Review of Evaluating daytime planetary boundary-layer height estimations resolved by both active and passive remote sensing instruments during the CHEESEHEAD19 field campaign
Anonymous Referee #1

Within this manuscript, the authors compare estimates of the convective planetary boundary layer height (PBLH) made from several different remote sensors with radiosonde-derived estimates of the PBLH. The authors discuss key differences in the ability of each remote sensor (or system) to accurately ascertain the PBLH, using the radiosonde estimates as truth. When significant differences are apparent, the authors discuss possible reasons appropriately. This includes a statistical comparison as well as a few case days that were examined in closer detail.

Overall, this manuscript is fairly well written and of interest to the Atmospheric Measurement Techniques reader community. While largely developed PBLH algorithms are used, the analysis presented herein further assess the strengths and limitations of each sensor and associated method to determine the PBLH, including new methods that have not been previously thoroughly evaluated. As such, this manuscript is acceptable to AMT pending minor revisions in which the following comments are adequately addressed.

Specific Comments

- Title: Suggest changing name from ‘Evaluating daytime planetary ...’ to ‘Evaluating convective planetary’... given the focus is on convective PBLH estimates with the parcel method, which would be inappropriate for stable PBLH estimates.
- Line 24 & 62: Should this be ‘Collaborative Lower Atmospheric Mobile Profiling Systems’? That is the name given for it on the NSSL website and seems more apt.
- Line 177: While radiosondes are used as the truthing dataset here and are treated as the ‘gold standard’ (which is fine), there are still limitations in radiosondes for use of
determining the PBLH. Most notably is that radiosondes provide a nearly instantaneous measurement and are only representative of the exact location it transited the BL / free troposphere interface. Thus, if the radiosonde transited this interface at a downdraft, the local PBLH estimate may be slightly low biased compared to the area-averaged PBLH. Conversely, if an updraft is present the local PBLH would be slightly displaced upward compared to the area averaged PBLH. Over a large number of profiles, these effects should average out thus not leading to a bias. Still, it may lead to some of the significant scatter seen in the comparison plots later on.

- Line 187 (and the entire manuscript): Highly recommend removing the radiosondes at 6 LT from the analysis. As the authors clearly state here, the parcel method is suited for convective conditions and the PBL is rarely convective at 6 LT, especially during the fall when sunrise is later.
- Line 221: Is 'TROPoe' an acronym for something?
- Line 254: Is there evidence that supports that inclusion of RASS or model data in the TROPoe retrieval does not improve PBLH estimates? A short statement on the impact would suffice, or an appendix if additional analysis is warranted.
- Line 269: What range gate size was used for the Doppler lidar measurements? This detail can help the reader interpret the results given the tradeoff between reduced range/sensitivity (for a shorter gate) and reduced ability to resolved small turbulent eddies (for a longer gate).
- Line 319: It would be helpful to give a brief 1-2 sentence description of the Mues et al (2017) algorithm to summarize how it work. How is each 16-s PBLH estimate established independently from the backscatter profile?
- Figure 5: Perhaps I’m misunderstanding something, but the BL-View 1-h mean PBLH doesn’t look like a mean of the individual 16-s PBLH estimates. This is particularly apparent between 7-8 LT at the Lakeland site, where most of the 16-s measurements show the BLH being around 400 m whereas the 1-h mean is at about 1 km. Please explain.
- Figure 6 discussion: Why does it look like there’s a high bias in the QC-scaled PBLHs that scales with the PBLH itself? This is particularly apparent at the Lakeland site, wherein many of the QC-scaled PBLH measurements are above the 1-1 line when the PBLH is greater than 2.5 km.
- Line 393: While I understand having an independent ‘expert’ dataset is useful for assessing the PBLH, the estimates are subjective. I myself would place the PBLHs differently in Fig.8, as the clouds between 10-15 LT appear to be cumulus clouds (could be verified with visible satellite imagery) meaning the PBLH would be higher than currently indicated and at least at the cloud base. One way to get around this subjectivity would be to have each coauthor provide independent estimates of the PBLH each hour that could be averaged together. Given these estimates are for only 2 IOPs (each a week long), this shouldn’t be a lot of work for each coauthor.
- Figure 9 (and elsewhere): It would be useful to provide the equation of the best-fit line for each subpanel. I assume the grey line shows this (it’s unclear since it’s not explained in the caption).
- Line 437: What exactly is meant by ‘time-matched”? It’s unclear how these comparisons are different from those that are not time-matched.
- Lines 527-553: This is a very long paragraph. Suggest breaking up into at least 2, perhaps 3 paragraphs for readability.
- Line 597: For this profile at 13:00, the PBLH in the radiosonde profile is relatively ambiguous compared to other times (on this day and for the summer case). This is due to the fact that there is a slightly stable layer between 1.5 and 2.5 km. Thus, small changes in measurements of the surface temperature (due to instantaneously measuring warm or cold anomalies) as well as use of the 0.5 K addition in the PBLH can make significant differences in the actual parcel-based PBLH estimate. This should be discussed.
Technical Corrections:

- Lines 22-23: These instruments should not be capitalized.
- Line 75: i.e. should be followed by a comma (i.e.,)
- Line 81: SNR should be defined as an acronym here where it first appears, not at line 90.
- Line 202: Should the -1 should be superscript?
- Line 499: Should CLAMPS be plural here (CLAMPSs)?