2nd Review of « Confronting the Boundary Layer Data Gap : Evaluating New and Existing Methologies of Probing the Lower Atmospher » by Bell et al.

by H. Luce (Toulon University)

After some necessary corrections, the authors improved the manuscript. I have still two important comments to make (that do not prevent its acceptation).

(1) **About the purpose of the manuscript** (comment 1). If the objective is to evaluate the performance of the Coptersonde with the available sensors onboard (IMET and IST sensors) for operational use, it makes sense. If not, considering the equipped Coptersonde as a whole, without clarifying the contributions of the sensors and Coptersonde separately provides lower value to this study. I think that the response given by the authors below should be rephrased and included in the text.

> *You are correct that the Greene et al. (2018, 2019) studies have worked to mitigate the effect of the UAS on the TU measurements, which would decouple the UAS technique from the sensors themselves. However, the recommendations made from these studies were done with fairly idealized setups. The purpose of continuing to evaluate the CopterSonde as a whole is to verify the results from the idealized situations while in a more "operational" mode, especially against a radiosonde that is as widely used as the RS92SGP.*

(2) **About outliers**. In statistics, an outlier is a data point that differs significantly from other observations. Thus, in principle, they may have their own distribution. Considering data points associated with the tail of a distribution as outliers is not necessarily true. They can be outliers or not because the threshold is necessarily arbitrary. Rejecting doubtful data when they differ two much is necessary for scientific works (ie, when physical processes are studied), but is questionable when technical performance is studied. These doubtful data (or, at least, some of them) are part of the dataset to analyze. If their number is very small, if their origin is well identified and if they are expected not to occur again in operational use, then they can be ignored (but, it is the case here?). Otherwise, their characteristics should be specified. With the rejection method applied by the authors, it is not possible to know if the rejected data occur sporadically or if it is a whole "package" of consecutive points. The consequences of the rejection process would not be the same.

I understand that this topic is not easy to tackle, but I think that it should be expanded a little bit by including some examples of rejected data in profiles and scatter plots and discussing the impacts of these rejected data on the statistics (e.g do they affect or not the mean differences, do they introduce biases, etc..)?

Minor comments:

1) In abstract, (and elsewhere in the text) "dew point"-> "dew point temperature"

2) Page 2, lines 30: the references Lawrence and Balsley (2013, https://doi.org/10.1175/JTECH-D-12-00089.1, Balsley et al. 2013, DOI 10.1007/s10546-012-9774-x) are in the reference list, but not in the text.

3) Page 7 table 1, caption: "flights.xs" ?