Comment on amt-2021-363

Anonymous Referee #3

Referee comment on "Evaluating daytime planetary boundary-layer height estimations resolved by both active and passive remote sensing instruments during the CHEESEHEAD19 field campaign" by James B. Duncan Jr. et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2021-363-RC2, 2022

The current manuscript about the Planetary Boundary Layer Height (PBLH) analyzes and compares retrievals from different methods and instruments during the CHEESEHEAD19 field campaign. This subject is within the agenda of AMT and is of high interest for the scientific community, since it is not common to have this number of instruments in a close range. The work focuses on the differences between methods and instruments and validates retrievals using collocated radiosondes for reference. Case studies of days with different cloud conditions offer a deeper insight of the inconsistencies between the retrievals. The manuscript is well written and all major issues of the methods and the results are discussed, hence I suggest to be accepted for publication after minor revisions.

We thank the Referee for the positive and constructive comments. We hope we have addressed all of the Referee's concerns and we think that our manuscript did benefit from the constructive comments made by all Referees.

Specific comments:

Introduction: I think some literature should be added, considering comparisons of retrievals from the instruments used in this study. Also, some discussion is expected about the different definitions of PBL and the known differences among the retrievals based on the variable in study.

As suggested, some literature and discussion has been added to the Introduction.

Section 2.1 Some discussion about the problems/ errors/uncertainties of the radiosondes retrievals should be added.

According to this Referee's and to Referee #1's comments, additional discussion on the limitations in the use of radiosondes for determining the PBLH has been added to the revised manuscript (Section 2.1, before the description of the various radiosonde methodologies): *"We note that limitations in using radiosondes for determining the PBLH include that they provide nearly instantaneous measurements and are only representative of the exact location transited at the interface between the boundary layer and the free troposphere above. Moreover, if the radiosonde transited a downdraft, the local PBLH estimate may be biased slightly low compared to the area-averaged PBLH. Conversely, if an updraft is present, the local PBLH estimate would be displaced slightly upward compared to the area-averaged PBLH. However, given the large number of profiles used in this study, the impacts of updrafts and downdrafts on PBLH estimate*

should average out and not lead to a bias; albeit, this may have contributed to some of the scatter in the comparison plots presented later on."

Figure 2. The errorbars and the outliers should be described at the caption. I am in doubt that this representation of the variation of each method is the most adequate, because ranges of more than 2km for PBLH can include all possible values. Probably a visualization of synchronous values would be more appropriate. Also, somewhere earlier in the manuscript, the sunrise/sunset LT for the 7 day IOP, in order to understand the low values at 6.00LT. If 6.00 LT is before sunrise or even shortly after, the parcel method is not applicable, since it is referring to convective conditions.

Description of the error bars used in the figure has been added to the caption of Fig. 2: *"The boxes show the interquartile range with the median indicated by the horizontal line and the whiskers extend to points that lie within 1.5 times the interquartile range of the lower and upper quartiles"*.
Also, according to all Referees' comments, we removed the radiosondes at 06:00 LT from the analysis, and we included sunrise and sunset times for the 2 IOPs in the manuscript as suggested by the Referee. At the beginning of Section 2.1 we included the text *"Since sunrise for 19-24 August and 23-28 September is approximately at 05:20 and 06:00 LT, respectively, the 06:00 LT sounding was not included in the analysis as a convective PBL would not yet have been present in the remote sensing observations. Sunset for the 19-24 August and 23-28 September is at around 19:00 and 18:00 LT, respectively."* As a consequence of this change, all figures have been reproduced and all statistical results have been updated in the revised version of the manuscript.
Since 6:00 LT was the time when larger differences in the number of available PBHL estimations between the different radiosonde methods were observed in Fig. 2, and since for the other radiosonde launch times the available number of PBLH estimations is pretty similar among the methods, we believe that our representation is now adequate.

Figure 4. The peak around 2.00LT should be discussed in the corresponding paragraph. It appears a variation in wind conditions during this time, that leads the algorithm to recognize a stratification at higher height.

Since the peak mentioned by the Referee happens at nighttime, in stable conditions, shear-induced turbulence could have caused increased values of vertical velocity variance, which is mainly used to determine the PBLH estimate during nighttime. As a matter of fact, the vertical velocity plot indicates stronger downdrafts at around this time which might be related to a vertical mixing process.

