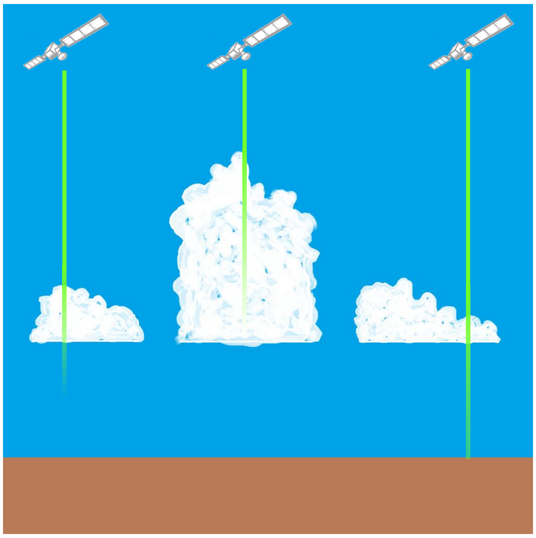

Figure 1. Schematic of CALIOP cloud base determination and evaluation strategy. {j_μ: Amazing figure by Christoph. The only way to make it even better would be to add a ceilometer.}

Table 1. CBASE cloud base statistics by decile of predicted uncertainty

pred.rmse	n	r	RMSE (m)	bias (m)	fit
(187,436]	2624	0.749	403.	−33.2	$y = 1.14x - 63.5$ m
(436,461]	2624	0.722	429.	−26.4	$y = 1.20x - 186.$ m
(461,478]	2624	0.725	460.	−26.3	$y = 1.25x - 256.$ m
(478,492]	2624	0.693	466.	−29.4	$y = 1.19x - 182.$ m
(492,506]	2635	0.616	512.	−0.0279	$y = 1.14x - 157.$ m
(506,517]	2721	0.564	550.	−12.7	$y = 1.14x - 157.$ m
(517,530]	2516	0.566	562.	−0.809	$y = 1.11x - 131.$ m
(530,550]	2624	0.566	569.	−29.9	$y = 1.15x - 150.$ m
(550,582]	2624	0.497	640.	−11.8	$y = 1.06x - 66.1$ m
(582,764]	2624	0.446	716.	13.1	$y = 0.977x + 14.6$ m

5 Conclusions

Summary. How useful is the cloud-base retrieval for various purposes? There are two main cases:
145 where accuracy is more important than efficiency, and the other way around. For surface radiative flux (hyperlinear in CBT), accuracy is more important. For geometric thickness, efficiency. For sub-adiabaticity, need to do the error propagation.

Acknowledgements. CALIOP VFM from ICARE. Computing (and data hosting?) at DKRZ. Funding.

