Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2018-43-AC1, 2018

Interactive comment on "Using CALIOP to estimate cloud-field base height and its uncertainty: the Cloud Base Altitude Spatial Extrapolator (CBASE) algorithm and dataset" *by* Johannes Mülmenstädt et al.

Johannes Mülmenstädt et al.

johannes.muelmenstaedt@uni-leipzig.de

Received and published: 9 August 2018

The manuscript is well organized and well written, and the data product is of interest for the scientific community. This manuscript is suitable for publication in this journal. However, I have several comments that should be addressed before the manuscript can be published.

We thank the reviewer for his or her insightful reading of the manuscript and useful suggestions for improvements.

#1: Figure 3: I suggest to give explanations and to add discussions: a) What is CALIOP cloud base height in the X-axis? Is it the mean value of the N relevant local CALIOP zi? Please clarify.

There is an entry for each local CALIOP cloud base estimate within the collocation criteria (hence the large values of n in the table). We have included this missing information in the caption.

b) Please specify whether height is above sea level or above ground level.

All heights are above ground (to avoid spurious correlation between satellite and ceilometer cloud base due to terrain height). We have rephrased the first mention of this in the manuscript from "For this comparison, we use z above ground level (AGL); ..." to "Throughout this work, we use z above ground level (AGL); ..." to make it clearer that this statement applies to the entire manuscript. We have also included "AGL" in the axis labels in all figures.

c) I don't see the 95% confidence intervals shaded in light red and in light blue.

Thanks for pointing this out. Due to the essentially infinite A-Train statistics, the width of the confidence intervals is smaller than the line width. Since this type of plot appears several times and we refer to the first instance's caption for a full description, we would like to keep the full description here. We have added a note that the confidence intervals are narrower than the line width in this figure.

d) Looking at the contours of the joint probability density for instance in the right hand-side plot (high), it looks like the agreement between ceilometer and CALIOP cloud base heights for bases larger than 1.5 km is better than indicated by the red and blue lines. Can you comment?

The straight-line fit that tries to accommodate high and low cloud bases at the same time is problematic when there is curvature in the relationship between satellite and ceilometer cloud base. This is why we also include the GAM regression (blue line), whose slope and intercept need not be constant. The GAM regression is still below the 1-to-1 line at $z_c \approx 1.5$ km because there is a nonnegligible population of clouds with low ceilometer base and high CALIOP base (see the bulge in the outermost density contour near 0 ceilometer base height and 1.5 km CALIOP base height). These cases mostly occur when the ceilometer reports multiple layers of fractional cloudiness, where the satellite perspective may cause the algorithm to select one of the higher layers as the "base" because they obscure the lowermost layer. We are investigating whether the number of layers reported by 2B-GEOPROF-LIDAR can be used as another predictor variable for cloud base uncertainty as a future improvement to CBASE.

e) It looks like Fig. 3 (high) has been obtained before discarding the classes of CALIOP profiles listed page 6, lines 1 to 9. If this is correct, it would be very informative to show scatter plots as in Fig. 3 (high), but after discarding these profiles.

That is correct. At the point in the text where Figure 3 is discussed, the additional requirements have not yet been introduced. We agree with the reviewer that it is interesting to see the improvement in the scatter as the selection criteria on CALIOP columns are tightened. We have included this plot as a new Figure 4, discussed at the point in the manuscript when the selection criteria are introduced.

#2: Page 4, Eq. (1): why are the authors using a new notation "E" for RMSE?

We defined E out of a general aversion to multiletter variable names. However, since we do not use E again (and use "RMSE" as column heading in the tables), we agree with the reviewer that it would be preferable not to introduce new notation. We have changed E to RMSE in the equation.

#3: Page 6, lines 22-28: a) Please explain how CALIOP z is converted to "z Above Ground" (which reference for the elevation maps?).

We subtract the elevation provided in the CALIOP VFM data files from the MSL cloud height. As of V4.10 (the version used here), the CALIOP L2

products use the "CloudSat DEM", according to https://eosweb.larc.nasa.gov/news/caliop-v410-l1-l2-release-announcement, which contains reduced artifacts compared to GTOPO30. This DEM is also mentioned in the CloudSat R05 1A-AUX description (ftp://ftp.cira.colostate.edu/ftp/Partain/1A-AUX_PDICD_5.0.doc), but we could not find a reference for it. We note that DEM errors may spuriously increase the apparent disagreement between satellite and ceilometer cloud base, but this effect should be very small to the spurious correlation enhancement that would occur if we used MSL heights. We have added the following passage to the manuscript: "To convert cloud base heights to AGL height, we subtract the surface elevation contained in the CALIOP VFM data files, which in turn comes from the CloudSat R05 surface digital elevation model."

b) For more clarity, it would be useful to use different notations for "Above Sea Level z" and "Above Ground Level z", for both the ceilometer and CALIOP.

