Earth Syst. Sci. Data Discuss., referee comment RC1
https://doi.org/10.5194/essd-2022-150-RC1, 2022

Comment on essd-2022-150

Anonymous Referee #1

Referee comment on "A merged continental planetary boundary layer height dataset based on high-resolution radiosonde measurements, ERA5 reanalysis, and GLDAS" by Jianping Guo et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2022-150-RC1, 2022

General comments:

This study developed a state-of-the-science method to derive a global-wide PBLH dataset merging in situ observations and reanalysis dataset, which has optimized the performance of a so-called "data fusion" technology and provided critical data for climate research. There are no obvious flaws in the methodology, and the final output is informative enough to compensate for the disadvantages of current atmospheric datasets existing as the spatial-temporal discrepancy. Despite the good structure and comprehensive analysis, the authors are required to answer or address the following questions or comments. After that, I think this manuscript can be accepted for publication.

Specific comments:

- Line 123: It is suggested that the authors explain a little bit more of the relationship by a gradient of terrain or lower-tropospheric stability induced underestimation of the PBLH.
- The title of the paper is ''…ERA5 reanalysis, and GLDAS''. However, GLDAS didn't occur until the last paragraph. It is suggested that the authors can add some descriptions of GLDAS.
- Line 154: please clarify if the interpolation is based on altitude or elevation.
- Line 158: It seems to me not correct to say spatially even coverage. The coverage in Australia is substantially not even especially in Figure 1d.
- Line 173: Any reference for the definition of LST?
- Line 207: how did the authors match the stational PBLH and gridded PBLH in the comparison?
- Line 259: Please specify clearly if all the data from 2011-2021 were included in the model training stage. Were they divided by the measuring time (e.g., 0000, 0006…)?
- A simple question: What is the merit of ~100/200 m improvement of PBLH (compared

with the raw method) considering the future application of this dataset? Any impacts on climate-scale studies?

Technical corrections:

- Line 99 and 116: the definition of ERA-5 should be moved ahead.
- Please keep it consistent by using either ERA5 or ERA-5 in the whole manuscript.
- Line 233: in the main text, the authors mentioned that Table 2 shows the correlation coefficients between PBLH and each variable, but the caption of Table 2 says that it is a correlation coefficient with PBLH bias between radiosonde and ERA5 reanalysis, which is easy to be misinterpreted. Please address.
- Line 242, please use subscripts or other notations to mark PBLH-M and PBLH-E in the equation. Otherwise, it will be easy to be recognized as a minus.