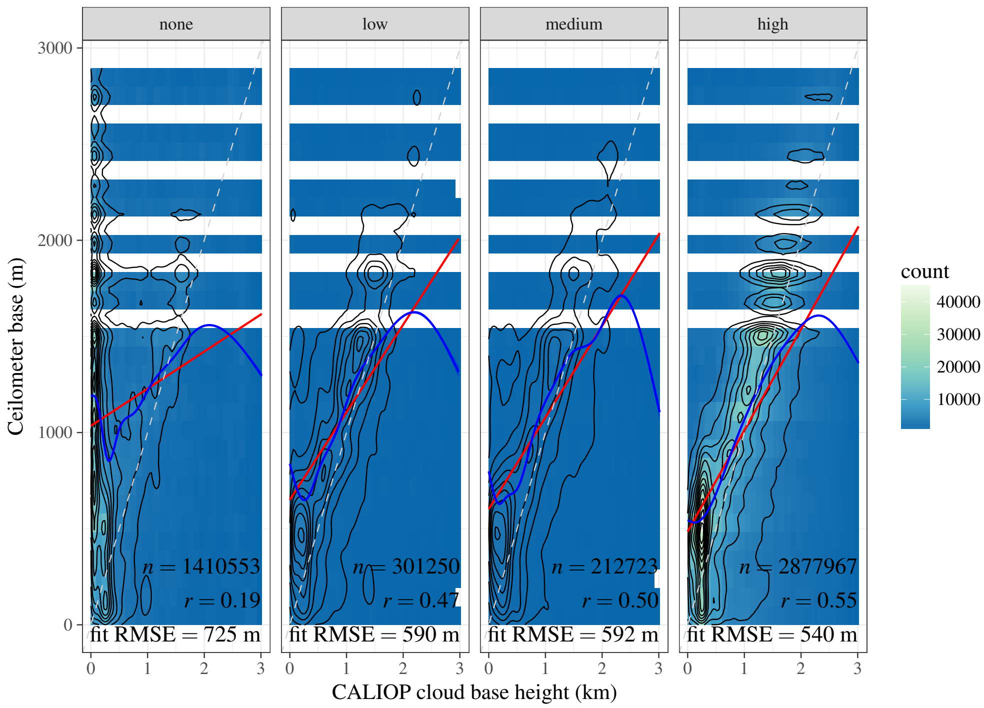

Figure 2. Scatter plots of CALIOP versus ceilometer local cloud base height faceted by the CALIOP VFM QA flag

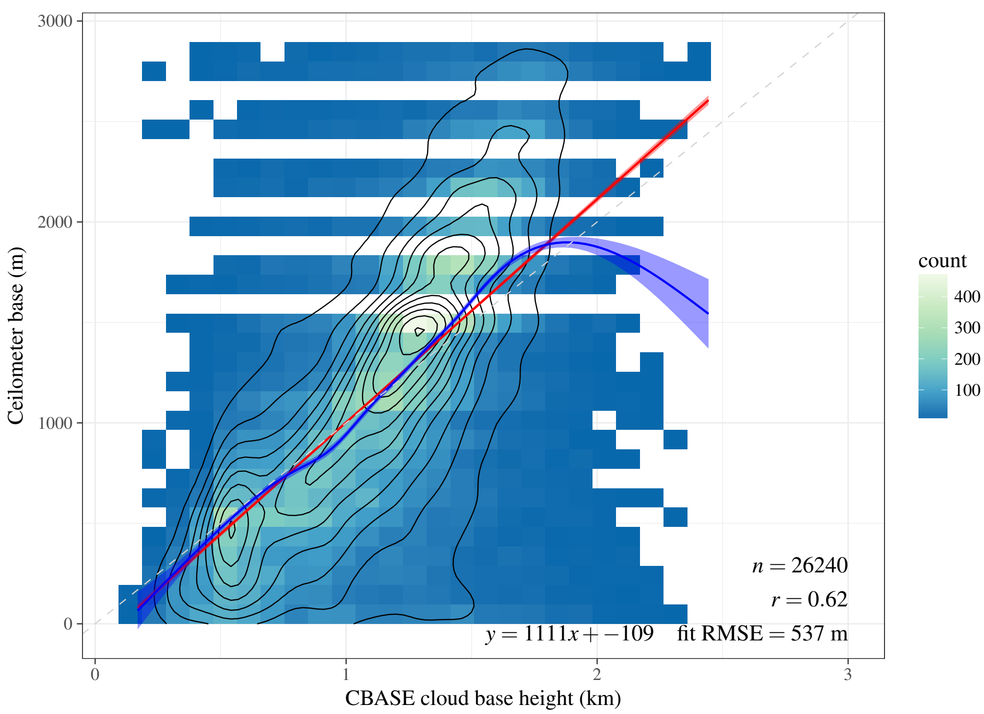

Figure 3. Scatter plot of CBASE versus ceilometer cloud base height

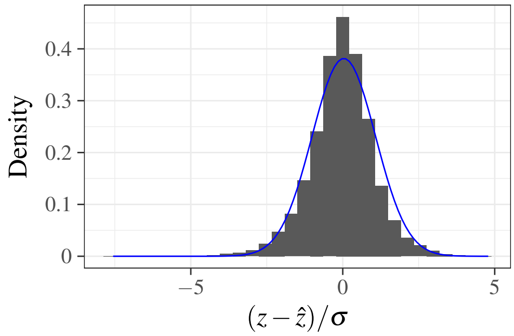

Figure 4. Distribution of cloud base error divided by predicted uncertainty; least-squares gaussian fit with mean 0.03 and standard deviation 1.05 is overlaid.

