

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.

Anonymous Referee #1

Received and published: 22 May 2018

A new technique is presented to retrieve water clouds base heights in the lower troposphere (base below about 3 km). The technique uses CALIOP base heights reported in the V4.10 VFM product and base heights from ground-based laser ceilometers that are part of the Automated Surface Observing Stations in the United States of America. The algorithm is trained using collocated (< 100 km and < 1 h) water clouds base heights from the CALIOP satellite and from the reference ground-based ceilometers in the year 2008 under selected conditions. The algorithm supplies estimated uncertainties shown to be in some regions smaller than currently assumed uncertainties,

C1

thereby offering the possibility to reduce the uncertainty in the computation of surface downwelling longwave radiation. Results as well as the data products are presented. 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.

#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 z_i ? Please clarify. b) Please specify whether height is above sea level or above ground level. c) I don't see the 95% confidence intervals shaded in light red and in light blue. 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? 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.

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

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

#4: Page 6, lines 29-30: a) Do you confirm that you are introducing a new notation for

C2

RMSE, which is now "sigma"? 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?

#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. b) Please define "n" in Eq. (3). c) In Eq. (4), is σ_i actually $\sigma_i(D_i,n_i,Dz_i)$, with D_i , n_i , and Dz_i associated to CALIOP local z_i ? How is n_i defined for a local z_i ? 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 z_{EE}) equal to zero (as suggested by statement line 23, page 7)?

#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 $\sigma_i(D_i,n_i,Dz_i)$ 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.

#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. 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. 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?

#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?

C3

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

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

C4