All heights are above ground (see our response above), so the "for this comparison" qualifier is misleading. We have changed it to "Throughout this manuscript".

c) I am not sure why the satellite z estimate is intrinsically biased high "due to this boundary". Do you mean that the technique requires the local CALIOP "Above Ground Level" zi to be positive?

Not the technique, but rather the physics, because there are no clouds below the surface. Consider the idealized case where the satellite provides an unbiased estimate of cloud base height with random error ϵ symmetric about 0. When $z \gg \epsilon$, the average of many satellite profiles will converge to the true cloud base height. However, when z is comparable to ϵ, the sample will preferentially contain cases where the random error is positive, because false detection of cloud below the surface is physically impossible. In that case, the average of many satellite profiles will not converge to the true cloud base height unless a correction is applied.

d) I could not figure out how these biases are corrected (lines 26-28). Please develop *and quantify. These bias corrections seem to be an important part of the algorithm training.*

We agree that the correction needs to be explained in greater detail. We have added a brief description of support vector machines in general and attempted to explain the correction method by analogy to a linear correction.

Quantification is nontrivial because the correction is a function of z_c, D, n, and Δz (and D, n, and Δz can be correlated). To reduce the dimensionality of this multivariate correction, we have used the training dataset (with its joint distribution of z_c, D, n, and Δz) to calculate an ensemble of correction factors that can be expected in a realistic sample of clouds. We have added discussion and a plot of these correction factors to the manuscript.

The correction is desirable because otherwise (a) the average over many CBASE estimates would not converge to the true cloud base height and (b) the bias would be a function of cloud base height, meaning a characterization of the bias based on one evaluation dataset would only apply to datasets with the same distribution of cloud base heights. These are both clearly undesirable features, which we considered important enough to avoid that we accepted the added complication of nonlinear correction functions.

#4: Page 6, lines 29-30: a) Do you confirm that you are introducing a new notation for RMSE, which is now "sigma"?

Yes, for the reason hopefully now better stated in the manuscript: “The quality metric we use is the root mean square error (RMSE); the category RMSE determined from comparison to ceilometer z then serves as the (sample) estimate of the predicted (population) standard deviation of the measurement error $z - \hat{z}$, i.e., the predicted column z uncertainty. We denote this uncertainty as σ_c.”

b) Can you elaborate? For instance: what is the range of values for sigma(D,n,Dz)?

What is the sensitivity of sigma to D, n, and Dz?

Due to the multivariate dependence on D, n, and Δz, this is hard to summarize succinctly. We have calculated one-dimensional projections of the full trivariate function onto each of the predictor variables, using the 2008 training dataset, similarly to what we have done in response to the question about the correction function. This additional figure is discussed in the revised manuscript.

In the R package that will be made public to accompany the final manuscript, the full predicted uncertainty table for each of the 125 categories in D, n, and Δz is available.

#5: Page 7, Equations 3 and 4: a) If I understand correctly, “z” in Eq. (3) is CBASE_z. Please clarify and use a specific notation for the different quantities. Indeed, “z” is used several times throughout the manuscript, but with different meanings.

Good point. We now consistently use z for the CBASE combined cloud base, z_c for the local column cloud base from CALIOP, and $\hat{z}$ for the ceilometer cloud base.

b) Please define “n” in Eq. (3).

Done; thanks for pointing out the oversight.

c) In Eq. (4), is sigma-i actually sigma-i(Di,ni,Dzi), with Di, ni, and Dzi associated to CALIOP local zi? How is ni defined for a local zi?

Thank you for pointing out this source of confusion, which we have corrected. The revised manuscript notes that D_i and Δz_i are indeed properties of the individual profiles, but n is a property of the group (and therefore carries no subscript).

d) Can you discuss the impact of the training? For instance: how do CBASE_z and CALIOP mean base height compare for the year 2008 used to train the algorithm? How did you train the algorithm to have mean (CBASE_z -ceilometer $\hat{z}$) equal to zero (as suggested by statement line 23, page 7)?