L323. The description of the method of selecting value based on the score is not described clearly. How the criterion of ±200m came up?

As suggested by the Referee we expanded this section describing the differences between our approach and the Mues et al. (2017) approach, providing more details. Also, the 200 m range restriction was introduced by Mues at al. (2017) to avoid unrealistic results in two-layer situations, as explained in their paper.

L330. The BL software provides the higher value of 4km in many cases of 16s retrievals, is there any physical explanation, considering the atmospheric conditions, for this result?

We have to clarify that not all hours when the cyan diamonds are missing correspond to PBLHs estimated from the BL-View having a higher value of 4 km agl. We made this now clearer in the revised text: *"The BL-View software has an upper limit for PBLH values of 4 km agl. Additionally, due to the proprietary QC methods imposed within the BL-View software, there are some hours when no PBLH value is provided."* Nevertheless, since the BL-View software is proprietary, we are not completely aware of how it operates, therefore we are not able to explain why it provides higher values of the QC-scaled approach for some hours (i.e., 14:30 LT at both sites). However, when all the days were visually inspected we noticed that when these (rare) differences occurred, the QC-scaled approach appeared to provide more reasonable PBLH estimates than the hourly BL-View output, so we were comfortable in employing the QC-scaled approach (which we had control on) in our study.

L350-355 The different response of the comparison between BL and QC for the two sites should be discussed. Is there some local or systematic effect that could explain the worst statistics for Lakeland?

Unfortunately, we are unable to pinpoint the reason for the different statistical comparisons between the BL-View and QC-scaled approach at the Lakeland and Prentice sites due to either local or systematic effects. This consideration was added to the text of the revised manuscript.

L390 I think the idea of an independent dataset selected manually by visual inspection of the recordings can be a valid reference for evaluating the retrievals. Why don't include more data from other instruments for creating this reference databases, specially in cases of sharp gradients?

We agree with the Referee that a synergetic use of the instruments presented in this study would be the optimal way to achieve accurate PBLHs, since the strengths of the various platforms would be combined together. We nevertheless believe that this would be a further step in the study and we would rather postpone it to a future analysis, as it would undoubtedly become rather complicated and require a lengthy description.

L420-425. I cannot see radiation information been used, only the cloud information. Please restate or explain. Also, the cloud fraction is the cause of different development of convective Boundary Layer, but the result is not immediate, since some time is needed to propagate the effect to the layer. Hence I suggest to investigate the possibility of correlating the with cloud fractions in a wider time window. More specifically a window including the previous time steps will correlate better due the delay response.

Regarding the cloud fraction information, we have clarified its use in the revised version of the manuscript rewording the text to *"Two RadSys stations were deployed at Lakeland and Prentice (data available from Riihimaki, et al., 2020a, b) providing complete surface irradiance measurements, which then enable RadFlux analysis (Long and Ackerman 2000; Long et al.*

2006) that produces derived cloud variables such as cloud fraction. This cloud fraction information was used to expand the PBLH evaluation under different cloud coverage conditions."

We agree with the Referee that the impact of the clouds on the convective boundary layer development can indeed be delayed, but we also think that it is difficult to establish what optimal time window to use that would be overall adequate for different cloud heights, different cloud types and cloud fractions, and for different cloud advections. Therefore, we decided to keep the window used to determine the cloud fraction as the same temporal window used to determine the PBLH estimates, but we included some text in the revised version of the manuscript mentioning the fact that this will indeed neglect taking into account the delayed impact of the clouds on the development of the convective boundary layer: *"The hour window used to determine the cloud fraction is the same temporal window (±30 min of the hour) used to determine the PBLH estimates. This neglects any possible delays in the impact of clouds on the development of the convective boundary layer, which is however difficult to quantify."*

Figure 9, The caption should explain what are the black and grey lines.

The wording *"Black lines represent the 1-to-1 line, and grey lines represent best-fit line"* has been included in all figure captions, when necessary (Figs. 6, 9, 11, 12, and 13). Also, the best-fit line equations have been included inside the panels.