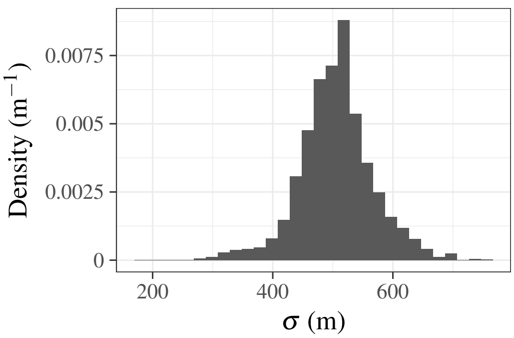

Figure 5. Distribution of predicted uncertainty

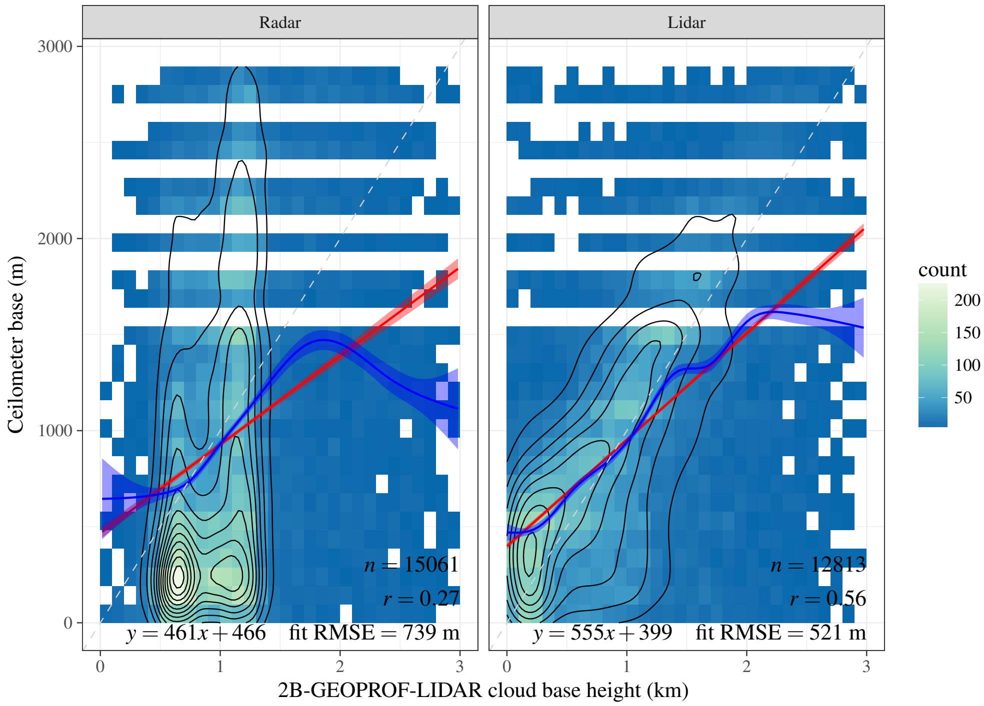

Figure 6. Scatter plot of 2B-GEOPROF-LIDAR versus ceilometer cloud base height. {j_μ: I *think* this uses Claudia's algorithm (minimum cloud base within $D_{\max}$), but I have to check.}

Figure 7. Geographic distribution of median cloud bases and median cloud thicknesses {j_μ: Make the figure less complex. One panel with overall CBH.}