Unsurprisingly, the relationship between the CBASE cloud base and ceilometer cloud base in the training dataset satisfies the constraints we designed into it: slope 1 and intercept 0 for linear regression, bias 0 and width 1 for $(z - \hat{z})/\sigma$. As this is more of a check that we do not have software bugs in the implementation, we did not consider it important enough to include in the manuscript, but we have added it now: "To check that the algorithm satisfies its design constraints, we have verified that linear regression between z and $\hat{z}$ has zero intercept and unit slope, and that the quantity $(z - \hat{z})/\sigma$ has zero mean and unit standard deviation."

The real test is whether we overtrained the algorithm to fixate on peculiarities of the training dataset, which is the reason for testing on a statistically independent validation dataset, the results of which we discuss at length.

The removal of the bias is accomplished by be bias correction procedure; see our response to the question about the SVM-based correction method, where we have clarified the manuscript.

#6: Page 8, lines 12-18: My understanding is that the algorithm training and the verification presented in Sect. 3.4 using the 2007 data set have been carried out with no distinction between nighttime and daytime data. Did you investigate whether sigma-i (Di,ni,Dzi) are the same for nighttime and daytime data? I wonder whether the differences between the nighttime and daytime CBASE_z highlighted here could be due in part to the fact that the algorithm training combines daytime and nighttime data.

This is a very good point. While testing this hypothesis exceeds our resources at the moment, we have added discussion of this possibility to the manuscript: "Training a potential future update of the algorithm on daytime and nighttime profiles separately may reduce σ."

#7: Page 8, lines 19-25: a) Figure 9 seems out of place. In my opinion, this discussion could be earlier in the manuscript. However, comparisons of 2B-GEOPROF-LIDAR and CBASE base altitudes would be informative.

Agreed. We have added a subsection to the validation section that focuses on radar and lidar cloud base heights in 2B-GEOPROF-LIDAR compared to the same set of ceilometer measurements used to validate CBASE. We have also added a figure comparing the lidar-only 2B-GEOPROF-LIDAR and CBASE cloud bases explicitly. Unlike the CBASE cloud bases (which have been corrected), the 2B-GEOPROF-LIDAR lidar-only cloud bases are, on average, underestimates of the ceilometer bases for very low clouds (< 0.5 km). The comparison of 2B-GEOPROF-LIDAR and CBASE similarly shows lower cloud base estimates by 2B-GEOPROF-LIDAR for very low clouds, and higher cloud base estimates for clouds above 1.5 km. The linear correlation between 2B-GEOPROF-LIDAR and CBASE is fairly good ($r = 0.79$).

b) Please describe the "underlying physical measurement" in 2B-GEOPROF-LIDAR that explain the similarity of lidar-only 2B-GEOPROF-LIDAR and CBASE cloud bases.

We simply meant that physically, both products are based on CALIOP attenuated backscatter, so the added value in CBASE comes from understanding the factors controlling profile-by-profile uncertainty. We have clarified the manuscript by explicitly stating this assumption.

c) Are you implying that for lidar-only cases, cloud bases reported in 2B-GEOPROF-LIDAR differ from those reported in the V4.10 CALIOP VFM and are in better agreement with CBASE? Does 2B-GEOPROF- LIDAR use V4.10 CALIOP data?

We did not intend to imply that; in fact, now that the reviewer has pointed it out, we realize we implicitly assumed that the 2B-GEOPROF-LIDAR and V4.10 CALIOP VFM cloud bases would be the same.

The version of 2B-GEOPROF-LIDAR we use (P2_R04_E02) uses VFM version 3 "or later", according to Mace and Zhang (2014), but the data files predate version 4, so presumably version 3.x was used. We have specified the 2B-GEOPROF-LIDAR version in the revised manuscript. Characterizing the differences between 2B-GEOPROF-LIDAR and the VFM would be outside the scope of this manuscript.

#8: Page 9, lines 20-21: did the authors investigate how CBASE_z and CALIOP base altitude compare for clouds that are thick enough to attenuate the lidar laser beam?

At the suggestion of the reviewer, we have included and discussed this figure in the revised manuscript (Figure 10); see the answer above to the question about 2B-GEOPROF-LIDAR and CBASE base altitudes would be informative.

#9 Page 9, line 27: was stated but not shown.

The statement in question is "The performance of CBASE z is similar to that of 2B-GEOPROF-LIDAR lidar-only z . . . ". We have clarified our intended meaning by adding "when validated against the same collocated ceilometer measurements". However, we agree with the reviewer's suggestion to include an explicit comparison of CBASE and lidar-only 2B-GEOPROF-LIDAR (see our answers above).

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2018-43, 2018.